1. Introduction

1.1. A couple of examples. As a preview, we write down two non-trivial examples by hand: the first is a representation of a finite group and the second, a representation of a Lie algebra. Some of the terms will be defined later. For now, we define only a few terms in ad-hoc manner. Below $k$ denotes a field.

Groups (resp. $k$-algebras) often appear as automorphisms (resp. endomorphisms) of objects. For example:

- If $X$ is a set, then the set of all bijections of $X$, denoted $\text{Aut}(X)$ is a group.
- If $V$ is a $k$-vector space, then the set of $k$-linear automorphisms $\text{Aut}(V)$ is a group.
- If $V$ is a $k$-vector space, then the set of all $k$-linear maps from $V$ to $V$, denoted $\text{End}_k(V)$, is a $k$-algebra.

Let $G$ be a group. A permutation representation (resp. a linear representation) of $G$ is a group homomorphism from $G$ to $\text{Aut}(X)$ (resp. $\text{Aut}(V)$) for some set $X$ (resp. vector space $V$). A permutation representation $G \to \text{Aut}(X)$ is the same as an action of $G$ on $X$. A linear representation $G \to \text{Aut}(V)$ is same as a action of $G$ on $V$ where the elements of $G$ act as linear maps.

Let $A$ be a nonzero $k$-algebra. For now, this just means $A$ is a ring containing $k$ in its center. A linear representation of $A$ is a $k$-algebra homomorphism $A \to \text{End}_k(V)$. Think of the elements of $A$ as acting on $V$ via linear maps.

1.1.1. A representation of $A_5$: We describe a “three dimensional irreducible representation” of the smallest non-abelian simple group $A_5$: the group of even permutations of 5 letters. Let $I$ denote an icosahedron in $\mathbb{R}^3$ with center at the origin.

Exercise: Show that the faces of the icosahedron can be colored with five colors $\{1, 2, 3, 4, 5\}$ such that for each vertex $v$, the five faces meeting at $v$ gets five distinct colors.

A consequence of this exercise is the following: Let $G$ be the set of rotations of $\mathbb{R}^3$ that takes $I$ to $I$. One verifies easily that there are sixty such rotations (including the identity). One verifies that each rotation permutes the five colors. The sixty rotation corresponds bijectively to the sixty even permutations of $\{1, 2, 3, 4, 5\}$. This way we get a bijection $G \to A_5$. Let $\rho : A_5 \to G$ be the inverse of this map. Composing rotations correspond to multiplying the permutations, so $\rho^{-1}$ is a group homomorphism; hence so is $\rho$. Since $G \subseteq SO(3) \subseteq GL_3(\mathbb{R})$, we get a three dimensional representation $\rho : A_5 \to GL_3(\mathbb{R})$. Each even permutation acts on $\mathbb{R}^3$ as a rotation that preserves $I$.

1.1.2. Representations of $\mathfrak{sl}_2(\mathbb{C})$: A complex Lie algebra is a complex vector space with a skew symmetric multiplication denoted $[X, Y]$ (called the Lie bracket) satisfying the “Jacobi identity”. Verify that the set of $2 \times 2$ trace zero complex matrices, denoted $\mathfrak{sl}_2(\mathbb{C})$ is a Lie algebra with the bracket $[X, Y] = XY - YX$. The Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ is three dimensional with basis

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and the multiplication rule is determined by

$$[E, F] = H, \quad [H, E] = 2E, \quad [H, F] = -2F. \quad (1)$$
Let $V$ be the vector space of all homogeneous complex polynomials in two variables $x$ and $y$. Define the linear differential operators $e, f, h \in \text{End}(V)$ by

$$e = -y \partial_x, \quad f = -x \partial_y, \quad h = y \partial_y - x \partial_x.$$ 

By direct computation, one verifies that

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f.$$ 

(2)

Here $[u, v]$ means $(uv - vu)$. It follows from this that $(E, F, H) \mapsto (e, f, h)$ defines a linear map from $\mathfrak{sl}_2(\mathbb{C})$ to $\text{End}(V)$ that preserves the brackets, i.e. is a homomorphism of Lie algebras. We say that the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ acts on $V$ as differential operators: $E|_V = e$, $F|_V = f$ and $H|_V = h$. Write $V = \bigoplus_{k=0}^{\infty} V_k$ where $V_k$ is the $(k + 1)$-dimensional subspace consisting of homogeneous polynomials of degree $k$. The action of $\mathfrak{sl}_2(\mathbb{C})$ preserves each $V_k$. So each $V_k$ is a subrepresentation. The action of $h$ on $V_k$ is diagonalizable with a basis of eigenvectors being the standard basis vectors $\{x^k, x^{k-1}y, x^{k-2}y^2, \ldots, xy^{k-1}, y^k\}$. The following picture shows the action of $e$ and $f$ on the one dimensional subspaces spanned by these eigenvectors:

Using this information, one can argue that $V_k$ does not have any proper subspace that is stable under the action of $\mathfrak{sl}_2(\mathbb{C})$, so each $V_k$ is an irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$. The representation $V$ thus decomposes into irreducible components as $V = \bigoplus_k V_k$. 

