Prepared for submission to JHEP

Groups and Representations Tutorial Class Supplementary Notes

Lucas T. Y. Leung[a]

[a] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK*

E-mail: lucas.leung@physics.ox.ac.uk

Abstract: This set of notes is my attempt on summarising and expanding on the topics and discussion points made during the tutorial classes. These only act as a set of supplementary notes and are not designed to replace the course notes or as solutions to the problem sheets. You should not share these notes without permission.

Contents

1 General References, Articles and Books

The canonical reference for the course is Fulton and Harris' *Representation Theory* [1]. This is a really good book for both mathematicians and theoretical physicists and it provides a rigorous treatment to all the topics covered in the course. It might be, however, useful to look elsewhere if this is a bit too mathematically rigorous for you.

But just in case you need more mathematical rigour — here are some books mainly catering to mathematicians. For general algebra Dummit and Foote is a standard [2]. For finite groups it might be helpful to look at James and Liebeck's 'Representations and Characters of Groups' [3]. For Lie algebras and Lie groups, have a look at Humphrey, Erdmann or Hall [4–6]. For some differential geometry background, Lee's 'Introduction to Smooth Manifolds' [7] is a very good introduction and reference.

For physicist-oriented books, the best book is Georgi's 'Lie Algebras in Particle Physics' [8]. I didn't like it when I was doing the course [1], but I later found out it is one of the best books on the subject! Fuchs and Schweigert's 'Symmetries, Lie algberas and Representations' [9] offers a comprehensive rigorous view on the subject which also has a lot of interesting topics in it. Finally Costa and Fogli's book is quite good as well [10] — I did use it for quite a bit when I was learning the subject.

Throughout the notes I will cite the relevant references and articles. You should go look at them if you want a bit more information on the topics I have selected here. My presentation is nowhere coherent or formal enough — it is just the kind of style I like [2]. I apologise if there are any typos/mistakes in the notes but do email me if you spot them!

[1] Well, okay, the Symmetries, Particles and Fields course in Cambridge.

[2] It's quite formal for most theoretical physicists. Sorry. I just like my weird sense of mathematical and physical intuition so you will find comments dotted around... however, I must say that the current version of notes is a lot more formal than most presentations you will find aimed at physicists.

2 Elements of basic group theory

In the first class we mainly discussed finite groups and their representations. The main references here are [1, 2]. In this section, I would like to highlight some important topics that you might not have encountered as part of your undergraduate course, namely the subgroup test and some facts about permutation groups.

2.1 Subgroup test

When marking I noticed a lot of you tried to show $H \leq G$ (H is a subgroup of G) by showing all the axioms, namely: Closure, Associativity, Identity and Inverse [3]. This is generally not too helpful when you are trying to show a number of groups as subgroups of a larger group - you will be doing the same tedious arguments again and again! So I want to introduce a test here that will shorten the argument by a bit [11].

Theorem 2.1 (Subgroup Test)**.** *Let H be a non-empty subset of a group G. Then H is a subgroup if and only if $\forall h_1, h_2 \in H$, $h_1 h_2^{-1} \in H$.*

Proof. The forward direction is obvious (from group axioms, check!). So let us show the converse. Assume $\forall h_1, h_2 \in H$, $h_1 h_2^{-1} \in H$. We see that $e \in H$, since $hh^{-1} \in H$. This also shows that $h^{-1} \in H$. Now take $h_2^{-1} \in H$, so $h_1(h_2^{-1})^{-1} \in H$, $\forall h_1, h_2 \in H$. This means H is closed under the multiplication inherited from G. Associativity clearly descends from G, so all axioms are satisfied. □

In the exam I would clearly state I am using the **subgroup test** if I am using this theorem — this is a trivial group theoretical result that you probably do not need to provide a proof for.

2.2 Permutation groups

In Q3 of the first problem sheet we were asked about the conjugacy classes of S_3. However, a lot of you simply explicitly calculated that these are in fact the only conjugacy classes of the group - this is doable for small finite groups, but in general not feasible for larger groups. Some people stated the fact that the conjugacy classes are all defined by the cycle type, but without proof. Here I want to show that this is true in general for S_n. First with the following lemma.

Lemma 2.1. *Let $\alpha, \beta \in S_n$, α a k-th cycle $(a_1 ... a_k)$. Then*

$$\beta\alpha\beta^{-1} = (\beta(a_1) ... \beta(a_k)) \,, \tag{2.1}$$

i.e. conjugation does not change the cycle type.

Proof. Consider $\beta(a_i)$, then $\beta^{-1}\beta(a_i) = a_i$, $\alpha(a_i) = a_{i+1 \text{ mod } k}$, and

$$\beta\alpha\beta^{-1}(\beta(a_i)) = (\beta(a_{i+1 \text{ mod } k})) \,. \tag{2.2}$$

[3]Kit: By the way, it will be a great idea if you name these during an exam!

Take $b_j \neq a_i$, $\forall i$. Then $\alpha(b_j) = b_j$ and

$$\beta\alpha\beta^{-1}(\beta(b_j)) = \beta(b_j) \; . \tag{2.3}$$

So $\beta\alpha\beta^{-1}$ fixes any number not in the form $\beta(a_i)$ for some i. Therefore, we deduce that

$$\beta\alpha\beta^{-1} = (\beta(a_1)...\beta(a_k)) \; . \tag{2.4}$$

□

This lemma tells us that the conjugacy classes are determined by cycle type. In fact, the cycle type uniquely determines the conjugacy classes. To see this we need:

Theorem 2.2. *Every permutation is uniquely expressible as a product of disjoint cycles.*

Proof. Given $g \in S_n$, the cyclic subgroup $\langle g \rangle \leq S_n$ generated by g acts on some set $X = \{1, ..., n\}$ decomposes X into distinct orbits:

$$O_x = \left\{ g^i x \mid i \in \mathbb{Z} \right\} \tag{2.5}$$

for choices of orbit representations of $x \in X$. For each x, let N_x be the order of g when restricted to the orbit $\langle g \rangle \cdot x$ and define the cycle

$$C_x = \left(x, gx, ..., g^{N_x - 1} x \right) \; . \tag{2.6}$$

Since distinct orbits are disjoint, the cycles are disjoint. Given $y \in X$, we choose an orbit representation x such that $y \in \langle g \rangle \cdot x$. Then $g \cdot y = C_x y$ and g is the product of the cycles C_x over the orbit representations x. □

So now we have the theorem:

Theorem 2.3. *The conjugacy classes of any S_n are uniquely determined by cycle type.*

Proof. This follows immediately from Theorem 2.2 but now we are considering a product of cycles (α in Theorem 2.2 is a single cycle)[4]. □

We can think of this as follows - the conjugation simply relabels the numbers inside the cycle when the element $\sigma \in S_n$ is written in cycle notation. So the cycle type is just the conjugacy class!

[4] If you want it more fleshed out here are the details: Write $\sigma = \alpha_1 ... \alpha_l$ for $\sigma \in S_n$. Then for $\tau \in S_n$, we have

$$\tau\sigma\tau^{-1} = (\tau\alpha_1\tau^{-1})...(\tau\alpha_l\tau^{-1}) \tag{2.7}$$

From Lemma 2.1, we have that $\tau\alpha_i\tau^{-1}$ also a k_i-cycle. Now for $i \neq j$, α_i and α_j are disjoint and $\tau\alpha_i\tau^{-1}$ and $\tau\alpha_j\tau^{-1}$ are also disjoint since τ is injective. Therefore, $\tau\sigma\tau^{-1}$ are also separated into disjoint cycles. To show uniqueness, we note that Theorem 2.2 states that every permutation can be uniquely described as a product of disjoint cycles. Since this cycle type is unchanged under conjugation, the conjugacy classes therefore uniquely correspond to all possible cycle types constructible in S_n.

3 More on Group Algebras

In the lectures **group algebras** are defined in a really concise manner - it might be difficult to really understand what it is though! Here I would like to highlight some facts about group algebras and how they are constructed. In fact, this will also allow us to define representations and present the proofs in a more concise manner.

This section is here mainly for your reference (and for your mathematical satisfaction)! Therefore, this is entirely out-of-syllabus.

3.1 Groups algebras - Definitions

First recall the definition in the lecture.

Definition 3.1. The **group algebra** A_G of group G is defined as the set of formal linear combinations:

$$v = \sum_{g \in G} v(g) g \tag{3.1}$$

where $v(g) \in \mathbb{C}$ all $g \in G$ are considered to be linearly independent.

This definition might appear to be a bit confusing - but if you look at it again this is actually not too bad to work with. The group algebra is effectively like 'a vector space' with the group elements, labelled as $\{e_g\}$, being the basis for this vector space. Let us make this notion a bit more precise by recalling some definitions from elementary algebra[5]:

Definition 3.2. A **ring** R is a set with two binary operations (addition and multiplication) such that the following statements are true.

(i) R is an abelian group with respect to addition.
(ii) Multiplication is associative and distributive over addition. i.e.

$$(xy)z = x(yz) \tag{3.2}$$

$$x(y+z) = xy + xz \;, \quad (y+z)x = yx + zx \tag{3.3}$$

(iii) For all $x, y \in R$, $xy = yx$.
(iv) There is an identity element $\exists 1 \in R$ such that for all $x \in R$, $x1 = 1x = x$. This identity element is unique.

Definition 3.3. Let R be a ring. An R**-module** is an abelian group M on which R acts linearly. More precisely, it is the pair (M, μ) where M is an abelian group and $\mu : A \times M \to M$ is a mapping such that (writing $\mu(r, m)$ as rm for $r \in R$, $m \in M$) the following axioms are satisfied.

(i) $r(x + y) = rx + ry$.
(ii) $(r + s)x = rx + sx$.
(iii) $(rs)x = r(sx)$.

[5] I say elementary - this might not be if you haven't seen these notions before! One can consult a basic algebra book for details, say [2, 11].

(iv) $1x = x$.

A module is kind of like a vector space using rings as the 'scalar'. If $R = \mathbb{Z}$ then the conditions are immediately satisfied — the $\mathbb{Z}$-modules are the same as abelian groups. If R is a field $\mathbb{K}$ [6], then an R-module is just a $\mathbb{K}$-vector space. Modules are therefore generalisations of vector spaces when we work over an arbitrary ring. We also note the following equivalent definition of a module (which will become important later).

Definition 3.4. An **R-module** M is an abelian group together with a ring homomorphism $R \to \mathrm{End}(M)$ where $\mathrm{End}(M)$ is the ring of endomorphisms of the abelian group M.

So why do we need to know something about modules? It turns out that group algebra, or **$\mathbb{F}G$-modules**, turns out to be representations of G. This requires as bit more explanation. Let's start with a definition of $\mathbb{F}G$-modules.

Definition 3.5. Let R be a commutative ring [7]. An **associative R-algebra** (or an R-algebra simply) is a ring S that is also an R-module such that the two additions (ring and module) are the same operation and scalar multiplication satisfies:

$$r \cdot (xy) = (r \cdot x)y = x(r \cdot y) \ . \tag{3.4}$$

Equivalently, let $f : R \to S$ be a ring homomorphism. Define a product

$$r \cdot s \equiv f(r)s \ , \tag{3.5}$$

with $r \in R$ and $s \in S$. This definition of scalar multiplication turns S into an R-module. The ring S equipped with this R-module structure is then an **R-algebra**.

There are a few more facts on module theory in Appendix A. In particular, we will need the following definition.AN

Definition 3.6. A **free R-module** is one which is isomorphic to an R-module of the form

$$\bigoplus_{i \in I} M_i \ , \tag{3.6}$$

where each $M_i \cong R$ behaving as an R-module.

Again, looks quite complicated but the idea is simple. A free R-module is one that has a basis - which is a subset of M being the generating set of M and also having linearly-independent elements. This means that for some element in the free R-module x, there exists unique non-zero elements $r_i \in R$ and $m_i \in M$ such that

$$x = \sum_i r_i m_i \ . \tag{3.7}$$

This is quite important - it means that you can effectively disentangle your complicated module into simple R-modules. Now with all of these defined, let us now look at what a group algebra is.

[6] We will use this notation for the rest of a section — if something is a field, it will be indicated by a double-lined letter.

[7] This means the multiplication operation defined for the ring is commutative, i.e. $xy = yx$ for $x, y \in R$.

Definition 3.7. The **group algebra** $R[G]$ **of a group** G **over a ring** R is the associative algebra whose underlying R-module is the free module over R on the underlying set of G and whose multiplication is given on these basis elements by group operation in G.

Let us unpack this definition - it looks a bit complicated. Firstly, there is an underlying R-module. This means that we are kind of working with a vector space [8] of some-sort. This underlying module is a free module - this means that any element x in this group algebra can be uniquely written in the form

$$x = \sum_{i \in I} r_i g_i \ . \tag{3.8}$$

This exactly matches up with what we have defined at the very start of the section (i.e. in Andre's notes). The group algebra is the 'formal' linear combination of group elements. Now the second point - multiplication of these elements are given by the group operation in G. Let us, for clarity, write the basis elements of the module as $\{e_g\}$ for the moment. This warrants the following algebraic structure:

$$e_g \cdot e_h = e_{gh} \ . \tag{3.9}$$

So it is not actually that difficult — $R[G]$ is kind of like a vector space with the following structure. Writing

$$x = \sum_{g \in G} r_g e_g \tag{3.10}$$

$$y = \sum_{h \in G} s_h e_h \tag{3.11}$$

Then in the group algebra we can write

$$\begin{aligned} x \cdot y &= \sum_{g,h \in G} r_g s_h e_g e_h \\ &= \sum_{g \in G} \sum_{h \in G} r_g s_h e_{gh} \\ &= \sum_{g \in G} \sum_{h \in G} r_h s_{h^{-1}g} e_g \end{aligned} \tag{3.12}$$

So the two definitions align. Phew!

3.2 $\mathbb{F}G$-modules and representations

Recall in the lectures we defined a representation as follows (for finite groups):

Definition 3.8. A **representation** of a finite group G on a finite-dimensional complex vector space V is a homomorphism $\rho : G \to GL(V)$ of G to the group of automorphisms of V.

[8] Recall that a vector space is a specific kind of modules.

It's quite clear what this means - a representation 'represents' how the group elements 'acts' on your vector space of choice V. Since any map $\gamma : V \to V \in GL(V)$ will be an automorphism - it is quite clear that this definition aligns with the idea of what a group element should do. We want to rephrase this using group algebras. How do we do that? Let us first make a definition.

Definition 3.9. Let V be a vector space over $\mathbb{F}$ and G be a group. Then V is an **$\mathbb{F}G$-module** if a multiplication $g \cdot v$ is defined to be satisfied for all $u, v \in V$, $\lambda \in \mathbb{F}$ and $g, h \in G$ the following statements.

(i) $gv \in V$.
(ii) $(gh)v = g(hv)$.
(iii) $1v = v$.
(iv) $\lambda gv = g(\lambda v)$.
(v) $g(u + v) = gu + vu$.

Technically you can generalise this notion to $R[G]$-modules - an algebra is a ring so the module in general will not be a vector space. But with this specific algebra (where the underlying ring is a field) this is indeed the case. Now let us make the following definition.

Definition 3.10. A **representation** of the group algebra on a vector space V is an algebra homomorphism $\rho : \mathbb{F}G \to \mathrm{End}(V)$.

When G is finite, we additionally have the following nice result.

Lemma 3.1. *If $\mathbb{F}$ is a field and G is a finite group, then an $\mathbb{F}G$-module is finitely-generated iff it has finite dimension as an $\mathbb{F}$-vector space.*

Proof. If V is generated as an $\mathbb{F}G$-module by the basis $\{v_1, \ldots, v_t\}$, then V is generated as an $\mathbb{F}$-vector space by $\{gv_i \,|\, g \in G,\, 1 \le i \le t\}$, and since G is finite we see that $\dim_{\mathbb{F}}(V) < \infty$. The converse is trivial. □

The fundamental connection between modules over group algebras and representation theory can be encapsulated in the following proposition.

Proposition 3.1. *If $\mathbb{F}$ is a field and G is a finite group, then there is a bijective corresponding between finitely-generated $\mathbb{F}G$-modules and linear actions of G on finite-dimensional $\mathbb{F}$-vector spaces, i.e.*

$$\left\{\begin{matrix} \textit{finitely-generated} \\ \mathbb{F}G\textit{-modules} \end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix} \textit{linear actions of } G \textit{ on} \\ \textit{finite-dimensional } \mathbb{F}\textit{-vector spaces} \end{matrix}\right\} . \tag{3.13}$$

Proof. Say V is a finitely-generated $\mathbb{F}G$-module. Then $\dim_{\mathbb{F}}(V)$ is finite by Lemma 5, and the map $G \times V \to V$ obtained by restricting the module structure map from $\mathbb{F}G \times V \to V$ is a linear action. To see this, note that for each $g \in G$ we obtain a map from V to V, denoted by $\varphi(g)$ as,

$$\varphi(g)(v) = g \cdot v \ , \tag{3.14}$$

with $g \cdot v$ the given action of the ring element g on the element v of V, the finitely-generated $\mathbb{F}G$-module. Clearly, this is linear, and the properties of a module ensures that φ is a group homomorphism, in particular $\varphi(g^{-1}) = \varphi(g)^{-1}$. Conversely for V a finite-dimensional $\mathbb{F}$-vector space on which G acts linearly with representation φ, V can be endowed with an $\mathbb{F}G$-module structure by defining,

$$\left(\sum_{g \in G} \alpha_g g \right) v = \sum_{g \in G} \alpha_g \varphi(g)(v) \ , \tag{3.15}$$

where $\left(\sum_{g\in G} \alpha_g g\right) \in \mathbb{F}G$ and $v \in V$. Here $\varphi : G \to GL(V)$ is a representation of G on the vector space V; and we can see how φ as a group homomorphism can be shown by,

$$(g_i g_j) \cdot v = \varphi(g_i g_j)(v) = \varphi(g_i)(\varphi(g_j)(v)) = g_i \cdot (g_j \cdot v) \ . \tag{3.16}$$

The processes are clearly mutually inverse of each other. □

Remark 3.1. Note that the content of the proof in Proposition 3.1 does not explicitly depend on the fact that the modules have to be finitely-generated. This means that the finite conditions in the previous proposition can be dropped.

We usually consider the case where the field $\mathbb{F}$ is chosen to be the field of complex numbers $\mathbb{C}$. Then, a representation $\rho : G \to \mathrm{Aut}(V)$ will extend by linearity to a map $\tilde{\rho} : \mathbb{C}G \to \mathrm{End}(V)$. This means that representations of $\mathbb{C}G$ will correspond exactly to representations of G and in particular, this sets up the bijection between the set of representations $\rho : G \to GL(V)$ and the set of $\mathbb{F}G$-module structures $\mathbb{F}G \times V \to V$ on V.

Let us perhaps illustrate this with a few examples.

Example 3.1 (Trivial module)**.** The field $\mathbb{F}$ can always be regarded as an $\mathbb{F}G$-module by defining $g\lambda = \lambda$ for all $g \in G$ and $\lambda \in \mathbb{F}$. This is known as the **trivial module**, and is of course bijectively related to the trivial representation.

Example 3.2 (Permutation modules)**.** Suppose G acts on a finite set X, and let $\mathbb{F}X$ be the set of all formal $\mathbb{F}$-linear combinations of elements of X. Then this set obviously has an $\mathbb{F}$-vector space structure with basis X. Define an $\mathbb{F}G$-moduel structure on $\mathbb{F}X$ by linearly extending the action of G on X, i.e. if $X = \{x_1, \ldots, x_n\}$, then for $g \in G$ and $\sum_i c_i x_i \in \mathbb{F}X$ then we have $g(\sum_i c_i x_i) = \sum_i c_i(g x_i)$. These are known as **permutation modules**.

In particular, we have the following special case:

Definition 3.11. The **regular representation** R_G corresponds to the left action of G on itself. Equivalently, R_G is the space of complex-valued functions on G where an element $g \in G$ acts on a function α by

$$(g \cdot \alpha)(h) = \alpha(g^{-1}h) \tag{3.17}$$

Proposition 3.2. *The left $\mathbb{C}G$-module given by $\mathbb{C}G$ itself corresponds to the regular representation. i.e. Taking the elements of G as a basis of $\mathbb{C}G$, then each $g \in G$ permutes these basis elements under the left regular permutation representation,*

$$g \cdot g_i = g g_i \ . \tag{3.18}$$

Proof. This is a simple proof - the action of an element on $\mathbb{C}G$ on itself is expressed in Equation 3.12. This is exactly the regular representation where this action is simply the lift by extending the definition of the regular representation above by linearity. $\square$

Remark 3.2. It is now trivial to check any definition of regular representations in your undergraduate course is exactly the same as the definition above. Simply set a basis for V, the representation space, and it should be obvious how Definition 3.11 matches up with the stated definition.

Of course, if U and V are $\mathbb{F}G$-modules, then $U \oplus V$, $U \otimes_{\mathbb{F}} V$, $\mathrm{Hom}_{\mathbb{F}}(U, \mathbb{F})$ and $\mathrm{Hom}_{\mathbb{F}}(U, V)$ are have natural $\mathbb{F}G$-module structures. Note that in this language equivalent representations can now be formalised.

Definition 3.12. Two representations of G are **equivalent** of the $\mathbb{F}G$-modules affording them are isomorphic modules. Representations which are not equivalent are called **inequivalent**.

Definition 3.13. Let R be a ring and M be a non-zero R-module. The module M is **irreducible** if its only submodules are 0 and M, otherwise M is **reducible**. The module M is **indecomposable** if M cannot be written as $M_1 \oplus M_2$ for any non-zero submodules M_1 and M_2, otherwise M is **decomposable**. M is **completely reducible** if it is a direct sum of irreducible submodules. A representation is accordingly named according to the properties of the $\mathbb{F}G$-module affording it.

Note that an irreducible module is by definition both indecomposable and completely reducible. Reducible representations are those with a corresponding matrix representation with matrices in block upper triangular form — to see this extend the G-invariant subspace $U \subset V$ to the whole space to get,

$$R_{\mathrm{red}}(g) = \begin{pmatrix} \varphi_1(g) & \psi(g) \\ 0 & \varphi_2(g) \end{pmatrix} . \tag{3.19}$$

Decomposable representations, on the other hand, can be written in the natural basis $V = U \oplus U'$ in block diagonal form,

$$R_{\mathrm{red}}(g) = \begin{pmatrix} \varphi_1(g) & 0 \\ 0 & \varphi_2(g) \end{pmatrix} . \tag{3.20}$$

There is also the very important notion of submodules.

Definition 3.14. If V is an $\mathbb{F}G$-module affording the representation φ, then a subspace U of V is **G-invariant** or **G-stable** if $g \cdot u \in U$ for all $g \in G$ and $u \in U$. Then, the **FG-submodules** of V are precisely the G-stable subspaces of V.

Another basic result of the representation theory of finite groups should be stated here.

Theorem 3.1 (Maschke's Theorem)**.** *Let G be a group, and suppose that the characteristic of $\mathbb{F}$ [9]is either zero or coprime to $|G|$. If U if an $\mathbb{F}G$-module and V is an $\mathbb{F}G$-submodule of U, then V is a direct summand of Y as $\mathbb{F}G$-modules.*

Proof. Since V is an $\mathbb{F}$-vector subspace of U we must have some subspace $W \subset U$ such that $U = V \oplus W$. Let $\pi : U \to V$ be the projection of U onto V along W such that π is the unique linear transformation that is the identity on V and zero on W. Define a linear transformation $\pi' : U \to U$ as,

$$\pi'(u) = \frac{1}{|G|} \sum_{g \in G} g\pi(g^{-1}u) \tag{3.21}$$

for $u \in U$. This definition is precise only when $|G| \neq 0$, hence the characteristic requirement. Now, as $V \subset U$, $gv \in V$ for any $v \in V$ and $g \in G$ so π' maps U into V. Also since $\pi|_V = \mathrm{id}_V$ so $g\pi(g^{-1}v) = gg^{-1}v = v$ and $\pi'|_V = \mathrm{id}_V$. So $U = V \oplus \ker \pi'$. It then remains to show that $\ker \pi'$ is an $\mathbb{F}G$-submodule of U. It suffices to show that π' is an $\mathbb{F}G$-module homomorphism, i.e. $\pi'(xu) = x\pi;(u)$ for any $x \in G$ and $u \in U$. Then,

$$\begin{aligned}
\pi'(xu) &= \frac{1}{|G|} \sum_{g \in G} g\pi(g^{-1}xu) \\
&= \frac{1}{|G|} x \left(\sum_{g \in G} x^{-1} g\pi(g^{-1}xu) \right) \\
&= \frac{1}{|G|} x \left(\sum_{g \in G} y\pi(y^{-1}u) \right) \\
&= x\pi'(u) \ ,
\end{aligned} \tag{3.22}$$

where we have reindexed $y = x^{-1}g$ on the third line. □

The reason to introduce Maschke's Theorem is to show that we can work with irreducible representations. To see this, note first the above results we have proved so far are independent of the *finite* condition, and note that,

Lemma 3.2. *An $\mathbb{F}G$-module is finitely-generated iff it is finite-dimensional.*

Proof. The proof is simple. If V is a finite dimensional vector space over $\mathbb{F}$ then it is clear that any $\mathbb{F}$ gives a finite set of generators over $\mathbb{F}G$ and hence V is finitely generated as an $\mathbb{F}G$-module. Conversely, if V is finitely generated as an $\mathbb{F}G$-module by $\{v_i\}$ then V is spanned by a vector space by the finite set $\{g \cdot v_i \,|\, g \in G,\, 1 \leq i \leq k\}$. □

Corollary 3.1. *If G is a finite group and $\mathbb{F}$ is a field whose characteristic does not divide $|G|$, then every finitely generated $\mathbb{F}G$-module is completely reducible.*

[9]The **characteristic** of a field $\mathbb{F}$ is defined as the smallest positive integer n such that n-summands of unity $1 + \cdots + 1 = 0$.

Proof. Let V be a finitely generated $\mathbb{F}G$-module. Since V is finite-dimensional over $\mathbb{F}$, we can proceed by induction on its dimension — if it is irreducible it is completely reducible and we are done. Otherwise, we use Maschke's Theorem to decompose $V = U \oplus W$ and by induction each of U and W are direct sums of irreducible submodules and so is V. $\square$

In particular, this means that for a representation $R : G \to GL(V)$ of finite degree, there is a basis V such that for each $g \in G$ the matrix of $R(g)$ with respect to this basis is block diagonal,

$$R(g) = \begin{pmatrix} R_1(g) & & \\ & \ddots & \\ & & R_m(g) \end{pmatrix} , \tag{3.23}$$

with R_i an irrep for $1 \leq i \leq m$.

Maschke's Theorem is extremely useful in establishing a classification of the representations of finite groups — it states that all representations, or $\mathbb{F}G$-modules, can be decomposed into a finite number of non-isomorphic irreducible $\mathbb{F}G$-modules, thus reducing the task of classifying all representations to only classifying irreducible representations.

3.3 Weddenburn-Artin Theorem

What we have seen so far from Maschke's Theorem is the following result — if $\mathbb{F}$ is a field where

$$|G| \bmod \mathrm{char}(\mathbb{F}) \neq 0 , \tag{3.24}$$

then every finitely-generated $\mathbb{F}G$-module can be written as a direct sum of finitely many simple (irreducible) modules. This, however, doesn't inform us anything about the structure of these simple modules. In particular, we want to know how the precise structure of these simple modules to construct the irreducible representations of an $\mathbb{F}G$-module — leading us to construct the irreps for a finite group G.

To do this we will need to classify all semisimple algebras. In particular, we will start by proving the semisimplicity of a certain class of algebras known as matrix algebras, and then ultimately show that all semisimple algebras lie in this class. We begin with some definitions.

Definition 3.15. Suppose S is a finite-dimensional $\mathbb{F}$-algebra. The set of $n \times n$ matrices $\mathcal{M}_n(D)$ with entries in D is a finite-dimensional $\mathbb{F}$-algebra with dimension $n^2 \dim_{\mathbb{F}} D$ known as the **matrix algebra over** D.

In particular, we denote $E_{ij}(\alpha)$ to be the matrix in $\mathcal{M}_n(D)$ with the only non-zero entry at (i, j) being α.

Definition 3.16. The algebra D is a **division algebra** if the non-zero elements of D form a group under multiplication.

Now we have a theorem.

Theorem 3.2. *Let D be a division algebra, and let $n \in \mathbb{N}$. Then any simple $\mathcal{M}_n(D)$-module is isomorphic with D^n and $\mathcal{M}_n(D)$ is isomorphic as $\mathcal{M}_n(D)$-modules with the direct sum of n copies of D^n. In particular, $\mathcal{M}_n(D)$ is a semisimple algebra.*

Proof. A non-zero submodule of D^n must contain some non-zero vector, which has some non-zero element x at the j-th position. If we premultiply this vector by $E_{jj}(x^{-1})$ then the submodule contains the j-th standard basis vector. Premultiplying by permutation basis then means this submodules contains every vector, so D^n is simple. Now suppose $C_k \subset \mathcal{M}_n(D)$ consisting of those matrices whose only non-zero entries appear in the k-th column. Then we must have $\mathcal{M}_n(D) \cong \bigoplus_{k=1}^n C_k$ as $\mathcal{M}_n(D)$-modules but each C_k is isomorphic as an $\mathcal{M}_n(D)$-module with D^n. Therefore, $\mathcal{M}_n(D)$ is a semisimple algebra and D^n is the unique simple $\mathcal{M}_n(D)$-module. □

This theorem suggests the following — the action of the matrices in the matrix algebra $\mathcal{M}_n(D)$ on a module must be isomorphic to how matrices act on the vector space with n-copies of D, D^n. The additional surprising property is that the matrices can be generated by the vectors D^n and it is a semi-simple algebra as it is a direct sum of the individual pieces D^n.

We additionally introduce two concepts about simple algebras, which are algebras which has only itself and the zero-ideal as the only two-sided ideals.

Lemma 3.3. *Simple algebras are semisimple.*

Proof. Let A be a simple algebra, and let Σ be the sum of all simple submodules of A. Let S be a simple submodule of A, and let $a \in A$. Then Sa is the image of S under the homomorphism $s \mapsto sa$ so Sa is either zero or simple. Then in either case, we have $Sa \subset \Sigma$ for any simple submodule S and any $a \in A$ so Σ is a right ideal in A and therefore a two-sided ideal. But A is simple so $\Sigma = A$ and A is a sum of simple A-modules so A is a semisimple algebra. □

Now we can discuss how matrix algebras and division algebras are related.

Theorem 3.3. *Let D be a division algebra, and let $n \in \mathbb{N}$. Then $\mathcal{M}_n(D)$ is a simple algebra.*

Proof. Let $0 \neq M \in \mathcal{M}_n(D)$, we must show that the principal two-sided ideal J of $\mathcal{M}_n(D)$ generated by M is equal to $\mathcal{M}_n(D)$. It suffices to show that J contains each $E_{ij}(1)$ since these matrices generate $\mathcal{M}_n(D)$ as an $\mathcal{M}_n(D)$-module. As $M \neq 0$, there are some $1 \leq r, s \leq n$ such that the (r, s)-entry of M is non-zero. We call this x. Then we can easily verify that $E_{ss}(1) = E_{sr}(x^{-1})ME_{ss}(1) \in J$. Now let $1 \leq i, j \leq n$, and let w and w' be the permutation matrices corresponding to the transpositions $(i\ s)$ and $(s\ j)$ respectively. Then $E_{ij}(1) = wE_{ss}(1)w' \in J$. □

We can extend this by utilising the structure of direct sums of semisimple algebras. The facts of the direct sums of algebras are found in Appendix A.2, and we will need the theorem below. Now we can combine this with the division algebra structure.

Theorem 3.4. *Let $r \in \mathbb{N}$. For each $1 \leq i \leq r$, we let D_i be a division algebra over $\mathbb{F}$, $n_i \in \mathbb{N}$ and let $B_i = \mathcal{M}_{n_i}(D_i)$. Let B be the external direct sum of the B_i. Then B is a semisimple algebra having exactly r isomorphism classes of simple modules and exactly 2^r two-sided ideals, namely every sum of the form $\bigoplus_{j \in J} B_j$ where J is a subset of $\{1, \ldots, n\}$.*

Proof. For each i, we can write $B_i = \bigoplus_{j=1}^{n_i} C_{ij}$ using Theorem 3.2 with C_{ij} mutually isomorphic simple B_i-modules. Since each C_{ij} is also simple as a B module, $B \cong \bigoplus_{i,j} C_{ij}$ as B-modules and B is therefore a semisimple algebra by Lemma A.4 and Proposition A.2 — where any simple B-module is isomorphic with some C_{ij}. But since $C_{ij} \cong C_{kl}$ as B-modules iff $i = k$, we have exactly r-isomorphism classes of simple B-modules. The two-sided ideals arguments follow easily. □

The key result of the discussion above — a direct sum of matrix algebras over division algebras is semisimple. The converse is in fact also true — this is the content of Wedderburn's Theorem. but to get there we will need to understand the proof by first developing some more concepts.

Definition 3.17. The **endomorphism algebra** of an A-module M, denoted $\mathrm{End}_A(M)$, is an $\mathbb{F}$-algebra with its multiplication given by the composition of mappings.

Definition 3.18. The **opposite algebra** A^{op} of an algebra A is the set A endowed with the usual addition and scalar multiplication but with the opposite multiplication,

$$a \cdot b = ba \ . \tag{3.25}$$

We can show that endomorphism algebras and opposite algebras are closely related.

Lemma 3.4. *Let B be an algebra. Then $B^{op} \cong \mathrm{End}_B(B)$.*

Proof. Let $\varphi \in \mathrm{End}_B(B)$ and $a = \varphi(1)$. Then $\varphi(b) = b\varphi(1) = ba$ for any $b \in B$ so φ is equal to the endomorphism ρ_a given by the right multiplication by a. i.e. $\mathrm{End}_B(B) = \{\rho_a \,|\, a \in B\}$ so $\mathrm{End}_B(B)$ and B are in bijective correspondence. Finally, note that,

$$(\rho_a\rho_b)(x) = \rho_a(xb) = xba = \rho_{ba}(x) = \rho a \cdot b(x) \ , \tag{3.26}$$

as required. □

This suggests that we can gain information about semisimple algebras by studying the properties of the endomorphism algebras of semisimple modules. We will come back to use this fact on opposite algebras — the only thing now is to remember that the opposite algebra is that the opposite algebra of a division algebra is also a division algebra. But let's move on and talk about A modules.

Lemma 3.5. *Let $S_1, \ldots, S_r$ be the distinct simple A-modules. For each i, let U_i be a direct sum of copies of S_i, and let $U = U_1 \oplus \cdots \oplus U_r$. Then we have that,*

$$\mathrm{End}_A(U) \cong \mathrm{End}_A(U_1) \oplus \cdots \oplus \mathrm{End}_A(U_r) \ . \tag{3.27}$$

Proof. Let $\varphi \in \mathrm{End}_A(U)$ and fix some i. Since every composition factor of U_i is isomorphic with S_i, the Jordan-Hölder theorem for modules tells us that the same is true of $\varphi(U_i)$ since $\varphi(U_i)$ is isomorphic with a quotient of U_i. Let us first suppose that $\varphi(U_i)$ were not contained in U_i. Then the image of $\varphi(U_i)$ in U/U_i under the natural map would be a non-zero submodule having S_i as a composition factor. But it follows from the hypothesis that the composition factors of U/U_i are exactly S_j with $j \neq i$, so S_i cannot be a composition factor of U/U_i or its submodules. Therefore for each i, we can define $\varphi_i \in \mathrm{End}_A(U_i)$ to be the restriction to U_i of φ. So we can define a map $\Gamma : \mathrm{End}_A(U) \to \mathrm{End}_A(U_1) \oplus \cdots \oplus \mathrm{End}_A(U_r)$ by setting $\Gamma(\varphi) = (\varphi_1, \ldots, \varphi_r)$. Then Γ is an A-module monomorphism, and let $(\varphi_1, \ldots, \varphi_r) \in \mathrm{End}_A(U_1) \oplus \cdots \oplus \mathrm{End}_A(U_r)$. We define $\hat{\varphi} \in \mathrm{End}_A(U)$ as: given $x \in U$, we write $x = x_1 + \cdots + x_r$ where $x_i \in U_i$ and define $\hat{\varphi}(x) = \varphi_1(x_1) + \cdots + \varphi_r(x_r)$. We then have $(\varphi_1, \ldots, \varphi_r) = \Gamma(\hat{\varphi})$ so Γ is surjective. □

Lemma 3.6. *If S is a simple A-module, then for any $b \in \mathbb{N}$ we have $\mathrm{End}_A(nS) \cong \mathcal{M}_n(\mathrm{End}_A(S))$.*

Proof. We can regard the elements of nS as being column vectors of length n with entries from S. Let $\Phi = (\varphi_{ij}) \in \mathcal{M}_n(\mathrm{End}_A(S))$. We define $\Gamma(\Phi) : nS \to nS$ by using,

$$\Gamma(\Phi)\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} = \begin{pmatrix} \varphi_{11}(s_1) + \cdots + \varphi_{1n}(s_n) \\ \vdots \\ \varphi_{n1}(s_1) + \cdots + \varphi_{nn}(s_n) \end{pmatrix} \tag{3.28}$$

We find that $\Gamma(\Phi)(a\boldsymbol{s} + \boldsymbol{t}) = a\,[\Gamma(\Phi)(\boldsymbol{s})] + \Gamma(\Phi)(\boldsymbol{t})$ for any $a \in A$ and $\boldsymbol{s}, \boldsymbol{t} \in nS$ since each φ_{ij} is an A-module homomorphism, then $\Gamma(\Phi) \in \mathrm{End}_A(nS)$. You can check that $\Gamma : \mathcal{M}_n(\mathrm{End}_A(S)) \to \mathrm{End}_A(nS)$ defined like this is an algebra monomorphism. Now let ψ be an element of $\mathrm{End}_A(nS)$. We define $\psi_{ij} : S \to S$ by a similar construction like,

$$\psi\begin{pmatrix} s \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \psi_{11}(s) \\ \psi_{21}(s) \\ \vdots \\ \psi_{n1}(s) \end{pmatrix}, \tag{3.29}$$

so find that $\psi_{ij} \in \mathrm{End}_A(S)$. Let $\Psi = (\psi_{ij}) \in \mathcal{M}_n(\mathrm{End}_A(S))$, then $\Gamma(\Psi) = \psi$ so Γ is surjective as required. □

Therefore, if S is a simple A-module, it follows immediately from Schur's Lemma that the non-trivial endomorphisms of S form a group under multiplication so $\mathrm{End}_A(S)$ is a division algebra. In particular, we can learn something more specific about the structure of $\mathrm{End}_A(S)$ when the ground field $\mathbb{F}$ is algebraically closed.

Lemma 3.7. *Suppose $\mathbb{F}$ is algebraically closed, and let S be a simple A-module. Then $\mathrm{End}_A(S) \cong \mathbb{F}$.*

Proof. Let $\varphi \in \mathrm{End}_A(S)$. We can see φ as an invertible $\mathbb{F}$-linear self-map of the finite-dimensional $\mathbb{F}$-vector space S and therefore it has a non-zero eigenvalue λ_φ. Suppose $I \in \mathrm{End}_A(S)$ is the identity element, then $\varphi - \lambda_\varphi I \in \mathrm{End}_A(S)$ has a non-zero kernel and therefore is not invertible so $\varphi = \lambda_\varphi I$ since $\mathrm{End}_A(S)$ is a division algebra. Then the map sending φ to λ_φ is an isomorphism from $\mathrm{End}_A(S)$ to $\mathbb{F}$. □

Lemma 3.8. *Let B be an algebra. Then $\mathcal{M}_n(B)^{op} \cong \mathcal{M}_n(B^{op})$ for any $n \in \mathbb{N}$.*

Proof. We define $\psi : \mathcal{M}_n(B)^{op} \to \mathcal{M}_n(B^{op})$ by letting $\psi(X)$ be the transpose X^T of the matrix X. This map ψ is clearly bijective. Let $X = (x_{ij})$ and $Y = (y_{ij})$ be elements of $\mathcal{M}_n(B)^{op}$. Then for any i and j, we have,

$$
\begin{aligned}
(\psi(X)\psi(Y))_{ij} &= \sum_{k=1}^{n} \psi(X)_{ik} \cdot \psi(Y)_{kj} \\
&= \sum_{k=1}^{n} x_{ki} y_{jk} = (YX)_{ji} = (X \cdot Y)^T_{ij} = \psi(X \cdot Y)_{ij} \ ,
\end{aligned} \tag{3.30}
$$

so we have $\psi(X \cdot Y) = \psi(X)\psi(Y)$ and hence ψ is an algebra homomorphism. □

Now we can put all of the pieces together to state Wedderburn's main structure theorem.

Theorem 3.5. *The algebra A is semisimple iff it is isomorphic with a direct sum of matrix algebras over division algebras.*

Proof. Suppose that the algebra A is semisimple. Then we can write A in the form $A = U_1 \oplus \cdot \oplus U_r$ where U_r is the direct sum of n_i copies of a simple A-module S_i and no two of the S_i are isomorphic. We then have,

$$
\begin{aligned}
A^{op} &\cong \mathrm{End}_A(A) \\
&\cong \mathrm{End}_A(U_1) \oplus \cdots \oplus \mathrm{End}_A(U_r) \\
&\cong \mathrm{End}_A(n_1 S_1) \oplus \cdots \oplus \mathrm{End}_A(n_r S_r) \\
&\cong \mathcal{M}_{n_1}(\mathrm{End}_A(S_1)) \oplus \cdots \oplus \mathcal{M}_{n_r}(\mathrm{End}_A(S_r)) \ ,
\end{aligned} \tag{3.31}
$$

and therefore,

$$
\begin{aligned}
A &\cong [\mathcal{M}_{n_1}(\mathrm{End}_A(S_1)) \oplus \cdots \oplus \mathcal{M}_{n_r}(\mathrm{End}_A(S_r))]^{op} \\
&\cong \mathcal{M}_{n_1}(\mathrm{End}_A(S_1))^{op} \oplus \cdots \oplus \mathcal{M}_{n_r}(\mathrm{End}_A(S_r))^{op} \\
&\mathcal{M}_{n_1}(\mathrm{End}_A(S_1)^{op}) \oplus \cdots \oplus \mathcal{M}_{n_r}(\mathrm{End}_A(S_r)^{op}) \ .
\end{aligned} \tag{3.32}
$$

This means that since the endomorphism algebra of a simple module is a division algebra and since the opposite algebra of a division algebra is also a division algebra any semisimple algebra is isomorphic with a direct sum of matrix algebras over division algebras. The converse is established in Theorem 3.4. □

Most importantly we have the following corollary.

Corollary 3.2. *The algebra A is simple iff it is isomorphic with a matrix algebra over a division algebra.*

Proof. Suppose that A is simple. Then A is semisimple and by Theorem 3.5 A is isomorphic with a direct sum of r matrix algebras over division $\mathbb{F}$-algebras and hence by Theorem 3.4 A has exactly 2^r ideals. But A is simple and therefore has exactly two ideals so we have $r = 1$ and therefore any simple algebra is isomorphic with a matrix algebra over a division algebra. The converse is established in Theorem 3.3. □

Additionally, semisimple algebras over algebraically closed fields have a more specific classification than those over arbitrary fields.

Theorem 3.6. *Suppose that field $\mathbb{F}$ is algebraically closed. Then any semisimple algebra is isomorphic with a direct sum of matrix algebras over $\mathbb{F}$.*

Proof. Follows from Lemma 3.7 and Theorem 3.5. □

The point of Wedderburn's Theorem is thus. It states that a semisimple algebra is indeed isomorphic with a direct sum of matrix algebras over division algebras, thus enabling us to write the group ring $\mathbb{F}G$ as follows. Since $\mathbb{F}G$ is semisimple, we have by Wedderburn's Theorem,

$$\mathbb{F}G \cong R_1 \times \cdots \times R_r \ , \tag{3.33}$$

where R_i is the ring of $n_i \times n_i$ matrices over some division ring D_i. Now if $\mathbb{F} = \mathbb{C}$, then $\mathbb{F}$ is algebraically-closed and we have $D_i = \mathbb{C}$ so

$$R_i = \mathcal{M}_{n_i}(\mathbb{C}) \ . \tag{3.34}$$

We can then prove the following statement.

Proposition 3.3. *The group order can be determined by the dimensions of the irreducible representations.*

$$\sum_{i=1}^{r} n_i^2 = |G| \ . \tag{3.35}$$

Proof. From Wedderburn's Theorem the decomposition gives each matrix ring $\mathcal{M}_{n_i}(\mathbb{F})$ having dimension n_i^2 over $\mathbb{F}$ and therefore R has dimensions $\sum_{i=1}^{r} n_i^2$ over $\mathbb{F}$. also since the centre of each $\mathcal{M}_{n_i}(\mathbb{F})$ is one-dimensional $Z(R)$ has dimension r over $\mathbb{F}$. □

We realise that this result comes not from character theory, but is a direct consequence of the structure of the matrix algebras. There is one final result to show.