2
2. Some background material

Reading assignment: Please read up the very basics of category theory as discussed in class. Basic familiarity with categories, functors and natural transformations and some first examples of these will be enough. For us, it is a very convenient language and we will use it freely.

2.1. A summary of some basics on modules. We recall a few facts about modules. The main purpose is to set up notation. This is not meant to be an introduction to modules. The only thing dealt with in some detail is the example of group algebra.

2.1.1. Conventions. We reserve the letter $G$ for a group, $R$ for a ring and $F$ for a field. All rings appearing in these notes have 1 and ring homomorphisms take 1 to 1 unless otherwise stated. All group actions are on the left unless otherwise stated. An $R$-module means a left $R$-module unless otherwise stated. Let $M$ be an abelian group. Note that the set of all abelian group homomorphisms from $M$ to $M$, denoted by $\text{End}(M)$, is an ring under pointwise addition and composition of functions with identity function being the unit element.

2.1.2. Definition. A left $R$-module is an abelian group $M$ together with a ring homomorphism $\rho : R \rightarrow \text{End}(M)$. Usually we write $\rho(r)m = r \cdot m = rm$ for $r \in R$ and $m \in M$. If $M$ and $N$ are left $R$-modules, then a $R$-module homomorphism $f : M \rightarrow N$ is an abelian group homomorphism such that $f(rm) = rf(m)$ for all $r \in R$ and $m \in M$. The category of all left $R$-modules is denoted by $R\text{-Mod}$. Think of $rm$ as the result of $r$ acting on $m$. So a left $R$-module gives a left action of $R$ on an abelian group by abelian group homomorphisms. Similarly one defines the category of all right $R$-modules $\text{Mod-R}$ when the ring acts on the right. Unless otherwise stated, an $R$-module will mean a left $R$ module. If $R$ is commutative, then of course there is no difference between left and right modules.

Let $M$ be a left $R$-module and $N \subseteq M$. We say $N$ is a submodule of $M$ if $N$ is a subgroup of $M$ and $rn \subseteq N$ for all $r \in R$ and $n \in N$. If $f : M \rightarrow N$ is a $R$-module homomorphism, then verify that $\text{ker}(f)$ is a submodule of $M$ and $\text{im}(f)$ is a submodule of $N$. Here the kernel and image have the same meaning as for abelian groups.

2.1.3. First Examples: A module over $\mathbb{Z}$ is simply an abelian group. A module over a field $F$ is nothing but a $F$-vector space. A ring $R$ is a left module over itself and a submodule is nothing but a left-ideal. Similarly for right ideals. A left module for a group algebra $k[G]$ (defined below) is nothing but a $k$-representation of $G$. these examples show that the notion of modules encompass and unify many important concepts. Further, $R\text{-Mod}$ is a very rich category and often study of this category is an useful way to understand $R$.

The next example is specially important for us.

2.1.4. The group algebra. The group ring (or group algebra) $R[G]$ is the set of all functions from $G$ to $R$ that are non-zero only at finitely many elements of $G$. Note that $R[G]$ is a left $R$-module with pointwise multiplication by $R$: if $f \in R[G]$, then $(rf)(g) = rf(g)$. Let
$e_g : G \to R$ be the function defined by $e_g(x) = 1$ if $x = g$ and $e_g(x) = 0$ if $x \neq g$. Let $a \in R[G]$. Then note that
\[ a = \sum_{g \in G} a_ge_g \]
and the sum on the right hand side is actually finite. Sometimes we abuse notation and write
\[ a = \sum_{g \in G} a_g. \]
So we think of elements of $R[G]$ as “free finite $R$-linear combinations of the elements of $G$”. This also identifies $G$ as a subset of $R[G]$. Define a multiplication on $R[G]$ by
\[ \left( \sum_{g \in G} a_g \right) \left( \sum_{g' \in G} b_{g'}g' \right) = \sum_{g,g' \in G} a_gb_{g'}gg' = \sum_{h \in G} \left( \sum_{g \in G} a_gb_{g^{-1}h} \right)h \]
Verify that all the sums are finite. This makes $R[G]$ into a unital ring with the unit id$_G$.

2.1.5. **Exercise.** Verify that the category of $k[G]$ modules is isomorphic to the category of $k$-representations of $G$.

Next we describe some methods of constructing new modules from old ones. You are surely familiar with these notions for abelian groups or vector spaces. The generalizations to modules are obvious and immediate.

2.1.6. **direct sum and direct product.** Let $\{M_i : i \in I\}$ be a family of $R$-modules indexed by a set $I$. The direct product $\prod_{i \in I} M_i$ consists of all tuples $(m_i)_{i \in I}$ such that $m_i \in M_i$, with addition and $R$-action defined componentwise. The direct sum $\bigoplus_{i \in I} M_i$ is the $R$-submodule of $\prod_{i \in I} M_i$ consisting of all $(m_i)_{i \in I}$ such that all but finitely many $m_i$ are equal to 0.

The next Exercise specifies the “universal properties” of direct sum and direct product.

2.1.7. **Exercise.** Write $M = \prod_{i \in I} M_i$ and let $\pi_i : M \to M_i$ be the natural projections.

(a) Show that $M$ satisfies the following property: *Given any $R$-module $N$ and morphisms $\{f_i : N \to M_i : i \in I\}$, there exists a unique morphism $f : N \to M$ such that $\pi_i \circ f = f_i$ for all $i$. (When we are talking about the category of $R$-modules, a morphism means a module homomorphism.)*

(b) Show that any module $P$ satisfies the property in part (a) is isomorphic to $\prod_i M_i$.

(c) State and prove the analogous exercise for $\bigoplus_{i \in I} M_i$. (Hint: Reverse all the arrows in part (a)).

2.1.8. **Generators.** Let $M$ be an $R$-module. Verify that the intersection of an arbitrary family of submodules of $M$ is again a submodule of $M$. Now, let $S$ be a subset of $M$. Let $N$ be the intersection of all submodules of $M$ that contain $S$. Verify that $N$ is the smallest submodule of $M$ that contains $S$. We say that $N$ is generated by $S$ or that $S$ is a set of generators of $N$. Verify that explicitly $N$ consists of all elements of the form
\[ r_1s_1 + \cdots + r_ks_k \]
where $r_j \in R$ and $s_j \in S$. We write $N = \sum_{s \in S} Rs$. If $S = \{s_1, \ldots, s_n\}$ then we write
\[ N = Rs_1 + Rs_2 + \cdots + Rs_n. \]
A cyclic module means a module generated by a single element.
2.1.9. **Sums and direct sums.** Let \( \{N_j : j \in J\} \) be a collection of submodules of \( M \). Then \( \sum_j N_j \) means the submodule of \( M \) generated by \( \bigcup_{j \in J} N_j \). Verify that explicitly \( \sum_j N_j \) consists of all finite sums of elements from the \( N_j \)'s. We say that \( M \) is the internal direct sum of the \( N_j \)'s if \( M = \sum_j N_j \) and whenever \( n_1 \in N_{j_1}, n_2 \in N_{j_2}, \ldots, n_k \in N_{j_k} \) for distinct \( j_1, \ldots, j_k \in J \) such that \( n_1 + \cdots + n_k = 0 \), then \( n_1 = \cdots = n_k = 0 \).

Now let \( \{M_j : j \in J\} \) be a family of \( R \)-modules. Let \( M = \bigoplus_{j \in J} M_j \). Verify that one can define inclusions \( i_j : M_j \to M \) in the obvious manner and that \( M \) is the internal direct sum of \( \{i_j(M_j) : j \in J\} \). We shall often identify \( M_j \) with \( i_j(M_j) \) and think of the \( M_j \)'s as submodule of \( \bigoplus_{j \in J} M_j \).

2.1.10. **Free module.** Let \( J \) be an arbitrary set. Then the free module \( M = \bigoplus_{j \in J} R \) generated by \( J \) is the direct sum of \( J \) copies of \( R \). Let \( i_j : R \to M \) be the inclusion of the \( j \)-th copy of \( R \) into \( M \). Write \( e_j = i_j(1_R) \). Then elements of \( M \) can be uniquely expressed in the form \( \sum_{j \in J} r_j e_j \) where \( r_j \in R \) and only finitely many \( r_j \) are nonzero. We say \( \{e_j : j \in J\} \) is a basis of the free module \( M \). Sometimes we think of \( e_j \)'s as a set of formal symbols indexed by the elements of \( J \) and the elements of the free module \( M \) are free and finite \( R \)-linear combinations of the \( e_j \)'s.

2.1.11. **Definition/Exercise.** Let \( M \) be an \( R \)-module. A subset \( S \) of \( M \) is called a basis of \( M \) if every element of \( M \) can be written uniquely as a finite \( R \)-linear combination of elements of \( S \). Verify that the following are equivalent:

1. \( S \) is a basis of \( M \).
2. \( M \) is the internal direct sum of the cyclic modules \( \{Rs : s \in S\} \).
3. \( M \) is isomorphic to the free \( R \)-module on the set \( S \).

2.1.12. **Definition** (quotient module). Let \( N \) be a submodule of an \( R \)-module \( M \). Verify that the abelian group \( M/N \) can be given the structure of \( R \)-module in obvious manner by defining

\[
r(m + N) = rm + N
\]

for \( r \in R \), \( m \in M \). This defines the quotient module \( M/N \). It comes with a natural projection \( \pi : M \to M/N \). Verify that \( \pi \) is a \( R \)-module homomorphism.

The isomorphism theorems for modules have the same statement as the isomorphism theorems for abelian groups. The same proofs go through with a little extra verification.

2.1.13. **Exercise.** [Isomorphism theorems] (a) Any module homomorphism \( f : M \to N \) induces an isomorphism \( \bar{f} : M/\ker(f) \to \text{im}(f) \) such that \( f = \bar{f} \circ \pi \).