Corollary 3.3. *The number of Wedderburn components in $\mathbb{F}G$, r (i.e. number of irreps), is equal to the number of conjugacy classes of G.*

Proof. From Proposition 3.3 we have $r = \dim_{\mathbb{C}} Z(\mathbb{C}G)$. Let $\mathcal{K}_i$ be the distinct conjugacy classes of G. Fro each conjugacy class $\mathcal{K}_i$ of G let

$$X_i = \sum_{g \in \mathcal{K}_i} g \in \mathbb{C}G \; . \tag{3.36}$$

Note that X_i and X_j $(i \neq j)$ have no common terms since they are linearly independent elements. Also X_i commutes with all group elements since $h^{-1}X_i h = X_i$ for arbitrary $h \in G$ so $X_i \in Z(\mathbb{C}G)$. Now we show that X_i form a basis of $Z(\mathbb{C}G)$, so $s = \dim_{\mathbb{C}} Z(\mathbb{C}G) = r$. Since X_i are linearly independent it remains to show that they span $Z(\mathbb{C}G)$. Let $X = \sum_{g \in G} \alpha_g g$ be an arbitrary element of $Z(\mathbb{C}G)$. Since $h^{-1}Xh = X$, we have,

$$\sum_{g \in G} \alpha_g h^{-1} g h = \sum_{g \in G} \alpha_g g \; . \tag{3.37}$$

So since the elements of G form a basis of $\mathbb{C}G$ the coefficients $\alpha_{hgh^{-1}} = \alpha_g$. Since h is arbitrary, every element in the same conjugacy class of a fixed group element g has the same coefficient in X, then X can be written as a linear combination of the X_is. $\square$

We have therefore shown that the group algebra $\mathbb{F}G$ or $\mathbb{C}G$ has exactly r distinct isomorphism types of irreducible modules and these have complex dimensions $n_1, \ldots, n_r$ so G has exactly r inequivalent irreducible complex representations of the corresponding degrees. In particular, r, the number of irreps, is equal to the number of conjugacy classes in G and $\sum_{i=1}^{r} n_i^2 = |G|$. The fact that these results come directly from the structure of semisimpleness of the group algebra shouldn't surprise you — in the lectures we have used character theory to show these results (as the proofs there are normally easier) but character theory works in the end because the existence of class functions exactly depends on this nontrivial structure of group algebras.

4 Character tables

Most of you did very well when tackling character theory questions. I just want to summarise some facts here and give you a general direction as to how to construct character tables for finite groups (you can probably generalise this to Lie groups - try it)!

Let us first recall the definition of a character table for finite groups. Throughout this section, we fix G to be a group and $\mathbb{F}$ to be an arbitrary field [10].

Definition 4.1. A **class function** is a function $f : G \to \mathbb{F}$ which is constant on the conjugacy classes of G. i.e. for $g, x \in G$, we have

$$f(g^{-1}xg) = f(x) \ . \tag{4.1}$$

Definition 4.2. Let R be a representation of G afforded by the $\mathbb{F}G$-module V [11]. The **character** χ_V of the representation V is the class function $\chi : G \to \mathbb{F}$ such that,

$$\chi(g) = \operatorname{tr} R(g) \ . \tag{4.2}$$

where $\operatorname{tr} R(g)$ is the trace of the matrix $R(g)$ with respect to some basis of V.

Definition 4.3. The **character table** $\mathcal{X}$ of a finite group is a table of character values with the list of representations of the r conjugacy classes along the top row and the list of irreducible characters down the first column. The entry in the table in the row χ_i and column g_j is $\chi_i(g_j)$.

Suppose g_i where $i = 1, \dots, r$ are the representatives of the r conjugacy classes of G. Conventionally we pick $g_1 = e$ so an example of a character table $\mathcal{X}(G)$ would look like the one in Table 4.1.

Note that the first row of the character table is always the *trivial representation*, i.e. where

	e	g_2	$\dots$	g_r
χ_1	1	1	$\dots$	1
χ_2	f_2	$\chi_2(g_2)$	$\dots$	$\chi_2(g_r)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
χ_r	f_r	$\chi_r(g_2)$	$\dots$	$\chi_r(g_r)$

Table 4.1: A character table for group G. Here g_i are the conjugacy classes (labelled by the element g_i), χ_j are the irreps and f_j are the degrees of G (dimension of the irrep χ_j).

all of the characters are 1. This is known as the *principal character of* G.

Definition 4.4. The **principal character** of a finite group G is the character of its trivial representation; i.e. for all conjugacy classes represented by element g_r, $\chi_P(g_r) = 1$.

[10] This is not the field in the physical sense, i.e, not an electric field, etc. but a mathematical field. A field is a ring with a multiplicative inverse.

[11] Apologies for the inverted definition — the full definition of the $\mathbb{F}G$-module is defined in the following section.

The definitions of the characters should be obvious to you. If they are not, you should consult Andre's lectures notes again.

4.1 Constructing the character table

It is useful to determine the character tables for finite groups to work out their representation contents. In particular, it is useful to know how to derive the character tables of some basic examples of finite groups.

To construct the character table of a finite group G, we can use the following steps.

1. Determine all the conjugacy classes of the group [12].
2. Use the fact that the number of irreducible representations of G is equal to the number of conjugacy classes (Theorem 2.8 of the lecture notes or Corollary 3.3).
3. Find the dimensions of the irreps. You can use the following formula,
$$\sum_i (\dim R_i)^2 = |G| \ , \tag{4.3}$$
as a guide (Theorem 2.3 of the lecture notes or from Proposition 3.3).
4. Start with the *principal character* of G to get one dimension-one irrep.
5. Use row and column orthogonality of characters to determine the rest of the unknown entries of the character table.

The above steps, of course, are just here for guidance — there is no need to follow them if you have other ways of determining the character table! In particular, there are times when you do not want or cannot use the orthogonality property of characters (for example, if there are not enough rows or columns). In this case there might be other clever ways to construct other dimension-one characters, e.g. taking determinants of an existing representation. Alternatively, you may also see if you can get anything from the regular representation [13]...

4.2 Character orthogonality

The key statement to proof in this subsection is something known as row and column orthogonality. You probably already know what *row orthogonality* is: it is just Theorem 2.1 of the lecture notes:
$$(\chi, \tilde{\chi}) = \frac{1}{|G|} \sum_{g \in G} \overline{\chi(g)} \tilde{\chi}(g) \ , \tag{4.4}$$
or for irreps R_i and R_j,
$$(\chi_i, \chi_j) = \delta_{ij} \ , \tag{4.5}$$

[12] You can probably use some smart argument with the symmetric group S_n - using Cayley's Theorem in [2] that all groups are isomorphic to subgroups of the symmetric group.

[13] Often times, if you are given a question in the exam or in the problem sheet to find the character table there will be other clues somewhere... such as the existence of a matrix representation given in the question.

where we have defined the inner product of class functions as, given α and β class functions on G,

$$(\alpha, \beta) = \frac{1}{|G|} \sum_{g \in G} \alpha(g) \overline{\beta(g)} \ . \tag{4.6}$$

What is *column orthogonality*? It is in fact a corollary of Theorem 2.1 of the lecture notes:

Corollary 4.1. *For irreps R_i, and conjugacy class c_I, we have*

$$[\chi_I, \tilde{\chi}_J] = \frac{|c_I|}{|G|} \sum_i \overline{\chi_i(g_I)} \tilde{\chi}_i(g_J) = \delta_{IJ} \tag{4.7}$$

with A, B indicating the different conjugacy classes and $|c_I|$ the number of elements in the conjugacy class labelled I.

Proof. From simple linear algebra. If we rescale the original character table by $\sqrt{\frac{\dim(R_i)}{|G|}}$ for each i (i.e. each row), we obtain a unitary matrix. The transpose of a unitary matrix is unitary, so the column vectors are orthonormal to each other. The correct scaling is deduced by the redefinition of the sum. An alternative proof is as follows. We define class functions ψ_s for $1 \leq s \leq k$ by

$$\psi_s(g_r) = \delta_{rs} \ , \tag{4.8}$$

where k is the number of conjugacy classes. As characters form a basis of the space of class functions[14], ψ_s is a linear combination of χ_i,

$$\psi_s = \sum_{i=1}^{k} \lambda_i \chi_i \ . \tag{4.9}$$

We know that for irreps $(\chi_i, \chi_j) = \delta_{ij}$. So we must have

$$\lambda_i = (\chi_i, \psi_s) = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_i(g)} \psi_s(g) \ . \tag{4.10}$$

By definition of ψ_s, it is 1 if g is conjugate to g_I and 0 otherwise. So this gives:

$$\lambda_i = \frac{\chi_i(g_I)|g_I|}{|G|} \ , \tag{4.11}$$

where $|g_I|$ is the number of elements in the conjugacy class represented by g_I. So substituting into the definition of ψ_s gives

$$\delta_{IJ} = \frac{|g_I|}{|G|} \sum_i \overline{\chi_i(g_I)} \chi(g_J) \ , \tag{4.12}$$

as required. □

Column orthogonality strictly speaking gives no new information - but it might speed up your calculation when you are trying to deduce the entire character table, as we have mentioned before.

[14] See, for example, [1].

4.3 Examples

Example 4.1 (Character Table of $\mathbb{Z}_2$)**.** The most simple example is the group $\mathbb{Z}_2$. There are only two elements, namely the trivial e and the flip b. The conjugacy classes are just the two elements separately in each. Using $n_1 + n_2 = 2$, where the principal character simply has characters $(1, 1)$, we realise that the other character must be $(1, -1)$, hence giving the following group table.

	e	b
χ_1	1	1
χ_2	1	-1

Table 4.2: Character table for $\mathbb{Z}_2$.

Distributed version:

Exercise 4.1. Try and find the character table of S_3, as per Q3 of Problem Sheet 1.

5 Young Tableaux and Projectors

In this section I would like to discuss about Young Tableaux and representation projectors. Young Tableaux is a vast subject - they are extremely useful in describing the representations of S_n. As a result, they are used in a variety of contexts such as describing tensors and projectors. Meanwhile, projectors are extremely important in physics as they are effectively the operators that allow take you to the relevant space of the irreducible representation. Here I aim to provide a summary of results regarding Young Tablaeux, particularly highlighting the use of Young Tablaeux in describing projectors. The main references for this section are [1, 12].

5.1 Young Diagrams and Tableaux

We first recall that the number of irreducible representations of S_n is the number of conjugacy classes, i.e. the number of partitions $p(n)$ of $n = \lambda_1 + ... + \lambda_k$. To each partition we will associate a *Young diagram*[15]:

Definition 5.1. A **Young diagram** is a collection of boxes, or cells, arranged in left-justified rows, with a weakly decreasing number of boxes in each row. Listing the number of boxes in each row gives a partition $p(n)$ of integer n, where it is also the total number of boxes. Conversely, as stated above, each partition correspond to a Young diagram.

Example 5.1. Consider the partition $11 = 6 + 4 + 1$, or (6,4,1). Then the corresponding Young diagram is

$$\ydiagram{6,4,1} \tag{5.1}$$

Definition 5.2. A **Young tableau** is a filling (i.e. entering numbers into boxes) of a Young diagram such that the numbers are,

(i) weakly increasing across each row, and
(ii) strictly increasing down each column.

A **standard Young tableau** is a tableau in which entries are numbers from 1 to n, n being the total number of boxes.

Example 5.2. Consider the partition (5,2,1). Then one of the possible corresponding Young tableau is:

$$\begin{ytableau} 1 & 2 & 5 & 6 & 8 \\ 3 & 4 \\ 7 \end{ytableau} \tag{5.2}$$

[15]This is also called a Ferrers diagram or a Young frame according to [1]. Haven't really seen these names before to be honest.

5.2 Projection Formulae

Before we continue, let us revisit the topic of projection formulae. Firstly, recall that we have the trivial projection operator that projects onto the direct sum of trivial factors in the decomposition. In the lectures, we have used this to obtain several results in character theory, namely to derive character orthogonality. For completeness I will recap this here.

Definition 5.3. We define the direct sum of trivial factors in a representation[16] V of a group G to be

$$V^G = \{v \in V \mid gv = v, \forall g \in G\} \ . \tag{5.3}$$

We want a way to find V^G explicitly - to do this we will need a G-module homomorphism[17] so we take the average of all these endomorphisms and set

$$\varphi = \frac{1}{|G|} \sum_{g \in G} g \in \mathrm{End}(V) \ , \tag{5.4}$$

which this will be G-linear as $\sum g = \sum hgh^{-1}$. We then have the following proposition.

Proposition 5.1. *The map φ is a projection of V onto V^G.*

Proof. First, suppose $v = \varphi(w) = \frac{1}{|G|}\sum gw$. Then for any $h \in G$ we have

$$hv = \frac{1}{|G|} \sum hgw = \frac{1}{|G|} \sum gw \ , \tag{5.5}$$

such that the image of φ is contained in V^G. Conversely, if $v \in V^G$, then $\varphi(v) = \frac{1}{|G|}\sum v = v$, so $V^G \subset \mathrm{Im}(\varphi)$ and $\varphi \circ \varphi = \varphi$. □

Remark 5.1. In particular, if we wish to know the number m of the copies of trivial representation appearing in the decomposition of V then we can just compute the trace of φ.

Now let us move on to derive a more general projection formula. Let us first have the following Proposition.

Proposition 5.2. *Let $\alpha : G \to \mathbb{C}$ be any function on the group G, and for any representation V of G set*

$$\varphi_{\alpha,V} = \sum \alpha(g) \cdot g : V \to V \ . \tag{5.6}$$

Then $\varphi_{\alpha,V}$ is a homomorphism of G-modules for all V if and only if α is a class function.

[16] In this section a representation $R : G \to \mathrm{End}(V)$ will be indicated by its representation space V. This is just a notation.

[17] Observe in general that if V is the representation of a finite group G then each G will give a map $\rho(g) : V \to V$ but this is in general not a G-module homomorphism as for general $h \in G$ we have $g(h(v)) \neq h(g(v))$.

Proof. We have

$$\begin{aligned}
\varphi_{\alpha,V}(hv) &= \sum \alpha(g)\cdot g(hv) \\
&= \sum \alpha(hgh^{-1})\cdot hgh^{-1}(hv) \\
&= h\left(\sum \alpha(hgh^{-1})\cdot g(v)\right) \\
&= h\left(\sum \alpha(g)\cdot g(v)\right) \\
&= h\left(\varphi_{\alpha,V}(v)\right) \ ,
\end{aligned} \tag{5.7}$$

where the second line follows from substitution and the fourth follows from α being a class function. This is just the condition that $\varphi_{\alpha,V}$ is G-linear. □

We can now use this to construct a projection formula. Let us suppose W is a fixed irreducible representation. Then for representation V, we can look at the weighted sum:

$$\psi = \frac{1}{|G|}\sum_{g\in G} \overline{\chi_W(g)}\cdot g \in \mathrm{End}(V)\ . \tag{5.8}$$

From Proposition 5.2, it is clear that ψ is a G-module homomorphism. Therefore, if V is an irreducible representation, we will then have $\psi = \lambda\,\mathrm{id}$ (id being the identity map) and therefore

$$\begin{aligned}
\lambda &= \frac{1}{\dim V}\operatorname{Tr}\psi \\
&= \frac{1}{|G|\dim V}\sum \overline{\chi_W(g)}\cdot\chi_V(g) \\
&= \begin{cases} \frac{1}{\dim V} & \text{if } V = W \\ 0 & \text{if } V \neq W \end{cases} .
\end{aligned} \tag{5.9}$$

Therefore, what we have is for arbitrary V, the projector

$$\psi_V = \frac{\dim W}{|G|}\sum_{g\in G}\overline{\chi_W(g)}\cdot g : V \to V \tag{5.10}$$

is the projection of V onto the factor consisting of the sum of all copies of W (compare with the case with the trivial factors, the projectors act on the sum of the spaces). Therefore, we have that

Definition 5.4. Let the decomposition for a general representation V as

$$V = \bigoplus V_i^{\oplus a_i}\ , \tag{5.11}$$

where V_i are the inequivalent irreps, the **projector** of V onto $V_i^{\oplus a_i}$ is

$$\pi_i = \frac{\dim V_i}{|G|}\sum_{g\in G}\overline{\chi_{V_i}(g)}\cdot g\ . \tag{5.12}$$

This means that the projector $\pi^{(i)} : R_{\text{reg}} \to R_i^{\oplus \Delta_i}$ of the regular representation onto the irreducible subspaces $R_i^{\oplus \Delta_i}$, where $\Delta_i = \dim R_i$ is the dimension of the irrep R_i, is then

$$\pi^{(i)} = \frac{\dim(R_i)}{|G|} \sum_{g \in G} \overline{\chi_i(g)} \cdot g \ , \tag{5.13}$$

where $|G|$ is the order of the group and $\chi_i(g)$ is the character of the irrep R_i and element g.

5.3 Young Tableaux as Projectors

A standard Young tableau is what is used to describe projection operators for the regular representation onto irreducible representations of S_n. To show this, let us associate

$$P = P_\lambda = \{g \in S_n \mid g \text{ preserves each row}\} \ , \tag{5.14}$$

together with

$$Q = Q_\lambda = \{g \in S_n \mid g \text{ preserves each column}\} \ , \tag{5.15}$$

where e_g is the corresponding element of g in the group algebra $\mathbb{C}S_n$. In the same space we can introduce the following terms:

$$a_\lambda = \sum e_g \ , \quad g \in P \ , \tag{5.16}$$

and

$$b_\lambda = \sum \operatorname{sgn}(g) e_g \ , \quad g \in Q \ , \tag{5.17}$$

λ here is a free index labelling the invariant subspaces of the regular representation. We can now define the following.

Definition 5.5. The **Young symmetriser** is defined to be

$$c_\lambda = a_\lambda \cdot b_\lambda \in \mathbb{C}S_n \ . \tag{5.18}$$

This gives the following Theorem.

Theorem 5.1. *Some scalar multiple of c_λ is idempotent, i.e. $c_\lambda^2 = n_\lambda c_\lambda$. $\operatorname{im}(c_\lambda)$ is the irreducible invariant subspace V_λ of S_n. Every irreducible representation of S_n can be obtained in this manner for unique partition (i.e. for a valid standard Young tableau).*

Proof. See §4.2 and Theorem 4.3 of [1]. □

I also state two very useful formula for computing the projection operators.

Lemma 5.1. *(Hook Length Formula) We have*

$$\dim V_\lambda = \frac{d!}{\prod(\textit{Hook lengths})} \ , \tag{5.19}$$

where the hook length of a box in a Young diagram is the number of squares directly below and to the right of the box, including the box itself once.

Proof. Again see §4, Hook Length Formula 4.12 of [1]. □

In fact we can then write

$$n_\lambda = \frac{d!}{\dim V_\lambda} = \prod(\text{Hook lengths}) \ . \tag{5.20}$$

This is the inverse of the coefficient c you typically need to find when you write down the projector [18] [19]:

$$\pi_\lambda = c\left(\sum_{\sigma\in R_\lambda}\sigma\right)\left(\sum_{\sigma\in C_\lambda}\text{sgn}(\sigma)\sigma\right) = c c_\lambda \tag{5.22}$$

How does this relate to the projector stated in the previous section, in Eq. (5.13)?

$$\pi^{(i)} = \frac{\dim(R_i)}{|G|}\sum_{g\in G}\overline{\chi_i(g)}\cdot g \tag{5.13}$$

The projector π_i projects onto the sum of irreducible representations R_i of the same dimensions, i.e. onto $R_i^{\oplus_i \Delta_i}$ where $\Delta_i = \dim R_i$; whereas π_λ in S_n projects onto the individual irreducible representations R_λ of the regular representation, the individual invariant subspaces. Labelling the individual irreps as $R_\lambda^{(i)}$, where λ is the index denoting the irrep within the irreps of the same dimensions and i labelling the irreps of different dimensions, we then have the following **completeness relations**:

$$\sum_i \pi^{(i)} = \text{id} \ , \tag{5.23}$$

together with

$$\sum_\lambda \pi_\lambda^{(i)} = \pi^{(i)} \tag{5.24}$$

We can check this explicitly in Q3 of the first problem sheet for the two two-dimensional irreducible representations of the symmetric group S_3. This gives us the following example.

Distributed version:

Exercise 5.1. Try and find the projectors of the two-dimensional irreps of S_3 in Q3 of Sheet 1 again. You should find that the projector π_{2D} in Eq. (5.13) is the sum of the projector into the 2 irreps (invariant subspaces) of the group algebra $\mathbb{C}G$ in Eq. (5.22), or,

$$\pi_{2\text{D}} = \pi_{2\text{D},1} + \pi_{2\text{D},2} \tag{5.25}$$

[18] See p.32 of notes.

[19] Here, to calculate coefficient c, a quick way is to use idempotency, i.e.

$$\pi_\lambda^2 = \pi_\lambda \ . \tag{5.21}$$

6 Spin Groups and Lie Algebra Isomorphisms

In Question 1 of the second sheet we encountered the following Lie algebra isomorphisms:

$$\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2) \tag{6.1}$$

$$\mathfrak{so}(6) \cong \mathfrak{su}(4) \tag{6.2}$$

Of course, in the question you are asked to construct an explicit basis to show that the Lie algebras are isomorphic[20]. But now we have the questions:

1. Are there any more Lie algebra isomorphisms like this? Is there an exhaustive list?

2. Is there an alternative proof to this, perhaps without using an explicit basis?

The answer to the first question is actually really straightforward - these are known as real form isomorphisms. Here is a list of these 'accidental' isomorphisms of complex classical Lie algebras:

$$\mathfrak{u}(1) \cong \mathfrak{so}(2) \tag{6.3}$$

$$\mathfrak{su}(2) \cong \mathfrak{so}(3) \tag{6.4}$$

$$\mathfrak{su}(2) \oplus \mathfrak{su}(2) \cong \mathfrak{so}(4) \tag{6.1}$$

$$\mathfrak{sp}(2) \cong \mathfrak{so}(5) \tag{6.5}$$

$$\mathfrak{su}(4) \cong \mathfrak{so}(6) \tag{6.2}$$

$$\mathfrak{so}(8) \cong \mathfrak{so}(8) \tag{6.6}$$

The last one (Equation 6.6) is in fact a nontrivial "triality" automorphism of $\mathfrak{so}(8)$ and it is somehow closely related to the exceptional Lie group G_2.

You can definitely see a pattern here - most of the nontrivial ones are of the form $\mathfrak{so}(n)$ where $n \in \{3, 4, 5, 6\}$. This allows me to answer the second question - of course, one can always check this by using some clever way of finding a basis and checking that these is an explicit Lie algebra isomorphism between them. Another way, of course, is to use the Dynkin formalism to show that the Lie algebras have the same Dynkin diagram (hence they must be isomorhpic). These two methods however are not basis-indepedent - the former explicitly involves a choice of basis, but the latter also implicitly uses the Cartan-Weyl basis. The trick to find a basis-independent method is to notice that this isomorphism is related to the theory of Clifford algebras and spin groups. Understanding this is of course way beyond our course, but it is in fact illuminating - you can establish the isomorphism of $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ by first finding an appropriate cover of a Lie group and translating it back to the Lie algebra level by taking the tangent map. This covering business is intrinsically related to how spins arise and in fact are the reason why the isomorphism arises from spin structures in the first place. To fully understand what is going on we will need to first establish some relevant concepts.

[20] I am not allowed to reproduce the answer here sorry - internal policy!

6.1 Universal Covers and Projective Representations

In the lectures we are introduced to the idea that $SU(2)$ is a *two-fold cover* of the manifold $SO(3)$. This is in fact the exact reason why we have spins. To explain this, we need to first understand the relationships between covers and projective representations.

6.1.1 Universal Coverings

We begin by defining what a covering map is [7].

Definition 6.1. A subset $U \subset X$ is **evenly covered** by π if U is connected and open, and each component $\pi^{-1}(U)$ is an open set that is mapped homeomorphically onto U by π.

Definition 6.2. A **covering map** is a continuous surjective map $\pi : \tilde{X} \to X$ such that $\tilde{X}$ is path-connected and locally path-connected[21], and every point $p \in X$ has an evenly covered neighbourhood. We call $\tilde{X}$ the covering space of X and X the base of the covering.

Of course, everything so far is in the topology context. To specialise this to smooth manifolds (which is what we want), we will need to restrict the definition to a very specific type of covering map[22].

Definition 6.3. Take E and M connected smooth manifolds with or without boundary. A map $\pi : E \to M$ is called a **smooth covering map** if π is smooth and surjective, and each point in M has a neighbourhood U such that each component of $\pi^{-1}(U)$ is mapped diffeomorphically onto U by π. We say U is evenly covered. We call M to be the **base manifold**, and E a **covering manifold of** M. If E is simply-connected, it is called the **universal covering manifold of** M.

Here *simply-connected* means every loop is path-isomorphic to a constant path [23]. We want to show that this universal covering exists and is in fact unique. I will here quote a few lemmas and theorems without detailed proof - the details can be found in the references [7, 14].

Theorem 6.1. *Suppose M is a connected smooth-n-manifold, and $\pi : E \to M$ is a topological covering map. Then E is a topological n-manifold, and has a unique smooth structure such that π is a smooth covering map.*

Proof. See Proposition 4.40 of [7]. □

[21] It might be surprising to see how path-connectedness does not generally imply locally path-connectedness. A counterexample is the topologist's sine curve, $y = \sin\left(\frac{1}{x}\right)$ for $x \in (0, \pi)$ together with closed arc connecting $(0, 0)$ and $(\pi, 0)$ where the space is path-connected but not locally path-connected at $(0, 0)$.

[22] By the way, if you are completely baffled by the definitions I just made, these are just mathematical details that you can skip (if you want, but I am weird so I will babble on). Alternatively you should pick up some topology books and start learning what topology is.

[23] Phrased in the language of the fundamental group at X, simply-connectedness simply means that the fundamental group of a manifold at every point $q \in M$ is the trivial group [7, 13].

Corollary 6.1. *If M is a connected smooth manifold, there exists a simply connected manifold $\tilde{M}$ - the* ***universal covering manifold of*** *M, and a smooth covering map $\pi : \tilde{M} \to M$. The universal covering manifold is unique such that for any other universal covering manifold $\tilde{M}'$ with projection map π', then there exists a diffeomorphism $\Phi : \tilde{M} \to \tilde{M}'$ such that $\pi' \circ \Phi = \pi$.*

Proof. This is Corollary 4.43 of [7]. Since a proof is not given there I will give a sketch of the proof. The first step is show that any connected and locally simply connected space admits a unique universal cover. You will need some sort of path-connectedness arguments (see Theorem 12.8 of [14]) - to show the path classes are lifted to the upper space, then check several topological requirements (path-connectedness, topologies, covering maps). Now by Theorem 6.1 you have the existence of a smooth covering manifold of M that is simply-connected. To show uniqueness, we need to find ϕ between any two universal covers that is a diffeomorphism - to show this find open sets such that you can find a surjective smooth submersion of $\pi\big|_{V^{-1}}$ must give you a smooth ϕ and ϕ^{-1} in both directions. □

Now it is straightforward to generalise this to Lie groups.

Theorem 6.2. *Let G be a connected Lie group. There exists a simply connected Lie group $\tilde{G}$, called the universal covering group of G, that admits a smooth covering map $\pi : \tilde{G} \to G$ that is also a Lie group homomorphism.*

Proof. See Theorem 7.7 of [7]. Essentially the idea is you now need to also do group axiom checks on the universal covering group. □

Theorem 6.3. *For any connected Lie group G, the universal covering group is unique in the following sense: if $\tilde{G}$ and $\tilde{G}'$ are connected Lie groups with corresponding smooth covering maps π and π', then there exists a Lie group homomorphism $\Phi : \tilde{G} \to \tilde{G}'$ such that $\pi' \circ \Phi = \pi$.*

Proof. Again - this is similar to the proofs done above. See Theorem 7.9 of [7]. □

What is the point of establishing all this notation? Turns out establishing these universal covering groups on the Lie groups we know and love, say $SO(3)$, are extremely important when it comes to constructing so-called spins!

6.1.2 Projective Representations

Let us first define what a projective representation is [15, 16].

Definition 6.4. Let G be a group and V a finite-dimensional vector space over a field F. A map $\rho : G \to GL(V)$ is a **projective representation** of G over F if there exists a mapping $\alpha : G \times G \to F^*$ such that the following two properties hold:

1. $\rho(x)\rho(y) = \alpha(x,y)\rho(xy) \quad \forall x, y \in G$
2. $\rho(1) = \mathrm{id}_V$

The two conditions imply that α satisfies the following properties:

- $\alpha(x,y)\alpha(xy,z) = \alpha(y,z)\alpha(x,yz) \quad \forall x,y,z \in G$
- $\alpha(x,1) = \alpha(1,x) = 1 \quad \forall x \in G$

We provide an alternative definition most useful to this discussion. We note that for quantum mechanical objects, it is most useful to look at unitary representations[24]. First recall the following:

Definition 6.5. Let $\mathsf{U}(V)$ be the group of invertible linear transformations of a finite-dimensional Hilbert space V over $\mathbb{C}$ that preserve the inner product. A finite-dimensional **unitary representation** of a matrix Lie group G is a continuous homomorphism of $\Pi : G \to \mathsf{U}(V)$ for some finite-dimensional Hilbert space V.

Definition 6.6. Let V be a finite-dimensional Hilbert space over $\mathbb{C}$. The **projective unitary group** over V, denoted $\mathsf{PU}(V)$ is then the quotient group

$$\mathsf{PU}(V) = \mathsf{U}(V)/e^{i\theta}I \tag{6.7}$$

where $e^{i\theta}I$ denotes the group of matrices in $U(1)I$, I being the identity matrix here.

Now let $Q : \mathsf{U}(V) \to \mathsf{PU}(V)$ be the quotient homomorphism and let $q : \mathsf{u}(V) \to \mathsf{pu}(V)$ be the associated Lie algebra isomorphism. We note that given an ordinary unitary representation $\Sigma : G \to \mathsf{U}(V)$, we can always form a projective representation $\Pi : G \to \mathsf{PU}(V)$ by setting $\Pi = Q \circ \Sigma$. This is equivalent to saying the following diagram commutes:

$$\begin{tikzcd} & G \arrow[dl, "\Sigma"'] \arrow[dr, "\Pi"] & \\ \mathsf{U}(V) \arrow[rr, "Q"] & & \mathsf{PU}(V) \end{tikzcd}$$

Note that not all projective representations arise in this fashion. I will state the following propositions without detailed proof.

Proposition 6.1. *If V is a finite-dimensional Hilbert space over $\mathbb{C}$, then $\mathsf{PU}(V)$ is isomorphic to a matrix Lie group. The associated Lie algebra homomorphism q defined above has the kernel $\{iaI\}$, so $\mathsf{pu}(V)$ is isomorphic to $\mathsf{u}(V)/\{iaI\}$.*

Proof. Consider the homomorphism $\Gamma : \mathsf{U}(V) \to GL(gl(V))$, such that for given $U \in \mathsf{U}(V)$, $\Gamma : U \mapsto C_U(X) = UXU^{-1}$. Then one can show that $\ker\Gamma = \{U(1)I\}$, so the image under this homomorphism is isomorphic to the quotient group $\mathsf{U}(V)/\{e^{i\theta}I\}$, compact, and closed, i.e. a matrix Lie group isomorphic to $\mathsf{PU}(V)$. To find the related Lie algebra homomorphism, we note that $c_X(Y) = [X,Y]$, with the kernel of c_X being the scalar multiples of I in $\mathsf{u}(V)$ - the group $\{iaI\}$. The map c_X therefore must map onto $\mathsf{pu}(V)$, giving the required isomorphism. $\square$

[24]Recall this is to make sure that the states are positive-definite so we can define a notion of probability on the Hilbert space.

Every finite-dimensional projective representation can be "de-projectivised" at the Lie-algebra level. To state this we have the following proposition.

Proposition 6.2. *Let $\Pi : G \to \mathsf{PU}(V)$ be a finite-dimensional projective unitary representation of a matrix Lie group G, and $\pi : \mathfrak{g} \to \mathsf{pu}(V)$ be the associated Lie algebra homomorphism. Then there exists a Lie algebra homomorphism $\sigma : \mathfrak{g} \to \mathsf{u}(V)$ such that $\pi(X) = q(\sigma(X)) \quad \forall X \in \mathfrak{g}$. So the following diagram commutes:*

$$\begin{tikzcd} & \mathfrak{g} \arrow[dl, "\sigma"'] \arrow[dr, "\pi"] & \\ \mathsf{u}(V) \arrow[rr, "q"] & & \mathsf{pu}(V) \end{tikzcd}$$

This σ is unique upon fixing that $\operatorname{tr}\sigma(X) = 0 \quad \forall X \in \mathfrak{g}$.

Proof. This proposition boils down to the fact that you can always fix $\sigma(X)$ to have trace zero by choosing for $Y \in \mathfrak{u}(1)$, pick $\sigma(X) = Y + cI$ where c is an appropriate pure-imaginary constant. Such σ therefore always exist. (See Proposition 16.46 of [16] for more details.) □

Now we can say the most important theorem in this subsection:

Theorem 6.4. *Suppose G is a matrix Lie group and $\tilde{G}$ is a universal cover of G with the covering map Φ. Then the following hold:*

1. *Let $\Pi : G \to \mathsf{PU}(V)$ be a finite-dimensional projective unitary representation of G. Then there is an ordinary unitary representation $\Sigma : \tilde{G} \to \mathsf{U}(V)$ of $\tilde{G}$ such that $\Pi \circ \Phi = Q \circ \Sigma$. Any such Σ is irreducible if and only if Π is irreducible. Σ is unique if we choose* $\det(\Sigma(A)) = 1,\ A \in \tilde{G}$.

2. *Let Σ be a finite-dimensional irreducible unitary representation of $\tilde{G}$. Then the kernel of the associated projective unitary representation $Q \circ \Sigma$ contains the kernel of the covering map Φ. Therefore $Q \circ \Sigma$ factors through G and gives rise to a projective unitary representation of G.*

Point 1 is equivalent to saying that the following box diagram commutes:

$$\begin{tikzcd} \tilde{G} \arrow[r, "\Pi"] \arrow[d, "\Phi"'] & \mathsf{U}(V) \arrow[d, "Q"] \\ G \arrow[r, "\Sigma"] & \mathsf{PU}(V) \end{tikzcd}$$

Proof. See Theorem 16.47 of [16]. The idea is actually really simple - we make use of Proposition 6.2 to find an ordinary representation of $\mathfrak{g}$ at the base level, and then simply lift it up and apply Lie's Theorem at the cover level. The second half the theorem rests on the fact that $\ker\Phi$ is a discrete normal subgroup $\tilde{G}$ and is therefore central. We can then show that $\Sigma(A)$, where $A \in \ker\Phi$ under Schur's lemma gives $\Sigma(A) = cI$ as it intertwines V to itself. The A is in the kernel of the associated projective representation $Q \circ \Sigma$. $\square$

Of course, we should be dealing with the infinite-dimensional case. Here of course the unitary representation needs to be defined slightly different[25] The main thing to note here is that we can no longer do the de-projectivisation by passing to the Lie algebra since there is no unique member we can choose - the notion of trace doesn't work for unbounded operators on the Hilbert space. Point 1 in Theorem 6.4 no longer works. However, if G is connected and "semi-simple", every projective unitary representation of G can be de-projectivised after passing to the universal cover. This is in fact the crucial reason why we need to study the universal covering manifolds! The spins intrinsically comes from this universal cover, and it is precisely since we are looking at the de-projectivised version of the representation that brought us there in the first place!

6.2 Spins in Lorentz group

We state without proof the following relation related to the proper orthochronous Lorentz group $SO^+(1,3)$:

$$SO^+(1,3) = \frac{SL(2,\mathbb{C})}{\mathbb{Z}_2} \tag{6.8}$$

In the lecture notes you have precisely looked at the representations of $SL(2,\mathbb{C})$ and noted that we can generated irreducible representations classified by two spins $(j_+, j_-) \in \frac{\mathbb{Z}}{2} \times \frac{\mathbb{Z}}{2}$. The reason for that, by the argument in the previous section, is the fact that the representations that we actually really want to look at are states in our quantum field theory, which are intrinsically projective in nature. By looking at projective unitary representations on $SO^+(1,3)$, we can therefore look at the universal cover $SL(2,\mathbb{C})$ and analyse its unitary (de-projectivised) representations. This is why we have half-spins, i.e. our lovely fermions in our theory!

6.3 Clifford Algebras and Gamma Matrices

With the intuition sketched out in the previous section, let us try and make things concrete. The following discussion mainly follows [1].

We follow from the discussion of the representations of $\mathfrak{so}_n\mathbb{C}$. Recall that tensor powers of the fundamental and anti-fundamental representations of $\mathfrak{su}_n\mathbb{C}$ completely characterises all the representations of $\mathfrak{su}_n\mathbb{C}$, but this is not the case for the representations of $\mathfrak{so}_n\mathbb{C}$. As a matter of fact, only half of the representations of $\mathfrak{so}_n\mathbb{C}$ arises in this way. As we have seen above, this could be attributed to the different underlying topology of the two groups - the missing representations are exactly the $\mathrm{Spin}_m\mathbb{C}$ reps where $\mathrm{Spin}_m\mathbb{C} \to SO_m\mathbb{C}$ is a double cover.

How do we generate this double cover then? The theory of projective representations, as described above, seems to be hard to access from a construction point-of-view - where do

[25] You will need some sort of a strong continuity homomorphism $\Phi : G \to U(\boldsymbol{H})$. You can read more about this in [16].

we even start? It turns out to be helpful to define something called *Clifford Algebras*. Each complex Clifford algebra contains an orthogonal Lie algebra as a subalgebra - it is isomorphic either to a matrix algebra or a sum of two. This determines one or two representations of the orthogonal Lie algebras, as we will see, which are exactly the representations that we are missing from the tensor product constructions.

So, first a definition. In this section we fix V to be a vector space and $B : V \times V \to \mathbb{K}$ to be a symmetric bilinear form on V (here $\mathbb{K}$ is a field). To construct a corresponding quadratic form Q, we can define

$$Q(x) = B(x, x) \tag{6.9}$$

such that (V, Q) is a **quadratic vector space over** $\mathbb{K}$. One can conversely reconstruct the symmetric bilinear form B from Q by polarisation, i.e.

$$B(x, y) = \frac{1}{2} \left(Q(x + y) - Q(x) - Q(y) \right) \tag{6.10}$$

so we hereby denote Q and B interchangeably, writing $Q(x) = Q(x, x)$.

Now we have the following definition.

Definition 6.7. Let A be an associative $\mathbb{K}$-algebra and (V, Q) be a quadratic vector space. A $\mathbb{K}$-linear map $\phi : V \to A$ is **Clifford** if $\forall x \in V$,

$$\phi(x)^2 = Q(x) 1_A \tag{6.11}$$

where 1_A is the unit of A.

Definition 6.8. The **Clifford algbera** $C = C(Q) = \mathrm{Cliff}(V, Q)$ is an associative algebra with unit 1 and is generated by V, with $\forall v \in V$

$$v \cdot v = Q(v, v) \cdot 1 \tag{6.12}$$

Equivalently (if the characteristic of k is not 2), we have $\forall v, w \in V$,

$$v \cdot w + w \cdot v = 2Q(v, w) \tag{6.13}$$

A note about construction. The Clifford algebra can be constructed quickly by taking the tensor algebra

$$T^\bullet(V) = \bigoplus_{n \geq 0} V^{\otimes n} \tag{6.14}$$

and setting

$$C(Q) = \frac{T^\bullet(V)}{I(Q)} \tag{6.15}$$

Here $I(Q)$ is the two-sided ideal generated by all elements of the form $v \otimes v - Q(v, v) \cdot 1$. Clearly $C(Q)$ satisfies the universal property. From this we can see that the dimension of C is 2^m where $m = \dim(V)$ and that the canonical mapping $V \to C$ is an embedding, with the basis of $C(Q)$ being the products $e_I = e_{i_1} \cdot e_{i_2} \cdot ... \cdot e_{i_k}$ where e_i are the basis of V. To see this, in particular, we can check the following.

Proposition 6.3. *There is a natural embedding $V \hookrightarrow C(Q)$ which is the image of $V = V^{\otimes 1}$ under the canonical projection*

$$\pi_q : T^\bullet(V) \to C(Q) \ , \tag{6.16}$$

and this is an injection.

Proof. Say that an element $\varphi \in T^\bullet(V)$ is of pure degree s if $\varphi \in V^{\otimes s}$. We want to show that any element $\varphi \in T^\bullet(V) \cap V$ is zero. Suppose this is not true. Then we can write $\varphi = \sum_i a_i \otimes (v_i \otimes v_i + Q(v_i)) \otimes b_i$ where we assume that a_i and b_i is of pure degree. Now since $\varphi \in V$ we must have that the expression is equal to zero, with the sum taken over those indices with $\deg a_i + \deg b_i$ maximal. Contracting with Q means $\sum_i a_i Q(v_i) \cdot b_i = 0$. Proceed with induction to show $\varphi = 0$. □

The Clifford Algebra has a universal property as follows. This also gives a categorical definition of Clifford Algebras.

Proposition 6.4. *The Clifford algebra can be defined to be the universal algebra with the following property: If A is any associative algebra with unit and a linear mapping $j : V \to A$ is given such that*

$$j(v) \cdot j(v) = Q(v,v) \cdot 1, \quad \forall v \in V \ , \tag{6.17}$$

or equivalently $\forall v, w \in V$, (given that k has a characteristic not equal to 2,)

$$j(v) \cdot j(w) + j(w) \cdot j(v) = 2Q(v,w) \cdot 1 \tag{6.18}$$

then there should be a unique homomorphism of algebras from $C(Q)$ to A extending j, i.e. j extends uniquely to a $\mathbb{K}$-algebra homomorphism $\tilde{j} : C(Q) \to A$, and $C(Q)$ is the unique associative $\mathbb{K}$-algebra with this property.

Proof. Any linear map $j : V \to A$ extends to a unique algebra homomorphism $\bar{j} : T^\bullet(V) \to A$. Now Eq. (6.17) implies that $\bar{j} = 0$ on $I(Q)$ so $\bar{j}$ descends to $C(Q)$. Suppose now B is an associative $\mathbb{K}$-algebra with unit and that $\iota : V \to B$ is an embedding with the property that any linear map $j : V \to A$ with the property in Eq. (6.17) extends uniquely to an algebra homomorphism $\tilde{j} : A \to B$. Then the isomorphism from $V \subset Q$ to $\iota(V) \subset B$ clearly induces an algebra isomorphism $C(Q) \xrightarrow{\cong} B$. □

The proposition above effectively states the following. Given an associative algebra with unit A, together wiht a Clifford map $i : V \to C(Q)$ such that for every Clifford map $\phi : V \to A$ there is a unique algebra morphism $\Phi : C(Q) \to A$ that makes the following triangle commute.

$$\begin{tikzcd} & V \arrow[dl, "i"'] \arrow[dr, "\phi"] & \\ C(Q) \arrow[rr, "\Phi"] & & A \end{tikzcd}$$

Categorically, the Clifford Algebra is an initial object in the category $\mathbf{Cliff}(V, Q)$, which has Clifford maps $\phi : V \to \cdot$ from a fixed vector space equipped with a quadratic form Q as objects. The morphism from $V \to A$ to $V \to A'$ is given by a commuting triangle

$$\begin{tikzcd} & V \arrow[dl] \arrow[dr] & \\ A \arrow[rr, "f"] & & A' \end{tikzcd}$$

with $f : A \to A'$ as a homomorphism of associative algebras. This initial object is unique up to a unique isomorphism. In other words, the Clifford algebra $C(Q)$ is universal for Clifford maps to associative algebras. The construction via tensor algebra as before implies the following statement. If $\phi : V \to A$ is a Clifford map and $\tilde{\Phi} : T^\bullet(V) \to A$ is the unique extension of ϕ to the tensor algebra, then $\tilde{\Phi}$ indeed annihilates the ideal $I(Q)$ and therefore factors through a unique map $\Phi : T^\bullet(V)/I(Q) \to A$ from the quotient. Therefore, we have a commutative diagram:

$$\begin{tikzcd} V \arrow[r] \arrow[dr, "i"'] \arrow[drr, dotted] & T^\bullet(V) \arrow[d, "\phi"] \arrow[dr, "\tilde{\Phi}"] & \\ & C(Q) \arrow[r] & A \end{tikzcd}$$

Here i is really injective as the ideal only comes into play for $V^{\geq \otimes 2}$.

6.3.1 Constructing Clifford algebras

The way we have been discussing about Clifford algebras is not very suitable for computations. Instead, we will discuss the way that Clifford introduced the algebras. This is the way Clifford algebras are still taught in physics courses, following Dirac.

Traditionally, the discussion of Clifford algebras started with Dirac matrices.

Definition 6.9. Suppose $\{e_i\}$ is a $\mathbb{K}$-basis for V, where $i = 1, ..., \dim V$. The vector space V is equipped with the symmetric bilinear form where $B(e_i, e_j) = B_{ij} = B_{ji}$. The **Clifford generators** Γ_i is the image of e_i under the map $i : V \to C(Q)$, which satisfy the relations

$$\Gamma_i \Gamma_j + \Gamma_j \Gamma_i = 2 B_{ij} \mathbb{1} \tag{6.19}$$

where $\mathbb{1}$ is the unit in the Clifford algebra $C(Q)$.

Following this, we can define Clifford algebras by using generators in the following manner.

Definition 6.10 (Clifford algebras - generators)**.** An associative algebra over field $\mathbb{K}$ with unity 1 is the **Clifford algebra** $C(Q)$ of a non-degenerate quadratic form Q on V if it contains V and $\mathbb{K} = \mathbb{K} \cdot 1$ as distinct subspaces such that the following three conditions hold:

(i) $v^2 = Q(v)$ for any $v \in V$.

(ii) V generates $C(Q)$ as an algebra over $\mathbb{K}$.

(iii) $C(Q)$ is not generated by any proper subspace of V.