(b) Let \( N_1, N_2 \) be submodules of \( M \). Then \( (N_1 + N_2)/N_2 \cong N_1/(N_1 \cap N_2) \). (Hint: Consider the maps \( N_1 \to N_1 + N_2 \to (N_1 + N_2)/N_2 \) where the first map is inclusion and the second is the natural projection. Verify that the kernel of the composition is \( N_1 \cap N_2 \). Apply part (a)).

(c) Let \( P \subseteq N \) be submodules of \( M \). Then \( (M/P)/(N/P) \cong (M/N) \).

2.1.14. **Exercise** Let \( G \) be a finite group and let \( A \) be a commutative ring. Let \( CC(G) \) be the set of conjugacy classes in \( G \). For each conjugacy class \( O \) in \( G \), define the element \( e_O \in A[G] \) by \( e_O = \sum_{g \in O} e_g \). Show that the center of \( A[G] \) is equal to \( \sum_{O \in CC(G)} Re_O \) (this is a direct sum).

Next time we will talk about tensor products of modules. Before that we want to take a moment to talk about universal properties.
2.1.15. **Definition** (initial object). An initial object in a category $C$ is an object $I$ such that for every object $X \in \text{ob}(C)$, there exists precisely one morphism $u_{I,X} : I \to X$. A terminal object in $C$ is an initial object in the opposite category $C^{op}$.

2.1.16. **Lemma.** An initial object of $C$, if it exists, is unique up to an unique isomorphism. More precisely, if $I$ and $I'$ are two initial objects in $C$, then there exists an unique isomorphism $u : I \to I'$ satisfying $u_{I',X} \circ u = u_{I,X}$ for all objects $X$ in $C$.

**Proof.** The uniqueness of $u$, if it exists, is clear since $I$ is an initial object. Also, since $I$ (resp. $I'$) is an initial object, there is an unique morphism $u : I \to I'$ (resp. $u' : I' \to I$). Now $u' \circ u$ is a morphism from $I$ to $I$. But so is $\text{id}_I : I \to I$. Since there is precisely one morphism from $I$ to $I$ we must have $u' \circ u = \text{id}_I$. For similar reasons, $u \circ u' = \text{id}_{I'}$. So $u$ and $u'$ are mutually inverse isomorphisms. Given $X \in \text{ob}(C)$, both $u_{I',X} \circ u$ and $u_{I,X}$ are morphisms from $I$ to $X$ and since $I$ is an initial object, these two morphisms must be equal. $\square$

Often it is convenient to define objects in abstract algebra by their universal mapping properties. This usually means that the objects being defined are initial object in an appropriate category. A consequence of this is that objects defined by universal property are always unique up to an unique isomorphism in the sense of the above lemma. As an example, we could define direct product of modules by universal property as follows:

The direct product of $\{M_i : i \in I\}$ is a module $M$ together with a family of morphisms $\{\pi_i : M \to M_i : i \in I\}$ satisfying the following property: Given any module $N$ and morphisms $\{f_i : N \to M_i : i \in I\}$, there exists a unique morphism $f : N \to M$ such that $f_i = \pi_i \circ f$ for all $i \in I$.

To realize the direct product $M$ as an initial object, we define a category $C$ as follows: An object of $C$ is a family of module homomorphisms $\{f_i : N \to M_i : i \in I\}$ for some module $N$. A morphism in $C$ from $\{f_i : N \to M_i : i \in I\}$ to $\{f'_i : N' \to M_i : i \in I\}$ is a module homomorphism $f : N \to N'$ such that $f_i = f'_i \circ f$ for all $i \in I$. Verify that this indeed defines a category $C$ and the above definition of direct product by universal property is simply saying that the direct product $M$ (or more precisely, the family $\{\pi_i : M \to M_i : i \in I\}$) is an initial object of $C$. 
A detailed proof of Chapter 2, Proposition 1.2 in the TIFR lecture notes

First let us recall a couple of definitions. Let $A$ be a ring. Let $M$ be an $A$-module. Let $\{M_i\}_{i \in I}$ be a family of submodules of $M$.

**Definition.** We say that $M$ is a sum of $\{M_i\}_{i \in I}$, denoted $M = \sum_{i \in I} M_i$, if each $x \in M$ can be written as a finite sum $x = x_1 + x_2 + \cdots + x_k$ where $x_1 \in M_{i_1}, x_2 \in M_{i_2}, \cdots, x_k \in M_{i_k}$, for some $i_1, \ldots, i_k \in I$. We say that $M$ is a direct sum of $\{M_i\}_{i \in I}$, denoted $M = \bigoplus_{i \in I} M_i$, if:

- $M$ is the sum of $\{M_i\}_{i \in I}$, and
- if $x_1 + x_2 + \cdots + x_k = 0$ for $x_1 \in M_{i_1}, x_2 \in M_{i_2}, \cdots, x_k \in M_{i_k}$ and for distinct elements $i_1, \ldots, i_k \in I$, then $x_1 = x_2 = \cdots = x_k = 0$.

**Exercise.** Show that $M = \bigoplus_{i \in I} M_i$ if and only if each $x \in M$ can be written uniquely as a finite sum $x = x_1 + x_2 + \cdots + x_k$ where $x_1 \in M_{i_1}, x_2 \in M_{i_2}, \cdots, x_k \in M_{i_k}$ for some $i_1, \ldots, i_k \in I$.

**Proposition 1.2.** [See page 14 of TIFR lecture notes] Let $N$ be a submodule of $M$ and let $\{S_i\}_{i \in I}$ be a collection of simple submodules of $M$ such that $M = N + (\sum_{i \in I} S_i)$. Then there is some subset $J$ of $I$ such that $M = N \oplus (\bigoplus_{j \in J} S_j)$.

**Proof.** Let $\mathcal{S}$ be the set of all subsets $I' \subseteq I$ such that the sum $(N + \sum_{i \in I'} S_i)$ is actually a direct sum. Note that $\mathcal{S}$ is nonempty since $\emptyset \in \mathcal{S}$. Order $\mathcal{S}$ by inclusion of subsets. We want to show that any ascending chain in $\mathcal{S}$ has an upper bound in $\mathcal{S}$. To this end, let $\{I_\lambda\}_{\lambda \in \Lambda}$ be a totally ordered subset of $\mathcal{S}$. Let $I_* = \bigcup_{\lambda \in \Lambda} I_\lambda$.

**Claim 1.** One has $I_* \in \mathcal{S}$, that is, $(N + \sum_{i \in I_*} S_i)$ is actually a direct sum.

**proof of claim 1.** Suppose $n + x_1 + x_2 + \cdots + x_k = 0$ for some $n \in N$ and some $x_1 \in S_{i_1}, x_2 \in S_{i_2}, \cdots, x_k \in S_{i_k}$ where $i_1, \ldots, i_k \in I_*$. Since $I_* = \bigcup_{\lambda \in \Lambda} I_\lambda$, it follows that $i_1 \in I_{\lambda_1}, \ldots, i_k \in I_{\lambda_k}$ for some $\lambda_1, \ldots, \lambda_k \in \Lambda$. Since $\{I_\lambda\}_{\lambda \in \Lambda}$ is totally ordered, $I_{\lambda_1}, \ldots, I_{\lambda_k}$ are all contained in one of them, say $I_{\lambda_t}$ (for some $1 \leq t \leq k$). So $i_1, \ldots, i_k$ all belong to $I_{\lambda_t}$. Since $I_{\lambda_t} \in \mathcal{S}$, the sum $(N + \sum_{i \in I_{\lambda_t}} S_i)$ is a direct sum. So the equation $n + x_1 + x_2 + \cdots + x_k = 0$ implies that $n = x_1 = x_2 = \cdots = x_k = 0$. This proves claim 1.

By Zorn’s lemma, $\mathcal{S}$ has a maximal element. Choose one and call it $J$. Let $M' = N + \sum_{j \in J} S_j$. Since $J \in \mathcal{S}$, the sum in the previous sentence is actually a direct sum.

**Claim 2.** $M' = M$.

**proof of claim 2.** If $S_i \subseteq M'$ for all $i \in I$, then we would have $M' = M$ (since $N \subseteq M$ and since $M = N + \sum_{i \in I} S_i$). So $S_i \not\subseteq M'$ for some $i \in I - J$. So $S_i \cap M'$ is a proper submodule of $S_i$. So $S_i \cap M' = 0$. But this means that $(N + \sum_{j \in J \cup \{i\}} S_j)$ is actually a direct sum, which implies $J \cup \{i\}$ is an element of $\mathcal{S}$, contradicting the maximality of $J$. This proves claim 2 and the proposition. 


3. Representations of finite group

3.1. Group action.

3.1.1. If $G$ is a group acting on $A$ (always on the left, unless otherwise stated), let

$$A^G = \{ a \in A : ga = a \ \text{for all} \ g \in G \}$$

denote the set of fixed points of the action.

(a) Let $G$ act on $X$ and $Y$. Verify that $G$ acts on $\text{Fun}(X,Y)$ by

$$ (g\phi)(x) = g(\phi(g^{-1}x)) \text{ for } g \in G, x \in X, \phi \in \text{Fun}(X,Y). $$

(b) Let $\text{Fun}_G(X,Y) = \text{Fun}(X,Y)^G$.

Note that $\text{Fun}_G(X,Y)$ consists of functions such that $\phi(gx) = g(\phi(x))$. These are called the $G$-equivariant maps, or $G$-homomorphisms.

(c) Suppose $X$ and $Y$ has some structure, e.g. of vector space or smooth manifold, and $G$ acts preserving these structures, e.g. by linear or smooth maps. In such situations, instead of looking at all functions $\text{Fun}(X,Y)$, we often consider the set of all maps preserving the structures, e.g. all linear maps or all smooth maps, denoted by $\text{Hom}(X,Y)$. Note that the above discussion goes through.