We can immediately how the Dirac matrices furnishes a representation of the Clifford algebra. We then define

$$\Gamma_{ij} = \frac{1}{2}\left(\Gamma_i\Gamma_j - \Gamma_j\Gamma_i\right) \tag{6.20}$$

as the product of two generators. More generally, we have

$$\Gamma_{i_1 \ldots i_p} = \frac{1}{p!}\sum_{\sigma \in S_p}(-1)^\sigma \Gamma_{i_{\sigma(1)}} \ldots \Gamma_{i_{\sigma(p)}} \tag{6.21}$$

where $(-1)^\sigma$ indicates the sign of the permutation in S_p. We then see that since $C(Q)$ is generated by V and the identity it is the linear span of $\mathbb{1}$, Γ_i, Γ_{ij},... in total there are $1 + n + C_2^n + \ldots C_n^n = 2^n$ monomials. So $\dim C(Q) = 2^{\dim V}$. This is the same dimension as the exterior algebra $\bigwedge^\bullet V$ so we can establish a vector space isomorphism between the two.

In particular, if we use an orthonormal basis to generate $C(Q)$, then the first condition in the above Definition 6.10 then becomes

$$\Gamma_i^2 = 1, \quad 1 \le i \le p, \tag{6.22}$$

$$\Gamma_i^2 = -1, \quad p < i \le p, \tag{6.23}$$

$$\Gamma_i\Gamma_j = -\Gamma_j\Gamma_i, \quad i < j. \tag{6.24}$$

whilst condition (iii) becomes

$$\Gamma_1 \ldots \Gamma_n \neq \pm 1 \tag{6.25}$$

This is important in constructing a representation of Clifford algebras in general dimensions. Typically, this is needed in the discussion of supersymmetry and supergravity (and spinors) in various dimensions. The construction typically involves a set of matrices called **Dirac matrices** or **Gamma matrices**, defined as matrix representations of the Clifford algebra in various dimensions. You should have seen **Pauli matrices** in your elementary quantum mechanics courses:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.26}$$

These matrices can generate a basis for Clifford algebras of arbitrary dimensions. Here we follow the discussion in [17–19]. We construct the Euclidean γ-matrices from Gamma matrices which are the basic building block of the matrix representations of the Clifford algebras. We define $(2k+1)$-matrices by the tensor products of k Pauli matrices to get a

$2^k \times 2^k$ matrix representation as follows:

$$\Gamma_1^{(k)} = \sigma_1 \otimes \underbrace{\sigma_0 \otimes ... \otimes \sigma_0}_{k-1} \ , \quad \Gamma_2^{(k)} = \sigma_2 \otimes \underbrace{\sigma_0 \otimes ... \otimes \sigma_0}_{k-1} \ ,$$
$$\Gamma_3^{(k)} = \sigma_3 \otimes \sigma_1 \otimes \underbrace{\sigma_0 \otimes ... \otimes \sigma_0}_{k-2} \ , \quad \Gamma_4^{(k)} = \sigma_3 \otimes \sigma_2 \otimes \underbrace{\sigma_0 \otimes ... \otimes \sigma_0}_{k-2} \ ,$$
$$\vdots$$
$$\Gamma_{2k-1}^{(k)} = \underbrace{\sigma_3 \otimes ... \otimes \sigma_3}_{k-1} \otimes \sigma_1 \ , \quad \Gamma_{2k}^{(k)} = \underbrace{\sigma_3 \otimes ... \otimes \sigma_3}_{k-1} \otimes \sigma_2 \ ,$$
$$\Gamma_{2k+1}^{(k)} = \underbrace{\sigma_3 \otimes ... \otimes \sigma_3}_{k} \tag{6.27}$$

The matrices listed above can be generated using the recurring relations:

$$\Gamma_M^{(k+1)} = \Sigma_M^{(k)} \otimes \sigma_0 \ , \quad M = 1, ..., 2k \tag{6.28}$$

$$\Gamma_{2k+i}^{(k+1)} = \Sigma_{2k+1}^{(k)} \otimes \sigma_i \ , \quad i = 1, 2, 3 \tag{6.29}$$

which gives

$$\left\{ \Gamma_M^{(k)}, \Gamma_N^{(k)} \right\} = 2\delta_{MN} \tag{6.30}$$

So then we have the following definition.

Definition 6.11 (Gamma matrices)**.** The **Gamma** or **Dirac matrices** are matrix representations of the Clifford algebras, i.e. the map: $\Gamma : C(Q) \to \mathrm{GL}(\mathbb{C}^{2^k})$ where we map the generators $e_M \mapsto \Gamma_M$. The representation is faithful when $d = 2k$ and non-faithful when $d = 2k + 1$ where $\Gamma(\epsilon) = \Gamma_1^{(k)} ... \Gamma_{2k+1}^{(k)} = i^k$.

We will find Gamma matrices extremely helpful later when we construct spinors in spaces of Euclidean and Lorentzian signatures.

6.3.2 $\mathbb{Z}_2$-grading

Let us return to the tensor construction of the Clifford algebras. Since the ideal $I(Q)$ is not homogeneous, $C(Q)$ does not inherit a $\mathbb{Z}$-grading from $T^\bullet(V)$. However, notice that the ideal $I(Q)$ is generated by elements of an even degree. This means the Clifford algebra does inherit a $\mathbb{Z}_2$ grading. To study this grading recall the following definitions from elementary algebra.

Definition 6.12. A **graded ring** is a ring that is decomposed into a direct sum of additive groups

$$R = \bigoplus_{n=0}^{\infty} R_n = R_0 \oplus R_1 \oplus R_2 \oplus ... \tag{6.31}$$

such that

$$R_m R_n \subseteq R_{m+n} \tag{6.32}$$

for all nonnegative integers m and n.

Definition 6.13. An associative algebra A over a ring R is **graded** if it is graded as a ring.

So we can now go back to Clifford Algebras. Consider the automorphism $\alpha : C(Q) \to C(Q)$ which sends $\alpha(v) = -v$ on V. Since $\alpha^2 = \mathrm{id}$, the ideal $I(Q)$ is generated by elements of an even degree, and hence Clifford algebra inherits a $\mathbb{Z}_2$ grading:

$$C(Q) = C^0(Q) \oplus C^1(Q) \tag{6.33}$$

where $C^i(Q) = \{\varphi \in C(Q) \,|\, \alpha(\varphi) = (-1)^i \varphi\}$ are the eigenspaces of α. Since α is a homomorphism, we have

$$C^i(Q) \cdot C^j(Q) \subset C^{i+j}(Q) \tag{6.34}$$

with the indices taken modulo 2. This $\mathbb{Z}_2$-grading plays an important role in the analysis and application of Clifford algebras. In particular, $C^0(Q)$ is often called $C^{\text{even}}(Q)$ and is a subalgebra of dimension 2^{m-1}, where as $C^1(Q)$ is often called $C^{\text{odd}}(Q)$.

6.3.3 Grading and exterior algebras

The $\mathbb{Z}_2$-gradedness of the Clifford algebra is very different from the graded nature of the tensor algebra which inherently has a $\mathbb{Z}$-graded structure. To see this, define $\tilde{\mathcal{F}}$ as

$$\tilde{\mathcal{F}}^r = \sum_{s \le r} V^{\otimes s} \; . \tag{6.35}$$

This has the property

$$\tilde{\mathcal{F}}^r \otimes \tilde{\mathcal{F}}^s \subset \tilde{\mathcal{F}}^{r+s} \; . \tag{6.36}$$

The tensor algebra therefore has a natural filtration

$$\tilde{\mathcal{F}}^0 \subset \tilde{\mathcal{F}}^1 \subset \ldots \subset T^\bullet(V) \; , \tag{6.37}$$

which makes the tensor algebra into a **filtered algebra**. Every filtered algebra has an associated graded algebra. For the tensor algebra with the canonical filtration described above, the associated graded algebra is described by

$$\tilde{\mathcal{G}}^p = \tilde{\mathcal{F}}^p / \tilde{\mathcal{F}}^{p-1} \tag{6.38}$$

Then $\tilde{\mathcal{G}}^\bullet$ is a graded algebra where the product map is defined by

$$\tilde{\mathcal{G}}^p \times \tilde{\mathcal{G}}^q \to \tilde{\mathcal{G}}^{p+q} \; . \tag{6.39}$$

The canonical filtration of the tensor algebra $T^\bullet(V)$ defines a natural filtration on the Clifford algebra $C(Q)$. Suppose $\pi_q : T^\bullet(V) \to T^\bullet(V)/I(Q)$ where $I(Q)$ is the ideal that generates the Clifford algebras. Then $\mathcal{F}^i = \pi_q(\tilde{\mathcal{F}}^i)$ naturally has a natural filtration,

$$\mathcal{F}^0 \subset \mathcal{F}^1 \subset \mathcal{F}^2 \subset \ldots \; , \tag{6.40}$$

and naturally the associated graded algebra $\mathcal{G}^r = \mathcal{F}^r / \mathcal{F}^{r-1}$ naturally inherits the filtration. We now have the following proposition.

Proposition 6.5. *For any quadratic form Q, the associated graded algebra of $C(Q)$ is naturally isomorphic to the exterior algebra $\bigwedge^\bullet V$.*

Proof. The map $\bigotimes^r V \xrightarrow{\pi_r} \mathcal{F}^r \to \mathcal{G}^r = \mathcal{F}^r/\mathcal{F}^{r-1}$ given by $v_{i_1}\otimes...\otimes v_{i_r} \mapsto [v_{i_1}...v_{i_r}]$ descends to a map $\bigwedge^r V \to \mathcal{F}^r$ by the property in Eq. (6.18). (Note that when the characteristic of $\mathbb{K}$ is 2 then we will have to use the other condition.) This map is surjective and gives a homomorphism of graded algebras $\bigwedge^\bullet V \to \mathcal{G}^\bullet$. It remains to check the map is injective. The kernel of $\bigotimes^r V \to \mathcal{G}^r$ consists of the r-homogeneous pieces of elements $\varphi \in I_q(V)$ of degree less than r. Any such φ can be written as a finite sum $\varphi = \sum a_i\otimes(v_i \otimes v_i + q(v_i))\otimes b_i$ where $v_i \in V$ and where we may assume that the a_i and b_i are of pure degree with $\deg a_i + \deg b_i \leq r-2$. The r-homogeneous part of φ is then of the form $\varphi_r = \sum a_i \otimes v_i \otimes v_i \otimes v_i$ where $\deg a_i + \deg b_i = r-2$ for each i. The image of φ in the exterior algebra is however zero as $v_i \wedge v_i = 0$. So the map $\bigwedge^r V \to \mathcal{G}^r$ is injective. □

Note that the proposition above gives a canonical vector space isomorphism that is compatible with the filtrations as follows,

$$\bigwedge\nolimits^\bullet V \to C(Q)\ . \tag{6.41}$$

The map in Eq. (6.41) is of course not an isomorphism of algebras unless $q = 0$. However the map is indeed canonical so we can discuss embeddings of the form $\bigwedge^r V \subset C(Q)$ for all $r \geq 0$. To see that the isomorphism is only true when $q = 0$, consider the $\mathbb{Z}_2$-grading on the tensor algebra defined with $T^\bullet(V) = T^\bullet(V)_0 + T^\bullet(V)_1$ where

$$T^\bullet(V)_0 = \bigoplus_{k\geq 0} V^{\otimes 2k}\ , \quad T^\bullet(V)_1 = \bigoplus_{k\geq 0} V^{\otimes 2k+1}\ . \tag{6.42}$$

where the $\mathbb{Z}_2$-grading is the reduction mod-2 of the $\mathbb{Z}$-grading of the tensor algebra as discussed above. This reduction makes the ideal I_q homogeneous, and hence the projection $T^\bullet(V) \to C(Q)$ restricts to projections $TV_i \to C_i$ for $i = 0, 1$. Note however that for $i = 1$ this is only a projection of vector spaces, since neither TV_1 nor C_1 are algebras.

Now the canonical filtration of the tensor algebra $T^\bullet(V)$ defines a filtration on $C(Q)$ as follows. By filtering $T^\bullet(V)_0$ and $T^\bullet(V)_1$ separately, i.e.

$$\mathcal{F}^{2k}T(V)_0 = \bigoplus_{l\leq k} V^{\otimes 2l}\ , \quad \mathcal{F}^{2k+1}T(V)_1 = \bigoplus_{l\leq k} V^{\otimes 2l+1} \tag{6.43}$$

such that

$$0 \subset \mathcal{F}^0 T(V)_0 \subset \mathcal{F}^2 T(V)_0 \subset ... \tag{6.44}$$
$$0 \subset \mathcal{F}^1 T(V)_1 \subset \mathcal{F}^3 T(V)_1 \subset ... \tag{6.45}$$

Now under the projections $TV_0 \to C_0$ and $TV_1 \to C_1$, we can similarly identify the filtrations of the Clifford algebra as

$$0 \subset \mathcal{F}^0 C_0 \subset \mathcal{F}^2 C_0 \subset ... \tag{6.46}$$
$$0 \subset \mathcal{F}^1 C_1 \subset \mathcal{F}^3 C_1 \subset ... \tag{6.47}$$

We will henceforth use the shorthand $\mathcal{F}^pC$ as $\mathcal{F}^pC_0$ and $\mathcal{F}^pC_1$ if p is even and odd respectively. Now we note that $\mathcal{F}^pC/\mathcal{F}^{p-2}C \cong \bigwedge^p V$ as the corrections in replacing xy by $-yx$ where $x, y \in V$ involve terms of degree 2 less. The corrections are 0 when $q = 0$, so we can identify $Cl(V, 0) \cong \bigwedge^\bullet V$, exactly as mentioned above.

It is possible to understand the relation between the Clifford and exterior algebras in a different way which does not involve filtrations. The bilinear form B defines a map $\flat : V \to V^*$ where $x \mapsto B(x, \cdot)$. The map $flat$ is an isomorphism if and only if B is non-degenerate. The inverse is typically defined as $\sharp$ so together with the map $\flat$ they are referred to as the musical isomorphisms induced from the inner product B. We can then define a linear map $\phi : V \to \operatorname{End} \bigwedge^\bullet V$ by

$$\phi(x)\alpha = x \wedge \alpha\iota_x\alpha \tag{6.48}$$

where ι_x is the unique odd derivation defined by $\iota_x 1 = 0$ and $\iota_x y = B(x, y)$ for $y \in V$. So on a monomial we have,

$$\iota_x \left(y_1 \wedge ... \wedge y_p\right) = \sum_{i=1}^{p} (-1)^{i-1} B(x, y_i) y_1 \wedge ... \wedge \hat{y}_i \wedge ... \wedge y_p \ , \tag{6.49}$$

where the hat denotes omission. Then we can extend this linearly to all of $\bigwedge^\bullet V$ as in the following lemma.

Lemma 6.1. *The map $\phi : V \to \operatorname{End} V$ in Eq. (6.48) is Clifford.*

Proof. For every $x \in V$ and $\alpha \in \operatorname{End} \bigwedge^\bullet V$, we have

$$\begin{aligned} \phi(x)^2\alpha &= \phi(x)\left(x \wedge \alpha - \iota_x\alpha\right) \\ &= x \wedge x \wedge \alpha - x \wedge \iota_x\alpha - Q(x)\alpha + x \wedge \iota_x\alpha + \iota_x\iota_x\alpha \\ &= -Q(x)\alpha \ , \end{aligned} \tag{6.50}$$

where $x \wedge x = 0 = \iota_x\iota_x$ and $\iota_x(x \wedge \alpha) = Q(x)\alpha - x \wedge \iota_x\alpha$. □

By the universality of Clifford algebras we can then extend this to the algebra homomorphism uniquely,

$$\Phi : C(Q) \to \operatorname{End} \textstyle\bigwedge^\bullet V \tag{6.51}$$

So composing this with the evaluation at $1 \in \bigwedge^\bullet V$ gives a linear map $\Phi_1 : C(Q) \to \bigwedge^\bullet V$. This map obeys $\Phi_1(1) = 1$ and if $x \in V$ then $\Phi_1(i(x)) = x$ where $i : V \to C(Q)$. Since i is injective from the construction of $C(Q)$, $\Phi_1 \circ i$ is also injective. By further computations, we then get

$$\Phi_1(i(x)i(y)) = x \wedge y - B(x, y) \ , \tag{6.52}$$

and

$$\Phi_1(i(x)i(y)i(z)) = x \wedge y \wedge z - B(x, y)z + B(x, z)y - B(y, z)x \ , \tag{6.53}$$

so Φ_1 surjects onto $\bigwedge^\bullet V$. This is a vector space isomorphism with the inverse map defined by

$$y_1 \wedge ... \wedge y_p \mapsto \frac{1}{p!} \sum_{\sigma \in S_p} (-1)^\sigma y_{i_{\sigma(1)}} \wedge ... \wedge y_{i_{\sigma(p)}} \tag{6.54}$$

which gives an explicit quantisation of the exterior algebra.

6.3.4 Clifford algebras as representations of $\mathfrak{so}_n\mathbb{C}$

It is clear from the construction that Clifford algebras are associative algebras. As a results it determines a Lie algebra with the bracket defined by the associative multiplication. How are Clifford algebras related to the representations of $\mathfrak{so}_m\mathbb{C}$? To see this we will first embed the Lie algebra $\mathfrak{so}(Q)$ inside the Lie algebra of the even part of the Clifford algebra and from there identify $C(Q)$ with one or two copies of matrix algebras.

Let us see how this works in practice. We first need to make an explicit isomorphism of $\bigwedge^2 V$ with $\mathfrak{so}(Q)$, which is defined as

$$\mathfrak{so}(Q) = \{X \in \mathrm{End}(V) \,|\, Q(Xv, w) = -Q(v, Xw)\ \forall v, w \in V\} \ . \tag{6.55}$$

Define the map

$$\varphi_{a\wedge b}(v) = 2\left(Q(b,v)a - Q(a,v)b\right) \ , \tag{6.56}$$

which gives the isomorphism $\phi : \bigwedge^2 V \to \mathfrak{so}(Q) \subset \mathrm{End}(V)$ with $a \wedge b \mapsto \varphi_{a\wedge b}$. One can check that the bracket on $\bigwedge^2 V$ makes this an isomorphism of Lie algebras with the Clifford algebra[26], allowing the map $\psi : \bigwedge^2 V \to C(V,Q)$ to be defined by

$$\psi(a\wedge b) = \frac{1}{2}(a\cdot b - b\cdot a) = a\cdot b - Q(a,b) \ . \tag{6.57}$$

This is an injective embedding, which shows that the following Lemma.

Lemma 6.2. *The mapping $\psi \circ \varphi^{-1} : \mathfrak{so}(Q) \to C(Q)^{even}$ embeds $\mathfrak{so}(Q)$ as a Lie subalgebra of $C(Q)^{even}$.*

Proof. See discussion above. □

The reason why the embedding only goes into the even part is because, simply, $C(Q)^{\mathrm{odd}}$ is indeed not an algebra. You can also see that in Eq. (6.57) the map is defined with elements of even degree. By looking at the basis elements we can then see that ψ is an embedding and the map exactly maps the exterior algebra to the even part of the Clifford algebra.

What remains is to identify the subalgebra of $C(Q)_0$ or $C(Q)^{\mathrm{even}}$, the image of $\mathfrak{so}(Q)$ as matrix algebras. Let us separate this into two cases.

Case 1: $n = \dim V$ is even.

We first decompose V into two n-dimensional isotropic spaces for Q,

$$V = W \oplus W' \ . \tag{6.58}$$

Then we have the following lemma.

Lemma 6.3. *The decomposition $V = W \oplus W'$ determines an isomorphism of algebras,*

$$C(Q) \cong \mathrm{End}(\textstyle\bigwedge^\bullet W) \tag{6.59}$$

where $\bigwedge^\bullet W = \bigwedge^0 W \oplus ... \oplus \bigwedge^n W$.

[26] This is done by checking the brackets on $[a\wedge b, c\wedge d]$ and $[a\cdot b, c\cdot d]$.

Proof. Let us try and construct the map $\varphi : C(Q) \to E = \mathrm{End}(\bigwedge^\bullet W)$. The map φ is the same as defining a linearly mapping $V \to E$ with the condition in Eq. (6.18). We must therefore construct maps $l : W \to E$ and $l' : W' \to E$ such that

$$l(w)^2 = 0 = l'(w')^2 \tag{6.60}$$

$$l(w) \circ l'(w') + l'(w') \circ l(w) = 2Q(w, w')I \;, \tag{6.61}$$

for any $w \in W$, $w' \in W$. For each $w \in W$, let $L_w \in E$ be the left multiplication by w on the exterior algebra $\bigwedge^\bullet W$,

$$L_w(\xi) = w \wedge \xi, \quad \xi \in \textstyle\bigwedge^\bullet W \;. \tag{6.62}$$

For any $\vartheta \in W^*$, let $D_\vartheta \in E$ be the derivation of $\bigwedge^\bullet W$ such that,

$$D_\vartheta(1) = 0 \tag{6.63}$$

$$D_\vartheta(w) = \vartheta \in \textstyle\bigwedge^0 W = \mathbb{C} \tag{6.64}$$

$$D_\vartheta(\zeta \wedge \xi) = D_\vartheta \zeta \wedge \xi + (-1)^{\deg(\zeta)} \zeta \wedge D_\vartheta(\xi) \;, \tag{6.65}$$

where $w \in W = \bigwedge^1 W$. i.e. Explicitly,

$$D_\vartheta(w_1 \wedge ... \wedge w_r) = \sum_i (-1)^{i-1} \vartheta(w_i) \left(w_1 \wedge ... \wedge \hat{w}_i \wedge ... \wedge w_r\right) \;. \tag{6.66}$$

Now we can set

$$l(w) = L_w \;, \quad l'(w') = D_\vartheta \;, \tag{6.67}$$

where $\vartheta \in W^*$ is defined by $\vartheta(w) = 2Q(w, w'), \quad \forall w \in W$. It is straightforward to show that the maps defined obeys the requirements, as well as for $\zeta \wedge \xi$ if they obey for ζ and ξ separately. The map is clearly an isomorphism and one can see that by its action of a basis. □

Now note that there exists a decomposition of the exterior powers into even and odd parts $\bigwedge^\bullet W = \bigwedge^{\mathrm{even}} W \oplus \bigwedge^{\mathrm{odd}} W$ where $C(W)^{\mathrm{even}}$ respects the splitting. From Lemma 6.3, we then have the isomorphism,

$$C(Q)^{\mathrm{even}} \cong \mathrm{End}(\textstyle\bigwedge^{\mathrm{even}} W) \oplus \mathrm{End}(\bigwedge^{\mathrm{odd}} W) \;. \tag{6.68}$$

Combining this with Lemma 6.2, we then have an embedding of Lie algebras,

$$\mathfrak{so}(Q) \subset C(Q)^{\mathrm{even}} \cong \mathfrak{gl}(\textstyle\bigwedge^{\mathrm{even}} W) \oplus \mathfrak{gl}(\bigwedge^{\mathrm{odd}} W) \;, \tag{6.69}$$

and we find that there are two representations of $\mathfrak{so}(Q) = \mathfrak{so}_{2n}\mathbb{C}$. We denote the two representations by,

$$S^+ = \textstyle\bigwedge^{\mathrm{even}} W \;, \quad S^- = \bigwedge^{\mathrm{odd}} W \;. \tag{6.70}$$

Proposition 6.6. *The representations $S^\pm$ are the irreps of $\mathfrak{so}_{2n}\mathbb{C}$ with highest weights $\alpha = \frac{1}{2}(L_1 + ... + L_n)$ and $\beta = \frac{1}{2}(L_1 + ... + L_{n-1} - L_n)$. More precisely, we have,*

$$S^+ = \Gamma_\alpha \;, \quad S^- = \Gamma_\beta \;, \quad \text{if } n \text{ is even} \;; \tag{6.71}$$

$$S^+ = \Gamma_\beta \;, \quad S^- = \Gamma_\alpha \;, \quad \text{if } n \text{ is odd} \;. \tag{6.72}$$

Proof. We need to show that the natural basis vectors $e_I = e_{i_1} \wedge ... \wedge e_{i_k}$ for $\bigwedge^\bullet W$ are weight vectors. Tracing through the isomorphisms, we find that $H_i = E_{i,i} - E_{n+i,n+i}$ in $\mathfrak{h} \subset \mathfrak{so}_{2n}\mathbb{C}$ corresponds to $\frac{1}{2}(e_i \wedge e_{n+i})$ in $\bigwedge^2 V$, which corresponds to $\frac{1}{2}(e_i \cdot e_{n+i} - 1)$ in $C(Q)$, and this maps to,

$$\frac{1}{2}\left(L_{e_i} \circ D_{2e_i^*} - I\right) = L_{e_i} \circ D_{e_i^*} - \frac{1}{2} I \in \mathrm{End}(\bigwedge^\bullet W)\ . \tag{6.73}$$

We can compute,

$$L_{e_i} \circ D_{e_i^*}(e_I) = \begin{cases} e_I & \text{if } i \in I \\ o & \text{if } i \notin I \end{cases} \tag{6.74}$$

So e_I spans a weight space with weight $\frac{1}{2}\left(\sum_{i\in I} L_i - \sum_{j\notin I} L_j\right)$. All such weights with given $|I| \mod 2$ are congruent by the Weyl group (they are equivalent up to transformations of the Weyl group), so $S^+ = \bigwedge^{\mathrm{even}} W$ and $S^- = \bigwedge^{\mathrm{odd}} W$ must be an irreducible representation. The highest weights are then straightforward to read off - the one for $\bigwedge^{\mathrm{even}} W$ is $\alpha = \frac{1}{2}\sum_i L_i$ if n is even and β if n is odd. □

Definition 6.14. The representations $S^\pm$ are the **half-spin representations** of $\mathfrak{so}_{2n}\mathbb{C}$, and $S = S^+ \oplus S^- = \bigwedge^\bullet W$ is called the **spin representation**. Elements of S are known as **spinors**.

We are going to come back to spinor representations in the next section.

Case 2: $n = \dim V$ **is odd.**

This time we decompose the space V as follows,

$$V = W \oplus W' \oplus U\ .,\tag{6.75}$$

where W and W' are n-dimensional isotropic spaces and U is a one-dimensional space perpendicular to them. For the standard Q on $\mathbb{C}^{2n+1}$ these are spanned by the first n, second n, and the last basis vector. We then have the following lemma.

Lemma 6.4. *The decomposition $V = W \oplus W' \oplus U$ determines an isomorphism of algebras,*

$$C(Q) \cong \mathrm{End}(\bigwedge^\bullet W) \oplus \mathrm{End}(\bigwedge^\bullet W')\ . \tag{6.76}$$

Proof. We can proceed exactly as the even case, as in Lemma 6.3. The only difference is with the element u_0 where $Q(u_0, u_0) = 1$. We send u_0 to the endomorphism that is the identity on $\bigwedge^{\mathrm{even}} W$ and minus the identity on $\bigwedge^{\mathrm{odd}} W$. This involution then skew commutes with all L_w and D_ϑ, which means the map $V \to E = \mathrm{End}(\bigwedge^\bullet W)$ determines an algebra homomorphism from $C(Q) \to E$. The map for $\mathrm{End}(\bigwedge^\bullet W')$ is similar but with the roles of W and W' reversed. The maps are isomorphic by checking the basis elements. □

From Lemma 6.4 we see that the subalgebra $C(Q)^{\mathrm{even}} \subset C(Q)$ ia mapped isomorphically onto the factors,

$$C(Q)^{\mathrm{even}} \cong \mathrm{End}(\bigwedge^\bullet W) \tag{6.77}$$

which gives a representation $S = \bigwedge^\bullet W$ of Lie algebras,

$$\mathfrak{so}_{2n+1}\mathbb{C} = \mathfrak{so}(Q) \subset C(Q)^{\text{even}} \cong \mathfrak{gl}(\bigwedge^\bullet W) = \mathfrak{gl}(S) \ . \tag{6.78}$$

So now we have the following proposition.

Proposition 6.7. *The representation $S = \bigwedge^\bullet W$ is the irrep of $\mathfrak{so}_{2n+1}\mathbb{C}$ with the highest weight*

$$\alpha = \frac{1}{2}(L_1 + ... + L_n) \ . \tag{6.79}$$

Proof. This is similar to the even case - each e_I is an eigenvector with weight $\frac{1}{2}\left(\sum_{i\in I} L_i - \sum_{j\notin I} L_j\right)$. All the weights are congruent by the Weyl group so it must be an irrep with highest weight α. □

We have therefore constructed the spin representations of $\mathfrak{so}_n\mathbb{C}$.

6.4 Classification of Clifford algebras

As we have seen above, Clifford algebras are extremely useful. There exists a classification that classifies real and complex Clifford algebras. The details of which could be found in [20] and I will refer to that for the full classification[27].

6.5 Spinor Representations and Clifford Algebras

Having discussed the spin representations of $\mathfrak{so}_n\mathbb{C}$, it is prudent to discuss its relation with the spinors in this section.

6.5.1 Pin and Spin Groups

First, let us define something known as pin and spin groups.

Definition 6.15. The **multiplicative group of units** in the Clifford algebra is defined to be the subset

$$C^\times(Q) = \left\{\phi \in C(Q) \,\middle|\, \exists\phi^{-1},\ \phi^{-1}\phi = \phi\phi^{-1} = 1\right\} \ . \tag{6.80}$$

This group contains all elements $v \in V$ with $Q(v) \neq 0$.

The group of units always acts naturally as automorphisms of the algebra, i.e. the **adjoint representation**,

$$\mathrm{Ad} : C^\times(Q) \to \mathrm{Aut}(C(Q)) \ , \tag{6.81}$$

which is given by,

$$\mathrm{Ad}_\phi(x) = \phi x \phi^{-1} \ . \tag{6.82}$$

Taking the derivation of this map gives the usual Lie bracket action $\mathrm{ad}_y(x) = [y, x]$. Hiterto we have assumed that the characteristic of the field could be any integer. Let us assume from now that the chracterisitc of the field $k \neq 2$. Then we have the following important proposition.

[27] If I have time in the future I will come back to add more information but this is not relevant to the current discussion.

Proposition 6.8. *Let $v \in V \subset C(Q)$ be an element with $Q(v) \neq 0$. Then $\mathrm{Ad}_v(V) = V$, and $\forall w \in V$, we have,*

$$-\mathrm{Ad}_v(w) = w - 2\frac{Q(v,w)}{Q(v)}v \ . \tag{6.83}$$

Proof. We have that $v^{-1} = -\frac{v}{Q(v)}$, so

$$-Q(v)\,\mathrm{Ad}_v(w) = -Q(v)vwv^{-1} = vwv = -v^2w - 2Q(v,w)v = Q(v)w - 2Q(v,w)v \ . \tag{6.84}$$

□

Naturally, this lead us to consider the subgroup of elements $\phi \in C^\times(Q)$ such that $\mathrm{Ad}_\phi(V) = V$. From Proposition 6.8 above, we see that the group contains all the elements $v \in V$ with $Q(v) \neq 0$, and when this happens the transformation Ad_v preserves the quadratic form Q,

$$(\mathrm{Ad}_v^* Q)(w) = Q(\mathrm{Ad}_v(w)) = Q(w) \ , \tag{6.85}$$

for all $w \in V$. We define $P(Q) \subset C(Q)$ to be the subgroup generated by the elements $v \in V$ with $Q(v) \neq 0$. Note that then there is a representation,

$$P(Q) \to O(V, Q) \ , \tag{6.86}$$

where

$$O(V, Q) = \{\lambda \in GL(V)\,|\,\lambda^* Q = Q\} \tag{6.87}$$

is the orthogonal group of the form Q.

We are now ready to explore the important subgroups of $P(Q)$.

Definition 6.16. The **Pin group** is the subgroup of $P(Q)$ generated by the elements $v \in V$ with $Q(v) = \pm 1$, i.e.

$$\mathrm{Pin}(V) = \left\{v_1 \cdot \ldots \cdot v_m \in C(Q) \,\middle|\, v_j \in S^{n-1} \subset \mathbb{F}^n, m \in \mathbb{N}\right\} \tag{6.88}$$

Definition 6.17. The Spin group is simiarly defined as,

$$\begin{aligned} \mathrm{Spin}(V) &= \mathrm{Pin}(n) \cap C^0(Q) && (6.89)\\ &= \left\{v_1 \cdot \ldots \cdot v_m \in C(Q) \,\middle|\, v_j \in S^{n-1} \subset \mathbb{F}^n, m \in 2\mathbb{N}\right\} && (6.90) \end{aligned}$$

Note that from the definition of the Pin group, the inverse element to $v_1 \cdot \ldots \cdot v_m$ is,

$$(v_1 \cdot \ldots \cdot v_m)^{-1} = (-v_m)\ldots(-v_1) \in \mathrm{Pin}(V) \tag{6.91}$$

Let us take a deep look at Eq. (6.83). Notice that the right hand side of the equation is basically the reflection $R_v(x)$ of the vector $x \in V$ where v is the vector marking the perpendicular direction of the reflection hyperplane. To remove this sign we therefore consider the following action.

Definition 6.18. The **twisted adjoint representation** is the map $\tilde{\text{Ad}} : C^\times(Q) \to GL(C(Q))$ where,

$$\tilde{\text{Ad}}_\phi(y) = \alpha(\phi) y \phi^{-1} \ . \tag{6.92}$$

For even elements ϕ, $\tilde{\text{Ad}}_\phi = \text{Ad}_\phi$. We also have $\tilde{\text{Ad}}_{\phi_1\phi_2} = \tilde{Ad}_{\phi_1} \circ \tilde{Ad}_{\phi_2}$. Explicitly,

$$\tilde{\text{Ad}}_v(w) = w - 2\frac{Q(v,w)}{Q(v)} v \ . \tag{6.93}$$

We state without proof the following result.

Theorem 6.5 (Cartan-Dieudonné). *Every $g \in O(V)$ is the product of a finite number of reflections $g = R_{u_1} \circ ... \circ R_{u_r}$ along the null lines where $Q(u_i) \neq 0$ and $r \leq \dim V$.*

Proof. See [21]. □

We note that the twisted adjoint action must define a group homomorphism $\tilde{\text{Ad}} : \text{Pin}(V) \to O(V)$. It follows from the Cartan-Diedonné Theorem 6.5 that $\tilde{\text{Ad}}$ is surjective. But what is the kernel of $\tilde{\text{Ad}}$?

Proposition 6.9. *Suppose V is finite dimensional and Q is non-degenerate. Then the kernel of the homorphism $\tilde{\text{Ad}} : \tilde{P}(V,Q) \to GL(V)$ is the group $\mathbb{K}^\times$ of non-zero multiples of 1. Here the group $\tilde{P}(V,Q)$ is defined as,*

$$\tilde{P}(V,Q) = \left\{ \phi \in C^\times(V,Q) \,\middle|\, \tilde{\text{Ad}}_\phi(V) = V \right\} \ , \tag{6.94}$$

where $P(V,Q) \subset \tilde{P}(V,Q)$.

Proof. See [20] for a complete proof. The proof is also outlined in [22]. □

[20] goes into a bit more detail in how you would define the homomorphism from the group $\tilde{P}(V,Q)$ to the orthogonal group $O(V)$ (see Propositions 2.5 and Corollary 2.6 of [20]). It also shows how the images $\tilde{\text{Ad}}(\text{Pin}(V,Q))$ and $\tilde{\text{Ad}}(\text{Spin}(V,Q))$ is a normal subgroup of $O(V)$ (see Proposition 2.8 of [20]). To summarise there are two exact short sequences.

Theorem 6.6. *Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and Q a non-denerate quadratic form on V. Suppose the field $\mathbb{K}$ of characteristic $\neq 2$ is **spin**, i.e. at least one of the two equations $t^2 = a$ and $t^2 = -a$ can be solved in $\mathbb{K}$ for each non-zero element $a \in \mathbb{K}^\times$. Then there are two short exact sequences.*

$$0 \to F \to \text{Spin}(V,Q) \xrightarrow{\tilde{\text{Ad}}} SO(V) \to 1 \ , \tag{6.95}$$

$$0 \to F \to \text{Pin}(V,Q) \xrightarrow{\tilde{\text{Ad}}} O(V) \to 1 \ , \tag{6.96}$$

where

$$F = \begin{cases} \mathbb{Z}_2 = \{1, -1\} & \text{if } \sqrt{-1} \notin \mathbb{K} \\ \mathbb{Z}_4 = \{\pm 1, \pm\sqrt{-1}\} & \text{otherwise} \end{cases} \tag{6.97}$$

The sequences above hold for general fields provided that $SO(V)$ and $O(V)$ are replaced by appropriate normal subgroups of $O(V)$ (since the map $\tilde{\text{Ad}}$ maps to normal subgroups of $O(V)$ in general and field $\mathbb{K}$, which is spin, solves the equation $t^2 Q(v) = \pm 1$ so every $v \in V^\times$ can be renormalised to have $Q(v) = 1$).

Proof. See [20] Theorem 2.9. The details of the field being **spin** is not relevant if we restrict to $\mathbb{K} = \mathbb{R},\ \mathbb{C}$ as both fields are spin. $\square$

The real case of the above Theorem 6.6 is reduced to the simple short exact sequences,

$$0 \to \mathbb{Z}_2 \to \mathrm{Spin}_{r,s} \to SO_{r,s} \to 1 \ , \tag{6.98}$$

$$0 \to \mathbb{Z}_2 \to \mathrm{Spin}_{r,s} \to SO_{r,s} \to 1 \ , \tag{6.99}$$

for all (r, s), where the subscripts denote the signature of the quadratic form Q. In particular, for $SO_n = SO_{n,0} = SO_{0,n}$, we have

$$0 \to \mathbb{Z}_2 \to \mathrm{Spin}_n \xrightarrow{\tilde{\mathrm{Ad}}} SO_n \to 1 \ , \tag{6.100}$$

where the map $\tilde{\mathrm{Ad}}$ acts as the universal covering map of SO_n for all $n \geq 3$. From this we see that the Pin and Spin groups are covering groups of the orthogonal group. We will now examine the representations of the Pin and Spin groups, which in term will give representations of the orthonal groups. We will see how this unifies with the picture we have taken in two subsections ago.

6.5.2 Spinor Representations

Before we begin let us recall what representations of an algebra is.

Definition 6.19. Suppose A is an associative algebra and $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. A **$\mathbb{K}$-representation** of A is a $\mathbb{K}$-linear homomoprhism $\rho : A \to \mathrm{End}_{\mathbb{K}}(E)$ for some $\mathbb{K}$-vector space E. Two $\mathbb{K}$-representations $\rho : A \to \mathrm{End}_{\mathbb{K}}(E)$ and $\rho' : A \to \mathrm{End}_{\mathbb{K}}(E')$ are **equivalent** if there exists a $\mathbb{K}$-linear isomorphism $f : E \to E'$ such that the following triangle commutes:

$$\begin{tikzcd} & A \arrow[dl, "\rho"'] \arrow[dr, "\rho'"] & \\ \mathrm{End}_{\mathbb{K}}(E) \arrow[rr, "\mathrm{Ad}\, f"] & & \mathrm{End}_{\mathbb{K}}(E') \end{tikzcd}$$

where $\mathrm{Ad}\, f : \mathrm{End}_{\mathbb{K}}(E) \to \mathrm{End}_{\mathbb{K}}(E')$ is defined as $\phi \mapsto f \circ \phi \circ f^{-1}$, so $f \circ \rho(a) = \rho'(a) \circ f$ for all $a \in A$.

We can now define the following representations.

Definition 6.20. A **pinor representation** of $\mathrm{Pin}(V)$ is the restriction of an irreducible representation of $C(Q)$. Similarly, a **spinor representation** of $\mathrm{Spin}(V)$ is the restriction of an irreducible representation of $C_0(Q)$

We here note that most representations of Clifford algebras (similar to Lie algebras) are reducible. The volume element plays a key role in determining irreducible representations. In particular, since the volume element is used in determining the classification of real and complex Clifford algebras [20], one can use it to determine the properties of irreducible real and complex representations of Clifford algebra. You can read about the details of the classfication and determination of irreducible representations in [20] and [22]. The crux of the discussion there is that one must distinguish between (broadly) even and odd-dimension V cases, similar to our discussion above. In particular, let us suppose Cl_n is the Clifford algbera of $\mathbb{R}^n$, and its complexification is $\mathbb{C}l_n = Cl_n \otimes_{\mathbb{R}} \mathbb{C}$. For the standard basis e_i of $\mathbb{R}^n$, define

$$z_j = \frac{1}{2}(e_{2j-1} - ie_{2j}) \in \mathbb{C}l_n \tag{6.101}$$

for $j = 1, ..., m$, $n = 2m$ and $\bar{z}_j$ be its conjugate. Then the span

$$\Sigma = \{z_{j_1} \cdot ... z_{j_k} \cdot \bar{z}_1 \cdot ... \cdot \bar{z}_m \,|\, k = 0, ..., m, \quad 1 \leq j_1 \leq ... \leq j_k \leq m\} \ , \tag{6.102}$$

defines a complex vector subspace of $\mathbb{C}l_n$ of dimension 2^m. The space that is known as the **spinor space** with its elements being **spinors**[28]. The spinor space is invariant under Clifford multiplications (multiplication by e_j). We further define $\Sigma^{\pm}$ to be the spans where k is even and odd respectively. The sign of the space is then the **chirality** of the space, and by setting the above notation $E = \Sigma$ in Definition 6.20, we now have reobtained our definitions of spinor representations in terms of action on spinor spaces. We then have the following propositions.

Proposition 6.10. *Suppose the complex spinor representation is* $\Delta_n^{\mathbb{C}} : \mathrm{Spin}_n \to GL_{\mathbb{C}}(\Sigma)$ *which is given by restricting an irreducible complex representation* $\mathbb{C}l_n \to \mathrm{Hom}_{\mathbb{C}}(\Sigma, \Sigma)$ *to* $\mathrm{Spin}_n \subset Cl_n^0 \subset \mathbb{C}l_n$*. Then for n add this definition is indenpedent of which irrep of* $\mathbb{C}l_n$ *is used, and that the representation* $\rho_n^{\mathbb{C}}$ *is irreducible. When n is even then there is a decomposition,*

$$\Delta_n^{\mathbb{C}} = \Delta_n^{\mathbb{C}+} \oplus \Delta_n^{\mathbb{C}-} \ , \tag{6.103}$$

i.e. into a direct sum of two inequivalent irreducible complex representations of Spin_n*.*

Proof. See Proposition 5.15 of [20], which uses the argument for the real case. For more information one can also look at [23]. □

We now see how this is exactly the same as how we have defined spinor representations to start with. This unifies the two pictures of spinor representations in [1] and in [20] - the subspaces in [1] are exactly the Spin(Pin) groups that are defined at the start of this subsection.

6.5.3 Weyl, Dirac, Pauli and Majorana Spinors

Physicists like calling spinor with names that confuse mathematicians. In this section I will clarify the relationship between the spin(or) representations we have seen in the previous section with the ones often defined in a hand-waving manner in the physics literature. In the following I will fix V to be a vector space where spinors live.

Case 1: $n = \dim V$ **is odd.**

[28] Note very carefully that the spinor space Σ dimensions has nothing to do with the dimensions of V, the vector space where we have defined the Clifford algebra. This is important for the classification of spinors in the next subsection.

I want to cover the odd case first as this is a bit easier. The spin representations are defined as $(n = p + q)$,

$$\Delta_n : \mathrm{Spin}_{p,q} \to GL_{\mathbb{C}}(\Sigma) , \tag{6.104}$$

where Σ is the spinor space. We note that we can design the map Δ_n to map to a subspace of the spinor space defined above. We then have the following 'definitions'.

Definition 6.21. For $p + q = \dim V$ odd, we have the following spinors.

- $\dim \Sigma = 1$: This is the **scalar representation**, the complexified spin group acts on the space of complex scalars $\mathbb{C}$.
- $\dim \Sigma = \dim V$: This is the **vector representation**, so the complexified spin group acts on the space of complexified vectors $\mathbb{C}^{\dim V}$.
- $\dim \Sigma = m$, so the action of the spin group is represented by complex matrices $M \in \mathrm{Mat}(m, \mathbb{C})$. This is called the **spin-$\frac{1}{2}$ representation**, the spinors are known as **Pauli spinors**.

Ah, that is exactly what we called Pauli spinors. The easiest case is to look at the spinor representations of $\mathfrak{so}_3$ which represents the Clifford algebra as Pauli spinors (and transformations as Pauli matrices).

Case 2: $n = \dim V$ is even.

Now we move on to the case where n is even. From the analysis of the odd case we have for each separate decomposition $\Sigma^{\pm}$ a scalar representation, a set of spin-$\frac{1}{2}$ representations and a vector representation. Now we need to consider the pair of irreducible representations.

Definition 6.22. For $p + q = \dim V$ even, we have the following spinors.

- $0 \oplus 0 = (0, 0)$: This is the **scalar representation** acting on the space $\mathbb{C} \oplus \mathbb{C}$.
- $\frac{1}{2} \oplus 0 = (\frac{1}{2}, 0)$: This is the **left spin-$\frac{1}{2}$ representation** which acts on the space of **left Weyl spinors**.
- $0 \oplus \frac{1}{2} = (0, \frac{1}{2})$: This is the **right spin-$\frac{1}{2}$ representation** which acts on the space of **right Weyl spinors**.
- $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$: This is the **Dirac representation** which acts on the space of **Dirac spinors**. This can be viewed as two Weyl spinors stacked on top of each other.
- $(\frac{1}{2}, \frac{1}{2})$: This is called the **vector representation**, where the spin group acts on the space of complexified vectors.

We note that the Weyl spinors are indeed the building blocks of spinor representations in this case. This addresses how the complicated spinor definitions are related to the spinors you have left in Quantum Field Theory courses [29]!

[29] Majorana spinors require the definition of a symplectic structure. I will not cover the details here but please refer to [24].

6.5.4 Spin Structures

To fully address the mathematical structures one needs to also define something known as spin structures and spin bundles. Since this is completely unrelated to our main discussion, I will leave the interested amongst you to have a go at [20, 22, 23]. This concludes our excursion into spinors!

7 Correspondence between Lie Groups and Lie Algebras

In the lectures we have established that the following facts.

1. The tangent map e of a group homomorphism is a Lie algebra homomorphism, and vice versa.

2. Lie groups can be recovered from Lie algebras using the exponential map.

Effectively what this says is that there is a natural correspondence between Lie groups and Lie algebras. However, from the Lie algebra isomorphism discussion we have seen that in fact the Lie algebra $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic to one another. We also happen to know that $SU(2)$ as a real Lie group is a universal (2:1) covering of the Lie group $SO(3)$. Is there a more precise statement of the facts as discussed in the lectures? In particular, what kind of manifold will I get if I exponentiate the $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$ real Lie algebras?

The answer lies again in the global construction of Lie groups. The statement made in the lectures follows Lie's original argument - the construction is *local*. In this section, we will discuss the consequences of the exponential map in the global setting, and fully complete the correspondence discussion set out in the notes.

7.1 Lie Brackets and Flows

The first question to address is why the tangent map e of a group homomorphism is a Lie algebra homomorphism. Additionally there is a side question - what is the need of defining a left-invariant vector field, when we can locally define vector fields at the tangent space of identity e?

To answer these two questions we must first go back and define what an integral curve, a flow and vector fields are. To do this I will go a bit deeper - is there a more general notion for tangent spaces at a point? Turns out there is mathematical notion for the 'collection of tangent spaces' T_pM of the manifold M, typically known as a tangent bundle. This will be very useful in our discussion (and typically covered in a 'second' course in differential geometry) so here is the definition. Note throughout this section we will fix M to be a smooth n-manifold[30].

Definition 7.1. The **tangent bundle of** M, denoted by TM, is *loosely* the disjoint union of the tangent spaces at all points of M,

$$TM = \coprod_{p\in M} T_pM \ , \tag{7.1}$$

where T_pM denotes the tangent space at $p \in M$.

More accurately, the **tangent bundle of** M TM is defined as a smooth topological space with the following data.

[30] The notion of smoothness is defined in Andre's notes. If it is not clear then have a look at [7].

- There is a natural surjective **projection map** $\pi : TM \to M$. Denoting the points of TM as the ordered pair (p, v) where $p \in M$ and $v \in T_pM$, then $\pi(p, v) = p$.

- Suppose $\{U_i\}_{i\in I}$ is an open cover of M (here I is an indexing set). For each i there exists a diffeomorphism
$$\Phi_i = \pi^{-1}(U_i) \to U_i \times \mathbb{R}^n \ , \tag{7.2}$$
such that $\mathrm{pr}_1 \circ \Phi_i = \pi$ and for all i, j the map $\Phi_i \circ \Phi_j^{-1}$ has the form $(U_i \cap U_j) \times \mathbb{R}^n \to (U_i \cap U_j) \times \mathbb{R}^n$ where
$$(x, v \in \mathbb{R}^n) \mapsto (x, g_{ji}(x)(v)) \tag{7.3}$$
for some smooth map $g_{ji} : U_i \cap U_j \to GL(n, \mathbb{R})$. This smooth map g_{ij} is precisely the transition map defined between the two charts.

There is of course the notion of a vector bundle, which we now can define below.

Definition 7.2. A **vector bundle of** M **of rank** k $\pi : E \to M$ is defined as a smooth topological space with the following data.

- There is a natural surjective **projection map** $\pi : E \to M$. Denoting the points of TM as the ordered pair (p, v) where $p \in M$ and $v \in E$, then $\pi(p, v) = p$.

- Suppose $\{U_i\}_{i\in I}$ is an open cover of M (here I is an indexing set). For each i there exists a diffeomorphism
$$\Phi_i = \pi^{-1}(U_i) \to U_i \times \mathbb{R}^k \ , \tag{7.4}$$
such that $\mathrm{pr}_1 \circ \Phi_i = \pi$ and for all i, j the map $\Phi_i \circ \Phi_j^{-1}$ has the form $(U_i \cap U_j) \times \mathbb{R}^n \to (U_i \cap U_j) \times \mathbb{R}^n$ where
$$(x, v \in \mathbb{R}^k) \mapsto (x, g_{ji}(x)(v)) \tag{7.5}$$
for some smooth map $g_{ji} : U_i \cap U_j \to GL(k, \mathbb{R})$.

The only thing that has changed is the dimension of the total space E and the transition maps g_{ji} which are no longer directly related to the coordinate transformations when we defined different charts. We can then define a vector field as follows.

Definition 7.3. A **section** on E is a map $s : M \to E$ such that $s \circ \pi = \mathrm{id}_M$. A section of TM is a **(rough) vector field on** M. If the restrictions of X to the coordinate chart on M is smooth and the component functions with respect to this chart is also smooth then the vector field is **smooth**.

Now we see how the notion of Lie brackets can be defined naturally on vector fields. To see this first note that (see Proposition 8.14 of [7]) if X is a smooth vector field, then for every $f \in C^\infty(M)$ function, then the function Xf is smooth on M. We can then apply another smooth vector field Y onto this function to get $f \mapsto YXf$ - but this does not in general satisfy the product rule so cannot be a vector field. The way to do this is to define something known as the Lie bracket of X and Y,
$$[X, Y]f = XYf - YXf \tag{7.6}$$

This operator now is a vector field and gives the map $[X,Y] : C^\infty(M) \to C^\infty(M)$. It is then a general question whether the set of vector fields $\{X_i\}$ form a Lie algebra (the answer is it does... but we need to wait). To see this let us return to what a flow is.

Definition 7.4. Suppose V is a vector field on M. An integral curve of V is a differentiable curve $\gamma : J \to M$ such that

$$\gamma'(t) = V_{\gamma(t)} \quad \forall t \in J \ . \tag{7.7}$$

The point $\gamma(0)$ is called the **starting point of** γ if $0 \in J$.

Integral curves are the curves generated by moving along the vector field - the curve is entirely defined by the vector field V so given a starting point you will know what the curve is. Another way to visualise this is as follows. Suppose for each point $p \in M$, V has a unique integral curve starting at p defined for all $t \in \mathbb{R}$. We can define a map $\theta_t : M \to M$ by sending $p \in M$ to the point which is obtained after transporting t units along the integral curve defined by V. This motivates the following definition.

Definition 7.5. A **flow domain** is an open neighbourhood U of $0 \times M$ in $\mathbb{R} \times M$ such that for each $p \in M$ the set $U \cap (\mathbb{R} \times \{p\})$ is connected (i.e. it is an open interval around 0)[31].

A **local flow** of a vector field V is a smooth map $\theta : U \to M$ where U is a flow domain, and satisfying the following two conditions.

- $\theta(0, \cdot) = \mathrm{id}_M$.
- $\frac{d}{dt}\theta(t,p) = V(\theta(t,p)), \forall (t,p) \in U$.

It is a **global flow** if $U = \mathbb{R} \times M$. For shorthand write θ^t for $\theta(t, \cdot)$.

Defined in the above manner, the flow θ as a continuous map then automatically satisfies $\theta^t \circ \theta^s = \theta^{t+s}$. In that sense $\theta^{-t} = (\theta^{-1})^t$. One can also define the map as one with the condition above (together with $\theta^0 = p$ are known as the group laws of flow.). V is then known as the **infinitesimal generator** of θ.

Definition 7.6. A vector field is **complete** if it admits a global flow.

With a vector field V and a local flow θ one can now define something known as the Lie derivative.

Definition 7.7. Given a smooth map between two manifolds X and Y, $F : X \to Y$, the **derivative of** F **at** p is the map $D_pF : T_pX \to T_{F(p)}Y$ given by $[\gamma] \to [F \circ \gamma]$, where γ is the integral curve for some vector field v, with the bracket indicating the tangent vector represented by curve γ. We denote D_p by F_* the **pushforward** by F on tangent vectors. Similarly, we define the dual map $(D_pF)^\vee : T_{F(p)}T \to T_pX$ to be the **pullback** F^*.

[31]This definition is not standard, and is credited to Dr. Jack Smith from whom I learnt most of these my differential geometry.

Definition 7.8. The **Lie derivative** of a vector field W, denoted $\mathcal{L}_V W$, is

$$\left.\frac{d}{dt}\right|_{t=0} (\theta^t)^* W \tag{7.8}$$

Here $\cdot^*$ indicates the pullback map.

The Lie derivative is defined as above as we want to pullback the vector transported to compare with the original vector at the origin (and not the other way - otherwise there will be a minus sign). We then note that we can define the Lie bracket of vector fields as follows.

Definition 7.9. The **Lie bracket** of vector fields V and W is defined as

$$[V, W] = \mathcal{L}_V W = -\mathcal{L}_W V \ . \tag{7.9}$$

You can check that this makes the space of vector fields on M, $\mathfrak{X}(M)$ into a Lie algebra as this definition of Lie bracket is now alternating and satisfies the Jacobi identity. This is surprising - the set of vector fields $\mathfrak{X}(M)$ forms a Lie algebra under the Lie bracket of vector fields. Why do we care so much about Lie groups then, if this structure is already prevalent in a typical smooth manifold X?

The reason why is because Lie groups have a lot more structure to a general smooth manifold. The notion of left-invariant vector fields are specifically defined only when you have a group action G on the manifold, and since this group action already exists for Lie groups, the notion of left-invariant vector fields naturally exist and hence one can naturally define a Lie algebra on the set of vector fields. For a generic smooth manifold, there may not exists such smooth right action that maps any point to another by a global diffeomorphism [32]. Of course, one can also take the viewpoint that the group of all diffeomorphisms of a smooth manifold M as an infinite-dimensional Fréchet Lie group - in that sense its Lie algbera is exactly $\mathfrak{X}(M)$ with its Lie bracket structure [25]. How do we connect all these notions with the 'local'-sense of how a Lie algebra is defined - as the Lie bracket structure of tangent vector fields at the identity? This point is explored in the lectures but it is worth emphasising once more — the left-invariant property of the vector fields means that once the vector field is determined we will known its value at each point $p \in M$. We then have the vector space isomorphism,

$$\mathcal{L}(G) \to T_e G, \quad \xi \mapsto \xi_e \ , \tag{7.10}$$

so the tangent vector at identity is uniquely defined [33]. This global to local property is unique to the Lie group structure. All that means is it is sufficient to understand the local tangent space (germ) of G to understand the properties of the manifold (bar some global properties as we will see soon).

[32] This point is explored in more detail in §20 of [7] where Lee talks about the homomorphism of $\mathfrak{g} \to \mathfrak{X}(M)$ if such a smooth action G on M exists.

[33] The inverse of this map of course is $\xi_e \mapsto g_* \xi_e = \xi_g$.

Now the remaining question to address is why the tangent map e of a group homomorphism is a Lie algebra homomorphism. This is again sketched out in the lecture note but it is worth emphasising again. There is a natural map $\Psi : G \to \mathrm{Aut}(G)$ [34], where $\Psi_g(h) = g \cdot h \cdot g^{-1}$. A homomorphism ρ respects the adjoint action of a group G on its tangent space T_eG at the identity and therefore by taking the differential of the characterisation $(d\rho)_e$ respects the adjoint action of the tangent space to G on itself. Of course, with the arguments above, we now see how this map relates to the Lie bracket definitions of left-invariant fields and hence everything is now consistent.

7.2 The exponential map and BCH Formula

We have see in the previous subsection that the correspondence is a local construction. The Lie group structure naturally endows a Lie bracket on the space of tangent vectors at identity (or the space of left-invariant vector fields by the arguments above) so it naturally has a Lie algebra structure. The local construction we discussed follows two important principles (see §8 of [1]):

1. Let G and G' be Lie groups, with G connected. A map $f : G \to G'$ is uniquely determined by its differential $df_e : T_eG \to T_eG'$ at the identity.

2. Let G and G' be Lie groups, with G connected and simply-connected. A linear map $T_eG \to T_eG'$ is the differential of a homomorphism $f : G \to G'$ if and only if it preserves the bracket operation:

$$df_e\left([X, Y]\right) = [df_e(X), df_e(Y)] \tag{7.11}$$

   Alternatively, a linear map between the Lie algebras $\alpha : \mathfrak{g} \to \mathfrak{g}'$ is the differential of a map $f : G \to G'$ of Lie groups if and only if α is a map of Lie algebras.

The natural question to ask now is how to perform the inverse, i.e. could we possibly reconstruct a Lie group given a Lie algebra? The answer is yes, and it is often done with the help of one-parameter subgroups and exponential maps which you have all encountered in the lecture notes. These are also discussed in §3.4 of the notes and also §8.3 of [1]. We note here that the exponential map is the unique map from $\mathfrak{g}$ to G taking 0 to e whose differential at the origin $(\exp_*)_0 : T_0\mathfrak{g} \to T_eG$ is the identity and the restrictions to the lines through the origin in $\mathfrak{g}$ are one-parameter subgroups of G. So in this so-called local construction we have shown the following — since the differential of the exponential map at the origin of $\mathfrak{g}$ is an isomorphism, the image of exp will contain a neighbourhood of the identity in G.

Before we continue let us see this more explicitly by constructing the actual correspondence in detail for matrix Lie groups. Using the standard power series for e^x:

$$\exp(X) = 1 + X + \frac{1}{2}X^2 + \frac{1}{3!}X^3 + \ldots . \tag{7.12}$$

[34] Here we look at the automorphisms of G by conjugation as there are no fixed points for multiplication m_g.

We want to show that the group structure of G is encoded in the Lie algebra. Explicitly we want to show that for $X, Y \in \mathfrak{g}$ in a sufficiently small neighbourhood near the origin we can write down the product $\exp(X) \cdot \exp(Y)$ as an exponential. By defining,

$$\log(g) = (g - I) - \frac{1}{2}(g - I)^2 + \frac{1}{3}(g - I)^3 + ... \in \mathfrak{gl}_n\mathbb{R} \,, \tag{7.13}$$

where $g \in G$, we want to set a new bilinear operation $*$ on $\mathfrak{gl}_n\mathbb{R}$ where

$$X * Y = \log\left(\exp(X) \cdot \exp(Y)\right) \,. \tag{7.14}$$

The formula of $X * Y$ in terms of X, Y and brackets is known as the **Baker-Campbell-Hausdorff formula** and gives the local identification of Lie algebras and Lie groups. To degree three it is

$$X * Y = X + Y + \frac{1}{2}[X, Y] \pm \frac{1}{12}[X, [X, Y]] \pm \frac{1}{12}[Y, [Y, X]] + ... \tag{7.15}$$

If G is connected the exponential map generates all of G (provided G is simply-connected, as we will see soon in the next subsection) so every finite-dimensional Lie algebra is the Lie algebra of a Lie group, by applying to an embedding of a Lie algebra $\mathfrak{g}$ into $\mathfrak{gl}_n$.

7.3 Injectivity and surjectivity of an exponential map

Let me emphasise that the exponential map is a local construction — it tells you that there is a diffeomorphism between the open neighbourhoods of $0 \in \mathfrak{g}$ and $e \in G$. In general, this map $\exp : \mathfrak{g} \to G$ fails to be injective or surjective, as you would expect.

Let us investigate this a bit deeper. Firstly, let us look at injectivity. It is immediately clear that if there exists any circle subgroups in the Lie group, say S^1 which has $\mathbb{R}$ as its Lie algebra. The generator in $\mathbb{R}$, say some t, must eventually loop back on itself in the image which causes the kernel of the exponential map to be some $\ker \exp = \mathbb{Z}$. Another simple counter-example to exp being injective is the group $SO(2)$, which Lie algebra given by the space of two-dimensional skew-symmetric matrices. In this case the exponential map is given by:

$$\exp : \begin{pmatrix} 0 & t \\ -t & 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \tag{7.16}$$

which clearly is not injective as we can identify $t \sim t + 2\pi$. In general, we have the following characterisation.

Proposition 7.1. *The following statements are equivalent.*

(1) exp *is injective.*

(2) exp *is surjective.*

(3) exp *is a real analytic diffeomorphism.*

(4) G is solvable, simply connected, and $\mathfrak{g}$ does not admit $\mathfrak{e}$ as a subalgebra of a quotient.

(5) G is solvable, simply connected, and $\mathfrak{g}$ does not admit $\mathfrak{e}$ or $\tilde{\mathfrak{e}}$ as a subalgebra.

(6) G has no closed subgroup isomorphic to either the circle $\mathbb{R}/\mathbb{Z}$, the universal covering $\widetilde{SL_2(\mathbb{R})}$, E the isometry group of the plane and its central extension $\tilde{E}$.

Here $\mathfrak{e}$ is the 3d algebra, isomorphic to the group of isometries of the plane, with basis (H, X, Y) and brackets $[H, X] = Y$, $[H, Y] = -X$, $[X, Y] = 0$. The central extension of $\mathfrak{e}$ is $\tilde{\mathfrak{e}}$ which adds the central generator Z with the non-trivial bracket $[X, Y] = Z$. Obviously, E and $\tilde{E}$ are the simply-connected solvable Lie groups associated to $\mathfrak{e}$ and $\tilde{\mathfrak{e}}$ respectively.

Proof. As we have seen from above, the injectivity of the exponential implies that there is no closed subgroup isomorphic to S^1 so the maximal compact subgroup in G is trivial and G is contractible [35]. Now, by Levi's decomposition, we can write a contractible Lie group in the form $R \rtimes \widetilde{SL_2(\mathbb{R})}^k$ with R a simply-connected solvable Lie group (as we must have simply-connected covers), but for an injective exponential map we must have $k = 0$, so $G = R$, a connected, simply-connected, contractible, solvable Lie group. This is not enough — any Lie algebra that has $\mathfrak{e}$ as subalgebra also includes rotations, which implies that $\exp(tH) = e$, for $H \in \mathfrak{e}$ the rotation operator. The proof for (1)-(4) being equivalent can be found in [26].

The equivalence between (4) and (5) is non-trivial and I leave the proof in [27]. It is clear that the simple connected Lie groups E and $\tilde{E}$ associated to $\mathfrak{e}$ and $\tilde{\mathfrak{e}}$ have a non-injective exponential map, so (1) implies (5) as well. Finally, the last statement comes naturally from all the statements above. □

In particular, what we have observed is this — if a Lie group G contains a subgroup isomorphic to the circle (hence tori) then the exponential map is automatically non-injective. In analysing compact Lie groups, it is often useful to define a torus as follows.

Definition 7.10. A Lie group T is a **torus** if T is a compact, connected abelian Lie group T isomorphic to the direct product of $k \in \mathbb{N}$ copies of the group $S^1 \cong U(1)$.

The Lie algebra of the torus $\mathfrak{t}$ is an abelian Lie algebra. In particular, we can observe that the exponential map $\exp : \mathfrak{t} \to T$ is locally a diffeomorphism, and has a kernel,

$$\ker(\exp) = \Lambda \,, \tag{7.17}$$

where Λ is the discrete subgroup to be modded out, identified as $T = \mathbb{R}^k/\Lambda$. Why is this a useful fact for analysing compact Lie groups? Turns out we have the following fact.

Proposition 7.2. *Let G be a non-trivial, compact, connected Lie group. Then G contains a positive dimensional torus.*

Proof. G is positive dimensional and has a non-zero Lie algebra $\mathfrak{g}$. If $X \in \mathfrak{t}$, then $\exp(tX)$ is a non-trivial one-parameter subgroup $A \subset G$, which is connected, positive dimensional and abelian. So its closure is a Lie subgroup and a positive dimensional torus. □

[35] A Lie group G is contractible if the identity map on G is null-homotopic.

Corollary 7.1. *There is a positive dimensional torus in G that is not properly contained in any other torus in G. This is known as the* ***maximal torus*** *in G.*

Proof. If T is a torus of maximal dimensional in G, then T is not properly contained in any other torus in G. This is since $T \subset T'$ for some T', then $|\mathfrak{t}'| > |\mathfrak{t}|$ so $\dim T' > \dim T$ and we arrive at a contradiction. □

The maximal torus is massively helpful in analysing the Weyl group, a subgroup of the isometry group generated by reflections through the hyperplanes orthogonal to at least one of the roots.

Why is this fact helpful? The logic goes as follows — for every Lie group which is compact and connected, there is a maximal torus which we know has a non-trivial kernel under the exponential map. This means that in these cases all the exponential maps are then not injective by definition. Of course, some exponential maps can be injective — a quick example is $\exp : \mathbb{R}^n \to \mathbb{R}^n$.

Now let us move on to surjectivity. Surjectivity of Lie groups obviously fail for disconnected groups since the Lie algebra is necessarily connected so the map under the exponential map must also be connected. A simple counter-example in this case is the Lorentz group $SO(1,3)$ which we know from the lectures have four disconnected components.

However, even for connected Lie groups, surjectivity can also fail. An example is given by Terrance Tao (in his blogpost), where he points out the following observation. Suppose the exponential map is surjective, then every group element $g \in G$ has a square root as $\exp(X/2)$ obviously exists for any $X \in \mathfrak{g}$. However, there exist elements in connected Lie groups without square roots, a simple example given by,

$$g = \begin{pmatrix} -4 & 0 \\ 0 & -1/4 \end{pmatrix} , \tag{7.18}$$

in the connected Lie group $SL_2(\mathbb{R})$, which square root do not exist as it must have an eigenvalue $\pm 2i$ and the other being $\pm i/2$ but since $h = \sqrt{g} \in SL_2(\mathbb{R})$, the complex eigenvalues must come in conjugate pairs, giving a contradiction.

The characterisation is given by the following proposition.

Proposition 7.3. *If G is a compact connected Lie group, then the exponential map is surjective.*

Proof. The central idea of the proof is to relate the exponential map in Lie theory to the one in Riemannian geometry. Firstly, every compact G can be given a Riemannian metric via a bi-invariant metric by averaging an arbitrary definite inner product on $\mathfrak{g}$ by the adjoint action of G using the Haar measure; which can then by translated by left translation to get a bi-invariant Riemannian structure on G. (Alternatively, one can also use the Peter-Weyl Theorem to embed $G \subset U(N)$ and then induce a metric on $\mathrm{Mat}_n(\mathbb{C})$.) Now we apply the Hopf-Rinow Theorem [36] and conclude that any two points are connected by at least one

[36] The theorem of Hopf-rinow states the following.

geodesic, so the Riemannian exponential map $\mathfrak{g} \to G$ formed by following geodesics from the origin is surjective. This requires noting that the group structure naturally defines a connection on the tangent bundle which is torsion-free and preserves the bi-invariant metric and therefore must agree with the Levi-Civita metric. $\square$

7.4 Functorial properties of exponential map

Having discussed the subtleties of the exponential map, we still have the following question — does every Lie algebra have a natural correspondence, under the exponential map, to a Lie group? In particular, the first principle that we have stated above, namely a map $f : G \to G'$ is uniquely determined by its differential $df_e : T_eG \to T_eG'$ at the identity, states that the following diagram commutes,

$$\begin{tikzcd} T_eG \arrow[r, "df_e"] \arrow[d, "\exp_G"'] & T_eG' \arrow[d, "\exp_{G'}"] \\ G \arrow[r, "f"] & G' \end{tikzcd}$$

The exponential map $\exp_G$ is a diffeomorphism between open neighbourhoods of $0 \in \mathfrak{g}$ and $e \in G$. Therefore the homomorphism f determined on the open subgroup $H \subset G$ is generated by some open subset $U \ni e$ in G, but since open subgroups are closed in any topological group (non-trivial cosets are open) so H is closed in G and $H = G$ by connectedness of G. This means given a Lie group homomorphism, we could always find a corresponding Lie algebra homomorphism which satisfies the commutative square.

This is not true in reverse. Given a Lie algebra map we may not be able to find a corresponding Lie group map under the exponential map. So what is the correspondence? To do this let us diverge and discuss centres and covers, following our discussion on covering spaces in §6.1.

Definition 7.11. The **centre** $Z(G)$ of a group G is defined as

$$Z(G) = \{z \in G \,|\, \forall g \in G, zg = gz\} \tag{7.19}$$

Proposition 7.4. *Let G be a Lie group, H be a connected manifold and $\phi : H \to G$ a covering space map. Let e' be an element lying over the identity e of G. Then there is a unique Lie group structure on H such that e' is the identity and ϕ is a map of Lie groups. The kernel of ϕ is the centre of H.*

Theorem 7.1. *Let M be a connected Riemannian manifold. The following are equivalent.*

(1) M is complete as a metric space.

(2) M is geodesically complete.

(3) There exists a point $p \in M$ such that $\mathcal{D}_p = T_pM$, $\mathcal{D}$ is the space of vector fields (space of sections on TM).

(4) A subset of M is compact iff it is closed and bounded.

Proof. The multiplication on G lifts uniquely to a map $H \times H \to H$ which takes $(e', e') \mapsto e'$. This product satisifes the group axioms (check). From Proposition 7.5 we can in fact do this for the universal cover H of G and then use the proposition to obtain the appropriate cover H. $\square$

Proposition 7.5. *Let H be a Lie group and $\Gamma \subset Z(H)$ a discrete subgroup of its centre. Then there is a unique Lie group structure on the quotient group $G = H/\Gamma$ such that the quotient map $q : H \to G$ is a Lie group map.*

Proof. This is straightforward from the axioms. $\square$

This motivates the following definition.

Definition 7.12. A Lie group map between two Lie groups G and H is an **isogeny** if it is a covering space map of the underlying manifolds.Then we say G and H are **isogenous**.

Isogeny generates an equivalence notion - by identifying Lie groups with their universal covers. This means that starting from a Lie group G, one can generate the universal covering space $\tilde{G}$ and then if the centre of the universal cover is discrete (as for all semisimple groups) $\tilde{(G)}/Z(\tilde{(G)})$ can be generated. This identifies

$$G \simeq \tilde{G} \simeq \tilde{G}/Z(\tilde{G}) \tag{7.20}$$

as an equivalence class in the set (category) of Lie groups.

Definition 7.13. Using the notation above, we call $\tilde{G}$ the **simply connected form** of the group G and $\tilde{G}/Z(\tilde{G})$, if it exists, the adjoint form of the group G.

We can now state the following theorem.

Theorem 7.2. *Let $\mathfrak{g}$ (over $k = \mathbb{R}, \mathbb{C}$) be a semisimple Lie algebra of finite dimension. Then there is a connected, simply connected Lie group G such that $\mathfrak{g}$ is the Lie algebra of G.*

Proof. This is Lie's Third Theorem and its modern version was first proved by Cartan. A discussion can be found in [28]. The actual proof is extremely complicated - the theorem can be proved by Ado's Theorem (as a Corollary, see Tao's webpage [29]) or topological [30] and cohomological proofs [31]. Also see [32–34]. $\square$

In fact, one can look at this from a functorial approach [32]. This consists of describing the assignment of the category of Lie groups G to the category of Lie algebras $\mathfrak{g}$ via the contravariant functor Lie:

$$\mathrm{Lie} : G \mapsto \mathfrak{g}\ . \tag{7.21}$$

The functorial properties of the exponential map is only true over the set of connected, simply-connected Lie groups such that $L(G) = \mathfrak{g}$. In the case of $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$, although they are isomorphic at the algebra level, $\mathfrak{su}(2) \cong \mathfrak{so}(3)$, under the exponential functor both will only be isomorphic if we take the universal cover over the two groups, $SU(2)$ and

$\widetilde{SO(3)} \cong SU(2)$ which is trivially the same. If the map between Lie groups $f : G \to G'$ is a discrete cover map, the descent to Lie algebra will then be an isomorphism.

I will end this section with a small example explicitly constructing the two-to-one map between $SU(2)$ and $SO(3)$.

Example 7.1 ($SU(2)$ and $SO(3)$ revisited)**.** To illustrate the point above let us revisit the $SU(2)$ and $SO(3)$ homomorphism, $\varphi : SO(3) \to SU(2)/\mathbb{Z}_2$. Suppose $A \in SU(2)$ and $R \in SO(3)$, where

$$A = \frac{1}{2}\operatorname{tr}(A)\mathbb{1} + \frac{1}{2}\operatorname{tr}(\boldsymbol{\sigma} A)\cdot \boldsymbol{\sigma} \ . \tag{7.22}$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli matrices. Then for a 2d complex vector $\mathbf{x}$ we then have,

$$\det(\mathbf{x}\cdot\boldsymbol{\sigma}) = -\mathbf{x}^2 \ . \tag{7.23}$$

Define a linear transformation $\mathbf{x} \to \mathbf{x}'$ as

$$\mathbf{x}'\cdot\boldsymbol{\sigma} = A\mathbf{x}\cdot\boldsymbol{\sigma}A^\dagger \ , \tag{7.24}$$

where $\mathbf{x}^{2'} = \mathbf{x}^2$. We can now use component-wise the $SO(3)$ transformation,

$$x'_i = R_{ij}x_j \ , \tag{7.25}$$

to get (using also $\operatorname{tr}(\sigma_i\sigma_j) = 2\delta_{ij}$),

$$\sigma_i R_{ij} = A\sigma_j A^\dagger \ , \tag{7.26}$$

i.e.

$$R_{ij} = \frac{1}{2}\operatorname{tr}\left(\sigma_i A \sigma_j A^\dagger\right) \ . \tag{7.27}$$

The converse can be constructed from $\sigma_j A^\dagger \sigma_j = 2\operatorname{tr}(A^\dagger)\mathbb{1} - A^\dagger$ to get

$$R_{jj} = |\operatorname{tr}(A)|^2 - 1, \quad \sigma_i R_{ij}\sigma_j = 2\operatorname{tr}\left(A^\dagger\right)A - \mathbb{1} \ , \tag{7.28}$$

which gives

$$A = \pm\frac{\mathbb{1} + \sigma_i R_{ij}\sigma_j}{2\,(1+R_{jj})^{\frac{1}{2}}} \ . \tag{7.29}$$

The sign $\pm$ then signifies the correspondence $\pm A \leftrightarrow R_{ij}$ which gives the map φ.

Distributed version

Exercise 7.1. Try repeating the same procedure of Example 7.1 for the map

$$\tilde{\varphi} : SO(3,1) \xrightarrow{\cong} SL(2,\mathbb{C})/\mathbb{Z}_2 \ . \tag{7.30}$$

8 Tensor Methods and Young Diagrams

I know a lot of you are extremely confused about Young diagrams and tensor methods. Here I am going to provide a pedagogical account of tensor methods, focussing on how you can utilise Young diagrams to produce irreducible tensorial representations of compact Lie groups. We will only focus on $SU(n)$-tensors in this short account [37], and you can look up on [1] for the other classical Lie groups. For completeness I will reproduce some results presented in the last class, but for efficient presentation of the material I will ignore most proofs (but I will try and point you to resources where you can look it up - they are not pretty in general!).

8.1 Tensor Irreps in $\mathrm{GL}(n)$ and $SU(n)$

First let us define tensors and tensor representations.

Definition 8.1. An r**-tensor representation** of a group G is a map $T : G \to GL(V^{\otimes r})$ such that T is a representation of G. A **rank** r G**-tensor** [38] is defined to be an element of the product space $V^{\otimes r}$. For any given basis $\{v_1, ...v_n\}$, we can expand a tensor in the form:

$$T = T_{i_1 i_2 ... i_r} v_{i_1} \otimes ... \otimes v_{i_r} \tag{8.1}$$

$T_{i_1 i_2 ... i_r}$ are then called the **components** of the tensor T with respect to the basis $\{v_i\}$.

Similarly, one can define a (p, q) tensor with dual vector spaces.

Definition 8.2. Take V^* to be the dual space of V (the set of linear maps $\varphi : V \to \mathbb{F}$ where $\mathbb{F}$ is a field). An (p, q)**-tensor representation** of a group G is a map $T : G \to GL(V^{\otimes p} \otimes V^{*\otimes q})$ such that T is a representation of G. A **rank** (p, q) G**-tensor** is defined to be an element of the product space $V^{\otimes p} \otimes V^{*\otimes q}$. For any given basis $\{v_1, ...v_p\}$ and dual basis $\{w^1, ..., w^q\}$, we can expand a tensor in the form:

$$T = T^{j_1 ... j_q}_{i_1 ... i_p} v_{i_1} \otimes ... \otimes v_{i_p} \otimes w^{j_1} \otimes ... \otimes w^{j_q} \tag{8.2}$$

In particular, we want to look at how tensor transforms under a group G. The most general case is the general linear group $\mathrm{GL}(n)$. For this, we see that a (p, q)-tensor in $\mathrm{GL}(n)$ will transform as

$$T^{j_1 ... j_q}_{i_1 ... i_p} \mapsto A_{i_1}{}^{k_1} ... A_{i_p}{}^{k_p} B^{j_1}{}_{l_1} ... B^{j_q}{}_{l_q} T^{l_1 ... l_q}_{k_1 ... k_p} \tag{8.3}$$

Here the matrices B are related to $A \in \mathrm{GL}(n)$, the fundamental transformation, as

$$B^T = A^{-1} \tag{8.4}$$

[37] We will in fact start with a discussion of GL_n tensors, but we will see later the invariant tensors of GL_n is exactly the same as $SU(n)$.

[38] We note that the tensors are not immediate representations of the groups themselves. They are in fact the objects that the representations transform on and forms space called the tensor space. The tensor space however is defined with respect to the group G; given a different group G the law of transformations on the space will be different and hence the space of tensors will be different.

Our central goal is to find irreducible representations of $\mathrm{GL}(n)$. From the above definitions, we can see that the tensors are in fact the objects that the matrices transform on. They therefore gives a representation of the invariant subspace that the irreducible representations of $\mathrm{GL}(n)$ (or in fact a matrix subgroup $G \subset \mathrm{GL}(n)$) act on.

Let us look at second-rank tensors to illustrate this point. Recall that the direct product of two vector spaces is a reducible space:

$$V \otimes V = \mathrm{Sym}^2 V \oplus \bigwedge^2 V \tag{8.5}$$

where $\mathrm{Sym}^2 V$ and $\bigwedge^2 V$ denotes the symmetric and antisymmetric part of the space. Suppose T_{ij} is a second-rank tensor and $\boldsymbol{p}$ is a permutation operator that switches the two indices of the second-rank tensor. Since $T_{ij} \in V \otimes V$, we can write the decomposition of Eq. (8.5) as

$$T_{ij} = T_{(ij)} + T_{[ij]} = (\mathrm{id} + \boldsymbol{p})\, T_{ij} \tag{8.6}$$

where id is just the identity map. The $GL(n)$-transformation acts on the tensor in the following manner,

$$T_{ij} \mapsto A_i{}^k A_j{}^l T_{kl} \ , \tag{8.7}$$

and we note in particular that the operator $\boldsymbol{p}$ commutes with this $GL(n)$-action. The reason for this is the product of matrices $A_i{}^k A_j{}^l$ is **bisymmetric**, the product indeed remains unchanged when the same permutation is applied to both $\mathrm{GL}(n)$ actions:

$$\begin{aligned} \boldsymbol{p} \cdot T_{ij} &= T_{ji} \\ &= A_j{}^k A_i{}^l T_{kl} \\ &= A_i{}^k A_j{}^l (\boldsymbol{p} \cdot T)_{kl} \end{aligned} \tag{8.8}$$

A similar calculation will show that the permutation operator $\boldsymbol{p}$ of any two indices indeed commute with the $\mathrm{GL}(n)$-action acting on the tensors since tensor transformation is bisymmetric. We conclude that the whole space of rank-r tensors is reducible into subspaces consisting of tensors of different symmetry properties. The rank-r tensors of a given symmetry forms the basis of an irreducible representation of $\mathrm{GL}(n)$ - this corresponds to the irreducible sub-block of the matrix product $A_i{}^k \ldots$ in the $\mathrm{GL}(n)$ transformations of tensors.

Definition 8.3. The rank-r tensors of a particular symmetry type forms a basis of irreps of $\mathrm{GL}(n)$; they are known as **irreducible tensors with respect to** $\mathrm{GL}(n)$.

We will henceforth discuss the irreps of a subgroup G of $\mathrm{GL}(n)$ by its representation as the invariant space in the tensor space, i.e. we will associate the irrep of G with the tensor itself.

The question is, how do we specialise this to the $SU(n)$ group? To do this we need to work out how irreducible representations of $\mathrm{GL}(n)$ can remain irreducible when we restrict to subgroups of $\mathrm{GL}(n)$. Firstly, the elements of the matrices of a $\mathrm{GL}(n)$ irrep in terms of the rank-r tensors are homogeneous polynomials of degree r in the elements A_{ij} of the transformation matrices [39]. Suppose the matrices are reducible when restricted to the subgroup $G \subset \mathrm{GL}(n)$. Since under a change of basis the matrix elements remain homogeneous polynomials of degree r in A_{ij}, say $p(A_{ij})$, then the restricted representation will be reducible for the subgroup $G \subset \mathrm{GL}(n)$ if there exists a certain set of homogeneous polynomials of degree r, $p_\alpha(A_{ij})$ which vanish for all $A \in G$ but not for all $B \in \mathrm{GL}(n)$.

Now we can try and see what happens when we restrict to $SU(n)$. First note the restriction to the unimodular group $\mathrm{SL}(n)$. Any matrix $A \in \mathrm{GL}(n)$ can be written as $A = \alpha B$ where $\alpha = (\det A)^{1/n}$, the polynomials $p_\alpha(A_{ij})$ are simply rescaled by a factor α^r for rank-r tensors so the vanishing of polynomials $p_\alpha(B_{ij})$ will require the original polynomials $p_\alpha(A_{ij})$ to vanish. Hence we have shown the following proposition.

Proposition 8.1. *The irreducible tensors and irreps of* $\mathrm{GL}(n)$ *and* $\mathrm{SL}(n)$ *are the same.*

Proof. See the vanishing of polynomial argument in the paragraph above. □

For $SU(n)$, recall that the elements of the Lie algebra of $SU(n)$, denoted by $\mathfrak{su}_n$, are simply the real form of the Lie algebra $\mathfrak{gl}_n\mathbb{C}$ [40]. Using the vanishing polynomial argument we can again conclude the following proposition.

Proposition 8.2. *The irreducible tensors and irreps of* $\mathrm{GL}(n)$ *and* $SU(n)$ *are the same.*

Proof. Consider the Lie algebra of $\mathrm{GL}(n)$, $\mathfrak{gl}_n$, and suppose X_{ij} be a basis of the $\mathfrak{gl}_n$ Lie algebra, so any element of $\mathfrak{gl}_n$ can be written as $\sum_{ij} \alpha_{ij} X_{ij}$ where $\alpha_{ij} \in \mathbb{C}$. For the Lie algebra of $SU(n)$, we will need $\alpha_{ij} \in \mathbb{R}$. Now suppose a representation of the Lie algebra basis element X_{ij} is given by the matrices x_{ij}. Then if a representation is reducible for $\mathfrak{su}_n$, we will need to find a basis in which the matrices $\alpha_{ij}x_{ij}$ are in reduced form for all $\alpha_{ij} \in \mathbb{R}$. Using linearity the linear forms $\alpha_{ij}x_{ij}$ must also vanish for any $\alpha_{ij} \in \mathbb{C}$ so the representation is also reducible for $\mathfrak{gl}_n$. The converse is the same, and we can then use the exponential map to lift the argument to the Lie groups equivalents. □

The above proposition therefore implies that we can analyse $SU(n)$-tensors the same way we analyse $\mathrm{GL}(n)$ tensors - they have the same irreducible structures. We will from now on use the two terms interchangeably.

In $SU(n)$, we distinguish between lower and upper indices using the dual basis as described above for $\mathrm{GL}(n)$ tensors. The two kinds of indices transform in under different representations in $SU(n)$, as illustrated explicitly in the following definition.

Definition 8.4. A (p, q) **tensor** in $\mathbf{n}^p \otimes \bar{\mathbf{n}}^q$ of $SU(n)$ carries p lower and q upper indices $\phi^{\nu_1 \dots \nu_q}_{\mu_1 \dots \mu_p}$. It transforms in $SU(n)$ as:

$$\phi^{\nu_1 \dots \nu_q}_{\mu_1 \dots \mu_p} \mapsto U_{\mu_1}{}^{\rho_1} \dots U_{\mu_p}{}^{\rho_p} \overline{U}^{\nu_1}{}_{\sigma_1} \dots \overline{U}^{\nu_q}{}_{\sigma_q} \phi^{\sigma_1 \dots \sigma_q}_{\rho_1 \dots \rho_p} \tag{8.9}$$

[39] To see this, take the list of the tensor components as a vector and work out the transformation matrix for a few simple cases.

[40] For a detailed discussion of complexifications of Lie algebras, see §9.

where $U \in SU(n)$ and $\overline{U}$ is the complex conjugate of U. An **invariant tensor** is a tensor that is invariant under an $SU(n)$ transformation. The **dual tensor** of a (p, q) tensor is a (q, p) tensor.

In $SU(n)$, there are only three invariant tensors.

Proposition 8.3. *The invariant tensors of $SU(n)$ are the Kronecker delta δ^{μ}_{ν} and the Levi-Civita tensors $\epsilon_{\mu_1 \dots \mu_n}$ and $\epsilon^{\nu_1 \dots \nu_n}$.*

Proof. Explicit calculation. □

What do we want to do with tensors? From the argument above, we saw that tensors are extremely useful in representation theory of Lie groups and Lie algebras as they can represent irreducible representations. We have two items on our wish-list:

1. Constructing new irreps from old ones.

2. Finding branching rules.

We will find that tensors and Young diagrams play a central role in achieving these goals. As mentioned above, we will mainly focus on $SU(n)$-tensors. There are some additional subtleties that must be dealt with when we talk about $\mathrm{SO}(n)$ and $\mathrm{Sp}(n)$ tensors - we will only briefly chat about this but this will not be the main focus on this section.

8.2 Building irreps using invariant tensors

Let us tackle our first goal of building irreducible representations. This is done by using the invariant tensors listed above for $SU(n)$ - namely the Levi-Civita tensors and the Kronecker Delta. The main procedure is as follows:

Proposition 8.4 (Building irreps using invariant tensors)**.** *To build irreducible representations using invariant tensors, one follows the following procedure:*

1. *Start with a particular (p, q) tensor. For tensor products, just multiply the two tensors together.*

2. *Now build all possible contractions with the invariant tensors.*

3. *Subtract these contractions from the overall tensor and repeat step 2 until the leftover tensor is also invariant.*

To see this in action let us look at an example:

Example 8.1. Let us try and decompose ϕ_{ij} in $SU(n)$. The only invariant tensors we can use at this order is ϵ^{ab}, so we apply this to project this component out:

$$\phi_{[ij]} = \frac{1}{2}\epsilon^{ab}\phi_{ab}\epsilon_{ij} \tag{8.10}$$

Subtracting this from the tensor gives the full decomposition for rank-2 tensors:

$$T_{ij} = T_{(ij)} + T_{[ij]} \tag{8.11}$$

The above example in fact works for the tensor product $\phi_i \otimes \psi_j$, which we can just write as the rank-2 tensor $T_{ij} = \phi_i \psi_j$.

From the example above, it is therefore, in principle, possible to carry out Clebsch-Gordon decomposition by simply using invariant tensors and contracting indices whenever possible. This is in general extremely annoying and algebraically-challenging. Therefore, we will instead exploit a property of the GL(n) irreps that allows us to ultilise Young diagrams to perform such computations.

Remark 8.1. SO(n) and Sp(n) tensors are a bit more annoying. For example, in the example above for the full decomposition of SO(n) one must also project out the trace part to get the full decomposition of the tensor.

8.3 Irreducible Representations of S_n and Young Diagrams

A brief interlude. I want to come back to our finite group discussion, and briefly return to symmetric groups S_n. For completeness, I am going to recall some facts from §2 and §5 related to symmetric groups S_n. First let us recall some facts about irreducible representations of S_n.

Proposition 8.5. *The number of irreducible representations of S_n is the number of conjugacy classes, which is the number $p(n)$ of partitions of n. The **partitions** of n is*

$$p(n) = \lambda_1 + ... + \lambda_k \, , \lambda_1 \geq ... \geq \lambda_k \geq 1 \, . \tag{8.12}$$

Proof. The number of irreps is clearly equal to the number of conjugacy classes [41]. It remains to show that the number of conjugacy classes is equal to the number of partitions of n. This is however a corollary of Theorem 2.3. □

To each partition of n we can associate a Young diagram, as defined in §5.

Definition 8.5. A **Young diagram** is a collection of boxes, or cells, arranged in left-justified rows, with a weakly decreasing number of boxes in each row. Listing the number of boxes in each row gives a partition $p(n)$ of integer n, where it is also the total number of boxes. Conversely, as stated above, each partition correspond to a Young diagram.

A Young diagram is sometimes called a Young frame or Ferrers diagram. To each diagram, we can number the boxes consecutively to generate Young tableau.

Definition 8.6. A **Young tableau** is a filling of a Young diagram such that the numbers are weakly increasing across each row but strictly increasing down each column. A **standard tableau** is a tableau in which entries are numbers from 1 to n, n being the total number of boxes.

Example 8.2. As an example let us look at the partitions of $n = 3$. Then we have the following partitions, Young diagrams and Standard Young Tableaux as illustrated in Table 8.1.

[41] If you for once doubt this statement, look at Theorem 2.8 of Andre's notes.