3.2. The representation category.

3.2.1. The center of the group algebra: Let $k$ be a field. Let $\text{Mod}_k$ be the category of $k$–vector spaces and linear maps. Let $G$ be a group. Let $k[G]$ be the group algebra of $G$, with standard basis denoted by $\{ e_g : g \in G \}$. Let $Z(k[G])$ be the center of the group algebra. Let $CC(G)$ be the set of conjugacy classes in $G$. For each conjugacy class $O$, define $e_O = \sum_{g \in O} e_g$. Verify that $e_O \in Z(k[G])$ and $\{ e_O : O \in CC(G) \}$ form a basis for $Z(k[G])$.

3.2.2. Definition (Representations). A linear representation of $G$ (defined over $k$) is a pair $(V, \rho)$ where $V$ is a $k$–vector space and $\rho$ is a homomorphism from $G$ to $\text{Aut}(V)$. Sometimes we denote a representation $(V, \rho)$ simply by $\rho$ in which case we write $V = V_\rho$. Sometimes it is also convenient to denote a representation $(V, \rho)$ simply by $V$; in this case we write $\rho(g) = g|_V \in \text{Aut}_k(V)$. We shall write

$$ \rho(g)v = g|_V v = gv $$

when there is no chance of confusion. A subspace $U$ of $V$ is called a subrepresentation if $gU \subseteq U$ for all $g \in G$.

3.2.3. Definition (Intertwiners). Let $V, V'$ be two representations of a group $G$ over a field $k$. A map $\alpha \in \text{End}(V, V')$ is called a map of $G$-representations if $\alpha \circ g|_V = g|_{V'} \circ \alpha$ for all $g \in G$. Such a map is called an intertwining operator or a $G$–linear map or a $G$–homomorphism. The set of $G$–homomorphisms is denoted by $\text{Hom}_G(V, V')$. So we have a
3.2.5. **Presentation.** Let \( V,W \) be \( G \)-representations. Let 
\[
\rho = \sum_{t \in G} f(t) \rho(t)
\]
(whenever the sum is finite). Then 
\[
\rho(g)^{-1} \rho_f(g)(v) = \sum_{t \in G} f(t) \rho(g^{-1} t g)(v) = \sum_{t \in G} f(gtg^{-1}) \rho(t)(v) = \rho(f)(v), \forall g \in G,
\]
(3) 
So \( \rho_f \) is an intertwining operator. Note that \( \rho_f = \rho(z_f) \) where \( z_f = \sum_{O \in \mathbb{C}C(G)} f(O)e_O \) is an element in \( Z(k[G]) \). So, in other words, if \( z \in Z(k[G]) \), then \( \rho(z) \) is an intertwining operator.

3.2.6. **First Examples:**
- Trivial representation: \( G \) acting on \( k \) by \( g|_k(x) = x \) for all \( g \in G, x \in k \).
- Regular representation: \( k[G] \) as a left \( k[G] \) module.
- Permutation representation: Let \( G \) act on a set \( X \). Then \( G \) acts linearly on \( \text{Fun}(X, k) = k^X \). Each element of \( G \) acts as a permutation matrix.
- A finite subgroup of the group of orthonal transformations of \( \mathbb{R}^2 \) (say dihedral group) or \( \mathbb{R}^3 \) (say symmetries of the cube) acting on \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \).

3.2.7. **Remark.** If \( V \) is a \( G \)-representation, then let \( V^G = \{ v \in V : g|_V v = v \text{ for all } g \in G \} \). Note that 
\[
\text{Hom}_{k[G]}(V, V') = \text{Hom}_k(V, V')^G.
\]

3.2.8. **The fundamental idempotent.** Define 
\[
\sigma_G = \sum_{g \in G} g \in k[G].
\]
Verify that \( g \sigma_G = \sigma_G \) and \( \sigma_G^2 = |G| \sigma_G \). Note \( k \sigma_G \) is a two sided principal ideal in \( k[G] \). Let \( G \) be a \( k \)-representation of \( V \). Verify that we have a map \( V \to V^G \) defined by \( \sigma \mapsto sv \). Now assume \( \text{char}(k) \) does not divide \( |G| \). Then 
\[
e_G = |G|^{-1} \sigma_G \in k[G]
\]
satisfies 
\[
ge_g e_G = e_G \text{ for all } g \in G \text{ and } e_G^2 = e_G.
\]
If \( v \in V^G \), then note that \( e_G v = v \). It follows that if \( \text{char}(k) \) does not divide \( |G| \), then one has a projection
\[
\text{av}_G : V \to V^G \text{ defined by } \text{av}_G(v) = e_G |V^G| v.
\]
In particular, if \( f \in \text{End}_k(V) \), then we can average over \( G \) to obtain a \( G \)-linear map \( \text{av}_G(f) \in \text{End}_k(V) \). Explicitly, one has
\[
\text{av}_G(f) = e_G |\text{End}_k(V)| f = |G|^{-1} \sum_{g \in G} g |\text{End}_k(V)| f = |G|^{-1} \sum_{g \in G} g |V^G| f = g \circ f \circ g^{-1}
\]

3.2.9. **Definition.** A \( G \)-representation is irreducible if it has no proper nonzero sub-representation, i.e. if it is a simple \( k[G] \) module. In other words these are the simple objects in the category \( \text{Rep}(G) \).

3.2.10. **Schur’s lemma:** Let \( V, W \in \text{Irr}(G) \). Then \( \text{Hom}_G(V,W) = 0 \) if \( V \) and \( W \) are not isomorphic. If \( V \cong W \), then each element of \( \text{Hom}_G(V,W) \) is an isomorphism. If \( k \) is algebraically closed, then \( \text{Hom}_G(V,W) \cong k \). More precisely, \( \text{Hom}_G(V,W) = \{ \lambda \cdot \text{id}_V : \lambda \in k \} \).

**Proof.** If \( \alpha \in \text{Hom}_G(V,W) \) then \( \ker(\alpha) \) and \( \text{Image}(\alpha) \) are sub-representations of \( V \) and \( W \). Assume \( \alpha \neq 0 \). Then irreducibility of \( V \) and \( W \) implies that \( \ker(\alpha) = 0 \) and \( \text{Image}(\alpha) = W \), that is \( W \) is an isomorphism. So we may assume \( V = W \) and \( \alpha \in \text{Aut}_k(V) \). Since \( k \) is algebraically closed there exists \( v \in V \) such that \( \alpha(v) = \lambda v \) for some \( \lambda \in k \setminus \{0\} \). Then \( \alpha - \lambda \text{id}_V \) is a \( G \)-map with nonzero kernel, hence \( \alpha - \lambda \text{id}_V = 0 \). \( \square \)

3.2.11. **Exercise** Let \( G \) be a finite group and \( k \) be a field. Consider \( k \) as the trivial \( k[G] \) module. Verify that \( \epsilon : k[G] \to k \) defined by \( \epsilon(\sum_{g \in G} a_g g) = \sum_{g \in G} a_g \) is a \( k[G] \) linear map and hence \( I := \ker(\epsilon) \) is a ideal in \( k[G] \), called the augmentation ideal.

If \( \text{char}(k) \nmid |G| \), then show that \( k[G] = k[\sigma_G] \oplus I \), so \( I \) is a direct summand of \( k[G] \) whose complement is the principal ideal \( k[\sigma_G] \).

On the other hand, if \( \text{char}(k) \mid |G| \), then note that \( \sigma_G \in I \). In this case, show that any proper left ideal of \( k[G] \) intersects \( I \) nontrivially and hence \( I \) is not a direct summand.

3.2.12. **Theorem** (Maschke’s theorem). Let \( G \) be a finite group and \( k \) be a field \( k[G] \) is a semisimple if and only if \( \text{char}(k) \) does not divide the order of \( G \).

**Proof.** Assume \( \text{char}(k) \) does not divide the order of \( |G| \). Let \( J \) be a left ideal in \( k[G] \). In particular, it is a \( k \)-vector subspace. So we have a onto \( k \)-linear map \( p : k[G] \to J \) such that \( p|_J = \text{id}_J \). Averaging over \( G \), we get a \( k[G] \) linear map \( \text{av}_G p : k[G] \to J \). Let \( x \in J \). If \( g \in G \), then \( gx \in J \) since \( J \) is a left ideal. So \( p(gx) = gx \). So
\[
(\text{av}_G p)(x) = |G|^{-1} \sum_{g \in G} g^{-1} p(gx) = |G|^{-1} \sum_{g \in G} g^{-1} gx = x.
\]

In other words, \( \text{av}_G p : k[G] \to J \) is a \( k[G] \) module onto homomorphism such that \( \text{av}_G p|_J = \text{id}_J \). Verify that, it follows that \( k[G] = J \oplus \ker(\text{av}_G p) \). So every left ideal in \( k[G] \) is a direct summand. This proves one implication. The other implication follows from the exercise above. \( \square \)

3.2.13. **Corollary.** Assume \( \text{char}(k) \nmid |G| \). Then any \( k[G] \) module is a direct sum of simple submodules. In other words, a representation of a finite group is a direct sum of irreducible sub-representations.
3.3. Complex representations of finite groups - Characters.

3.3.1. The setup: In this subsection, we assume that $G$ is a finite group and $k = \mathbb{C}$. A function $f : G \to \mathbb{C}$ is called a class function if it is constant on the conjugacy classes of $G$. Let $\text{cl}(G)$ be the set of class functions of $G$. If $G$ has $r$ conjugacy classes, then $\text{cl}(G)$ is a complex vector space of dimension $r$.