Partition	Young Diagram	Standard Young Tableaux
(3)	$\ydiagram{3}$	$\begin{ytableau} 1 & 2 & 3 \end{ytableau}$
$(2,1)$	$\ydiagram{2,1}$	$\begin{ytableau} 1 & 2 \\ 3 \end{ytableau}$ $\begin{ytableau} 1 & 3 \\ 2 \end{ytableau}$
$(1,1,1)$	$\ydiagram{1,1,1}$	$\begin{ytableau} 1 \\ 2 \\ 3 \end{ytableau}$

Table 8.1: The Young diagrams and all possible Standard Young Tableaux for dimension-3.

Recall that Young tableaux can be used to describe projection operators for the regular representation which will then give the irreducible representations of S_n, which we have already seen in §5. Given a standard Young tableau, we can define two subgroups of the symmetry group,

$$P = P_\lambda = \{g \in S_n \mid g \text{ preserves each row}\} \ , \tag{8.13}$$

$$Q = Q_\lambda = \{g \in S_n \mid g \text{ preserves each column}\} \ . \tag{8.14}$$

Then in the group algebra $\mathbb{C}S_n$ we can introduce two elements corresponding to these two subgroups. In particular, we define

$$a_\lambda = \sum_{g \in P} e_g, \quad b_\lambda = \sum_{g \in Q} \mathrm{sgn}(g) e_g \ . \tag{8.15}$$

The actions of $a_\lambda, b_\lambda \in \mathbb{C}S_n \to \mathrm{End}(V^{\otimes n})$ are the following subspaces.

$$\mathrm{Im}(a_\lambda) = \mathrm{Sym}^{\lambda_1} V \otimes ... \otimes \mathrm{Sym}^{\lambda_k} V \subset V^{\otimes n} \ , \tag{8.16}$$

$$\mathrm{Im}(b_\lambda) = \textstyle\bigwedge^{\mu_1} V \otimes ... \otimes \bigwedge^{\mu_l} V \subset V^{\otimes n} \ , \tag{8.17}$$

where μ is the conjugate partition to λ obtained by flipping the diagram along the 45° line. We can now define the following object.

Definition 8.7 (Young Projectors)**.** The **Young symmetriser** is defined as

$$c_\lambda = a_\lambda \cdot b_\lambda \in \mathbb{C}S_n \ . \tag{8.18}$$

Take $A = \mathbb{C}S_n$ as the group ring of S_n. If c_λ is a Young symmetriser, then the correesponding representation is $V_\lambda = Ac_\lambda$. We have the following proposition.

Proposition 8.6. *An element of S_n can be written in at most one way as a product $p \cdot q$ where $p \in P$ and $q \in Q$.*

Proof. $P \cap Q = \{1\}$. So no commutations are possible. □

This means that the form of c_λ can be written as

$$c_\lambda = \sum_{g=p\cdot q} \pm e_g \ , \tag{8.19}$$

with the coefficient ± 1 being the sign of q. Now we have the following lemma.

Lemma 8.1. *We have the following three statements.*

(1) For $p \in P$, $p \cdot a = a \cdot p = a$.

(2) For $q \in Q$, $(\mathrm{sgn}(q)q) \cdot b = b \cdot (\mathrm{sgn}(q)q) = b$.

(3) For all $p \in P$, $q \in Q$, $p \cdot c \cdot (\mathrm{sgn}(q)q) = c$ and up to multiplication by a scalar, c is the only such element in A.

Proof. The first two statements are obvious. So look at the third statement. Let us suppose $\sum_g n_g e_g$ satisfies the condition in (3). Then $n_{pgq} = \mathrm{sgn}(q) n_g$ for all g, p, q and $n_{pq} = \mathrm{sgn}(q) n_1$ where n_1 is the coefficient for the fully anti-symmetrised irrep. It is enough to show $n_g = 0$ when $g \notin PQ$. For such g it suffices to find a transposition t such that $p = t \in P$ and $q = g^{-1} t g \in Q$ so then $g = pgq$ and $n_g = -n_g$. Now suppose $T' = gT$ is the tableau obtained by replacing each entry i of T by $g(i)$. The claim is that there are two distinct integers that appear in the same row of T and in the same column of T' where the transposition t acts on. Take $p_i \in P$ and $q_1' \in Q' = gQg^{-1}$ such that $p_1 T$ and $q' T'$ have the same first row. Repeat this for the rest of the tableau to get $p \in P$ and $q' \in Q'$ such that $pT = q'T'$. Then $pT = q'gT$ so $p = q'g$. Defining $q = g^{-1}(q')^{-1} g \in Q$ we get $g = pq$. So if we don't have such distinct pair of integers t then we can write $g = pq$. □

Before we continue let me define what it means by partitions being lexicographically-ordered.

Definition 8.8. The partitions are ordered **lexicographically** when $\lambda > \mu$ if the first non-vanishing $\lambda_i - \mu_i$ is positive.

Finally we have the following lemma.

Theorem 8.1. *The following two statements are true.*

(1) If $\lambda > \mu$ then for all $x \in A$, $a_\lambda \cdot x \cdot b_\mu = 0$. In particular, if $\lambda > \mu$, then $c_\lambda = c_\mu = 0$.

(2) For all $x \in A$, $c_\lambda \cdot x \cdot c_\lambda$ is a scalar multiple of c_λ. In particular, $c_\lambda \cdot c_\lambda = n_\lambda c_\lambda$ for some $n_\lambda \in \mathbb{C}$.

In particular, that statement implies some scalar multiple of c_λ is idempotent, i.e.

$$c_\lambda^2 = n_\lambda c_\lambda \ , \tag{8.20}$$

and the image of c_λ is an irrep V_λ of S_n, where the multiplication on $\mathbb{C}S_n$ acts on the right. Every irrep of S_n can be obtained in this way for a unique partition.

Proof. First prove (1). Take $x = g \in S_n$. Since $g \cdot b_\mu \cdot g^{-1}$ is the element constructed from gT', where T' is the tableau used to generate b_μ, then it suffices to show that $a_\lambda \cdot b_\mu = 0$. Note $\lambda > \mu$ implies that there are two integers in the same row of T and the same column of T'. If t is the transposition of these integers, then $a_\lambda \cdot t = a_\lambda$, $t \cdot b_\mu = -b_\mu$ so we must have $a_\lambda \cdot b_\mu = a_\lambda \cdot t \cdot t \cdot b_\mu = -a_\lambda \cdot b_\mu$, as required. Part (2) immediately follows from Lemma 8.1. $\square$

This theorem gives a direct correspondence between conjugacy classes in S_n and irreducible representations of S_n - which is not true for general groups! We will eventually see that the image of the symmetrisers c_λ in $V^{\otimes n}$ provide essentially all the finite-dimensional irreps of $\mathrm{GL}(V)$.

So now we have two separate unrelated things. On the one hand, we have $\mathrm{GL}(n)$ tensors which effectively furnish irreps of $\mathrm{GL}(n)$; on the other hand, we have irreducible representations of S_n which can be analysed using Young diagrams. Let us now make the connection between the two kinds of objects.

8.4 Schur-Weyl Duality

Let us briefly return to the group $\mathrm{GL}(n)$. We have seen in the first section how we can decompose a rank-2 $\mathrm{GL}(n)$-tensor T_{ij} into two irreducible parts - namely the symmetric and anti-symmetric part.

$$T_{ij} = T_{(ij)} + T_{[ij]} = (\mathrm{id} + \boldsymbol{p})\, T_{ij} \ . \tag{8.6}$$

Here $\boldsymbol{p} = (12)$ in cycle notation, the permutation of the two indices. The operator $\boldsymbol{p}$ is obviously an element of the symmetric group S_2 so this begs the question - are the irreps of $\mathrm{GL}(n)$ related to S_n?

It turns out the symmetric group S_d does act on $V^{\otimes r}$ on the right and the action, as discussed above, commutes with the left action of $\mathrm{GL}(V)$. It is therefore possible to relate the irreducible representations of $\mathrm{GL}(n)$ with irreducible representations of the symmetric group S_n. This turns out to be an extremely powerful statement - one can now use the machinery from the representation theory of finite symmetric groups in analysing the representation theory of $\mathrm{GL}(n)$ and its subgroups!

Let us briefly illustrate this principle by looking at Eq. (8.6) again. To obtain tensors of a particular symmetry type, we apply Young symmetrisers to the indices of the general rank-r tensor $T_{i_1 \dots i_r}$. For example, there are two types of rank-2 tensors, namely the symmetric type

$$T_{(ij)} \longleftrightarrow \boxed{i}\boxed{j} \, , \tag{8.21}$$

and the antisymmetric type

$$T_{[ij]} \longleftrightarrow \begin{array}{|c|}\hline i \\ \hline j \\ \hline\end{array} \, . \tag{8.22}$$

We see that the two tensors are obtained by applying the Young symmetrisers $c_\pm = e \pm t$ where t indicates the transposition between the two indices respectively. This is true for rank-r tensors in general, with one exceptional caveat. Given a Young symmetriser of a particular partition λ of integer n, the Young symmetriser acts on the right on the product of spaces $V^{\otimes n}$ to obtain an invariant subspace $\mathbb{S}_\lambda V^{\oplus r_\lambda}$ where r_λ is the number of standard Young tableaux obtainable from the Young diagram corresponding to the partition λ.

We now state the most important theorem in this section - the theorem allows us to relate Young diagrams with irreducible tensors we have been discussing about. Define the image of the Young symmetriser c_λ on $V^{\otimes n}$ by $\mathbb{S}_\lambda V$,

$$\mathbb{S}_\lambda V = \mathrm{im}\left(c_\lambda|_{V^{\otimes n}}\right) \ , \tag{8.23}$$

which is a representation of $\mathrm{GL}(V)$. Then we have the following theorem and corollary.

Theorem 8.2. *The following statements about the action of c_λ on $V^{\otimes n}$ are true.*

(1) Let $k = \dim V$. Then $\mathbb{S}_\lambda V$ is zero if $\lambda_{k+1} \neq 0$. If λ corresponds to the partition $\lambda = (\lambda_1 \geq ... \geq \lambda_k \geq 0)$, then

$$\dim \mathbb{S}_\lambda V = \prod_{1 \leq i \leq j \leq k} \frac{\lambda_i - \lambda_j + j - i}{j - i} \ . \tag{8.24}$$

(2) Let m_λ be the dimension of the irreducible representation V_λ of S_n corresponding to λ. Then

$$V^{\otimes n} \cong \bigoplus_\lambda \mathbb{S}_\lambda V^{\otimes m_\lambda} \ . \tag{8.25}$$

(3) Each $\mathbb{S}_\lambda V$ is an irreducible representation of $\mathrm{GL}(V)$.

Proof. See §8.12 for the proof. □

In particular we have the following corollary.

Corollary 8.1 (Schur-Weyl Duality)**.** *If $c \in \mathbb{C}S_n$ and $(\mathbb{C}S_n) \cdot c = \bigoplus_\lambda V_\lambda^{\oplus r_\lambda}$ as representations of S_n, then there is a corresponding decomposition of $\mathrm{GL}(n)$-spaces:*

$$V^{\otimes n} \cdot c = \bigoplus_\lambda \mathbb{S}_\lambda V^{\oplus r_\lambda} \tag{8.26}$$

For the details of the proofs and extended statements of the theorem and corollary I refer to §8.12 and [1]. We proceed by assuming that this is a known result, and apply this to analyse tensor representations. This is amazing. With the Schur-Weyl duality we can now proceed and analyse tensor irreps! Recall that $\mathrm{GL}(n)$ irreducible tensors are the same as $SU(n)$ irreducible tensors. There is one extra caveat that we need to consider when we apply the Schur-Weyl duality to $SU(n)$ tensors however. Although we have shown that the invariant tensors are the same, the irreducible tensors in $SU(n)$ that descends from $\mathrm{GL}(n)$ might not be independent.

Proposition 8.7. *The partition λ and λ' where $\lambda = (\lambda_1 \geq ... \geq \lambda_n \geq 0)$ and $\lambda' = (\lambda_1 + s \geq ... \geq \lambda_n + s \geq 0)$ of unimodular subgroups of $\mathrm{GL}(n)$, such as $\mathrm{SL}(n)$ and $SU(n)$, are equivalent.*

Proof. In particular, note that every column of length-n simply corresponds to the factor $(\det A)$. Since $\det A = 1$ for unimodular groups, the two partitions are equivalent. $\square$

In particular, we only need to consider patterns which have fewer than n rows. There is a second equivalence to relate conjugate representations.

Proposition 8.8. *The partition λ and λ' where $\lambda = (\lambda_1 \geq \ldots \geq \lambda_n \geq 0)$ and $\lambda' = (\lambda_1 - \lambda_n \geq \ldots \geq \lambda_1 - \lambda_2 \geq 0)$ of unimodular subgroups of $\mathrm{GL}(n)$, such as $\mathrm{SL}(n)$ and $SU(n)$, are equivalent.*

Proof. This can be seen from the existence of the Levi-Civita tensors, which relates the two partitions by acting on each of the columns of indices to obtain the conjugate indices. $\square$

The second equivalence is what we call *conjugation*. In fact, we are going to use this property extensively in our analysis of tensor irreps of $SU(n)$. In practice, the conjugate tensor irrep is obtained by looking for the Young diagram that completes the $n \times d$ rectangle, where n is the one in $SU(n)$ and d is the number of columns in the original partition. For example, the following two partitions are equivalent in $SU(4)$,

$$\yng(5,3,2,1) \longleftrightarrow \yng(4,3,2) . \tag{8.27}$$

8.5 Tensors and Young Diagrams

Now that we have established the correspondence between tensor irreps and Young diagrams (and got the ugly maths out of the way), we can apply this technology to analyse tensors in $SU(n)$. The general rule to produce irreducible tensors is as follows:

1. Write out all partitions of r in Young diagram form for producing irreducible tensors of rank-r.

2. Write out all standard Young tableaux allowed by the rules. Each of that corresponds to an irreducible tensor.

We clearly see from the above example that Young diagrams correspond to the type of irreducible tensor you are constructing (i.e. the irrep), whilst the number of standard Young tableaux you can generate from that Young diagram would be the dimension of that irrep (see Corollary 8.1). It is often not feasible to write down all standard Young tableaux. Instead, there is a way to calculate the dimension of the tensor irrep as stated by the following proposition.

Proposition 8.9. *The dimension of an irrep in $GL(n, \mathbb{C})$ corresponding to a Young diagram can be calculated using the following algorithm.*

1. *First draw two copies of the Young diagram.*

2. *For the first diagram, start with n at the upper-left corner and assign numbers - for any box to the right add one, for any box below a box that already has a number in it minus one from that of the box above.*

3. *Compute Hook lengths using the second Young diagram.*

4. *Multiply all numbers in each diagram and divide the product of the numbers in the second diagram from that of the first - this is the dimension of the irrep.*

Proof. Not too hard but skipped - this is a direct consequence of Theorem 8.2 and dealt with in Exercise 6.4 of [1]. □

Example 8.3. Let us compute the dimension of

$$\begin{array}{|c|c|}\hline & \\ \hline & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array},$$

in $SU(3)$. The two diagrams are

$$\begin{array}{|c|c|}\hline 3 & 4 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \quad \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 1 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array},$$

where the first diagram corresponds to the n assignment and the second diagram is the Hook length diagram (Young tableaux with Hook lengths at each entry). Multiplying all the numbers in each diagram and dividing the numbers from 1 to 2 gives the dimension $4!/3 = 8$, as shown in the lecture notes (this is the adjoint representation of $SU(3)$).

In $SU(n)$ there is a caveat - as discussed above you are allowed dual partitions (Proposition 8.8). In particular we have the following corollary on our rules on applying Young diagrams to $SU(n)$ tensors.

Proposition 8.10. *The dual irreducible representation $\bar{\boldsymbol{r}}$ of $\boldsymbol{r}$ of $SU(n)$ can be obtained by drawing a "dual" diagram such that summing the number of boxes for each column for the two diagrams gives n.*

Proof. Direct consequence of Proposition 8.8. □

Let us look at a few examples.

Example 8.4. In $SU(4)$ the anti-fundamental representation $\bar{\mathbf{4}}$ is

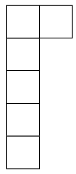
$$\begin{array}{|c|}\hline \\ \hline \\ \hline \\ \hline\end{array},$$

corresponding to the tensor $\phi^{\mu} = \epsilon^{\mu\nu\rho\sigma}\phi_{\nu\rho\sigma}$. The fundamental representation of course is

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

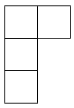
with corresponding tensor ϕ_{μ}. We can see that adding up the first column gives precisely 4 boxes.

Example 8.5. Since we have the Clebsch-Gordon decomposition $\mathbf{n} \otimes \bar{\mathbf{n}} = \mathbf{1} + \mathbf{adj}$, the adjoint representation of $SU(n)$ is always has the form - $(n-1)$ boxes in the first column and 1 box in the second column, i.e. for example for adjoint of $SU(6)$ this is


,

Remark 8.2. From Proposition 8.7 and the above argument you can see for $SU(n)$ the number of boxes per column cannot exceed n. If the number of boxes in a column is precisely n, this means this is equivalent to the singlet representation. You can alternatively see from the Young projector - you cannot anti-symmetrise $(n+1)$ objects non-trivially if you only have n unique objects.

Remark 8.3. When you are writing for an explicit form for the tensors, note carefully how you compose anti-symmetrisation and symmetrisation. Sometimes is best to write everything out, for example in $SU(4)$, we have the correspondence[42]:



$$\longleftrightarrow \phi^{\lambda}_{\mu} = \epsilon^{\lambda\nu\rho\sigma}(\phi_{\nu\rho\sigma\mu} + \phi_{\nu\rho\mu\sigma}) \ . \tag{8.28}$$

To do the tensor identification, it might be helpful to label the Young diagram using indices. In the above case, we can label the diagram using,

$$\begin{array}{|c|c|} \hline \sigma & \mu \\ \hline \nu & \multicolumn{1}{c}{} \\ \cline{1-1} \rho & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array} \ , \tag{8.29}$$

so we have symmetrised $\mu \leftrightarrow \sigma$ in the last equality in Eq. (8.28). The upper dual index can be used to replace $n-1$-indices, in this case we have swapped $[\nu\rho\sigma]$ with an upper index λ.

Distributed version:

Exercise 8.1. Try Q4 of the second sheet again and find the Young tableaux and their associated tensors for the irreps $\mathbf{1}$, $\mathbf{5}$, $\bar{\mathbf{5}}$, $\mathbf{15}$, $\mathbf{10}$ and $\mathbf{24}$ of $SU(5)$.

This establishes the necessary ingredients in relating tensor irreps to Young diagrams.

8.6 Clebsch-Gordon decomposition from Young tableaux

Having described the tensor irreps in terms of Young diagrams and tableaux, we can now try to formulate tensor products and their Clebsch-Gordon decompositions in the same language. Recall the Clebsch-Gordon decomposition is the following:

$$\mathbb{S}_{\lambda}(V) \otimes \mathbb{S}_{\mu}(V) = \bigoplus N_{\lambda\mu\nu} \mathbb{S}_{\nu}(V) \ , \tag{8.30}$$

[42] Note that you can use ( ) and [ ] as symmetrising and antisymmetrising symbols, but be careful about how you write them so they are not confusing. So for this example $\phi_{(\nu[\rho)\mu\sigma]}$ is allowed.

the integers $N_{\lambda\mu\nu}$ describing how many times the irrep $\mathbb{S}_\nu(V)$ appears in the tensor product. Here we will illustrate a practical algorithm to realise this decomposition - as mentioned in the notes, this is known as the Littlewood-Richardson Rule (or the Littlewood-Richardson Theorem, see [1] for details). I am just going to state the algorithm here without proof.

Proposition 8.11. *(Clebsch-Gordon decomposition from Young tableaux) To tensor A_{n-1} representations for two Young tableaux λ and μ proceeds as follows:*

1. *Write the first Young tableaux λ* [43] *with as in the first row, bs in the second and so on. For example:*
$$\begin{array}{|c|c|c|}\hline a & a & a \\ \hline b & b \\ \cline{1-2} c \\ \cline{1-1}\end{array}$$

2. *Now attach boxes of λ to μ one by one from the first row (labelled as) onwards. We need to make sure that no two letters of the same type appear in the same column*[44] *and the result is always a valid Young tableau.*

3. *For each of Young tableaux obtained in the previous step, read all letters from right to left and top to bottom. (So you read from right to left along the first row, and then in the same direction in the second row, and so on.) This sequence, say aaabbcabc, must form a* ***lattice permutation***[45]*, i.e. we count the total number of every letter encountered at each box, and we require the number of as to be more than or equal to the number of bs (and any subsequent letter), and so on. Otherwise we discard that Young tableau.*

4. *The number of the same diagrams that are obtained from this method are integers $N_{\lambda\mu\nu}$, which encodes the multiplicity of the irrep in the tensor product.*

Proof. We need the so-called Littlewood-Richardson Theorem. You can find this in [1]. $\square$

This probably deserves an example:

Example 8.6. Let us look at the tensor product:

$$\square \otimes \begin{array}{|c|c|}\hline a & a \\ \hline b \\ \cline{1-1}\end{array}\,. \tag{8.31}$$

Here I am deliberately decomposing the larger Young diagram to illustrate the rule - in practice you should always choose the smaller one. Now we do the second step:

$$\left(\begin{array}{|c|c|}\hline & a \\ \hline\end{array} \oplus \begin{array}{|c|}\hline \\ \hline a \\ \hline\end{array} \right) \otimes \begin{array}{|c|}\hline a \\ \hline b \\ \hline\end{array}\,. \tag{8.32}$$

[43] I would canonically choose the one with the smaller size as my first Young tableaux.

[44] This is to ensure that the indices that are originally symmetrised do not get anti-symmetrised, i.e. cancellation due to antisymmetry.

[45] A lattice permutation is a string composed of positive integers in which every prefix (substring before that integer) contains at least as many positive integers i as integers $i+1$. In our case just map $a \leftrightarrow 1$, $b \leftrightarrow 2$, etc.

Apply the second a:

$$\left(\begin{ytableau} \; & a & a \end{ytableau} \oplus \begin{ytableau} \; & a \\ a \end{ytableau} \oplus \begin{ytableau} \; & a \\ a \end{ytableau} \oplus \begin{ytableau} \; \\ a \\ a \end{ytableau} \right) \otimes \begin{ytableau} b \end{ytableau} \,. \tag{8.33}$$

Now eliminate the invalid diagrams and the degenerate [46] ones:

$$\left(\begin{ytableau} \; & a & a \end{ytableau} \oplus \begin{ytableau} \; & a \\ a \end{ytableau} \right) \otimes \begin{ytableau} b \end{ytableau} \,. \tag{8.34}$$

Add the third box:

$$\begin{ytableau} \; & a & a & b \end{ytableau} \oplus \begin{ytableau} \; & a & a \\ b \end{ytableau} \oplus \begin{ytableau} \; & a & b \\ a \end{ytableau} \oplus \begin{ytableau} \; & a \\ a & b \end{ytableau} \oplus \begin{ytableau} \; & a \\ a \\ b \end{ytableau} \,. \tag{8.35}$$

Now according to the third rule, the first and third diagrams are invalid so we take them away. We are therefore left with:

$$\begin{ytableau} \; \end{ytableau} \otimes \begin{ytableau} \; & \; \\ \; \end{ytableau} = \begin{ytableau} \; & \; & \; \\ \; \end{ytableau} \oplus \begin{ytableau} \; & \; \\ \; & \; \end{ytableau} \oplus \begin{ytableau} \; & \; \\ \; \\ \; \end{ytableau} \,. \tag{8.36}$$

In $SU(3)$, this means that we have the decomposition (following the rule for computing dimensions):

$$\mathbf{3} \otimes \mathbf{8} = \mathbf{15} \oplus \mathbf{6} \oplus \mathbf{3} \,, \tag{8.37}$$

recalling that three boxes in a column is just the singlet (i.e. $\mathbb{C}$).

8.7 Branching by index decomposition

Now we turn to branching - this is very important as we can generate subgroup representations from existing representations in the group. First recall what branching is.

Definition 8.9. Say H in a subgroup of a group G. A **branching rule** or **branching** is the restricted representation decomposition of irrep $R^{(G)}$ the form:

$$R^{(G)} \to R^{(H)} = \bigoplus n_s R_s^{(H)} \,, \tag{8.38}$$

where now $R_s^{(H)}$ are the irreps of H and n_s indicates the multiplicity of the irrep $R_s^{(H)}$ in the branching.

Branching is well-defined in the context of Lie groups. In fact, one can show that the Littlewood-Richardson Rule above allows you to give a restriction on a sub-representation of an irrep. This requires the use of the so-called Pieri's formula [1]. Here I will present some practical way of doing branching using tensors.

The first method is *index decomposition*. The idea is simple - we rewrite our tensor indices as the ones in the subgroup and deduce all possible combination of these new indices under appropriate symmetrisation and anti-symmetrisation. The algorithm is as follows:

[46] Degenerate diagrams are the diagrams where the labels are written in the same way.

1. First write the tensor with respect to the Young diagram. We use explicitly the indices μ.
2. Replace the μ indices with indices specific to the subgroup, say a and i.
3. Compute all possible combinations and take away degenerate terms. Compute the relevant dimensions (no index - singlet field).

Example 8.7. Let us do this for one $SU(5)$ irrep and consider the canonical splitting $SU(5) \to SU(3) \times SU(2)$. We denote μ, ν as $SU(5)$ indices and replace that with $a, b, ...$ and $i, j, ...$, the indices of $SU(3)$ and $SU(2)$. Now the branchings can be easily worked out by specialising the $SU(5)$ indices on a tensor to all possible combinations of $SU(3)$ and $SU(2)$ indices. Recall for the fundamental representation $\phi_\mu \sim \mathbf{5}$, so the branching

$$\phi_\mu \mapsto (\phi_a, \phi_i) \ , \tag{8.39}$$

or

$$\mathbf{5} \mapsto (\mathbf{3}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{2}) \ . \tag{8.40}$$

Now let us look at the **15** representation. This splits into [47]

$$\phi_{(\mu\nu)} \mapsto \left(\phi_{(ab)}, \phi_{ai}, \phi_{(ij)}\right) \ , \tag{8.41}$$

or

$$\mathbf{15} \mapsto (\mathbf{6}, \mathbf{1}) \oplus (\mathbf{3}, \mathbf{2}) \oplus (\mathbf{1}, \mathbf{3}) \ . \tag{8.42}$$

8.8 Branching using Young tableaux

There is of course another way of doing branching. Having established the one-to-one correspondence between tensor irreps and Young diagrams, one can actually try and split up possible Young diagrams from an existing Young diagram to generate branching tensors. The algorithm for branching for $SU(n+m) \to SU(n) \times SU(m)$ again is as follows.

1. Start with the Young diagram corresponding to the $SU(n+m)$-tensor irrep that you want to branch off from.
2. Generate all possible splitting possible from the Young diagram. In particular, we consider all pairs of diagrams, (Y_1, Y_2), where Y_1 and Y_2 denote the Young diagrams of the irreducible representations in $SU(n)$ and $SU(m)$ respectively, and tensor them up in $SU(n+m)$ using the Clebsch-Gordon decomposition in §8.6 rule by treating the boxes indexed from $1, \ldots, n$ and $n+1, \ldots, n+m$. The number of times (multiplicity) that the pair (Y_1, Y_2) appears in the branching rule, $n_{y1,y2}$, is given by the multiplicity of the original irrep in the decomposition $Y_1 \otimes Y_2$.
3. For simple cases, one can split the Young diagram corresponding to the original irrep into sub-diagrams and count the number of dimensions on both sides such that they match. If the dimensions do not add up, we additionally consider diagrams that we have missed in the above step.

[47] Noting by antisymmetry that $\phi_{ai} = -\phi_{ia}$ so the two tensors (ϕ_{ai}, ϕ_{ia}) get shoved into the same irrep.

4. Finally label the splitting using dimensions.

Technically, what we are actually computing is the embedding $SU(n) \times SU(m) \subset SU(n+m)$. The number of times the original representations appears is then the total number of times that the representations R_n and R_m of $SU(n)$ and $SU(m)$ respectively will appear in the branching [8]. It is much easier to see this in action so let us do that for the $SU(5)$ tensors again.

Example 8.8. Let us look at the branching of the fundamental rep of $SU(5)$ into $SU(3)\times SU(2)$, $\mathbf{5} \mapsto (\mathbf{3},\mathbf{1}) \oplus (\mathbf{1},\mathbf{2})$. To do this we start with one box and try and split the diagram (we cannot) so we have, with the two boxes indicating the irreps in $SU(3)$ and $SU(2)$ respectively,

$$\ydiagram{1} \mapsto (\ydiagram{1}\,,\bullet) \oplus (\bullet\,,\ydiagram{1}) \ , \tag{8.43}$$

with the first and second entries the $SU(3)$ and $SU(2)$ representations respectively so we get precisely the splitting. Now look at $\mathbf{15}$. This splits into:

$$\ydiagram{2} \mapsto (\ydiagram{2}\,,\bullet) \oplus (\ydiagram{1}\,,\ydiagram{1}) \oplus (\bullet\,,\ydiagram{2}) \ , \tag{8.44}$$

which is precisely the required splitting

$$\mathbf{15} \mapsto (\mathbf{6},\mathbf{1}) \oplus (\mathbf{3},\mathbf{2}) \oplus (\mathbf{1},\mathbf{3}) \ . \tag{8.45}$$

Example 8.9. Let us do a more complicated example. Consider the braching rule of the adjoint representation of $SU(4)$, $\mathbf{15}$, under $SU(4) \to SU(2)\times SU(2)$. The splitting rule in fact only gives four diagrams,

$$\ydiagram{2,1,1} \mapsto \left(\ydiagram{1,1}\,,\ydiagram{2}\right) \oplus \left(\ydiagram{2}\,,\ydiagram{1,1}\right) \oplus \left(\ydiagram{2,1}\,,\ydiagram{2}\right) \left(\ydiagram{1}\,,\ydiagram{2,1}\right) \ . \tag{8.46}$$

Notice that in this case the dimensions on the right only add up to $\mathbf{14}$. In fact, the missing diagram comes from the tensor rule,

$$\ydiagram{1,1} \otimes \ydiagram{1,1} \ ,$$

and we therefore have the branching rule,

$$\ydiagram{2,1,1} \mapsto \left(\ydiagram{1,1}\,,\ydiagram{2}\right) \oplus \left(\ydiagram{2}\,,\ydiagram{1,1}\right) \oplus \left(\ydiagram{2,1}\,,\ydiagram{2}\right) \left(\ydiagram{1}\,,\ydiagram{2,1}\right) \oplus \left(\ydiagram{1,1}\,,\ydiagram{1,1}\right) \ . \tag{8.47}$$

In tensor notation, this corresponds to the splitting, with α the $SU(4)$ indices and a and i the indices of the two $SU(2)$ respectively,

$$\begin{aligned}\psi_{(\alpha[\delta)\beta\gamma]} &\mapsto \psi_{(i[j)ab]} \oplus \psi_{(a[b)ij]} \oplus \psi_{(a[b)ci]} \oplus \psi_{(i[j)ka]} \oplus \psi_{(a[i)jb]} \\ &= (\phi_{[ab]}\,,\,\tilde{\phi}_{(ij)}) \oplus (\phi_{(ab)}\,,\,\tilde{\phi}_{[ij]}) \oplus (\phi_{(a[b)c]}\,,\,\tilde{\phi}_{i}) \oplus (\phi_{a}\,,\,\tilde{\phi}_{(i[j)k]}) \oplus (\phi_{[ab]}\,,\,\tilde{\phi}_{[ij]}) \ ,\end{aligned} \tag{8.48}$$

after taking the symmetrisation and antisymmetrisation into account.

8.9 Branching with $U(1)$ factors

We want to now start considering branching that includes $U(1)$ factors. We will begin with a specific case where we are looking at a branching,

$$SU(n+1) \to SU(n) \times U(1) \ . \tag{8.49}$$

The extra $U(1)$ is added in to make it into a maximal subgroup. To proceed, we consider the subgroup $SU(n) \times U(1)$. Let us think about what happens in the tensor indices language. Let us start with the fundamental representation $\mathbf{n}$ and considers what happens to the tensor:

$$\phi_a \to \phi_i \oplus \phi' \ , \tag{8.50}$$

Here we note that a is the $SU(n+1)$ index and i is the $SU(n)$ index. The ϕ' indicates the tensor in $U(1)$ - there are no indices associated to $U(1)$ in this language! This means that for every box in the Young diagram (noting that each box indicates an index), it will either be assigned to the fundamental rep $\mathbf{n}$ of $SU(n)$ or the trivial rep $\mathbf{1}$. Therefore, for a Young diagram with n boxes, j will transform like $\mathbf{n}$ and $n-j$ will transform like singlets in $SU(n-1)$.

Now, to determine the $U(1)$ charge, we need to work out the transformations of the fundamental representations $\mathbf{n}$ and the trivial rep in $U(1)$. Let us clear consider a more general problem of the following.

Proposition 8.12. *Consider the Lie group $SU(n+m)$ which branches into $SU(n) \times SU(m)$. Then there is a unique $U(1) \subset SU(n+m)$ subgroup which commutes with $SU(n) \times SU(m)$. The matrices that commutes with $SU(n) \times SU(m)$ are diagonal.*

Proof. Suppose V is an $SU(n+m)$ arbitrary matrix and $U \in SU(n) \times SU(m) \subset SU(n+m)$. We can write these matrices in block form:

$$V = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \ , \tag{8.51}$$

$$U = \begin{pmatrix} U_n & 0 \\ 0 & U_m \end{pmatrix} \ . \tag{8.52}$$

Now the commutation condition $[V, U] = 0$ implies that ,

$$\begin{cases} U_n A = A U_n \\ U_n B = B U_m \\ U_m C = C U_n \\ U_m D = D U_m \end{cases} \ . \tag{8.53}$$

Now setting $U_m = \mathbb{1}_m$ in the second and third equations give $B = C = 0$. By Schur's Lemma [48], the first and fourth equation gives

$$A = \lambda_n \mathbb{1}_n \ , \tag{8.54}$$

$$D = \lambda_m \mathbb{1}_m \ , \tag{8.55}$$

for $\lambda_n, \lambda_m \in \mathbb{C}$. Now for

$$V = \begin{pmatrix} \lambda_n \mathbb{1}_n & 0 \\ 0 & D = \lambda_m \mathbb{1}_m \end{pmatrix}$$

to be an $SU(n+m)$ matrix, we require $|\lambda_n| = 1 = |\lambda_m|$ (from unitarity) and

$$\lambda_n^n \lambda_m^m = 1 \ . \tag{8.56}$$

This allows us to determine the form of the matrix, for example, with the solution:

$$\begin{cases} \lambda_n = e^{-im\alpha} \\ \lambda_m = e^{in\alpha} \end{cases} \ . \tag{8.57}$$

To show this is the unique $U(1)$ group (with freedom in choosing arbitrary charges), note that if we start with a different solution in Equation 8.57 then we will get the same answer up to a root of unity transformation. $\square$

The conclusion from the above proposition is that the for a branching $SU(n+m) \to SU(n) \times SU(m)$, the fundamental representation $\mathbf{n} + \mathbf{m}$ breaks into $\mathbf{n}$ and $\mathbf{m}$ with the $U(1)$ charge m and $-n$ respectively. In our present case $SU(n+1) \to SU(n)$, we set $m = 1$ and note that $SU(1)$ does not exist - we can assign the normalised charge $\frac{1}{n+1}$ to the fundamental rep $\mathbf{n}$ and $-\frac{n}{n+1}$ to the trivial rep $\mathbf{1}$. Note here that if the group is $SU(m)$ where $m = 1$, then we only allow symmetric Young diagrams (i.e. horizontal diagrams).

Example 8.10. Let us look at an example where $SU(3) \to SU(2) \times U(1)$. The fundamental rep breaks like

$$\mathbf{3} \to \mathbf{2} \oplus \mathbf{1} \ , \tag{8.58}$$

For a Young diagram representing a tensor with n indices, we then have n boxes which then branches with m boxes that transform like doublets and $n - m$ boxes that transform like singlets. The total $U(1)$ charge is then

$$q = \frac{m}{3} - \frac{2(n-m)}{3} = -\frac{2n}{3} + m \ . \tag{8.59}$$

Let us look at a particular example. Consider $\mathbf{adj}$ of $SU(3)$.

$$\ydiagram{2,1} \to \left(\ydiagram{2,1}\ , \bullet\right) \oplus \left(\ydiagram{1,1}\ , \ydiagram{1}\right) \oplus \left(\ydiagram{2}\ , \ydiagram{1}\right) \oplus \left(\ydiagram{1}\ , \ydiagram{2}\right) \ . \tag{8.60}$$

Then using Equation 8.59 to compute charges gives

$$\mathbf{8} \to \mathbf{2}_1 \oplus \mathbf{1}_0 \oplus \mathbf{3}_0 \oplus \mathbf{2}_{-1} \ . \tag{8.61}$$

[48] In last year's notes this is incorrectly referred to the proof in Q4 of the first problem sheet. The generalised Schur's Lemma doesn't apply here - but the proof is of a similar spirit regardless.

To generalise this to the case $SU(n+m) \to SU(n) \times SU(m)$ is simple. We follow the rule in the previous subsection for the branching and then assign the $U(1)$ charges according to the rule set out in Proposition 8.12. Let us now consider some examples.

Example 8.11. Let us consider the splitting of $SU(6) \to SU(4) \times SU(2)$. Using the argument in Proposition 8.12 would give the following commuting matrix:

$$U = \mathrm{diag}(e^{i\alpha}, e^{i\alpha}, e^{i\alpha}, e^{i\alpha}, e^{i\beta}, e^{i\beta}) \; . \tag{8.62}$$

Imposing the special unitary condition, i.e. $\det(U) = 1$, means that $2\alpha + \beta = 0 \mod 6$. We can canonically pick the scaling that gives the following diagonal matrix:

$$U = \mathrm{diag}(e^{-i\alpha}, e^{-i\alpha}, e^{-i\alpha}, e^{-i\alpha}, e^{2i\alpha}, e^{2i\alpha}) \; . \tag{8.63}$$

Remark 8.4. If there is a splitting to more than two groups, you will in general obtain more than one possible assignments of $U(1)$ charges. For example, in the any splitting of a Lie group into two Lie subgroups case, you will always have an $\mathbb{R}^*$ ambiguity in assigning charges. This is normally picked to have a consistent charge with the literature (so probably the simplest integer combination/with a normalisation like in the illustrated $SU(3)$ case).

Now to generate the required $U(1)$ charges for the tensor splitting, we associate a charge[49] to each of the indices in the subgroup. For branching $SU(n+m) \to SU(n) \times SU(m)$ we look at the fundamental:

$$\mathbf{n} + \mathbf{m} \to \mathbf{n}_m \oplus \mathbf{m}_{-n} \; . \tag{8.64}$$

So we associate to each $SU(n)$ index a charge of m and to each $SU(m)$ index a charge of $-n$. We then add up the charges.

Example 8.12. For example, we associate a -1 charge for every $SU(4)$ index and 2 for every $SU(2)$ **lower** index in the $SU(6)$ splitting in Example 8.12. Then in the index decomposition language, we can immediately obtain:

$$\bar{\mathbf{6}} \mapsto (\bar{\mathbf{4}}, \mathbf{1})_1 \oplus (\mathbf{1}, \mathbf{2})_{-2} \; , \tag{8.65}$$

by using the splitting:

$$\psi^\mu \mapsto (\psi^a, \psi^i) \; . \tag{8.66}$$

Also doing that for the two index tensor case:

$$\psi_{[\mu\nu]} \mapsto \left(\psi_{[ab]}, \psi_{ai}, \psi_{[ij]}\right) \; , \tag{8.67}$$

we get:

$$\mathbf{15} \mapsto (\mathbf{6}, \mathbf{1})_{-2} \oplus (\mathbf{4}, \mathbf{2})_1 \oplus (\mathbf{1}, \mathbf{1})_4 \; , \tag{8.68}$$

remembering the conjugate the charge for anti-fundamental indices [50].

[49] As in the usual case in the literature.

[50] Remember there is some funky stuff going on when you apply the raising and lowering operators (Levi-Civita tensors). They act non-trivially to the $U(1)$ charges as they add a non-zero number of indices (n) to your tensor expression. This means that ϕ^μ is technically not the same as $\phi_{[\mu\nu\rho\sigma]}$!

In the Young diagram language, this is equivalent to associating a number to each of the boxes in the splittings. This will allow you to compute the charge by just summing the numbers in the Young diagram.

Example 8.13. Let us continue with the $SU(6)$ splitting started in Example 8.13. To each box in $SU(4)$ and $SU(2)$ we associate a charge -1 and $+2$. Simply sum up in each case ($\bar{\mathbf{6}}$ and $\mathbf{15}$):

$$\begin{array}{|c|}\hline \\ \hline \\ \hline \\ \hline \\ \hline \\ \hline \end{array} \mapsto \left(\begin{array}{|c|}\hline a \\ \hline a \\ \hline a \\ \hline \end{array} , \begin{array}{|c|}\hline 2 \\ \hline 2 \\ \hline \end{array} \right) \oplus \left(\begin{array}{|c|}\hline a \\ \hline a \\ \hline a \\ \hline a \\ \hline \end{array} , \begin{array}{|c|}\hline 2 \\ \hline \end{array} \right) , \tag{8.69}$$

$$\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array} \mapsto \left(\begin{array}{|c|}\hline a \\ \hline a \\ \hline \end{array} , \bullet \right) \oplus \left(\begin{array}{|c|}\hline a \\ \hline \end{array} , \begin{array}{|c|}\hline 2 \\ \hline \end{array} \right) \oplus \left(\bullet , \begin{array}{|c|}\hline 2 \\ \hline 2 \\ \hline \end{array} \right) , \tag{8.70}$$

where I have used a to indicate the charge -1. You can see the charges match up with

$$\bar{\mathbf{6}} \mapsto (\bar{\mathbf{4}}, \mathbf{1})_1 \oplus (\mathbf{1}, \mathbf{2})_{-2} , \tag{8.65}$$

$$\mathbf{15} \mapsto (\mathbf{6}, \mathbf{1})_{-2} \oplus (\mathbf{4}, \mathbf{2})_1 \oplus (\mathbf{1}, \mathbf{1})_4 . \tag{8.68}$$

8.10 Branching using Symmetries

In this section let us further discuss branching but using symmetry arguments. We have seen that for $SU(5)$ the antisymmetric representation branches under $SU(3) \times SU(2)$ as [51]

$$\mathbf{10} \mapsto (\bar{\mathbf{3}}, \mathbf{1}) \oplus (\mathbf{3}, \mathbf{2}) \oplus (\mathbf{1}, \mathbf{1}) . \tag{8.71}$$

The question is now as follows: can we reproduce that using symmetry arguments? Let us note first that the **10** rep is the antisymmetric representation, i.e.

$$\mathbf{10} = (\mathbf{5} \otimes \mathbf{5})_A , \tag{8.72}$$

where the subscript A denotes the antisymmetric part of the multiplication, i.e. after the Clebsch-Gordon decomposition is performed. The fundamental rep **5** has the branching

$$\mathbf{5} \to (\mathbf{3}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{2}) , \tag{8.73}$$

under $SU(3) \times SU(2)$. We can now use symmetry arguments to obtain the branching of **10** in $SU(5)$ - we simply take the antisymmetric part as follows:

$$\begin{aligned} \mathbf{10} = (\mathbf{5} \times \mathbf{5})_A &= [((\mathbf{3}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{2})) \otimes ((\mathbf{3}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{2}))]_A \\ &= [(\mathbf{3}, \mathbf{1}) \otimes (\mathbf{3}, \mathbf{1})]_A \oplus [(\mathbf{3}, \mathbf{1}) \otimes (\mathbf{1}, \mathbf{2}) \oplus (\mathbf{1}, \mathbf{2}) \otimes (\mathbf{3}, \mathbf{1})]_A \oplus [(\mathbf{1}, \mathbf{2}) \otimes (\mathbf{1}, \mathbf{2})]_A \\ &= (\bar{\mathbf{3}}, \mathbf{1}) \oplus (\mathbf{3}, \mathbf{2}) \oplus (\mathbf{1}, \mathbf{1}) . \end{aligned} \tag{8.74}$$

[51] This is in fact one of the questions on the example sheet. You should now be able to reproduce this result using the Young diagram method.

Note that to evaluate the second term one must regard the two **5** representations to be unique (distinguishable). This means that in order to evaluate the term we must have, say identified the first **5** as A and the last one as B and noted that [52],

$$(A, B) = \frac{1}{2}\left[(A, B) + (B, A)\right] + \frac{1}{2}\left[(A, B) - (B, A)\right] \ . \tag{8.75}$$

Taking the antisymmetric part of this hence gives us one copy [53] of **5** or in our case, after taking the product, $(\mathbf{3}, \mathbf{2})$. We then obtain the result as obtained from alternative methods.

This concludes our discussion on tensor representations.

8.11 $SU(5)$ Matter Multiplet Branching

In this subsection, we will explicitly work out the $SU(5)$ matter multiplet branching rules as sketched out in the notes using the methods described in this section.

Distributed version:

Exercise 8.2. Try Q4 of the second sheet again! You should now be able to proceed up to part (d) with ease. In particular, you should be able to compute the branching of $\bar{\mathbf{5}}$ and **10** and the corresponding $U(1)$ charges. Using tensor methods to construct singlets for (e) is the easiest way to proceed, for example the singlet in $\mathbf{5} \otimes \bar{\mathbf{5}}$ is $\epsilon^{abcde}\phi_a\psi_{abcde}$. In particular, use the notations $\mathbf{10} \sim \psi_{[\mu\nu]}$, $\bar{\mathbf{5}} \sim \chi^\mu$, $\mathbf{5} \sim \bar{H}_\mu$ and $\bar{\mathbf{5}} \sim H^\mu$, and the branching rule $\psi_{[\mu\nu]} = \left(u_{[\alpha\beta]}, Q_{\alpha i}, e_{[ij]}\right)$, $\chi^\mu = \left(d^\alpha, L^i\right)$, $H_\mu = (t_\alpha, h_i)$, $\bar{H}^\mu = \left(\bar{t}_\alpha, \bar{h}_i\right)$. What is the physical significance of the last part of the question?

8.12 Proof of Schur-Weyl Duality

In this section we will provide a proof to Theorem 8.2. We will need a bit of machinery about semisimple algebras. Take G to be any finite group, U a right module over the group algebra $A = \mathbb{C}G$ and let

$$B = \mathrm{Hom}_G(U, U) = \{\varphi : U \to U \,|\, \varphi(v \cdot g) = \varphi(v) \cdot g \ \forall u \in U, \ g \in G\} \tag{8.76}$$

Note that B acts on U from the left and commutes with the right action of A. B is known as the commutator algebra. If $U = \bigoplus_i U_i^{\oplus n_i}$ is an irreducible decomposition with U_i nonisomorphic irreducible right A-modules, then by Schur's Lemma we have

$$B = \bigoplus_i \mathrm{Hom}_G\left(U_i^{\oplus n_i}, U_i^{\oplus n_i}\right) = \bigoplus_i M_{n_i}(\mathbb{C}) \tag{8.77}$$

where $M_{n_i}(\mathbb{C})$ is the ring of $n_i \times n_i$ complex matrices. If W is any left A-module, the tensor product

$$U \otimes_A W = \frac{U \otimes_{\mathbb{C}} W}{\text{subspace generated by } \{va \otimes w - v \otimes aw\}} \tag{8.78}$$

is a left B-module by acting on the first factor $b \cdot (v \otimes w) = (b \cdot v) \otimes w$. Now we have the following lemma.

[52] The bracket notation here has nothing to do with the bracket indicating branching of groups like above!
[53] Well, one effective antisymmetric copy.

Lemma 8.2. *Let U be a finite-dimensional right A-module.*

(i) For any $c \in A$, the canonical map $U \otimes_A Ac \to Uc$ is an isomorphism of left B-modules.

(ii) If $W = Ac$ is an irreducible left A-module, then $U \otimes_A W = Uc$ is an irreducible left B-module.

(iii) If $W_i = Ac_i$ are the distinct irreducible left A-modules, with m_i the dimension of W_i, then

$$U \cong \bigoplus_i (U \otimes_A W_i)^{\oplus m_i} \cong \bigoplus_i (Uc_i)^{\oplus m_i} \tag{8.79}$$

is the decomposition of U into irreducible left B-modules.

Proof. By the semisimplicity of all representations of G, Ac is a direct summand of A as a left-A module. First prove (i). Consider the commutative diagram

$$\begin{tikzcd}
U \otimes_A A \arrow[r, "\cdot c"] \arrow[d] & U \otimes_A Ac \arrow[r, hook] \arrow[d] & U \otimes_A A \arrow[d] \\
U \arrow[r, "\cdot c"] & U \cdot c \arrow[r, hook] & U
\end{tikzcd}$$

The vertical maps are the maps $v \otimes a \mapsto v \cdot a$. Now the left horizontal maps are surjective, the right ones are injective, and the outside vertical maps are isomorphisms, so the middle vertical map must be isomorphic. Now for (ii), consider first the case where U is an irreducible A-module, so $B = \mathbb{C}$. It suffices to show that $\dim U \otimes_A W \leq 1$. Identify A with a direct sum $\bigoplus_{i=1}^r M_{m_i}\mathbb{C}$ of r matrix algebras (refer to Proposition 3.29 of [1]). We can now identify W with a minimal left ideal of A. Any minimal ideal in the sum of matrix algebras is isomorphic to one which consists of r-tuples of matrices zero except for one factor, in which they are all zeros except for one column. Similarly we can identity U with a minimal right ideal of A, so same as above but with column switched out for row. Therefore any non-zero $U \otimes_A W$ can be identified with the matrices which are zero except in one row and column the same factor, otherwise it will be zero. For the general case, by Maschke's theorem we can write $U = \bigoplus_i U_i^{\oplus n_i}$ irreducible right A-modules, then $U \otimes_A W = \bigoplus_i (U_i \otimes_A W)^{\oplus m_i}$. Now since A is the left regular representation of G (by definition of group algebra), $U_i \otimes W = \mathbb{C}$ iff U_i and W correspond to the same factor of A, as above. So $U \otimes_A W = \bigoplus_i (U_i \otimes_A W)^{\oplus n_i} = \mathbb{C}^{\oplus n_k}$ for some k, and this is visibly irreducible over $B = \bigoplus_i M_{n_i}(\mathbb{C})$. Part (iii) simply follows from $A \cong \bigoplus_i W_i^{\oplus m_i}$ which gives

$$U \cong U \otimes_A A \cong \bigoplus_i (U \otimes_A W_i)^{\oplus m_i} \ . \tag{8.80}$$

□

We apply this lemma with $U = V^{\otimes n}$ which is a right $\mathbb{C}S_n$-module. Set $G = S_n$. The lemma above tells us how to decompose $V^{\otimes n}$ into a B-module where B is the algebra of all endomorphisms of U that commute with all permutations of the factors. So we now have:

Lemma 8.3. *The algebra B is spanned as a linear subspace of* $\mathrm{End}(V^{\otimes n})$ *by* $\mathrm{End}(V)$*. A subspace of $V^{\otimes n}$ is a sub-B modules iff it is invariant by* $\mathrm{GL}(V)$*; in other words a B-submodule is the same as a* $\mathrm{GL}(V)$*-submodule.*

Proof. If W is any finite-dimensional vector space, then $\mathrm{Sym}^n W$ is the subspace of $W^{\otimes n}$ spanned by all $w^n = n!w \otimes ... \otimes w$. Apply this to $W = \mathrm{End}(V)$ proves the first statement as $\mathrm{End}(V^{\otimes} n) = (V^*)^{\otimes n} \otimes V^{\otimes n} = W^{\otimes n}$, which is compatible with the actions of $G = S_n$. Since $\mathrm{GL}(V)$ is dense in W which is either in the Euclidean or Zariski topology, the second statement also follows. □

The proof of Theorem 8.2 then follows from the two lemmas. In particular, we define,

$$\mathbb{S}_\lambda V = V^{\otimes n} \otimes_A V_\lambda \ , \tag{8.81}$$

where V_λ denotes the irrep corresponding to a partition λ. Then this is irreducible as a $\mathrm{GL}(V)$-module by Lemma 8.2. Further writing $A = \bigoplus_\lambda (V_\lambda)^{\oplus m_\lambda}$ as a left-semisimple decomposition we then have

$$V^{\oplus n} = V^{\oplus n} \otimes_A A = \bigoplus_\lambda \left(V^{\otimes n} \otimes_A V_\lambda\right)^{\oplus m_\lambda} = \bigoplus_\lambda \mathbb{S}_\lambda V \ , \tag{8.82}$$

which exactly gives the Schur-Weyl Duality in Corollary 8.1.

8.13 $O(n)$ and $SO(n)$ tensors

When we now further restrict to $SO(n)$ tensors, in addition to symmetrisations a new operation called *contraction* will appear which commutes with orthogonal transformations [35]. Note that orthogonal transformations satisfy,

$$a_{ij} a_{ik} = a_{ij} a^T_{ki} = \delta_{kj} \ . \tag{8.83}$$

Therefore, given a rank-r tensor one can construct a rank-$(r-2)$ tensor by contracting two indices,

$$T_{i_3 \ldots i_r} = T_{iii_3 \ldots i_r} = \delta_{i_1 i_2} T_{i_1 \ldots i_r} \ , \tag{8.84}$$

and this action commutes with the transformation of the tensor,

$$T'_{i_3 \ldots i_r} = a_{ij_1} a_{ij_2} a_{i_3 j_3} \ldots a_{i_r j_r} T_{j_1 \ldots j_r} = a_{i_3 j_3} \ldots a_{i_r j_r} T_{j_3 \ldots j_r} \ . \tag{8.85}$$

We have the following proposition.

Proposition 8.13. *Every $O(n)$-tensor can be decomposed uniquely into a traceless tensor for which all pair traces are zero and where pair traces are taken, i.e.*

$$T_{i_1 \ldots i_r} = T^0_{i_1 \ldots i_r} + \sum_{\alpha, \beta \ \textit{distinct}} \delta_{i_\alpha i_\beta} T_{i_1 \ldots i_{\alpha-1} i_{\alpha+1} \ldots i_{\beta-1} i_{\beta+1} \ldots i_r} \ . \tag{8.86}$$

Proof. Consider subspace Σ_Φ of the tensors in,

$$\Phi_{i_1 \dots i_r} = \sum_{\alpha,\beta\,\text{distinct}} \delta_{i_\alpha i_\beta} T_{i_1 \dots i_{\alpha-1} i_{\alpha+1} \dots i_{\beta-1} i_{\beta+1} \dots i_r} \ . \tag{8.87}$$

Then we need,

$$(T^0, \Phi) = 0 \ , \tag{8.88}$$

which can be checked by selecting all pair traces. □

Under orthogonal transformations the invariant subspace Σ_Φ transform amongst itself so we deduce that for every $GL(n)$-tensor, and hence every Young diagram, this corresponds to two tensor irreps — a completely traceless $T^0_{i_1 \dots i_r}$, as well as the invariant subspace Σ_Φ. However, there is a caveat, as given by the following statement.

Proposition 8.14. *The traceless tensors corresponding to Young diagrams in which the sum of the lengths of the first two columns exceeds n in $O(n)$ must be identically zero.*

Proof. Notice how the action of contraction picks two indices in the diagram and sum over all possibilities. In the case where the numbers of the first two columns satisfy $a + b > n$, the symmetrisation and anti-symmetrisation of the tensor can always bring the indices into the first two columns and since there are more than n indices we can arrange the indices such that we obtain some diagram of the form,

$$\begin{array}{|c|c|}\hline i & \dots \\ \hline i & \dots \\ \hline \dots \\ \cline{1-1}\end{array} = 0 \ . \tag{8.89}$$

By anti-symmetrisation in the first column this tensor is zero, hence giving the result. □

Let us illustrate this with a few examples. We first illustrate the method using tensors in $SO(n)$.

Example 8.14. For second-rank tensors, we can write the decomposition,

$$T_{ij} = \frac{1}{n} T_{kk} \delta_{ij} + \left(T_{ij} - \frac{1}{n} T_{kk} \delta_{ij} \right) = \Phi_{ij} + T^0_{ij} \tag{8.90}$$

Then T^0_{ij} is a tensor with zero trace.

Example 8.15. For third-rank tensors, we can write the decomposition,

$$T_{ijk} = T^0_{ijk} + H_k \delta_{ij} + K_j \delta_{ik} + L_i \delta_{jk} \ , \tag{8.91}$$

Then T^0_{ijk} is a tensor with zero trace, and solving the zero-trace conditions give the following equations,

$$H_j = \frac{1}{n^2 + n - 2} \left[(n+1) T_{iij} - T_{iji} - T_{jii} \right] \tag{8.92}$$

$$K_j = \frac{1}{n^2 + n - 2} \left[-T_{iij} + (n+1) T_{iji} - T_{jii} \right] \tag{8.93}$$

$$L_j = \frac{1}{n^2 + n - 2} \left[-T_{iij} - T_{iji} + (n+1) T_{jii} \right] \tag{8.94}$$

using Einstein's notation. Suppose we want to find $\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3}$ in $SO(3)$. Then the procedure above must be repeated for all types of tensors with different irreps in $GL(n)$, i.e.

$$\left\{ \ydiagram{3} \,,\ \ydiagram{2,1} \,,\ \ydiagram{1,1,1} \right\} . \tag{8.95}$$

This decomposition in fact gives seven tensors, giving the decomposition,

$$\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3} = \mathbf{7}\oplus\mathbf{5}\oplus\mathbf{5}\oplus\mathbf{3}\oplus\mathbf{3}\oplus\mathbf{3}\oplus\mathbf{1} . \tag{8.96}$$

Example 8.16. Let us look at a few examples showing Proposition 8.14 in effect. For the case $n = 2$, we can look at the tensor of the symmetry,

$$\ydiagram{2,1} \to \left\{ \begin{ytableau} 1 & 1 \\ 2 \end{ytableau} \,,\ \begin{ytableau} 1 & 2 \\ 2 \end{ytableau} \right\} , \tag{8.97}$$

which has the two independent non-zero components as listed. But tracelessness requires us to take away the components,

$$0 = \begin{ytableau} 1 & 1 \\ 2 \end{ytableau} - \left(\begin{ytableau} 1 & 1 \\ 2 \end{ytableau} + \begin{ytableau} 2 & 2 \\ 2 \end{ytableau} \right) , \tag{8.98}$$

$$0 = \begin{ytableau} 1 & 2 \\ 2 \end{ytableau} - \left(\begin{ytableau} 1 & 2 \\ 2 \end{ytableau} + \begin{ytableau} 1 & 1 \\ 1 \end{ytableau} \right) , \tag{8.99}$$

$$\tag{8.100}$$

due to the antisymmetry of the system. So traceless tensors are indeed zero in this case, as predicted by the proposition.

There is an additional effect in $SO(n)$ Young diagram methods known as associate diagrams.

Definition 8.10. Permissible diagrams in Proposition 8.14 are paired into **associate diagrams** Y and Y' as follows. The length of the first column of Y, a satisfies, $a \leq n/2$ whilst the length of the first column of Y' satisfies $a' = n - a$ with all the other columns of Y an Y' having the same length. Diagrams that are associate to themselves are **self-associate**.

Example 8.17. The following diagrams are associate for $n = 3$.

$$\ydiagram{5} \leftrightarrow \ydiagram{5,1} , \tag{8.101}$$

$$\ydiagram{1} \leftrightarrow \ydiagram{1,1} , \tag{8.102}$$

Associate diagrams have a sign difference under orthogonal transformations. For example, under improper transformations (inversions) in $SO(3)$, we have,

$$\ydiagram{1} \mapsto \ydiagram{1} \,,\quad \ydiagram{1,1} \mapsto -\ydiagram{1,1} . \tag{8.103}$$

Here the diagram with one box corresponds to a vector, whilst an antisymmetric tensor in $O(3)$ corresponds to the axial vector. Under proper transformations however there is no sign difference so in $SO(n)$ we will need to additionally identify the associate diagrams as the same irrep.

A quick comment about $Sp(n)$ tensors — the formalism is basically the same in $SO(n)$ with a few differences.

- The symplectic group $Sp(n)$ corresponds to linear transformations which preserves a skew-symmetric bilinear form, so if $g_{ij} = -g_{ji}$, $a \in Sp(n)$ gives $a_{ik}g_{ij}a_{jl} = g_{kl}$. a is unimodular and satisfies $\det(a) = 1$ so there is no distinction between signs of transformations.

- The contraction in symplectic groups satisfy,

$$\epsilon_{kl}a_{ki}a_{lj} = a_{ij} \ , \tag{8.104}$$

  and similar contraction rules apply for $Sp(n)$-tensors as in $O(n)$-tensors. The decomposition into traceless tensors and traceful tensors is the same as in the $SO(n)$ case.

- Traceless tensors that correspond to Young diagrams having more than $\lfloor \frac{n}{2} \rfloor$ are zero, as this corresponds to move than $\lfloor \frac{n}{2} \rfloor$ pairs of symmetric indices being antisymmetrised.

The details for $Sp(n)$-tensors can be found in §10.8-9 of [35].

9 Complexification and Real Forms

In this section I would like to go back to the Cartan-Weyl formalism and discuss some intricacies about the Killing form. We know that the Killing form can act as a bilinear form (effectively an inner product) in weight space. But we also remarked that compact Lie algebras in fact have a negative-definite Killing form. An inner product is by definition a positive-definite bilinear form. This section aims to disentangle this seeming contradiction.

9.1 The Killing Form

First let us recall some facts about the Killing form. Recall that the Killing form is defined as follows.

Definition 9.1. Define the **adjoint** linear operator $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ as follows:

$$\mathrm{ad}_X : Y \mapsto [X, Y] \tag{9.1}$$

With this the **Killing form** is then defined as follows:

$$\Gamma(X, Y) = \mathrm{tr}\left(\mathrm{ad}_X \circ \mathrm{ad}_Y\right) \tag{9.2}$$

where $\circ$ denotes the composition of maps.

We note that this definition is basis-independent - one can evaluate the Killing form with respect to a specific basis which yields a real number.

The Killing form is a bilinear symmetric operation. For a semi-simple Lie algebra $\mathfrak{g}$, we have the following two theorems:

Theorem 9.1. *The Killing form is non-degenerate if and only if $\mathfrak{g}$ is a semi-simple Lie algebra.*

Proof. See Theorem 4.3 of notes. □

Theorem 9.2. *For any nonzero $X \in \mathfrak{g}$, $\mathfrak{g}$ is a compact semisimple Lie algebra if and only if $\Gamma(X, X) < 0$ (i.e. the Killing form is negative-definite).*

Proof. There are two proofs to this. Here I will only present one of them. Define

$$g_{ij} = \Gamma(X_i, X_j) = \mathrm{Tr}\big(\mathrm{ad}_{X_i} \circ \mathrm{ad}_{X_j}\big) \tag{9.3}$$

where $\{X_i\}$ are a basis for Lie algebra $\mathfrak{g}$. Writing the matrix $\{g_{ij}\}$ as $\mathcal{G}$, we can then choose the coordinates on the Lie group manifold (using Sylvester's Theorem, remembering that we are operating the real basis)

$$\mathcal{G} = \mathrm{diag}(1, 1, ..., 1, -1, ..., -1) \tag{9.4}$$

where 1 appears r times and -1 s times. By diagonalising, we can write

$$\mathcal{G}' = S^T \mathcal{G} S \tag{9.5}$$

where $n = r + s$ is the dimension of $\mathcal{L}$. We note that the metric has no zeros along the diagonal due to the non-degeneracy of the Killing form, from Theorem 9.1. Now let us write Ad_g, the adjoint representation of the Lie group G as A. Recall that the matrix Lie group $\mathrm{O}(r, s)$ is defined by

$$\mathrm{O}(r,s) = \left\{A \in \mathrm{Mat}_{r+s}(\mathbb{R}) \,\middle|\, A^T \mathcal{G} = \mathcal{G} A^{-1}\right\} \tag{9.6}$$

Now we show the following. Using $\{X_i\}$ as the Lie algebra basis, and x^k the coordinate with respect to the basis $\{X_k\}$. Writing a general element $X \in G$ as

$$X = x^k X_k \tag{9.7}$$

Then using

$$\mathrm{Ad}_g(X_k) = g X_k g^{-1} = A^j_k X_j \tag{9.8}$$

then we have

$$\mathrm{Ad}_g(X) = x^k A^j_k X_j \tag{9.9}$$

Therefore, writing out $\Gamma(X, Y) = \Gamma\left(\mathrm{Ad}_g(X), \mathrm{Ad}_g(Y)\right)$ we get

$$\mathcal{G} = A^T \mathcal{G} A \tag{9.10}$$

i.e. exactly the definition of the matrix Lie group $\mathrm{O}(r, s)$. Similarly, writing the adjoint representation of the Lie algebra $\mathfrak{g}$ ad_X using the matrix B, and $\mathrm{ad}(X) = x^k B^j_k X_j$ and the relation

$$\Gamma\left(\mathrm{ad}_Z(X), Y\right) = -\Gamma\left(X, \mathrm{ad}_Z(Y)\right) \tag{9.11}$$

we obtain

$$B = -\mathcal{G}^{-1} B^T \mathcal{G} \tag{9.12}$$

We now compare this with the matrix Lie algebra $\mathfrak{so}(r, s)$:

$$\mathfrak{so}(r,s) = \left\{A \in \mathrm{Mat}_{r+s}(\mathbb{R}) \,\middle|\, A^T \mathcal{G} = \mathcal{G} A^{-1}\right\} \tag{9.13}$$

So we conclude the following - the adjoint representation of G and $\mathfrak{g}$ must be subalgebras of $\mathrm{O}(r, s)$ and $\mathfrak{so}(r, s)$ respectively. Now consider the case where $\mathcal{G} = \pm \mathbb{1}_n$. Then

$$\begin{cases} A^T A = \mathbb{1}_n \\ B^T = -B \end{cases} \tag{9.14}$$

So the adjoint reps of G and $\mathfrak{g}$ are subalgebras of $\mathrm{O}(n)$ and $\mathfrak{so}(n)$. In particular, for $X \in \mathfrak{g}$ we have

$$\Gamma(X, X) = \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_X) = \mathrm{Tr}\big(B^2\big) = -\,\mathrm{Tr}\big(BB^T\big) = -\sum_{ij} B_{ij} B_{ij} < 0 \tag{9.15}$$

The final sign is replaced by an equal sign only where $X = 0$ identically. Therefore $\mathcal{G} = -\mathbb{1}_n$ and we have a negative-definite Killing form. We deduce the following - a negative-definite Killing form implies that the adjoint group corresponding to the matrices of the adjoint representation of $\mathfrak{g}$ constitutes a closed subgroup of $\mathfrak{so}(n)$. Therefore the identity-connected component of the automorphism group of $\mathfrak{g}$ is a subgroup of a compact group and therefore is also compact - i.e. by definition $\mathfrak{g}$ is a compact semisimple Lie algebra [36] (Note that the adjoint map $\mathrm{Ad} : G \to GL(\mathfrak{g})$ has kernel that is the centre[54] of G (as the elements commute through). The Lie algebra $\mathfrak{g}$ is semi-simple, the centre is then finite so this shows that $G/Z(G)$ is compact. To show that G is compact, we will need use the Hopf-Rinow theorem to show that $(G, -\Gamma)$ is indeed a complete Riemannian manifold together with Bonnet-Myers Theorem. I refer to [37] for details.) The converse of this is obvious. (For a Lie group G the lie algebra must be compact.) □

With this now I would like to present two main questions.

1. For a semisimple compact Lie algebra, we have shown in Theorem 9.2 that the Killing form is negative-definite. It is obvious that to obtain a positive-definite bilinear form (i.e. an inner product) on root space one can take the negative of the Killing form. But in Theorem 4.6 of the notes for complex semisimple Lie algebras we somehow seem to be able to define an inner product directly using

$$(\alpha, \beta) = \Gamma(H_\alpha, H_\beta) \tag{9.17}$$

   How does this work? Is this a typo for the real compact Lie algebra case, or is the situation different for complex semisimple Lie algebras?

2. The derivation for complex semi-simple Lie algebras show that the roots are real in the subspace $\mathfrak{h}_0$ of the real form of the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$. However, we know that ad_X is skew-symmetric with respect to the bilinear form (Killing form) and therefore has imaginary eigenvalues $i\alpha(X)$ (or that $\alpha(X)$ are pure imaginary). Is this a contradiction?

To answer this we will need to understand what complexification and real forms are for Lie algebras.

9.2 Complexification, Realification and Real Forms

Let us go back to the basics. First consider a vector space V.

Definition 9.2. Fix a vector space V over field k. We define the vector space $V^{\mathbb{K}} := V \otimes_k \mathbb{K}$ as follows: we let $1 \otimes \cdot\, c$ be the k-linear map of $V \otimes_k \mathbb{K}$ to itself where $\cdot\, c$ denotes multiplication by $c \in \mathbb{K}$. This denotes scalar multiplication in $V^{\mathbb{K}}$. Then $V^{\mathbb{K}}$ is the **vector space in extension field** $\mathbb{K}$. Alternatively, starting with a vector space W over field $\mathbb{K}$, we can restrict the definition of scalar multiplication to scalars in k to get a vector space over k, denoted W^k. This is called the **vector space in restriction field** k.

[54] The **centre** of a group is defined as

$$Z(G) = \{z \in G \,|\, \forall g \in G, zg = gz\} \tag{9.16}$$

Definition 9.3. For the case $k = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$ in the above definition, and V being a real vector space, the complex vector space $V^{\mathbb{C}}$ is called the **complexification** of V. If W is a complex vector space, the real vector space $W^{\mathbb{R}}$ is the **realification** of W. Note that

$$(V^{\mathbb{C}})^{\mathbb{R}} = V \oplus iV \tag{9.18}$$

as real vector spaces[55] with i denoting a real linear transformation multiplication by i. This shows that the two operations are not inverses to each other, $(V^{\mathbb{C}})^{\mathbb{R}}$ has twice the real dimension of V, and $(W^{\mathbb{R}})^{\mathbb{C}}$ has twice the complex dimension of W.

Definition 9.4. Suppose W are V are complex and real vector spaces respectively. If

$$W^{\mathbb{R}} = V \oplus iV \tag{9.19}$$

we say V is a **real form** of the complex vector space W.

Proposition 9.1. *Any real vector space is a real form of its complexification.*

Proof. Obvious from Equation 9.18. $\square$

Now for Lie algberas. To impose a Lie algebra structure on the $\mathbb{K}$ vector space $\mathfrak{g} = (\mathfrak{g}_0)^{\mathbb{K}}$ the obvious thing is to define a quadra-linear map $\mathfrak{g}_0 \times \mathfrak{g}_0 \times \mathbb{K} \times \mathbb{K} \to \mathfrak{g}_0 \otimes_k \mathbb{K}$ given by

$$(X, Y, a, b) \mapsto [X, Y] \otimes ab \in \mathfrak{g}_0 \otimes_k \mathbb{K} \tag{9.20}$$

This extends to the k-linear map on $\mathfrak{g} \times \mathfrak{g}$ so we have the bracket product on $\mathfrak{g}$. The converse is similar - we can restrict to a subfield $\Bbbk$ by restricting to scalar multiplication for elements in $\Bbbk \subset k$. Then we have the definitions:

Definition 9.5. The complex Lie algebra $(\mathfrak{g}_0)^{\mathbb{C}}$ is the **complexification** of real Lie algebra $\mathfrak{g}_0$. Similarly if

$$\mathfrak{g}^{\mathbb{R}} = \mathfrak{g}_0 \oplus i\mathfrak{g}_0 \tag{9.21}$$

then $\mathfrak{g}_0$ is a **real form** of the complex Lie algebra $\mathfrak{g}$.

Proposition 9.2. *Any real Lie algebra is a real form of its complexification. The conjugation of a complex Lie algebra $\mathfrak{g}$ with respect to a real form is a Lie algebra isomorphism of $\mathfrak{g}^{\mathbb{R}}$ with itself.*

Proof. Also clear from definitions. $\square$

With these definitions it is immediately clear what we have to do. We have, both in the lectures and in the classes, developed a consistent method to analyse complex semisimple Lie algebras. Now given such a Lie algebra $\mathfrak{g}$, one can ask what is the real Lie algebra $\mathfrak{g}_0$ that is the real form of the original complex Lie algebra $\mathfrak{g}$. Once we have a consistent correspondence, we can then look at the consequences of that and the classifications of real forms.

Here I will illustrate a particular example - finding the real form of the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$.

[55]We typically drop the $(\cdot)^{\mathbb{C}}$

Example 9.1. Consider the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. We want to find a real Lie algebra $\mathfrak{g}_0$ such that

$$\mathfrak{g}_0 \otimes_{\mathbb{R}} \mathbb{C} = \mathfrak{sl}_2(\mathbb{C}). \tag{9.22}$$

Note that the subset of $\mathfrak{sl}_2(\mathbb{C})$ of non-semisimple matrices is a proper subalgebratic variety and therefore it cannot contain a real subspace of $\mathfrak{sl}_2(\mathbb{C})$. So we can find a real $H \in \mathfrak{g}_0$ such that $\mathrm{ad}(H)$ acts semisimply on $\mathfrak{g}_0$ and is an element of the Cartan subalgebra. Now we consider the eigenspaces of $\mathrm{ad}(H)$ acting on $\mathfrak{g}$. Note that there must be an eigenvalue-0 space $\mathfrak{h}_0 = \mathbb{R} \cdot H$ spanned by H. The remaining two eigenvalues must then sum to zero, which leaves two possibilities.

1. $\mathrm{ad}(H)$ has real eigenvalues $\pm\lambda$. Rescaling by a real scalar gives $\lambda = 2$. The resulting decomposition of the vector space $\mathfrak{g}_0$ into eigenspaces is then

$$\mathfrak{g}_0 = \mathfrak{h}_0 \oplus \mathfrak{g}_2 \oplus \mathfrak{g}_{-2} \tag{9.23}$$

   Choosing $X \in \mathfrak{g}_2$ and $Y \in \mathfrak{g}_{-2}$ and again rescaling to give $[X, Y] = H$ gives

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{9.24}$$

   This is the basis of the real form $\mathfrak{sl}_2(\mathbb{R})$.

2. $\mathrm{ad}(H)$ has complex eigenvalues $i \pm \lambda$, where λ is a nonzero real number. Set $\lambda = 1$. The decomposition is now

$$\mathfrak{g}_0 = \mathfrak{h}_0 \oplus \mathfrak{g}_{\{1,-1\}} \tag{9.25}$$

   Choose basis B and C on $\mathfrak{g}_{\{1,-1\}}$ such that

$$\begin{cases} [H, B] & = C \\ [H, C] & = -B \end{cases} \tag{9.26}$$

   Note we now have two choices

$$[B, C] = \pm H \tag{9.27}$$

   If the negative sign is taken then $\mathfrak{g}_0 \cong \mathfrak{sl}_2(\mathbb{R})$ and we have

$$H = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{9.28}$$

   is the basis. If the positive sign is taken then

$$\mathfrak{g}_0 \cong \mathfrak{su}_2 = \left\{ M \,\middle|\, A^\dagger = -A,\ \mathrm{tr}(A) = 0 \right\} \subset \mathfrak{sl}_2(\mathbb{C}) \tag{9.29}$$

   and this has basis

$$H = \begin{pmatrix} \frac{i}{2} & 0 \\ 0 & \frac{i}{2} \end{pmatrix} \quad B = \begin{pmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} \quad C = \begin{pmatrix} 0 & \frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix} \tag{9.30}$$

The take is as follows. Given a real form $\mathfrak{g}_0 \subset \mathfrak{g}$ of a complex semisimple Lie algbera $\mathfrak{g}$ we can always find a real subalgebra $\mathfrak{h}_0 \subset \mathfrak{g}_0$ such that $\mathfrak{h}_0$ is the real form of the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. $\mathfrak{h}_0$ is then called the **Cartan subalgebra** of $\mathfrak{g}_0$. In general (rank two or higher), the root $\alpha \in \Delta$ of $\mathfrak{g}$ on $\mathfrak{h}_0$ need not be all real or purely imaginary - it can be a generic complex number. The values of roots depend heavily on the choice of $\mathfrak{h}_0$.

9.3 Split Forms and Compact Real Forms

The task is now simple. We want a statement of the classification of the simple real Lie algberas which complexifications are the classical/exceptional Lie algberas. We then want to go the other way, i.e. starting from a complex semisimple Lie algebra $\mathfrak{g}$ we want to construct real forms of $\mathfrak{g}$. Here I will not provide a classification theorem - one can refer to [1] for a full classification (in §26). Instead, I would like to present two distinguished real forms possessed by all complex semisimple Lie algebras. This will ultimately lead us back to answering the question I set out at the start of the section.

Split and compact forms represent the two extremes of behaviour of the decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha} \mathfrak{g}_\alpha \tag{9.31}$$

with respect to a real subalgebra $\mathfrak{g}_0 \subset \mathfrak{g}$. We will first give the definitions of the two real forms.

Definition 9.6. Let $\mathfrak{g}$ be a complex semisimple Lie algebra. The **split form** of $\mathfrak{g}$ is a real form $\mathfrak{g}_0$ such that there exists a Cartan subalgebra $\mathfrak{h}_0 \subset \mathfrak{g}_0$ (where $\mathfrak{h}_0^{\mathbb{C}} = \mathfrak{h}$) whose action on $\mathfrak{g}_0$ has all real eigenvalues. This means that all the roots $\alpha \in \Delta \subset \mathfrak{h}^*$ of $\mathfrak{g}$ assume all real values on the subspace $\mathfrak{h}_0$. We then have the direct sum decomposition:

$$\mathfrak{g}_0 = \mathfrak{h}_0 \oplus \bigoplus_{\alpha} \mathfrak{j}_\alpha \tag{9.32}$$

where the one-dimensional eigenspaces[56] $\mathfrak{j}_\alpha$ for the action of $\mathfrak{h}_0$. Each pair $\mathfrak{j}_{\pm\alpha}$ generates a subalgebra isomorphic to $\mathfrak{sl}_2(\mathbb{R})$.

The existence of the split form can be shown as follows. To construct a real (even rational) form $\mathfrak{g}_0$ of Lie algebra $\mathfrak{g}$ we start with generators X_{α_i} for the simple positive roots spaces of α_i and complete this to a stardard basis $(X_{\alpha_i}, Y_{\alpha_i})$ and $H_i = [X_{\alpha_i}, Y_{\alpha_i}]$ for the corresponding $\mathfrak{sl}_2(\mathbb{C})$ subspace. We then take $\mathfrak{g}_0$ to be the real subalgebra generated by these elements - the Cartan subalgebra $\mathfrak{h}_0 \subset \mathfrak{g}_0$ is then the real span of H_i. From this we have the following claim.

Proposition 9.3. *The split form construction uniquely characterises the real $\mathfrak{g}_0$ of $\mathfrak{g}$.*

Proof. We note that from the construction described above once $\mathfrak{h}$ is fixed for $\mathfrak{g}$, the real subalgebra $\mathfrak{h}_0$ is uniquely determined as the span of the H_α for all roots $\alpha \in \Delta$. The uniqueness is guaranteed as this is the only real form $\mathfrak{g}_0$ of $\mathfrak{g}$ which has a Cartan subalgebra $\mathfrak{h}_0$ acting on $\mathfrak{g}_0$ with all real eigenvalues. □

Remark 9.1. $\mathfrak{g}_0$ is detmined up to isomorphism and is sometimes called the natural real form of $\mathfrak{g}$.

There exists another more important real form in our discussion.

[56] Note that $\mathfrak{j}_\alpha = \mathfrak{g}_\alpha \cap \mathfrak{g}_0$.

Definition 9.7. Let $\mathfrak{g}$ be a complex semisimple Lie algebra. The **compact (real) form** of $\mathfrak{g}$ is a real form $\mathfrak{g}_0$ such that there exists a Cartan subalgebra $\mathfrak{h}_0 \subset \mathfrak{g}_0$ where all roots $\alpha \in \Delta \subset \mathfrak{h}^*$ of $\mathfrak{g}$ assume all purely imaginary values on the subspace $\mathfrak{h}_0$. We then have the direct sum decomposition:

$$\mathfrak{g}_0 = \mathfrak{h}_0 \oplus \bigoplus_{\alpha} \mathfrak{l}_\alpha \tag{9.33}$$

where $\mathfrak{l}_\alpha$ are two-dimensional eigenspaces[57] on which $\mathfrak{h}_0$ acts by rotation and generates a subalgebra isomorphic to $\mathfrak{su}_2$.

Compact real forms have the following properties.

Proposition 9.4. *Suppose $\mathfrak{g}$ is a complex semisimple Lie algebra and $\mathfrak{g}_0 \subset \mathfrak{g}$ is a real form of $\mathfrak{g}$. Let $\mathfrak{h}_0$ be a Cartan subalgebra of $\mathfrak{g}_0$ and the complexification $\mathfrak{h} = \mathfrak{h}_0 \otimes \mathbb{C}$ the corresponding Cartan subalgebra of $\mathfrak{g}$. Then the following are equivalent.*

1. *Each root $\alpha \in \Delta \subset \mathfrak{h}^*$ of $\mathfrak{g}$ assumes purely imaginary values on $\mathfrak{h}_0$ amd for each root α the subalgebra of $l_\alpha = (\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}) \cap \mathfrak{g}_0 \subset \mathfrak{g}_0$ is isomorphic to $\mathfrak{su}(2)$.*

2. *The restriction of $\mathfrak{g}_0$ of the Killing form of $\mathfrak{g}$ is negative-definite.*

3. *The real Lie group G_0 with Lie algebra $\mathfrak{g}_0$ is compact.*

Proof. This is the crux of the discussion so I will present the full proof. (1) $\implies$ (2): Consider the Killing form on $H \in \mathfrak{h}_0$. Since the roots are purely imaginary we have

$$\Gamma(H, H) = \sum_{\alpha \in \Delta} (\alpha(H))^2 < 0 \tag{9.34}$$

Since the subspaces $\mathfrak{l}_\alpha$ are orthogonal to each other with respect to Γ it remains to show that $\Gamma(Z, Z) < 0$ for an arbitrary $Z \in \mathfrak{l}_\alpha$. Let X and Y be generators of $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ respectively with $H = [X, Y]$ a standard basis for $\mathfrak{sl}_2(\mathbb{C})$. Take generators of the algebra of $\mathfrak{l}_\alpha$ as

$$\begin{cases} H' &= iH \\ U &= X - Y \\ V &= iX + iY \end{cases} \tag{9.35}$$

Setting

$$Z = aU + bV = (a + ib)X + (-a + ib)Y \tag{9.36}$$

then we must have

$$\begin{aligned} \mathrm{ad}(Z) \circ \mathrm{ad}(Z) = (a + ib)^2 \mathrm{ad}(X) \circ \mathrm{ad}(X) + (a - ib)^2 \mathrm{ad}(Y) \circ \mathrm{ad}(Y) \\ - (a^2 + b^2) (\mathrm{ad}(X) \circ \mathrm{ad}(Y) + \mathrm{ad}(Y) \circ \mathrm{ad}(X)) \end{aligned} \tag{9.37}$$

Tracing over eliminates the non-cross terms so we are left with

$$\mathrm{tr}\,(\mathrm{ad}(Z) \circ \mathrm{ad}(Z)) = -2(a^2 + b^2)\, \mathrm{tr}\,(\mathrm{ad}(X) \circ \mathrm{ad}(Y)) \tag{9.38}$$

[57] Note that in this case $\mathfrak{l}_\alpha = (\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}) \cap \mathfrak{g}_0$.

Now $\mathrm{ad}(X) \circ \mathrm{ad}(Y)$ acts by multiplication by $\frac{(n-\lambda)(n+\lambda-2)}{4} \geq 0$ on the λ-eigenspace for H (this is the factor multiplying the relevant element after raising and lowering using the operators in $\mathfrak{sl}_2\mathbb{C}$, see p. 150 of [1]). So we have proven the claim. (2) $\implies$ (3): First note that the **adjoint form**, defined as $G/Z(G)$ where $Z(G)$ is the centre of G. If G is connected then the image of the adjoint represetnation $\mathrm{Ad} : G \to Aut(\mathfrak{g})$ is the adjoint form (this is clear from the fact that $\ker(\mathrm{Ad}) = Z(G)$). Now the adjoint form G_0 is the connected component of the identity of the group $Aut(\mathfrak{g}_0)$. In particular it is a closed subgroup of the adjoint group of $\mathfrak{g}$ and acts faithfully on the real vector space $\mathfrak{g}_0$ (so preserves Γ). From Theorem 9.2 it is clear that if Γ is negative definite then G_0 is a closed subgroup of $SO_n(\mathbb{R})$ so is compact. (3) $\implies$ (1): Now suppose G_0 is compact. Then we can construct the group-averaging positive definite inner product on $\mathfrak{g}_0$,

$$(X, Y) = \int dg\mu(\rho(g)X, \rho(g)Y) \ , \tag{9.39}$$

where ρ is a representation of G and μ is an appropriate measure such that this is invariant under action of G. Then for any $X \in \mathfrak{g}_0$ $\mathrm{ad}(X)$ is represented by a skew-summetric matrix with respect to an orthogonal basis of $\mathfrak{g}_0$. Then $\Gamma(X, X) \leq 0$ and the eigenvalues of $\mathrm{ad}(X)$ are pure imaginary and $\alpha(\mathfrak{h}_0) \subset i\mathbb{R}$ so $\bar{\alpha} = -\alpha$ for root $\alpha \in \Delta$ so (1) follows. $\square$

Proposition 9.5. *Every semisimple complex Lie algebra has a unique compact form.*

Proof. Define the **conjugate linear involution** action as $\sigma : \mathfrak{g} \to \mathfrak{g}$ of a complexification $\mathfrak{g}$ of real $\mathfrak{g}_0$ as for $X \in \mathfrak{g}_0$ and $z \in \mathbb{C}$

$$X \otimes z \mapsto X \otimes \bar{z} \tag{9.40}$$

This is conjugate linear, Lie bracket-preserving and $\sigma^2 = 1$. Now $\mathfrak{g}_0$ is the fixed subalgebra of σ and conversely. Starting with the split form $\mathfrak{g}_{0,s}$ of $\mathfrak{g}$. Given a basis

$$\{H_i \in \mathfrak{h},\ X_\alpha \in \mathfrak{g}_\alpha,\ Y_\alpha \in \mathfrak{g}_{-\alpha}\} \tag{9.41}$$

we can define a unique Lie algebra automorphism φ of $\mathfrak{g}$ taking

$$\begin{cases} H_i & \mapsto -H_i \\ X_\alpha & \mapsto Y_\alpha \end{cases} \tag{9.42}$$

φ is a complex linear involution that preserves $\mathfrak{g}_{0,s}$. Note that $\sigma\varphi = \varphi\sigma$ so the fixed part $\mathfrak{g}_{0,c}$ is another real form of $\mathfrak{g}$ with Cartan subalgebra $\mathfrak{h}_{0,c} = \mathfrak{h}^{\sigma\varphi} = i\mathfrak{h}_{0,s}$ (note that here the i appears as upon applying the automorphism φ which changes the sign of H_i the fixed algebra of $\mathfrak{g}$ under $\sigma\varphi$ is now the purely imaginary part - i.e. complex conjugation and negative leaves pure imaginary numbers invariant in signs). Since the restriction of the Killing form to $\mathfrak{h}_{0,s}$, it is obvious that the restriction to $\mathfrak{h}_{0,c}$ is negative definite and therefore $\mathfrak{g}_{0,c}$ is the compact form of $\mathfrak{g}$. This construction of $\mathfrak{g}_{0,c}$ from $\mathfrak{g}_{0,s}$ is reversible so the compact form is hence unique by Proposition 9.3. $\square$

So what have we shown? In the discussion of complex semisimple Lie algebras we have implicitly, when discussing the root system, taken the split real form to obtain real roots. This allows us to analyse the root system in the Euclidean plane. On the other hand, if we solely consider *real* compact Lie algebras, we are forced to take effectively the compact form of the complexified Lie algebra as that is what we obtain from the Lie algebra analysis. This means that by complexifying the real compact Lie algebra one can in fact return to our discussion of the root systems, and therefore reproduce all the results as illustrated in the notes and classes. This also illustrates the extra minus sign in the Killing form. When we choose the inner product on the root space of compact real forms (and hence on real compact Lie algebras), there is an extra sign taking care of the fact that all roots are pure imaginary.

Complexification in fact is purely a mathematical tool. The physics, at the end of the day, is illustrated by a real Lie algebra. This is why the discussion of real Lie algebras and real forms of complex Lie algebras is so important.

10 Cartan-Weyl Decomposition, Root Diagrams and Weight Diagrams

This is a bit of a mixed botched section. I want to address a couple of issues and questions regarding Lie algberas in general, so have thrown basically everything I want to chat about into one giant section. Hopefully this makes sense in the grand scheme of things. I implore you to read this section by subsections (so read the ones that interest you) - the subsections are standalone explanations except when I explicitly refer to something I have talked about before.

10.1 Metric tensor and coordinates

In this subsection I would like to first address a few issues concerning the metric tensor G. Recall first from the notes the following:

Definition 10.1. The **metric tensor** G is defined as

$$G_{ij} = \frac{(\alpha_j, \alpha_j)}{2} \left(A^{-1}\right)_{ij} \tag{10.1}$$

This allows us to express the Killing form of two weights w, w':

$$(w, w')_D = a'_i G_{ij} a_j \tag{10.2}$$

in terms of their Dynkin label.

We discuss two issues at hand concerning G.

10.1.1 Root diagram coordinates

Let us consider B_2 algebra. The Cartan matrix is

$$A_{B_2} = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} \tag{10.3}$$

Let us recall the following:

Proposition 10.1. *The Dynkin labels of the simple roots are the rows of the Cartan matrix.*

Proof. See notes - this is more like a statement and you can check it yourself [58]. □

Now let us suppose you are given the simple roots: $\alpha_1 = (1, 0)$, $\alpha_2 = (-1, 1)$. The coordinates of the simple roots here now are the coordinates in the root diagram, in a Euclidean space. We can now use the Euclidean inner product between the simple roots to compute the Cartan matrix - you should see that this returns the same Cartan matrix above. What is going on? Why can we simply take $(\cdot, \cdot)$ as the Euclidean inner-product [59]?

We note that $(\cdot, \cdot)_D$ in Definition 10.1 is actually an inner product with respect to the

[58] If you cannot let me know.

[59] If you are not sure why one can simply take the Euclidean product here, have a look at [8–10]. I believe there will also be a proof in [1, 5]

Dynkin basis. We know from the discussion of the root systems of Lie algebras that upon restricting the Killing form of a complex semisimple Lie algbera to the $\mathfrak{h}^*_{\mathbb{R}} \in \mathfrak{h}^*$ using $\{H_\alpha\}$ as the dual basis it becomes a positive-definite biliner form. This allows us to take it as a metric, and operate it on a Euclidean basis. We can then convert between the Euclidean inner product from the Dynkin basis inner product, as described by the metric tensor using a diagonalisation and a rescale of eigenvalues. To do this note first that the weights in the adjoint basis (excluding the zero weights) are precisely the roots. Let us take two of these roots and consider their inner product in Dynkin basis

$$(\alpha, \beta)_D = a_i G_{ij} b_j \tag{10.4}$$

Now treat the metric tensor as a positive-definite bilinear form. We can pick a basis such that this bilinear form has only 1s on the diagonal matrix δ. This means we need the orthogonal matrix and a rescaling:

$$G \mapsto O^T D O = O^T \sqrt{D} \delta \sqrt{D} O \tag{10.5}$$

Now we can consider the root basis. We note that if we effect the following transform:

$$a \mapsto O^{-1} \sqrt{D}^{-1} a \tag{10.6}$$

This effectively maps the Dynkin basis to the Euclidean basis of the root diagram... we have come full circle and got the result we wanted [60]!

10.1.2 δ as half-sum of positive roots

We saw that δ emerged a few times (e.g. the Freudenthal formula (Eq (6.20) of notes), the Weyl formula (Eq (6.21) of notes) and the quadratic Casimir formula (Theorem 6.4 of notes)). We know that $\delta = (1, 1, ..., 1)$ for A_n - but is this true in general?

Proposition 10.2. *The **Weyl vector** δ, defined as the half-sum of all positive roots:*

$$\delta = \frac{1}{2} \sum_{\alpha \in \Delta_+} a_\alpha \tag{10.7}$$

where a_α indicates the Dynkin label of the positive root α, is equal to the sum of fundamental weights.

Proof. Firstly, you can define some element s_α as the reflection of the root on your arbitrarily chosen hyperplane. Then we must have

$$s_\alpha(\delta) = \delta - \alpha \tag{10.8}$$

and s_α permutes the positive roots apart from root α. You can also see that

$$(\delta, \alpha) = \frac{1}{2}(\alpha, \alpha) \tag{10.9}$$

[60] As far as I know this discussion cannot be found anywhere - so here you go!

Now, suppose ω_i to be the set of fundamental weights such that

$$\frac{2(\omega_i, \alpha_j)}{(\alpha_j, \alpha_j)} = \delta_{ij} \tag{10.10}$$

Now if you write $\delta = a_i \omega_i$, you will get

$$(\delta, \alpha_j) = a_j(\alpha_j, \alpha_j)/2 \tag{10.11}$$

If you then solve this for a_j, you should then get:

$$\delta = \frac{2(\delta, \alpha_j)}{(\alpha_j, \alpha_j)}\omega_j = \sum_i \omega_i \tag{10.12}$$

using Equation 10.9. □

The next obvious question would be why δ? What is significant about this quantity? Turns out it is related to some flag varieties (see this MO post for more details). This δ is sometimes called the "Weyl vector" (see for example here), and a related question is whether δ falls into the root lattice and what the significance of that is. There is a nice summary here.

There is a lot more to the theory of the Weyl vector. The whole industry of Weyl symmetry and chambers for example, is an important aspect in the theory of Lie algebras. For more information do check out [16].

10.2 Cartan-Weyl Decomposition - revisited

Another random aspect I would like to go back to is the Cartan-Weyl decomposition. Is there a way to understand the Cartan-Weyl basis, in particular, can we write the Killing form in this basis? Recall that the Cartan-Weyl decomposition is for a Lie algebra $\mathfrak{g}$ to be written as

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Lambda} \mathfrak{l}_\alpha \ , \tag{10.13}$$

where Λ is the root space. Take κ to be the Killing form defined in §4.2 of the notes as the natural inner product in $\mathfrak{g}$. Then we know that:

1. $\kappa\big|_{\mathfrak{h}\times\mathfrak{h}}$ is non-degenerate.
2. For all $\alpha \in H'$, the root set of the Lie algebra $\mathfrak{g}$, there exists a unique $H_\alpha \in \mathfrak{h}$ such that for all $H \in \mathfrak{h}$,
$$\kappa(H_\alpha, H) = \alpha(H) \ . \tag{10.14}$$
3. $\kappa(H, E_\alpha) = 0$ where $[H, E_\alpha] = \alpha(H)E_\alpha$.
4. $\kappa(E_\alpha, E_\beta) = 0$ for all $\alpha, \beta \in H'$ and $\underline{\alpha} \neq -\underline{\beta}$.
5. $\kappa(E_\alpha, E_{-\alpha}) \neq 0$.

I am basically just rewriting Theorem 4.6 of the notes. But note that now we have the following structure for κ in matrix form.

$$\kappa \xrightarrow{\text{in basis}} \begin{array}{cc} & \begin{array}{cccccc} H & E_\alpha & E_{-\alpha} & E_\beta & E_{-\beta} & \cdots \end{array} \\ \begin{array}{c} H \\ E_\alpha \\ E_{-\alpha} \\ E_\beta \\ E_{-\beta} \\ \vdots \end{array} & \begin{pmatrix} \text{non-degenerate} & & & & & \\ & 0 & 1 & & & \\ & 1 & 0 & & & \\ & & & 0 & 1 & \\ & & & 1 & 0 & \\ & & & & & \ddots \end{pmatrix} \end{array} \tag{10.15}$$

We note the following points.

- In the $\mathfrak{h}$ (CSA) I can pick the basis $\{H_i, i = 1, ..., \operatorname{rk}\mathfrak{g}\}$. But to construct the $\mathfrak{su}(2)_\alpha$ subalgebra it is often easier to construct H_α using the root structure. In general however $\{H_\alpha\}$ span $\mathfrak{h}$ but it is not a basis as in general $\{H_\alpha\}$ are linearly-dependent.
- We can rescale the E_α to get a normalised basis, which is what I have done above. Of course it is possible to rediagonalise $\mathfrak{l}_\alpha \oplus \mathfrak{l}_{-\alpha}$ to get a diagonal basis, but this then obscures the $\mathfrak{su}(2)_\alpha$ interpretation.

This is all I want to say about Cartan-Weyl basis for now.

10.3 Weyl group and chambers

A small interlude here. I want to define something known as the Weyl group of the root system, which is a symmetry group of the root space under some natural involution action of the roots. This section is quite mathematical so you are welcomed to skip if it bothers you too much - but it will establish some of the notations that will be used later on.

The first thing to define is the Weyl group. Here we fix $\mathfrak{g}$ a Lie algebra and $\mathfrak{h}$ its Cartan subalgebra.

Definition 10.2. Define an involution $W_\alpha : \mathfrak{h} \to \mathfrak{h}$ by

$$W_\alpha(H) = H - 2\frac{(\alpha, H)}{(\alpha, \alpha)}\alpha\ , \tag{10.16}$$

with $(\cdot, \cdot)$ denoting the usual inner product of the root system Δ. The **Weyl group** of the root system Δ, denoted $\mathfrak{W}$, is then the subgroup of $GL(\mathfrak{h})$ generated by the involutions W_α where $\alpha \in \Delta$.

The geometric meaning of the involution W_α is the reflection about the hyperplane orthogonal to α. We then have the following theorem.

Theorem 10.1. *The action of $\mathfrak{W}$ on the roots preserves the root space Δ.*

Proof. The action of reflection W_α maps a root into a root by the definition of a root system [4]. If this is unclear, have a look at Theorem 7.26 of [16]. $\square$

Corollary 10.1. *The Weyl group is finite.*

Proof. The roots span $\mathfrak{h}$ so each $w \in \mathfrak{W}$ can be determined by its action on the root space Δ. In that sense $\mathfrak{W}$ can be though of as a subgroup of the permutation group on the roots, which is finite. $\square$

With this we can now define a base (note that the root system is based in Euclidean space).

Definition 10.3. A subset B of Δ is called a **base** if:

1. B is a basis of the Euclidean space E.
2. Each root β can be written as $\beta = \sum_{\alpha \in B} k_\alpha \alpha$ with integral coefficients k_α all non-negative or all non-positive.

In this sense the roots in B are known as **simple**. The **height** of a root is then defined as

$$h(\beta) = \sum_{\alpha \in B} k_\alpha \, . \tag{10.17}$$

This is often how the simple roots are defined [1, 4, 16]. Now we introduce a bit of terminology.

Definition 10.4. The hyperplanes P_α where $\alpha \in \Delta$ partitions E into finite many regions. The connected components of

$$E - \bigcup_\alpha P_\alpha \tag{10.18}$$

are called the open **Weyl chambers** of E.

Each component $\gamma \in E - \cup_\alpha P_\alpha$ is called **regular**.

There is an important theorem that states the following.

Theorem 10.2. *Let $\gamma \in E$ be regular. Then the set $S(\gamma)$ of all indecomposable roots (that cannot be written as a non-negative sum of other roots) in $\Delta^+(\gamma)$ is a base of Δ and every base is obtainable in this manner.*

Proof. See §10.1 of [4]. $\square$

The key point from the theorem is the following statement.

Corollary 10.2. *The Weyl chambers are in natural one-to-one correspondence with bases.*

Proof. This is clear by looking at $\mathcal{W}(\gamma)$, the Weyl chamber containing γ. Then $\mathcal{W}(\gamma) = \mathcal{W}(\gamma')$ means γ, γ' lie on the same side of each hyperplane P_α. $\square$

Definition 10.5. The **fundamental Weyl chamber relative to base** S is the Weyl chamber $\mathcal{W}(S) = \mathcal{W}(\gamma)$ where $S = S(\gamma)$.

The Weyl group then sends one Weyl chamber onto another. In fact $\mathfrak{W}$ permutes the bases of the root system Δ (see §10.3 of [4]).

How does this help us analyse root and weight systems? The key point to note is the following statement.

Proposition 10.3. *Every vertex of the convex hull of the weights of the representation $V = \oplus_\alpha V_\alpha$ is conjugate to α under the Weyl group $\mathfrak{W}$.*

Proof. The weights of V will contain a string of weights of the form

$$\alpha, \alpha + \beta, ..., \alpha - (-(\alpha, \beta))\beta \;, \tag{10.19}$$

the last term being exactly the term $\mathfrak{W}_\beta(\alpha)$ so it is conjugate to α. □

This seems to be not that helpful but this is now important.

Proposition 10.4. *Suppose $\mathcal{W}$ is a Weyl chamber and $e \in E$ is a point in the Euclidean space. Then there exists exactly one point in the $\mathfrak{W}$-orbit of e such that it lies in the closure of the Weyl chamber $\bar{\mathcal{W}}$.*

Proof. The key points of the proof is to use the fact that the Weyl group acts transitively on the set of open Weyl chambers [4]. See also Proposition 8.29 of [16]. □

We now define the following.

Definition 10.6. A point $\mu \in E$ is **strictly dominant relative to** S iff $\mu \subset \mathcal{W}(S)$. It is **dominant relative to** S if $\mu \subset \bar{\mathcal{W}}(S)$. From Proposition 10.4 for all $\mu \in E$ there exists $w \in \mathfrak{W}$ such that $w \cdot \mu$ is dominant.

This dominant point μ is in fact how we define the highest weight of the system. Typically the fundamental weights are fixed by the Cartan matrix (see below §10.5), and hence the Weyl chamber is fixed (this is known as the fundamental Weyl chamber associated to the fundamental weights). Then every point $\mu \in E$ that satisfies for all $\alpha \in \Delta$,

$$(\mu, \alpha) = 2\frac{(\mu, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z} \tag{10.20}$$

called an **integral element** is then a highest weight. I don't have time to go to all the details here, but you should look at §8 and §9 of [16] for the details (and it will be worth it).

Working in progress: Need to add in hand-drawn diagrams and related theories about Weyl characters.

10.4 Calculating root diagrams

With all the machinery set-up we are now ready to tackle root diagrams. Recall that root systems are often the key in analysing the Lie algebra content - constructing the root system of a a Lie algebra is quite straightforward. In this subsection I would like to present an algorithm to compute the root system of a simple Lie algebra $\mathfrak{g}$, given the simple roots of the Lie algebra $\mathfrak{g}$. The goal is to compute and classify Lie algebras using the root system.

Let's therefore suppose I am given a set of simple positive roots $\{\alpha_i\}$ with $\alpha_i \in H'$, H' the root space (see notes, this is the space of functionals such that for $T \in \mathfrak{g}$, $\mathrm{ad}(H)(T) = \alpha(H)(T)$). The Cartan matrix is given by,

$$A_{ij} = \frac{2\underline{\alpha}_i \cdot \underline{\alpha}_j}{\underline{\alpha}_j \cdot \underline{\alpha}_j} \;, \tag{10.21}$$

with no sum on i, j. In doing so we have secretly moved to the split form of the Lie algebra so the roots are real and can be evaluated in a Euclidean root space. The dot product here therefore is the Euclidean product. From the arguments in §10.1 and in the notes, we know that we can always pick the inner product $(\alpha, \beta) = \alpha_i \beta_j (\kappa|^{-1}_{\mathfrak{h}\times\mathfrak{h}})^{ij}$ and rotate the basis by an orthogonal transformation such that the inner product is Euclidean. It is now possible to generate the root diagram by using the following algorithm.

Proposition 10.5. *The folloiwing is an algorithm to generate all roots S from th set of simple roots Δ.*

1. *Let $S = \Delta$.*
2. *For each pair of distinct roots, $\alpha \in \Delta$ and $\beta \in S$, determine the maximum and minimum integers $n_\pm$ such that $\underline{\beta} + n_\pm \underline{\alpha}$ are roots. The useful results are:*
   - *(i)* $n_+ + n_- = -\frac{2\underline{\alpha}\cdot\underline{\beta}}{|\underline{\beta}|^2}$.
   - *(ii) If $\underline{\gamma}$ is a root, $\lambda\underline{\gamma}$ the constant multiple of $\underline{\gamma}$ is a root if and only if $\lambda = \pm 1$.*
   - *(iii) If $\underline{\alpha}, \underline{\beta} \in \Delta$, then $\underline{\alpha} - \underline{\beta}$ is not a root.*
   - *(iv) Any root string has at most length of 4.*
   - *(v) For $\underline{\alpha}, \underline{\beta} \in \Delta$, the $\underline{\alpha}$-string through $\underline{\beta}$*

     $$S_{\underline{\alpha},\underline{\beta}} = \{\underline{\beta} + \rho\underline{\alpha} \in S \,|\, \rho \in \mathbb{Z}\} \tag{10.22}$$

     has length

     $$1 - \frac{2\underline{\alpha}\cdot\underline{\beta}}{|\underline{\alpha}|^2} \in \mathbb{N} \;. \tag{10.23}$$
   - *(vi) If $\underline{\alpha}\cdot\underline{\beta} \lesseqgtr 0$ then $\underline{\alpha} \pm \underline{\beta}$ is a root.*
3. *Repeat (2) until exhaustion. The positive root set is S.*
4. *The full root set is then $S \cup (-S)$.*

Proof. The proof is basically just ultilising Proposition 6.1 of the notes together with a lot of root system facts. I will refer to [38] for the details.

The proof of (i) is a bit more complicated (Proposition 3 of [38]). The idea is to consider the vector space

$$V = \bigoplus_{n\in\mathbb{Z}} \mathfrak{g}_{\underline{\beta}+n\underline{\alpha}} \, , \tag{10.24}$$

and notice that this space is reducible (since this space is one-dimensional and invariant under the $\mathfrak{su}(2)_\alpha$ action). This immediately means that there exists integers p, q with $p \leq 0 \leq q$ such that $\underline{\beta} + n\underline{\alpha}$ is a root iff $p \leq n \leq q$. Now constrain using the action of $[H_\alpha, \cdot]$.

(iv) and (vi) comes from Theorem 6.1 of the notes. (ii), (iii) and (v) follows from the intermediate steps in constructing root strings, see [38] for details (Definition 3, Theorem 1 and Lemma 12). The multiple property is proved in §8.3 of [4]. $\square$

This is a bit complicated to explain without going into full details so let us illustrate this with an example.

Example 10.1. Let us look at $B_2 \cong \mathfrak{so}(5)_\mathbb{C}$. The simple roots in this case is $\underline{\alpha}_1 = (1,0)$ and $\underline{\alpha}_2 = (-1,1)$. Starting with $\Delta = \{\underline{\alpha}_1, \underline{\alpha}_2\}$, we have $\underline{\alpha}_1 \cdot \underline{\alpha}_1 = 1$, $\underline{\alpha}_1 \cdot \underline{\alpha}_2 = 2$ and $\underline{\alpha}_2 \cdot \underline{\alpha}_2 = -1$ (remember we are using the Euclidean inner product here). So setting $S = \Delta$, we can then look at the $\underline{\alpha}_1$-string through $\underline{\alpha}_2$, $\underline{\alpha}_2 + n\underline{\alpha}_1$.

- $\underline{\alpha}_2 - \underline{\alpha}_1$ is not a root by (iii) of Proposition 10.5. So $n_- = 0$.
- From (i) and (v) of Proposition 10.5 we have $n_+ = 2$.
- So we add $\underline{\alpha}_2 + \underline{\alpha}_1$ and $\underline{\alpha}_2 + 2\underline{\alpha}_1$ to set S.

Now we look at $S = \{\underline{\alpha}_1, \underline{\alpha}_2, \underline{\alpha}_1 + \underline{\alpha}_2, 2\underline{\alpha}_1 + \underline{\alpha}_2\}$, then we have:

- $\underline{\alpha}_1$-strings through any $s \in S$ will not lead to new roots since $2\underline{\alpha}_1$ is not a new root by (ii) of Proposition 10.5 and the others is the same as the root string $S_{\underline{\alpha}_1,\underline{\alpha}_2}$.
- $S_{\underline{\alpha}_2,\underline{\alpha}_1}$ gives $n_- = 0$ and $n_+ = 1$, so generates $\underline{\alpha}_1 + \underline{\alpha}_2$, which we already have.
- $S_{\underline{\alpha}_2,\underline{\alpha}_1+\underline{\alpha}_2}$ is redundant again.
- $S_{\underline{\alpha}_2,\underline{\alpha}_1+2\underline{\alpha}_2}$ gives $n_+ = 0 = n_-$ since both $2\underline{\alpha}_1$ and $2(\underline{\alpha}_1 + \underline{\alpha}_2)$ are not roots.

So we deduce by exhaustion that there are no more roots. The full root set Φ is therefore,

$$\Phi = \{\pm\underline{\alpha}_1, \pm\underline{\alpha}_2, \pm(\underline{\alpha}_1 + \underline{\alpha}_2), \pm(2\underline{\alpha}_1 + \underline{\alpha}_2)\} \tag{10.25}$$

giving the root diagram in Figure 10.1.

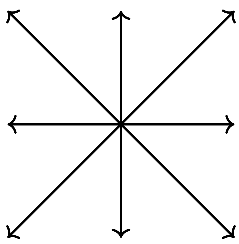


Figure 10.1: Root Diagram of B_2.

Distributed version:

Exercise 10.1. Repeat the analysis of the root system of B_2 to the other three rank-2 Lie algebras, A_2, D_2 and G_2. In each case,

1. Write down the Cartan matrix and metric tensor. Invert to find the simple roots $\underline{\alpha}_1$ and $\underline{\alpha}_2$ in the Euclidean basis.
2. Find the root system.
3. Draw the root system and verify these are exactly as illustrated in p.118 of the notes.

10.5 Weight diagrams

Now we move on to weight diagrams. In particular I want to introduce an algorithm that replaces the one on p.127 of the notes (as I find it quite confusing to check what "all the weights" is) to find all the weights given the highest weight Λ.

Proposition 10.6. *Let us suppose d_Λ is the irreducible representation of $\mathfrak{g}$ with highest weight Λ. Let $|\Lambda\rangle$ be the Dynkin index of the highest weight. To generate all the weights, we use the following algorithm.*

1. *Set the set of weights $S = \{\underline{\Lambda}\}$.*
2. *For each element $\underline{\mu} \in S$, write,*
$$\underline{\mu} = \sum \mu_i \underline{\omega}_i \,, \tag{10.26}$$
where $\underline{\omega}_i$ are the fundamental weights of $\mathfrak{g}$. Then we construct all weights of the form
$$\underline{\mu} - \sum_{i=1}^{r} m_i \underline{\alpha}_i \tag{10.27}$$
where $\underline{\alpha}_i$ are the simple roots from S for all $0 \leq m_i \leq \lambda_i$ where λ_i is the Dynkin index of $\underline{\mu}$ and $m_i \in \mathbb{N}$. We explicitly use the Dynkin-index basis of the simple roots $\underline{\alpha}_i$.
3. *Exhaust the previous step to generate the full weight system.*

Proof. This is known as the highest weight system and it is covered in most Lie algebra books, see for example, [9, 16]. □

Again let us illustrate this with an example.

Example 10.2. Consider G_2 with the Cartan matrix,

$$A_{G_2} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \tag{10.28}$$

and consider its fundamental irrep where $|\Lambda\rangle = |1,0\rangle$. Then go through the algorithm.

- Start with $|1,0\rangle$. We can apply $E_{-\underline{\alpha}_1}$ once to get
$$|1,0\rangle \xrightarrow{-\alpha_1} |-1,1\rangle \ .$$
- Dynkin index of the second entry is $+1$ so apply $E_{-\underline{\alpha}_2}$ once to get
$$|-1,1\rangle \xrightarrow{-\alpha_2} |2,-1\rangle \ .$$
- Dynkin index of the first entry is $+2$ so apply $E_{-\underline{\alpha}_1}$ twice to get
$$|2,-1\rangle \xrightarrow{-\alpha_1} |0,0\rangle \xrightarrow{-\alpha_1} |-2,1\rangle \ .$$
- $|0,0\rangle$ terminates so we can only apply $E_{-\underline{\alpha}_2}$ to $|-2,1\rangle$ once to get
$$|-2,1\rangle \xrightarrow{-\alpha_2} |1,-1\rangle \ .$$
- Finally apply once of $E_{-\underline{\alpha}_1}$ to get
$$|1,-1\rangle \xrightarrow{-\alpha_1} |-1,0\rangle \ .$$

At this point we need to check degeneracies,. But there are none with the same $\underline{H} = (H_1, H_2)$ eigenvalues so there is no degeneracy. This is the **7** rep of G_2.

To draw the weight diagram we can rotate it back into standard Euclidean basis. The fundamental weights are defined as

$$\underline{\mu}_i = B_{ij}\underline{\alpha}_j \ , \tag{10.29}$$

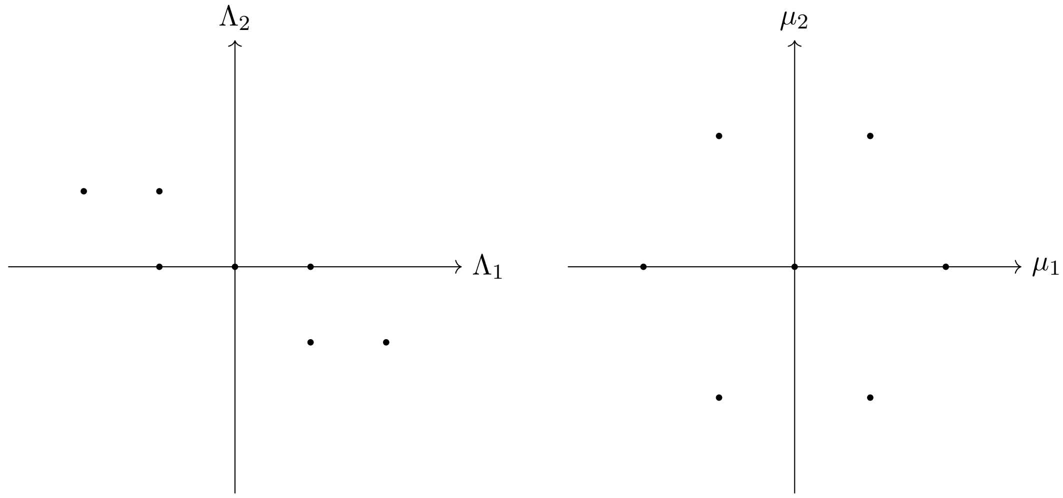
where B_{ij} is the inverse of the Cartan matrix. So using

$$\underline{\alpha}_1 = (1,0) \tag{10.30}$$
$$\underline{\alpha}_2 = \frac{1}{2}(-3,\sqrt{3}) \tag{10.31}$$

We get

$$\underline{\mu}_1 = \frac{1}{2}(1,\sqrt{3}) \tag{10.32}$$
$$\underline{\mu}_2 = (0,\sqrt{3}) \ . \tag{10.33}$$

The weight diagram of **7** of G_2 is then sketched out in Figure 10.5.



(a) Weight diagram of the fundamental representation of G_2 in the Dynkin basis. (b) Weight diagram of the fundamental representation of G_2 in the Euclidean basis.

Figure 10.2: The weight diagram of the fundamental representation of G_2 in different basis. Note that the two weight diagrams are not scaled with respect to each other.

Distributed version:

Exercise 10.2. Repeat for the irrep of G_2 defined by the highest weight $|\underline{\Lambda}\rangle = |0, 1\rangle$. Plot the weight diagram. Compare this with the G_2 root diagram. What is this represestation of G_2?

Distributed version:

Exercise 10.3. Recall that in Theorem 6.3 of the notes we have the Weyl formula:

$$\dim(\lambda) = \prod_{\alpha \in \Delta_+} \frac{(\lambda + \delta, \alpha)}{(\delta, \alpha)} , \tag{10.34}$$

with $\delta = \frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha$ or with Dynkin labels all being one. Calculate the dimension of the spinor representation **16** of $\mathfrak{so}_{10}$ in Q4 of Sheet 3 using this formula now that we have all the relevant elements defined.

10.6 Projection matrices and branching

The final little piece of information concerns the analysis of branching rules,

$$R^{(G)} \to R^{(H)} = \bigoplus_s R_s^{(H)} , \tag{10.35}$$

where $R_s^{(H)}$ are irreps of H (see Eq. (1.38) of the notes). It is often easier to analyse the embedding of H in G in terms of a **projection matrix** that takes the roots and weights of $\mathfrak{g}$ into $\mathfrak{h}$, the Lie algebras of G and H respectively. The existence of these matrices can be trivially proven as the roots of $\mathfrak{h}$ must be identified with a subset of that of $\mathfrak{g}$ since by definition of branching $H \subset G$ as a subgroup.

It is perhaps interesting to look at the effect of the Weyl group defined in §10.3 on the projection matrices. Clearly, the Weyl-reflected root diagrams give the same embedding but the coordination of the weight space of G will now be different. This normally means the projection matrix will just have different entries, as we will need to fix some initial coordinates to analyse the projection matrices of G onto H.

The question now becomes how we choose such a coordination. This is basically done by requiring the highest weight of an irreducible representation of $\mathfrak{g}$ to be projected to that of $\mathfrak{h}$, say the fundamental rep. Let us look at some examples to illustrate this point.

Example 10.3. Let us analyse the branching projection matrix of $SU(3) \to SU(2)\times U(1)$. To do this we first normalise the $U(1)$ charge to $3Y$, Y being the hypercharge so P is an integer matrix. We want to project

$$P\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ u_1 \end{pmatrix} , \tag{10.36}$$

where a_i, b_i and u_i are the weights of $SU(3)$, $SU(2)$ and $U(1)$ respectively. In particular we have also nominally set b_1 to be $2I_3$ of $SU(3)$. Apply this to

$$\mathbf{3} \to \mathbf{2}_{-1} \oplus \mathbf{1}_{-2} \tag{10.37}$$

we then are able to fix

$$P = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} . \tag{10.38}$$

It is now obvious to check that this projection matrix also works for the adjoint rep $\mathbf{8}$ of $SU(3)$, the branching rule exactly gives

$$\mathbf{8} \to \mathbf{3}_0 \oplus \mathbf{2}_1 \oplus \mathbf{2}_{-1} \oplus \mathbf{1}_0 . \tag{10.39}$$

Distributed version:

Exercise 10.4. Work out the projection matrix of P_5 for the branching rule of $SU(5) \to SU(3) \times SU(2)$. You may need the branching rules

$$\mathbf{5} \to (\mathbf{1},\mathbf{3}) \oplus (\mathbf{2},\mathbf{1}) \tag{10.40}$$
$$\bar{\mathbf{5}} \to (\mathbf{1},\bar{\mathbf{3}}) \oplus (\mathbf{2},\mathbf{1}) \tag{10.41}$$
$$\mathbf{10} \to (\mathbf{1},\mathbf{1}) \oplus (\mathbf{2},\mathbf{3}) \oplus (\mathbf{1},\bar{\mathbf{3}}) \tag{10.42}$$
$$\bar{\mathbf{10}} \to (\mathbf{1},\mathbf{1}) \oplus (\mathbf{2},\bar{\mathbf{3}}) \oplus (\mathbf{1},\mathbf{3}) . \tag{10.43}$$

Also work out the $U(1)$ dual vector. Verify all of this with the data you are given in Q3 of Sheet 3.

Working in progress: Include discussion on how weight diagrams are related to root diagrams. Also a discussion on the dual vector operator.

11 Subalgebras, Subgroups and Dynkin Diagrams

In this short section I would like to highlight why the algorithm provided in §6.3 in the notes on finding subalgebras of Lie algebras $\mathfrak{g}$ using Dynkin and extended Dynkin diagrams. Subalgebras and subgroups are often important as in physics symmetry breaking often requires one to find the subgroup of an existing group under some physical processes.

11.1 Subalgebras and Extended Dynkin Diagrams

Let us first, for a complete discussion, recall the definitions of a subalgebra.

Definition 11.1. A **Lie subalgebra** $\mathfrak{h}$ of Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h} \subset \mathfrak{g}$ that is closed under the bracket operation, i.e. if $X, Y \in \mathfrak{h}$, then $[X, Y] \in \mathfrak{h}$.

Similarly, we have

Definition 11.2. A **Lie subgroup** H of Lie group G is a subset $H \subset G$ that is also a group.

We recall from the notes the following definitions.

Definition 11.3. A **maximal subalgebra** is a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ such that there are no other subalgebras contained between $\mathfrak{g}$ and $\mathfrak{h}$. i.e. A subalgebra $\mathfrak{j} \subset \mathfrak{g}$ such that

$$\mathfrak{h} \subset \mathfrak{j} \subset \mathfrak{g} \tag{11.1}$$

does not exist. Here we use the $\subset$ notation to denote **proper** containment., i.e. the containment excludes the trivial containment $\mathfrak{g} \subseteq \mathfrak{g}$.

We also define the notion of a regular subalgebra.

Definition 11.4. A **regular subalgebra** $\mathfrak{g}'$ of a Lie algebra $\mathfrak{g}$ is a subalgebra with the property that given the root space decomposition of $\mathfrak{g}$ as

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha \tag{11.2}$$

we can write the root space decomposition of $\mathfrak{g}'$ as

$$\mathfrak{g}' = \mathfrak{h}' \oplus \bigoplus_{\alpha \in \Delta'} \mathfrak{g}_\alpha \tag{11.3}$$

such that $\mathfrak{h}' \subset \mathfrak{h}$ as a subalgebra and $\Delta' \subset \Delta$ is a subset of the roots.

Let us illustrate this with an example.

Example 11.1. Consider $A_3 \oplus A_1 \subset A_5$, or $\mathfrak{g}' = \mathfrak{sl}_4 \oplus \mathfrak{sl}_2 \subset \mathfrak{sl}_6 = \mathfrak{g}$. We can take the Cartan subalgebra $\mathfrak{h}'$ to be the diagonal matrices so clearly $\mathfrak{h}' \subset \mathfrak{h}$. We note that $\dim(\mathfrak{h}') + 1 = \dim(\mathfrak{h})$ as there is a one-dimensional subspace of $\mathfrak{h}$ generated by the elements

$$H = \mathrm{diag}(1, 1, 1, 1, -2, -2) \tag{11.4}$$

which indicates the extra $\mathfrak{u}(1)$ factor[61]. The root vectors of $\mathfrak{g}'$ are just the root vectors of $\mathfrak{g}$ with non-zero elements on the two diagonal blocks. We have therefore shown that $\mathfrak{g}'$ is a regular subalgebra of $\mathfrak{g}$.

Example 11.2. Let us look at a counter example - a non-regular subalgebra. Consider $\mathfrak{g}' = A_2 \oplus A_1 \subset A_5 = \mathfrak{g}$. Using the Pauli matrices $\{\sigma_i\}_{i=1,2,3}$ and the Gell-Mann matrices $\{\lambda_j\}_{j=1,\ldots,8}$ as basis of $\mathfrak{sl}_2$ and $\mathfrak{sl}_3$ respectively and defining

$$\begin{cases} \sigma_0 = \mathbb{1}_2 \\ \lambda_0 = \mathbb{1}_3 \end{cases} \tag{11.5}$$

then every 6×6 matrix can be written as

$$\{\sigma_i \otimes \lambda_j\}_{i=0,\ldots,3;\, j=0,\ldots,8} \tag{11.6}$$

Note that $[\sigma_0 \otimes \lambda_i, \sigma_j \otimes \lambda_0] = 0$. So the Cartan subalgebra of $\mathfrak{g}'$ is generated by

$$\{\sigma_3 \otimes \lambda_0, \sigma_0 \otimes \lambda_3, \sigma_0 \otimes \lambda_8\} \tag{11.7}$$

So $\mathfrak{h}' \subset \mathfrak{h}$. However, looking at the root vectors:

$$\{\sigma_\pm \otimes \lambda_0, \sigma_0 \otimes t_\pm, \sigma_0 \otimes u_\pm, \sigma_0 \otimes v_\pm\} \tag{11.8}$$

where $t_\pm = \frac{1}{2}(\lambda_1 \pm \lambda_2)$, $u_\pm = \frac{1}{2}(\lambda_4 \pm \lambda_5)$ and $v_\pm = \frac{1}{2}(\lambda_6 \pm \lambda_7)$, we see that $\Delta' \subsetneq \Delta$. So this is not a regular subalgebra of $\mathfrak{g}$.

In the lecture notes, we have set out an algorithm that allows us to compute regular subalgebras using Dynkin diagrams [40]. In [40], two notions of subalgebras are introduced.

Definition 11.5. ***R*-subalgebras** are subalgebras that are contained in some regular subaglebra of Lie algebra $\mathfrak{g}$. Otherwise the subaglebra is known as an ***S*-subalgebra**.

Before we continue again let us recall the definitions of simple and semisimple in the context of Lie algebras.

Definition 11.6. An **abelian** Lie algebra is a Lie algebra that satisfies $[\mathfrak{g}, \mathfrak{g}] = 0$.

Definition 11.7. A **simple** Lie algebra is a Lie algebra that contains no proper ideals and is not abelian.

Definition 11.8. A direct sum of simple Lie algebras is known as a **semisimple** Lie algebra.

Definition 11.9. A **reductive** Lie algebra is the direct sum of a simple and an abelian Lie algebra.

[61] The reason why this is not in $\mathfrak{h}'$ is intuitively quite clear. $\mathfrak{g}'$ contains a tensor sum sot eh two algebras A_3 and A_1 should be independently generated. If this element is contained in $\mathfrak{h}'$ then the two algebras are no longer linearly independent of each other.

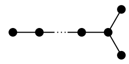
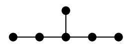
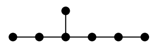
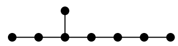
Semi-simple Lie algebra	Dynkin Diagram
A_n	
B_n	
C_n	
D_n	
E_6	
E_7	
E_8	
F_4	
G_2	

Table 11.1: The list of Dynkin diagrams for semi-simple Lie algebras.

We would like to classify all reductive subalgebras. This requires us to look for all maximal regular subalgebras as well as maximal S-subalgebras. We note that a maximal subalgebra may not be semisimple - here we restrict to the cases where all maximal subalgebras are semisimple. We state without proof the following result.

Lemma 11.1. *Every reductive subalgebra of a Lie algebra $\mathfrak{g}$ that contains an abelian ideal is an R-subalgebra.*

Proof. See Dynkin's original paper [40]. □

As it turns out the R-subalgebra case is much more easily dealt with and from now on we will focus on this (i.e. identifying regular subalgebras). Let us begin by considering $\mathfrak{g}' \subset \mathfrak{g}$ which is a regular subalgebra and $\Delta' \subset \Delta$ be the set of roots of $\mathfrak{g}'$ (where $\Pi' \subset \Delta'$ is the set of simple roots). Recall that if $\alpha', \beta' \in \Pi'$, then $\alpha' - \beta' \notin \Delta'$. (In fact $\alpha' - \beta' \notin \Delta$ otherwise this will lead to a contradiction where $\alpha' - \beta' \in \Delta'$.) This means the following:

Proposition 11.1. *To find regular subalgebras of $\mathfrak{g}$, we set to find sets of $\Pi' \in \Delta$ such that if $\alpha', \beta' \in \Pi'$ then*

$$\alpha' - \beta' \notin \Delta \tag{11.9}$$

Then the subalgebra will be generated by the set of elements $\{e_{\pm\alpha'}, h_{\alpha'}\}$ where $\alpha' \in \Pi'$.

Proof. This is clear from the analysis above. □

To do this Dynkin therefore introduced the notion of Extended Dynkin Diagrams (see Figures 11.1 and 11.2) with the help of adding in a new root.

Definition 11.10. An **extended Dynkin diagram** is formed with the root set together with the most negative root. For the root set Δ of Lie algebra $\mathfrak{g}$, the new root set is

$$\tilde{\Delta} = \Delta \cup \{-\theta\} \tag{11.10}$$

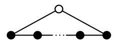
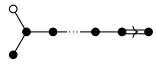
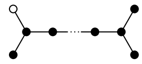
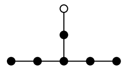
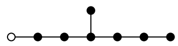
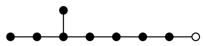
Semi-simple Lie algebra	Extended Dynkin Diagram
$\tilde{A}_n$	
$\tilde{B}_n$	
$\tilde{C}_n$	
$\tilde{D}_n$	
$\tilde{E}_6$	
$\tilde{E}_7$	
$\tilde{E}_8$	
$\tilde{F}_4$	
$\tilde{G}_2$	

Table 11.2: The list of extended (affine) Dynkin diagrams for semi-simple Lie algebras.

where θ is the highest (positive) root of $\mathfrak{g}$. This root system is called the **extended Π-system** where as the original root system is knwon as the **Π-system**.

Using this we can now restate Theorem 6.5 of the notes.

Theorem 11.1. *Up to isomorphism all maximal semisimple subalgebras of $\mathfrak{g}$ are obtained by consider all root subsets*

$$\tilde{S} \subset \Delta \cup \{-\theta\} \tag{11.11}$$

where θ is the highest root of $\mathfrak{g}$. In particular we have the following algorithm regarding the use of extended Dynkin diagrams.

1. *Regular semisimple maximal subalgebras $\mathfrak{g}'$ can be obtained by considering the Dynkin diagram that has one node removed from the extended Dynkin diagram of $\mathfrak{g}$, with the following exceptions that are not maximal*
   (a) *F_4, $\Delta \cup \{-\theta\} \setminus \{\alpha_3\}$: $A_3 \oplus A_1 \hookrightarrow B_4 \hookrightarrow F_4$.*
   (b) *E_7, $\Delta \cup \{-\theta\} \setminus \{\alpha_3\}$: $A_3 \oplus A_3 \oplus A_1 \hookrightarrow D_6 \oplus A_1 \hookrightarrow E_7$.*
   (c) *E_8, $\Delta \cup \{-\theta\} \setminus \{\alpha_3\}$: $A_3 \oplus D_5 \hookrightarrow D_8 \hookrightarrow E_8$.*
   (d) *E_8, $\Delta \cup \{-\theta\} \setminus \{\alpha_5\}$: $A_5 \oplus A_2 \oplus A_1 \hookrightarrow E_6 \oplus A_2 \hookrightarrow E_8$.*
   (e) *E_8, $\Delta \cup \{-\theta\} \setminus \{\alpha_6\}$: $A_7 \oplus A_1 \hookrightarrow E_7 \oplus A_1 \hookrightarrow E_8$.*
2. *Reductive semisimple maximal subalgebras $\mathfrak{g}'' \oplus \mathfrak{u}(1)_{\mathbb{C}}$ is obtained by removing one node from the Dynkin diagram of $\mathfrak{g}$.*

Proof. First consider the regular semisimple maximal subalgebras. The extended Π-system of Lie algebra $\mathfrak{g}$ has the following property. Choosing the lowest root $\alpha_0 = -\theta$, where θ is the highest root, means that $\alpha_0 - \alpha_j$ for any root α_j is not a root. This means that

$$\frac{2(\alpha_0, \alpha_j)}{(\alpha_0, \alpha_0)} \in \mathbb{Z} \tag{11.12}$$

$$\frac{2(\alpha_0, \alpha_j)}{(\alpha_j, \alpha_j)} \in \mathbb{Z} \tag{11.13}$$

so the extended Π-system obeys the same relations as a Π-system with an extra linear relation between the roots - the system is linearly-dependent. Now removing one root from the extended Π-system leaves a linearly-independent system, with the roots $\{\alpha_i\}$ obeying the relation

$$\frac{2(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)} = -(n_+ - n_-) \tag{11.14}$$

This means that the resulting algebra is still a Lie algebra, and the roots are simple roots of a regular maximal subalgebra of $\mathfrak{g}$. The system may not be indecomposable so the resulting algebra may not be semisimple, but we have restricted to the semisimple case (i.e. discard all non-semisimple algebras obtained from this method). The exceptional cases are discovered later - one can follow the discussions in [9, 39, 41]. My choice of root conventions follow [9]. For the reductive case, notice that these are parabolic subalgebras and are algebraically generated by the Cartan-Weyl basis of $\mathfrak{g}$ without e_α. Removing $e_{-\alpha}$ as well from the Cartan-Weyl basis[62], we see that there is an extra h_α left over which generates an additional $\mathfrak{u}(1)_\mathbb{C}$ subalgebra. Removing the set $\{e_{\pm\alpha}, h_\alpha\}$ gives the semisimple Lie algebra which is a subalgebra of this reductive subalgebra of the form $\mathfrak{g}'' \oplus \mathfrak{u}(1)_\mathbb{C}$. □

To obtain S-subalgebras one must follow a different approach - this is discussed in §8.4 of [9].

We also note that promoting the subalgebra to the subgroup level requires us to again consider global arguments as discussed in earlier sections and in [1]. In particular, discrete quotients may affect our results. As an example, the maximal subalgebra of E_8 is $E_6 \oplus A_2$, but in the group level the maximal subgroup of E_8 is $E_6 \times SU(3) \setminus (\mathbb{Z}/3\mathbb{Z})$.

11.2 Symmetry Breaking in Physics

Symmetry breaking is ubiquitous in physics. Often times due to some physical effect a high-energy gauge group G will be broken down to a smaller symmetry H. The principle is as follows:

Proposition 11.2. *Let G be the high-energy gauge group in which the subgroup $G' \subset G$ is broken. Then the low-energy symmetry will be the commutant of G', i.e. the group H where $G \supset G' \times H$ times the abelian part that commutes with G'.*

Proof. This is related to Goldstone's Theorem. See Advanced Quantum Field Theory. □

[62] Or Chevalley basis.

If the centre of the group G is discrete, or that the maximal subgroup has a discrete quotient, then its effects will be picked out by global topological effects that are perturbatively inaccessible. In particular in the perturbative regime we can operate solely on the algebra level.

We are told that the world we know is govern by symmetries. Symmetries naturallly provide invariants known as conserved quantities of the space by Noether's theorem, which in the end allow us easier access to the system as these conserved quantities, such as energy and momentum, provide physical laws that systems must obey provided that the symmetry exists. However, it is not true that all symmetries (that we have invented) are obeyed for all the possible physical models. One can, for example, imagine writing down a theory that contains terms that do not respect the symmetry. On the other hand, one can also imagine that the symmetry is 'broken' by some mechanism, such as a phase transition, such that the physical state that appears at first seems to do not respect the original symmetry of the theory. What we have just described above are the two 'senses' of the symmetry breaking in physics, summarised below:

- **Explicit symmetry breaking**. This is when the equations of motion explicitly violates the symmetry being considered. Typically, if the theory has a Lagrangian description, this is illustrated by an explicit set of terms in the Lagrangian that vary under the symmetry being considered. A common example is explicit supersymmetry breaking, where terms that violate supersymmetry (SUSY) are written down as 'small violations' of a supersymmetric theory (for example, say the MSSM) and its phenomenological effects are analysed.

- **Spontaneous broken symmetry**. This is typically manifested when there is a degeneracy in the vacua [63]

In your AQFT course you will encounter the latter type of broken symmetry and study that in more detail. In particular, the Higgs mechanism is in fact an explicit example of spontaneous broken symmetry where part of the Standard Model gauge symmetry is broken to the $U(1)$ gauge group,

$$SU(2)_L \times U(1)_Y \to U(1)_{\text{EM}} \ . \tag{11.15}$$

This mechanism is in fact the reason why the quarks and leptons acquire a mass via something known as the Yukawa term - as the Higgs boson acquires a VEV, the Yukawa term provides a Dirac-type coupling to the mass matrices and hence the massless quarks and charge leptons acquire the necessary mass terms. Another typical example is in string theory - to study Heterotic $E_8 \times E_8$ string compactifications, one picks the first E_8 as the gauge group and choose some relevant vector bundle with an associated gauge group $G \subset E_8$ such that the low-energy gauge group is the commutant of G in E_8. A typical

[63] For an extremely insightful physics discussion, see [42]. The symmetry is not realised as symmetry transformations of the physical states of the theory, and in particular they transform the vaccuum state. However, the full theory still contains the full symmetry being considered.

choice is a rank-5 line bundle which gives $G \simeq SU(5)$ so the commutant is some $SU(5)$ with a bundle of abelian factors (which may have phenomenological consequences) [43, 44]. I will leave this discussion here for you to ponder and think about after your AQFT and Strings II courses.

> Working in progress: There will be a more rigorous discussion for symmetry breaking later.

12 Casimirs

In the lectures and problem sheets we have encountered something known as the quadratic Casimir. The quadratic Casimir turns out to be very useful at identifying irreducible representations of $\mathfrak{su}_3$. Is that the only invariant combination we can find in a Lie algebra - are there any other objects of a similar property that we can construct? We will aim to answer this question in this subsection.

To understand the theory of Casimirs, let us take a step back and look at the structure of Lie algebras and look at a more general structure. We will deduce new Casimir operators in that manner.

12.1 Universal Enveloping Algebra

First let us recall what the definition of an algebra is (for completeness).

Definition 12.1. An **algebra** is a ring R together with the action of a field $\mathbb{F}$ on R such that multiplication and addition of the ring is compatible with the field action.

An associative algebra is an algebra that satisfies the associative property.

As alluded before the Lie algebra is an algebra (clearly, hence the name). It is important to note that it is possible to construct a Lie algebra from any associative algebra $\mathfrak{U}$ by simoply considering the same vector space and defining the Lie bracket as the commutator with respect to the associative product in $\mathfrak{U}$, i.e. if $a, b \in \mathfrak{U}$, we can construct the Lie bracket,

$$[a, b] = a \cdot b - b \cdot a \ , \tag{12.1}$$

where $\cdot$ is the associative product in $\mathfrak{U}$. The important statement is that we can go in the opposite direction - given a Lie algebra $\mathfrak{g}$, it is possible to construct an associative algebra $\mathfrak{U}$ which contains $\mathfrak{g}$ as a subspace and for which the commutator that naturally exists in $\mathfrak{U}$ will reproduce the Lie bracket for $\mathfrak{g}$ when restricted to $\mathfrak{g}$. This motivates the following theorem.

Theorem 12.1. *For any Lie algebra $\mathfrak{g}$, there exists an associative algebra with identity denoted $U(\mathfrak{g})$ together with the map $\iota : \mathfrak{g} \to U(\mathfrak{g})$ such that the following properties hold:*

- $\iota([X, Y]) = \iota(X)\iota(Y) - \iota(Y)\iota(X)$ *for all $X, Y \in \mathfrak{g}$.*
- $\iota(X)$ *generates the algebra $U(\mathfrak{g})$, so it is itself the smallest subalgebra with identity containing all $\iota(X)$.*
- *(Universal property) Suppose $\mathcal{A}$ is an associative algbera with identity and $j : \mathfrak{g} \to \mathcal{A}$ is a linear map such that $j([X, Y]) = j(X)j(Y) - j(Y)j(X)$ for all $X, Y \in \mathfrak{g}$. then there exists a unique algebra homomorphism $\phi : U(\mathfrak{g}) \to \mathcal{A}$ such that $\phi(1) = 1$ and $\phi(\iota(X)) = j(X)$ for all $X \in \mathfrak{g}$. In other words, the map j splits through the homomorphism ϕ; a universal enveloping algbera $U(\mathfrak{g})$ is a universal initial object such that $\iota_U : \mathfrak{l} \to U(\mathfrak{l})$ in the category of enveloping algberas of $\mathfrak{l}$.*

The pair $(U(\mathfrak{g}), \iota)$ is known as a ***universal enveloping algbera of*** $\mathfrak{g}$*.*

Proof. Noting that the tensor product on vector spaces is associative, define the **tensor algebra** $T(\mathfrak{g})$ over $\mathfrak{g}$ as

$$T(\mathfrak{g}) = \bigoplus_{k=0}^{\infty} \mathfrak{g}^{\otimes k} \ . \tag{12.2}$$

Each element of $T(\mathfrak{g})$ has finite linear combinations of elements in $\mathfrak{g}^{\otimes k}$ for each k in the sum. $T(\mathfrak{g})$ is then an associative algebra with identity if we define the product of two elements simply by using the tensor product, i.e. for $u_I = u_{i_1} \otimes ... \otimes u_{i_k}$ and $v_I = v_{j_1} \otimes ... \otimes v_{j_l}$ we have

$$u_I \cdot v_J = u_I \otimes v_J \tag{12.3}$$

and extend the product by linearity. Then the identity element $1 \in \mathbb{C} = \mathfrak{g}^{\otimes 0}$ is the multiplicative identity for $T(\mathfrak{g})$, and the associativity of the tensor product assutures that $T(\mathfrak{g})$ is an associative algebra.

Now we need to show that $T(\mathfrak{g})$ has the universal property. We construct the map $\phi : T(\mathfrak{g}) \to \mathcal{A}$ for any associative algbera with identity $\mathcal{A}$ by restricting ϕ to $\mathfrak{g}^{\otimes k}$ to be the unique linear map $\mathfrak{g}^{\otimes k} \to \mathcal{A}$ such that

$$\phi(X_i \otimes ... \otimes X_k) = j(X_1)...j(X_k) \ , \tag{12.4}$$

for any linear map $j : \mathfrak{g} \to \mathcal{A}$. It is straightforward to check that ϕ is an algbera homomorphism and if ϕ is to be an algbera homomorphism that agrees with j on $\mathfrak{g}$ then this is the unique form.

Now we construct $U(\mathfrak{g})$ as a quotient of $T(\mathfrak{g})$. We take the smallest **two-sided ideal** of $T(\mathfrak{g})$, $\mathfrak{j}$, where for all $\alpha \in T(\mathfrak{g})$ and $\beta \in \mathfrak{j}$, $\alpha\beta, \beta\alpha \in \mathfrak{j}$, to contain all the elements of form

$$X \otimes Y - Y \otimes X - [X, Y] \ , \tag{12.5}$$

where $X, Y \in \mathfrak{g}$. So $\mathfrak{j}$ is the space of elements of the form

$$\sum_{j=1}^{N} \alpha_j (X_j \otimes Y_j - Y_j \otimes X_j - [X-j, Y_j])\beta_j \ , \tag{12.6}$$

where $X_j, Y_j \in \mathfrak{g}$ and $\alpha_j, \beta_j \in T(\mathfrak{g})$. We form the quotient vector space $T(\mathfrak{g})/\mathfrak{j}$ and we can check that the kernel of ϕ will contain all the elements of the form in Eq.(12.5) and is a two-sided ideal so $\ker(\phi) \supset \mathfrak{j}$. Hence the map $\phi : T(\mathfrak{g}) \to \mathcal{A}$ factors through $U(\mathfrak{g})$ which gives the desired homomorphism. The uniqueness of $U(\mathfrak{g})$ is guaranteed by the fact that it is spanned by products of elements of $\mathfrak{g}$. □

12.2 Casimirs and Invariant Tensors

With the universal enveloping algebra defined, we can now construct something known as the Casimir which belongs to the centre of $U(\mathfrak{g})$.

Definition 12.2. A **Casimir element** is a distinguished element of the centre of the universal eveloping algebra $\mathcal{Z}(U(\mathfrak{g}))$ of Lie algebra $\mathfrak{g}$.

A prototype example of a Casimir element is the quadratic Casimir, as defined in the lectures as follows.

Definition 12.3. The **quadratic Casimir operator** of a semi-seimple Lie algebra $\}$ is defined as

$$\mathcal{C}_2 = \kappa^{ij} T_i T_j \ , \tag{12.7}$$

where κ^{ij} is the Killing form and T_i are the generators of the Lie algebra $\mathfrak{g}$. It is straightforward to show that the value of the quadratic Casimir operator is independent of the choice of the basis (see Proposition 10.5 of [16]). To check that $\mathcal{C}_2$ indeed is a distinguished element of $\mathcal{Z}(U(\mathfrak{g}))$ is also straightforward - it is straightforward to check $[X_i, \mathcal{C}_2] = 0$ for all $X_j \in \mathfrak{g}$, so since $U(\mathfrak{g})$ is generated by the elements of $\mathfrak{g}$ we then see that $\mathcal{C}_2$ commutes with every element of $U(\mathfrak{g})$ and hence is in the centre of $U(\mathfrak{g})$.

It is now straightforward to generate this notion to different orders known as the 'higher-order Casimir operators'.

Definition 12.4. The **higher-order Casimir operators** form a distinguished basis of the centre $\mathcal{Z}(U(\mathfrak{g}))$, given by homogeneous polynomials,

$$\mathcal{C}_n = d_{a_1 \ldots a_n} T^{a_1} \ldots T^{a_n} \tag{12.8}$$

where T^a are the generators of $\mathfrak{g}$ with $d_{a_1 \ldots a_n}$ being the suitable invariant tensors of the adjoint representation. The **order** n of the Casimir operator $\mathcal{C}_n$ is the order n of the polynomial.

Recall that representation matrices in the adjoint representation are just the structure constants. Since the Casimir operators are distinguished elements of the centre of the universal enveloping algebra $U(\mathfrak{g})$, they are homogeneous polynomials in the generators of $\mathfrak{g}$ and therefore constitute a maximal set of algebraiically indepedent elements of the centre. It also follows that they are invariant tensors of the adjoint representation (as they can be represented by structure constants effectively).

Proposition 12.1. *The Casimir operators are in one-to-one correpondence with the set of tensors that generate the space of adjoint representation invariant tensors.*

Proof. To show this let us look at the polynomial,

$$P = c\mathbb{1} + \sum t_a T^a + \sum_{a,b} t_{ab} T^a T^b + \ldots \ , \tag{12.9}$$

where T^a are the generators of $\mathfrak{g}$. Supppose this polynomial belongs to the centre of $U(\mathfrak{g})$. Then it is sufficient that all the coefficients $t_{a_1 \ldots a_n}$ are invariant tensors of the adjoint representations, as the G-invariance of the element requires the coefficients to all be in the adjoint representation such that the transformation leaves it in the universal enveloping algbera $U(\mathfrak{g})$ [64]. The converse is also true. □

[64] To further see this point, look at §17 of [9] or [45].

We have deduced that all Casimir operators of $\mathfrak{g}$ are of the form

$$\mathcal{C}_n = d_{a_1 \ldots a_n} T^{a_1} \ldots T^{a_n} \tag{12.8}$$

To see this explicitly, note that for the quadratic Casimir, which is,

$$\mathcal{C}_2 = \kappa_{ab} T^a T^b \ , \tag{12.7}$$

we have that the Killing form $\kappa_{ab} \propto \sum_{c,d} f^{ac}{}_d f^{bd}{}_c$. We however note that although other elements of the centre $\mathcal{Z}(U(\mathfrak{g}))$ can be obtained by taking the tensors $\mathrm{tr}(\mathrm{ad}_{T^i} \circ \ldots \circ \mathrm{ad}_{T^n})$, it is not necessarily true that they are all algebraically indepedent so they do not typically provide Casimir operators. There is however an independent high-order Casimir operator for each order (which for the case A_n you can see [9]).

There is clearly a lot more to the theory of invariant tensors and Casimirs. I implore you to read §14 and §17 of [9] for more on that.

Working in progress: Include discussion on how indices and Casimir operators are used. Section on affine Lie algebras at some point.

13 The Standard Model

In Questions 2-3 of Sheet 3 we have been looking at the weight systems and branchings of $SU(5)$. We have shown after (some) work:

$$\mathbf{10} \oplus \bar{\mathbf{5}} \to (\mathbf{3},\mathbf{2})_{1/3} \oplus (\bar{\mathbf{3}},\mathbf{1})_{-4/3} \oplus (\mathbf{1},\mathbf{1})_{2} \oplus (\mathbf{1},\mathbf{2})_{-1} \oplus (\bar{\mathbf{3}},\mathbf{1})_{2/3} \tag{13.1}$$

where $U_Y(1)$ charge is attached as a subscript. This branching is precisely the representation for one standard model family. To see this, let us first recall what the Standard Model is.

The Standard Model is a chiral quantum field theory that is gauged under local transformations of the Standard Model gauge group:

$$G_{SM} = SU(3)_C \times SU(2)_L \times U(1)_Y \tag{13.2}$$

The $SU(3)_C$ being the colour gauge symmetry, $SU(2)_L$ the electroweak symmetry on left-handed particles and $U(1)$ typically known as the hypercharge. We denote the fundamental gauge fields as G^A_μ, W^a_μ and B_μ, which after spontaneous symmetry breaking we will write G^A_μ, $W^\pm_\mu$, Z_μ and A_μ. The Lagrangian for the kinetic and self-interaction terms of the gauge bosons is then:

$$\mathcal{L}^{\text{gauge}} = -\frac{1}{4}\left(G^A_{\mu\nu}\right)^2 - \frac{1}{4}\left(W^a_{\mu\nu}\right)^2 - \frac{1}{4}\left(B_{\mu\nu}\right)^2 - \Theta_G G^A_{\mu\nu}\tilde{G}^A_{\mu\nu} - \Theta_W W^a_{\mu\nu}\tilde{W}^a_{\mu\nu} - \Theta_B B_{\mu\nu}\tilde{B}_{\mu\nu} \tag{13.3}$$

Note that here the topological θ-terms are allowed *a priori*. The fermionic matter content can be summarised below:

$$\begin{aligned}
(\mathbf{3},\mathbf{2},\tfrac{1}{6}):&\quad Q^i_L = \left\{ \begin{pmatrix} u_L \\ d_L \end{pmatrix} \begin{pmatrix} c_L \\ s_L \end{pmatrix} \begin{pmatrix} t_L \\ b_L \end{pmatrix} \right\} \\
(\bar{\mathbf{3}},\mathbf{1},\tfrac{2}{3}):&\quad u^i_R = \{u_R, c_R, t_R\} \\
(\bar{\mathbf{3}},\mathbf{1},-\tfrac{1}{3}):&\quad d^i_R = \{d_R, s_R, b_R\} \\
(\mathbf{1},\mathbf{2},-\tfrac{1}{2}):&\quad L^i_L = \left\{ \begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix} \begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix} \begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix} \right\} \\
(\mathbf{1},\mathbf{1},-1):&\quad e^i_R = \{e_R, \mu_R, \tau_R\} \\
(\mathbf{1},\mathbf{1},0):&\quad \nu^i_R = \{\nu_{e,R}, \nu_{\mu,R}, \nu_{\tau,R}\}
\end{aligned} \tag{13.4}$$

where the right-handed neutrinos can be neglected from the model. There are two pieces in the Lagrangian for the fermions. The first is the kinetic term:

$$\mathcal{L}^{\text{kinetic}}_F = i\bar{L}^i_L \not{D} L^i_L + i\bar{Q}^i_L \not{D} Q^i_L + i\bar{e}^i_R \not{D} e^i_R + i\bar{\nu}^i_R \not{D} \nu^i_R + i\overline{U}^i_R \not{D} u^i_R + i\bar{d}^i_R \not{D} d^i_R \tag{13.5}$$

The covariant derivative is written as:

$$D_\mu = \partial_\mu - ig_s G^A_\mu T^A - igW^a_\mu T^a - ig'YB_\mu \tag{13.6}$$

The Yukawa couplings are

$$\mathcal{L}_F^{\text{Yukawa}} = -y_{ij}^d \bar{Q}_L^i H d_R^j - y_{ij}^u \bar{Q}_L^i \tilde{H} u_R^j + y_{ij}^e \bar{L}^i H e_R^j \left(+ y_{ij}^\nu \bar{L}_L^i H \nu_R^j \right) + (\text{h.c.}) \tag{13.7}$$

The Higgs field is a complex scalar doublet in the system:

$$H = \begin{pmatrix} H_+ \\ H_0 \end{pmatrix} \tag{13.8}$$

with indices $(\mathbf{1}, \mathbf{2}, \frac{1}{2})$, with Lagrangian

$$\mathcal{L}^{\text{Higgs}} = D_\mu H (D^\mu H)^\dagger + m^2 |H|^2 - \lambda |H|^4 \tag{13.9}$$

So the Standard Model Lagrangian is given by, in general,

$$\mathcal{L}_{SM} = \mathcal{L}^{\text{gauge}} + \mathcal{L}_F^{\text{kinetic}} + \mathcal{L}_F^{\text{Yukawa}} + \mathcal{L}^{\text{Higgs}} \tag{13.10}$$

There is of course a lot more that goes on with the Standard Model — its physical behaviours, its mathematical properties and potential extensions (typically known as Beyond the Standard Model theories). We won't have time to go through all of this in detail but do refer to [46, 47] or for a wonderful set of lectures note [48].

14 Grand Unified Theories

Having clarified what the Standard Model is, let us have a look at what Grand Unified Theories are. The idea of Grand Unified Theories is to propose an underlying gauge group which contains $G_{SM} = SU(3) \times SU(2) \times U(1)$ as a subgroup. The G_{GUT} must hence be at least rank 4, and must contain:

$$G_{GUT} \supset G_{SM} = SU(3)_c \times SU(2)_L \times U(1)_Y \tag{14.1}$$

The motivation comes from the unification of the gauge couplings into a single one. Consider the one-loop renormalisation equation for gauge couplings.

$$\frac{1}{\alpha_a(Q^2)} = \frac{1}{\alpha_a(M^2)} + \frac{b_a}{4\pi} \log \frac{M^2}{Q^2} \tag{14.2}$$

with

$$b = -\frac{11}{3}N + \frac{2}{3}C(R)n_f + \frac{1}{3}C(R)n_s \tag{14.3}$$

For the Standard Model this is

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{22}{3} \\ -11 \end{pmatrix} + \frac{4}{3}N_{gen} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + N_{Higgs} \begin{pmatrix} \frac{1}{10} \\ \frac{1}{6} \\ 0 \end{pmatrix} \tag{14.4}$$

where N_{gen} and N_{Higgs} is the number of generations and Higgs respectively. We can look at the evolution of the of the couplings. For non-SUSY (supersymmetric) one-loop renormalisations, the couplings do not match precisely together at high energies. However, for MSSM models, the couplings exactly match - the gauge couplings join together to become one single gauge coupling at some energy scale Λ_{GUT}. This is what we are really considering when we talk about Grand Unified Theories - these are supersymmetric theories that represent the “unified” theory when the gauge couplings of QCD and EW join together at high energies.

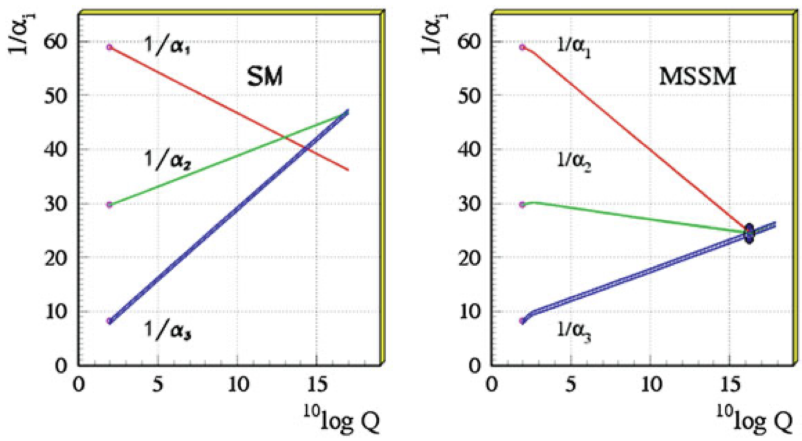


Figure 14.1: The evolution of the inverse couplings α_i with respect to the logarithmic energy scale $\log Q$. Figure taken from [49].

We in addition have to assume that there are no further relevant degrees of freedom in the intermediate scales. This is the so-called *desert-hypothesis*. We note that the full GUT multiplets do not modify the running of the relative gauge couplings at one-loop accuracy, so their introduction does not spoil the unification of couplings. Here we consider two prominent GUT groups - $SU(5)$ and $SO(10)$.

Working in progress: Updates on GUT theories.

14.1 $SU(5)$ GUT Group

$SU(5)$ is the unique compact simple Lie group of rank 4 admitting complex representations. It contains G_{SM} as a maximal subgroup. Let us look at some properties of this GUT group.

Properties of $SU(5)$ GUT:

1. The number of gauge bosons is $5^2 - 1 = 24$. There are 12 Standard Model and 12 extra gauge bosons labelled $(X_r^{\pm}, Y_r^{\pm})$ transforming as $SU(2)_L$ doublets and $SU(3)_c$ triplets.

2. The abelian subgroup commuting with the generators of $SU(3) \times SU(2)$ is

$$T_{24} = \sqrt{\frac{3}{5}} \begin{pmatrix} -\frac{1}{3} & & & & \\ & -\frac{1}{3} & & & \\ & & -\frac{1}{3} & & \\ & & & \frac{1}{2} & \\ & & & & \frac{1}{2} \end{pmatrix} = \sqrt{\frac{3}{5}} \frac{Y}{2} \tag{14.5}$$

3. Each quark-lepton generation fits nicely in the reducible representation $\bar{\mathbf{5}} \oplus \mathbf{10}$. The generators are:

$$T_A = \begin{pmatrix} \frac{1}{2}\lambda_A & 0 \\ 0 & 0 \end{pmatrix} \tag{14.6}$$

for $A = 1, ..., 8$ and

$$T_A = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2}\sigma_{A-20} \end{pmatrix} \tag{14.7}$$

for $A = 21, 22, 23$. The generators labelled $A_9, ..., 20$ correspond to the $SU(5)/G_{SM}$ generators. The explicit embedding is given by

$$\bar{\mathbf{5}} = \begin{pmatrix} d_1^c \\ d_2^c \\ d_3^c \\ e^- \\ \nu_e \end{pmatrix} \tag{14.8}$$

and

$$\mathbf{10} = \begin{pmatrix} 0 & u_3^c & u_2^c & u_1 & d_1 \\ & 0 & u_1^c & u_2 & d_2 \\ & & 0 & u_3 & d_3 \\ & & & 0 & e^+ \\ & & & & 0 \end{pmatrix} \tag{14.9}$$

where $d^c = D_R^1$, $(\nu_e, e^-) = L^1$, $u^c = U_R^1$, $(u,d) = Q_L^1$, $e^+ = E_R^1$. This assignment is free of anomalies.

4. The fermionic part of the Lagrangian is

$$\mathcal{L}_f = \bar{\mathbf{5}}_\alpha^\dagger (i\bar{\sigma}_\mu D^\mu)_\alpha{}^\beta \bar{\mathbf{5}}_\beta + \frac{1}{2}\mathbf{10}^{\alpha\beta\dagger}(i\bar{\sigma}_\mu D^\mu)_{\gamma\delta}^{\alpha\beta}\mathbf{10}^{\gamma\delta} \tag{14.10}$$

where $D^\mu = \partial^\mu + ig_5 T_A A_A^\mu$.

5. To break $SU(5)$ to the Standard Model spontaneously we add the scalar field "GUT-Higgs" transforming in the adjoint **24**:

$$\mathcal{L}_\Sigma = \mathrm{tr}(D_\mu\Sigma)^2 - \mathrm{tr}\big(\Sigma^2 - M_G^2\big) - (\lambda H^\dagger \Sigma H + \bar{\lambda}\bar{H}\Sigma\bar{H}^\dagger) - m_H^2 H^\dagger H - m_{\bar{H}}^2 \bar{H}^\dagger \bar{H} \tag{14.11}$$

Under $SU(5)$ this transforms as:

$$\Sigma \mapsto U\Sigma U^\dagger \tag{14.12}$$

and acquires a large VEV

$$\langle\Sigma\rangle = \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & -3 & \\ & & & & -3 \end{pmatrix} v \tag{14.13}$$

so the X, Y gauge masses $M_{X,Y}^2 \simeq \alpha_{GUT} v^2$.

The $SU(5)$ GUT theory has many implications[65]:

1. Charge Quantisation: $\mathrm{tr}\, Q_{EM} = 0$ since $Q_{EM} \in SU(5)$ so

$$Q_{d^c} = -\frac{1}{3}Q_{e^-} \tag{14.14}$$

2. Gauge Coupling Relation: Because of the unification of gauge couplings we obtain relations between couplings. For example, we get:

$$g_1^2 = \frac{3}{5}g_2^2 \tag{14.15}$$

[65] I am too lazy to use the Feynman diagram package, might update this later please forgive me...

3. Fermionic Mass Relations: SU(5) relates Yukawa couplings of d and e^-, giving

$$\frac{m_s}{m_d} = \frac{m_\mu}{m_e} \tag{14.16}$$

This is obviously incompatible (you can check this) with the observed values. Georgi and Jarlskog [50] used $SU(5)$ Clebschs to fix this in an ingenious way - you can find out more in [49].

4. Nucleon Decay: X and Y gauge bosons allows baryon number violating decay modes, for example: $p \to \pi^0 e^+$.

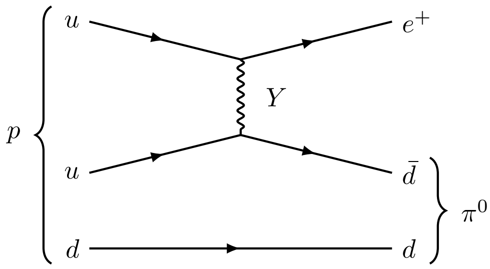


Figure 14.2: Feynman diagram of the decay $p \to \pi^0 e^+$.

This is suppressed by the large M_Y mass. The simplest $SU(5)$ GUT is therefore excluded as

$$\tau_{p\to\pi^0 e^+} \simeq 4 \times 10^{29\pm.7}\text{years} < \tau_{p\to\pi^0 e^+,\text{bound}} = 6.6 \times 10^{33}\text{years} . \tag{14.17}$$

A detailed calculation can be found in [49].

14.2 $SO(10)$ GUT Group

The $SO(10)$ is one of the two rank-5 compact simple Lie groups admitting complex representations. One family of Standard Model quarks and leptons fits nicely in the irrep **16**:

$$\psi_{16} = (\nu_e\ u_1\ u_2\ u_3;\ e^-\ d_1\ d_2\ d_3;\ d_3^c\ d_2^c\ d_1^c\ e^+;\ u_3^c\ u_2^c\ u_1^c \nu_R) \tag{14.18}$$

including the right-handed neutrino. $SO(10)$ in fact contains $SU(5) \times U(1)$ as a maximal subgroup, with the splitting (which you computed last sheet)

$$\mathbf{45} \to \mathbf{24} \oplus \mathbf{10} \oplus \bar{\mathbf{10}} \oplus \mathbf{1} \tag{14.19}$$

Breaking of the $SO(10)$ can proceed in many steps, one of them being, for example:

$$SO(10) \to SU(5) \times U(1) \to G_{SM} \tag{14.20}$$

This requires adjoint scalars Φ_{45} and also extra scalars ϕ_{16} in the **16** irrep. The minimal embedding of the electroweak Higgs doublet is in a multiplet of scalars H_{10} in **10**. Since

$$\mathbf{10} \to \mathbf{5} \oplus \bar{\mathbf{5}} \tag{14.21}$$

which gives (H_u, H_d), the Yukawa terms are

$$\mathcal{L} = Y^{ij} \bar{\psi}^i_{16} \psi^j_{16} H_{10} + h.c. \tag{14.22}$$

so we have unification of Yukawa couplings at the GUT scale. Of course, this is not really physical as this implies that there are no CKM mixings. One can consider improved models in Georgi [51].

15 The sequel: AQFT

We are not done - in fact, this is only the beginning. In fact, there is a lot more to symmetries, Lie algebras and physics. This course is just the teaser to let you peek at that wonderful mystery land of theoretical and mathematical physics. There are many, many more things to chat about, such as affine Lie algebras, gauge bundles, axions, string theory...

You are lucky though. The natural sequel to this course (and the culmination of the QFT course together with all the knowledge you have learnt so far) in the Advanced Quantum Field Theory course in Hilary Term. In particular, you will encounter three main ideas:

1. **Functional methods in QFT.** There is of course, a similar formulation of QFT as the path integral formulation of quantum mechanics. The path integral gives one access to information that cannot be accessed from the perturbative Hamiltonian formulation of QFT, such as the theory of anomalies which are fundamental in understanding low-energy effective field theories of high energy theories.

2. **Renormalisation.** This is a huge topic (and a big mess) that I have completely skipped in these set of notes [66]. Renormalisation in QFT typically covers two main ideas - regularisation, the idea of removing infinities by introducing cut-offs in a theory, and renormalisation, the idea of removing dependencies on the cut-offs you have introduced in regularisation to compute physical quantities. The actual theory takes about an entire lecture course to explain (and possibly a whole lifetime to really comprehend) but I will leave it here [67].

3. **Gauge theories.** Gauge theory is an extremely difficult topic. One can simply view gauges as redundacies and this is in fact a very useful way of thinking about it in the context of quantum field theory, but the 'correct' mathematical definition should be a connection on a gauge bundle. This probably takes about two lecture courses to explain [68], so I will just leave it for a later date or a later note.

There is way too much to do and learn but you will survive. The list of things you would want to learn just grows. I, for now, should stop typing this and return to actual research to learn more...

[66] But, who knows, maybe I will have something on the side to supplement this supplementary note.

[67] If you are interested, the best sets of notes I have found (on top of canonical QFT textbooks) are [52, 53].

[68] One on the physics notion and how you encounter and deal with them in QFTs, one on the mathematical formulation (connection theory) and then one to wrap everything together.

A Module Theory

Here are some more relevant results from elementary module theory.

Definition A.1. An R-module M is **finitely-generated** if every element of M can be written as an $R0$-linear combination of elements of some finite subset of M.

Note however that for modules, the minimal generating sets for a given module may have different numbers of elements. For the following fix R a ring and M a module.

Definition A.2. Let $N \subset M$. N is a **submodule** of M if $rn \in N$ for all $r \in R$ and $n \in N$.

Every module has at least two submodules, itself R and the zero module $\{0\}$. These are sometimes called **trivial submodules**.

Definition A.3. A module having no non-trivial submodules is called **simple**.

Definition A.4. The quotient R-module M/N is defined with the structure $r(m+N) = rm + N$.

We can of course also compose modules.

Definition A.5. The **direct sum** of two modules $N_1 \oplus N_2$ is the R-module

$$N_1 + N_2 = \{x + y \,|\, x \in N_1, y \in N_2\} \ , \tag{A.1}$$

with $N_1 \cap N_2 = 0$. The **direct product** is the R-module with structure,

$$r(m,n) = (rm, rn) \ . \tag{A.2}$$

Definition A.6. A **composition series** of an R-module M is a descending series of submodules of M which terminates in the zero submodule in which each successive quotient is a simple module.

Definition A.7. Suppose M and N are R-modules and $\varphi : M \to N$ is a group homomorphism. Then φ is an **R-module homomorphism** if $\varphi(rm) = r\varphi(m)$. The **kernel** of φ, $\ker \phi \subset M$ is the set of elements of M mapped under φ to the additive identity of N, the **image of** φ is a submodule of N, $\operatorname{im} \varphi \subset N$. The **cokernel** is a submodule of N defined as $\operatorname{coker}(\varphi) = N/\operatorname{im}\varphi$.

With the above definitions we can restate Schur's Lemma in module form.

Lemma A.1 (Schur's Lemma)**.** *Any non-zero homomorphism between simple R-modules is an isomorphism.*

Proof. Let M and N be simple R-modules and $\varphi : M \to N$ a R-module homomorphism. Since $\ker \varphi$ is a submodule of M and $\operatorname{im} \varphi$ is a submodule of N, if φ is a non-zero map, then $\ker \varphi = 0$ and $\operatorname{im} \varphi = N$ so φ is an isomorphism. □

Definition A.8. The **homomorphism**-module $\mathrm{Hom}_R(M, N)$ is the set of all R-module homomorphisms from M to N.

Let us define two things from ring theory for our purposes.

Definition A.9. The **opposite ring** R^{op} is the abelian group R with the multiplication rule of R reversed, i.e. $a \cdot b = ba$ for all $a, b \in R$.

Definition A.10. A module M is an (R, S)**-bimodule** if we have,

$$r(ms) = (rm)s \ , \tag{A.3}$$

for every $r \in R$, $s \in S$ and $m \in M$.

Tensor products are defined formally using bimodules.

Definition A.11. Let M be an (R, S)-bimodule, U be an R-module and N be an S-module. A set map $f : M \times N \to U$ is **balanced** if for $m_i \in M$, $n_i \in N$, $r \in R$ and $s \in S$:

- $f(m_1 + m_2, n) = f(m_1, n) + f(m_2, n)$.
- $f(m, n_1 + n_1) = f(m, n_1) + f(m, n_2)$.
- $f(ms, n) = f(m, sn)$.
- $f(rm, n) = rf(m, n)$.

A **tensor product over** S **of** M **and** N, $M \otimes_S N$ is one equipped with a balance map, $\eta : M \times N \to M \otimes_S N$, with the following universal property. If U is an R-module, and $f : M \times N \to U$ is a balanced map, then there is a unique R-module homomorphism $\alpha : M \otimes_S N \to U$ such that $f = \alpha \circ \eta$.

It can seen that we have the following proposition.

Proposition A.1. *Let R be a ring with unit and let M be an R-module. Then M and $R \otimes_R M$ are isomorphic R-modules.*

Proof. The map $f : R \times M \to M$ is balanced and induces an R-module homomorphism, $\alpha : R \otimes_R M \to M$ from $f(r, m) = rm$ to $\alpha(r \otimes m) = rm$. α is invertible with $\alpha^{-1} : m \mapsto 1 \otimes m$. □

A.1 Semi-simple and simple modules

Recall that simple modules are the modules with only the trivial submodules, namely itself and the zero module $\{0\}$. These are the building-blocks of more complicated module. In particular, similar to Lie algebras, we can define,

Definition A.12. A **semisimple** module is module that is a direct sum of simple modules.

Then we have the following statement.

Lemma A.2. *The following statements are equivalent about an A-module M.*

1. *Any submodules M is a direct summand of M.*
2. *M is semisimple.*
3. *M is a sum of simple submodules.*

Proof. The only non-trivial statement is implying (1) from (3). To see this, let $N \subset M$ be a submodule and V a submodule that is maximal among all submodules of M that intersects M trivially. Suppose there is some simple submodule S of M that is not contained in $N+V$. Since $S \cap (N+V) = 0$ by simplicity of S, we can let $n \in N \cap (V+S)$, but $s = n - v = 0$ implies $n = v$ which means $N \cap (V+S)$ and it contradicts the maximality of V. So $M = N \oplus V$ since $N \cap V = 0$. □

Lemma A.3. *Submodules and quotient modules of semisimple modules are semisimple.*

Proof. Use the property of semi-simpleness of the components of the module M, $M = \sum_i S_i$, with $M/N = \eta(M) = \sum_i \eta(S_i)$ and each S_i is isomorphic with a quotient module of S_i and is therefore zero or simple. □

Definition A.13. The algebra A is **semisimple** if all non-zero A-modules are semisimple.

Now we have the following basic results on semisimple algebras.

Lemma A.4. *The algebra A is semisimple iff the A-module A is semisimple.*

Proof. Suppose M is an A-module generated by $\{m_1, \dots, m_r\}$. Let A^r denote the direct sum of r copies of A. Define a map $f : A^r \to M$, $(a_1, \dots, a_r) \mapsto a_1 m_1 + \cdots + a_r m_r$ which is an A-module epimorphism. So M is isomorphic with a quotient module of the semisimple module A^r so it is therefore semisimple by Lemma A.3. The converse is trivial. □

Proposition A.2. *Let A be a semisimple algebra and suppose as A-modules $A \cong S_1 \oplus \cdots \oplus S_r$ with S_i simple submodules of A. Then any simple A-module is isomorphic with some S_i.*

Proof. Define an A-module homomorphism $\varphi : A \to S$ where S is a simple submodule of A by $\varphi(a) = as$ for some non-zero $s \in S$. φ is surjective since S is simple. For each i, we let $\varphi_i : S_i \to S$ be the restriction of $\varphi|_{S_i}$. Then φ_i must be non-zero for some i so $\varphi_i : S_i \to S$ must be an isomorphism by Schur's Lemma. □

Proposition A.3. *Suppose that A is a semisimple algebra, and let $S_1, \dots, S_r$ is a collection of simple A-modules such that every simple A-module is isomorphic with exactly one S_i. Suppose M is an A-module, and write $M \cong \bigoplus_i n_i S_i$ for some non-negative integers n_i. Then the n_i are uniquely defined.*

Proof. There is a composition series of $\bigoplus_i n_i S_i$ having $n_1 + \cdots + n_r$ terms, in which each S_i appears n_i times as a composition factor. The result then follows from the Jordan-Hölder theorem for modules. □

The Jordan-Hölder theorem for modules states the following. If there are two composition series for the same module M, then the two composition series of M must be of the same length and are equivalent.

A.2 Direct sums

Here are some additional facts about direct sums of modules which might be useful in the main text. In this section we fix A_i, $i = 1, \ldots, r$ to be algebras.

Definition A.14. The **external direct sum** of the A_i is the algebra A whose underlying set is the Cartesian product of the A_i with addition, multiplication and scalar multiplication operations defined component-wise. We write,

$$A = A_1 \oplus \cdots \oplus A_r \; . \tag{A.4}$$

For M an A_i-module, we can define an A-module as simply acting the relevant component a_i on $m \in M$. The simpleness of M as an A_i-module then translates to an A-module. Clearly, if B is an algebra with ideals B_i and $B = \bigoplus_i B_i$ as vector spaces. Then B is isomorphic with the external direct sum $\bigoplus_i B_i$ by the map,

$$b = b_1 + \cdots + b_r \mapsto (b_1, \ldots, b_r) \; . \tag{A.5}$$

Definition A.15. The B defined above is the **internal direct sum** as algebras of the B_i.

Note that since B_i and B_j are ideals, we have,

$$(b_1 + \cdots + b_r) \cdot (b_1' + \cdots + b_r') = b_1 b_1' + \cdots + b_r b_r' \; . \tag{A.6}$$

Lemma A.5. *Let $B = B_1 \oplus \cdots \oplus B_n$ be a direct sum of algebras. Then the two-sided ideals of B are exactly the sets of the form $J_1 \oplus \cdots \oplus J_n$, where J_i is a two-sided ideal of B_i for each i.*

Proof. Let J be a two-sided ideal of B and let $J_i = J \cap B_i$ for each i, clearly, $\bigoplus_{i=1}^n J_i \subset J$. Let $b \in J$ then $b = b_1 + \cdots + b_n$ for $b_i \in B_i$ for each i. We fix some i and let e_i be the element of B with the only non-zero entry the identity of B_i. Then $b_i = be_i \in J \cap B_i = J_i$. Therefore, $b \in \bigoplus_{i=1}^n J_i$. The converse is straightforward. □

Acknowledgements

I sincerely thank my fellow tutors and friends Kit Fraser-Taliente, Thomas Harvey, Riccardo Senese and Ethan Carragher[69] for their helpful suggestions and corrections! Any corrections and suggestions are welcomed - please send them to lucas.leung@physics.ox.ac.uk.

[69] Kit, Thomas, Riccardo and Ethan were tutors in previous years.

References

[1] W. Fulton and J. Harris, *Representation Theory.* Springer Science and Business Media, 12 2013.

[2] D. S. Dummit and R. M. Foote, *Abstract algebra.* John Wiley and Sons, 2004.

[3] G. James and M. Liebeck, *Representations and Characters of Groups.* Cambridge University Press, 2001.

[4] J. Humphreys, *Introduction to Lie Algebras and Representation Theory.* Graduate texts in mathematics, Springer, 1972.

[5] K. Erdmann and M. J. Wildon, *Introduction to Lie algebras.* Springer, 2006.

[6] B. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction.* Graduate Texts in Mathematics, Springer International Publishing, 2015.

[7] J. M. Lee, *Introduction to Smooth Manifolds.* Springer Verlag, Cop, 2000.

[8] H. Georgi, *Lie Algebras In Particle Physics.* CRC Press, 05 2018.

[9] J. Fuchs and C. Schweigert, *Symmetries, Lie algebras and representations : a graduate course for physicists.* Cambridge University Press, 2003.

[10] G. Costa and G. Fogli, *Symmetries and group theory in particle physics : an introduction to space-time and internal symmetries.* Springer, 2012.

[11] P. J. Cameron, *Introduction to algebra.* Oxford University Press, 2008.

[12] W. Fulton, *Young tableaux : with applications to representation theory and geometry.* Cambridge Univ. Press, 2003.

[13] A. Hatcher, *Algebraic topology.* Cambridge: Cambridge University Press, 2002.

[14] J. J. Lee, *Introduction to topological manifolds.* Springer Verlag, Cop, 2011.

[15] L. Nachbin, "Chapter 3: Introduction to projective representations," in *Continuation of the Notas de Matemàtica*, vol. 177 of *North-Holland Mathematics Studies*, pp. 105–131, North-Holland, 1993.

[16] B. C. Hall, *Quantum Theory for Mathematicians.* No. 267 in Graduate Texts in Mathematics, Springer New York, 2013.

[17] D. Z. Freedman and A. Van Proeyen, *Supergravity.* Cambridge, UK: Cambridge Univ. Press, 5 2012.

[18] M. Rausch de Traubenberg, "Clifford algebras in physics," in *7th International Conference on Clifford Algebras and their Applications in Mathematical Physics*, 6 2005.

[19] J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond.* Cambridge Monographs on Mathematical Physics, Cambridge University Press, 12 2007.

[20] H. B. Lawson and M. L. Michelsohn, *Spin geometry.* 1998.

[21] D. J. H. Garling, *Clifford algebras: An introduction.* 2011.

[22] J. Figueroa-O'Farrill, "Lectures on spin geometry," 2010.

[23] C. Bär, "Lectures on spin geometry," 2011.

[24] J. Figueroa-O'Farrill, "Majorana spinors," 2010.

[25] R. S. Hamilton, "The inverse function theorem of Nash and Moser," *Bull. Amer. Math. Soc.*, vol. 7, pp. 65–222, 1982.

[26] J. Dixmier, "L'application exponentielle dans les groupes de lie résolubles," *Bulletin de la Société Mathématique de France*, vol. 85, pp. 113–121, 1957.

[27] M. Saitô, "Sur certains groupes de Lie résolubles," *Scientific Papers of the College of General Education, University of Tokyo*, vol. 7, no. 1, pp. 1–11, 1957.

[28] J. Serre, *Lie Algebras and Lie Groups: 1964 Lectures Given at Harvard University*. No. no. 1500 in Lecture Notes in Mathematics, Springer, 1992.

[29] T. Tao, "Ado's theorem," 2011.

[30] J. Ebert, "Van est's exposition of cartan's proof of lie's third theorem."

[31] J. Duistermaat and J. Kolk, *Lie Groups*. Universitext, Springer Berlin Heidelberg, 1999.

[32] N. Bourbaki, *Lie Groups and Lie Algebras: Chapters 1-3*. Bourbaki, Nicolas: Elements of mathematics, Springer, 1989.

[33] M. M. R. Moreira, "On lie's third theorem."

[34] R. L. Fernandes and M. Crainic, "Lectures on integrability of lie brackets," 2006.

[35] M. Hamermesh, *Group theory and its application to physical problems*. New York, NY: Dover, 1989.

[36] A. Knapp, *Lie Groups Beyond an Introduction*. Progress in Mathematics, Birkhäuser Boston, 2002.

[37] M. M. Alexandrino and R. G. Bettiol, "Introduction to lie groups, isometric and adjoint actions and some generalizations," 2010.

[38] J. Gutowski, "Part ii: Classification of semi-simple lie algebras.," 2011.

[39] R. Slansky, "Group Theory for Unified Model Building," *Phys. Rept.*, vol. 79, pp. 1–128, 1981.

[40] E. B. Dynkin, "Semisimple subalgebras of semisimple Lie algebras," *Trans. Am. Math. Soc. Ser. 2*, vol. 6, pp. 111–244, 1957.

[41] R. Cahn, *Semi-Simple Lie Algebras and Their Representations*. Dover Publications, 2014.

[42] S. Weinberg, *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge University Press, 8 2013.

[43] L. E. Ibáñez and A. Uranga, *String theory and particle physics : an introduction to string phenomenology*. Cambridge University Press, 2012.

[44] L. B. Anderson, J. Gray, A. Lukas, and E. Palti, "Heterotic Line Bundle Standard Models," *JHEP*, vol. 06, p. 113, 2012.

[45] A. J. Mountain, "Invariant tensors and casimir operators for simple compact lie groups," *Journal of Mathematical Physics*, vol. 39, p. 5601–5607, Oct. 1998.

[46] P. Langacker, *The Standard Model and Beyond*. Taylor & Francis, 2017.

[47] C. P. Burgess and G. D. Moore, *The standard model: A primer*. Cambridge University Press, 12 2006.

[48] F. Quevedo and A. Schachner, "Cambridge Lectures on The Standard Model," 9 2024.

[49] S. Raby, *Supersymmetric Grand Unified Theories.* Cham Springer International Publishing, 2017.

[50] H. Georgi and C. Jarlskog, “A New Lepton - Quark Mass Relation in a Unified Theory,” *Phys. Lett. B*, vol. 86, pp. 297–300, 1979.

[51] H. Georgi and D. V. Nanopoulos, “Masses and Mixing in Unified Theories,” *Nucl. Phys. B*, vol. 159, pp. 16–28, 1979.

[52] T. J. Hollowood, “6 Lectures on QFT, RG and SUSY,” in *38th British Universities Summer School in Theoretical Elementary Particle Physics*, 9 2009.

[53] J. a. F. Melo, “Introduction to Renormalisation,” 9 2019.