3.3.2. Characters: The character $\chi_V$ of a representation $V$ is the class function defined by $\chi_V(g) = \text{Tr}(g|_V)$. Clearly, one has,

$$\chi_{V \oplus V'} = \chi_V + \chi_{V'} \quad \text{and} \quad \chi_{V \otimes V'} = \chi_V \cdot \chi_{V'}.$$

So $V \mapsto \chi_V$ is a ring homomorphism from $K(G)$ to $\text{cl}(V)$. Further, $\chi_{V^*} = \bar{\chi}_V$. \(^1\) So

$$\chi_{\text{Hom}_k(V,V')} = \chi_{V^* \otimes V'} = \bar{\chi}_V \cdot \chi_{V'}.$$

3.3.3. Let $R$ be a ring. Let $\mathcal{C}$ be a full subcategory of $R\text{-mod}$. Assume that the isomorphism classes of objects of $\mathcal{C}$ form a set. Then the Grothendieck group of $\mathcal{C}$, denoted $K(\mathcal{C})$ is defined to be the abelian group with one generator $[V]$ for each isomorphism class of objects of $\mathcal{C}$ and relations relation $[V] = [V'] + [V'']$ whenever there is an short exact sequence $0 \to V' \to V \to V'' \to 0$ in $\mathcal{C}$.

Let $G$ be a finite group. Let $\mathcal{C}[G]\text{-mod}_{\text{fin}}$ be the category of finite dimensional complex $G$-representations. We write $K(G) = K(\mathcal{C}[G]\text{-mod}_{\text{fin}})$. Explicitly, the elements of $K(G)$ are represented by finite dimensional complex $G$-representations. If $V$ and $V'$ are isomorphic representations then $[V] = [V']$. Since $\mathcal{C}[G]$ is semisimple, one has $[V] = [V'] + [V'']$ if and only if $V = V' \oplus V''$. Note that $K(G)$ is a commutative ring with the product defined by the tensor product of $G$-representations and the identity element given by the trivial $G$-representation. This is called the Grothendieck ring of $G$.

Verify that if $V$ and $W$ are isomorphic finite dimensional $G$-representations then $\chi_V = \chi_W$. So (4) implies that the character of a representation defines a ring homomorphism

$$\chi : K(G) \to \text{cl}(G) \text{ defined by } [V] \mapsto \chi_V.$$

3.3.4. The orthogonality of characters: Consider

$$e_G = |G|^{-1} \sum_{g \in G} g \in Z(k[G]), \quad e_G^2 = e_G.$$

Let $(V, \rho)$ be a representation of $G$. We have seen that $\rho(e_G)$ is the projection onto the subspace fixed by $G$. So $\text{Tr}(\rho(e_G)) = \dim(V^G)$.

Define an inner product on $\text{cl}(G)$ by

$$\langle f, f' \rangle = |G|^{-1} \sum_{g \in G} \overline{f(g)} f'(g).$$

Let $V, V' \in \text{Rep}(G)$ and $V'' = \text{Hom}_k(V, V')$. One has

$$\langle \chi_V, \chi_{V'} \rangle = |G|^{-1} \sum_{g \in G} \chi_{V''}(g) = \text{Tr}(e_G|_{V''}) = \dim((\text{Hom}_k(V, V'))^G) = \dim(\text{Hom}_{k[G]}(V, V')).$$

\(^1\)To prove this, choose a positive definite $G$-invariant hermitian form on $V$, so that each $g \in G$ acts as a unitary transformation on $V$. Fix $g \in G$. Then there is a basis $\{e_i\}$ of eigenvectors of $g|_V$, so that $ge_i = \lambda_i e_i$ and $|\lambda_i| = 1$. Let $\{e^i\}$ be the dual basis of $V^*$. Then $(ge^i)(e_j) = e^i(g^{-1}e_j) = \lambda_i \delta_{ij} = (\lambda_i e^i)(e_j)$ for all $j$, so $ge^i = \lambda_i e^i$ so the eigenvalues of $g|_{V^*}$ are the complex conjugates of the eigenvalues of $g|_V$. 

11
So by Schur’s lemma the irreducible characters are an orthonormal subset in the space of class functions.

3.3.5. Character’s determine representation. Let \( \text{Irr}(G) \) be the set of of irreducible \( G \)-representations (upto isomorphism). Let \( V \in \text{Rep}(G) \). For \( U \in \text{Irr}(G) \), Let
\[
 m_U(V) = \dim(\text{Hom}_G(V, U)) = \langle \chi_V, \chi_U \rangle.
\]

By complete reducibility, we can write \( V \simeq \bigoplus_{U \in \text{Irr}(G)} m_U U \). Taking trace and using orthogonality of characters, we get \( m_U = m_U(V) \). So \( V \) is determined by the coefficients \( (m_U(V) : \rho \in \text{Irr}(G)) \) and these numbers are determined by the character of \( V \). So the isomorphism class of the representation is uniquely determined by \( \chi_V \). The number \( m_U(V) \) is called the multiplicity of the irrep \( U \) in \( V \). Let \( V = \bigoplus_{U \in \text{Irr}(G)} m_U U \). Then \( \chi_V = \sum_U m_U \chi_U \). Orthonormality of characters imply that \( |\chi_V|^2 = \sum_U m_U^2 \). So \( V \) is an irrep if and only if \( |\chi_V|^2 = 1 \).

3.3.6. Theorem. (a) Let \( r \) be the number of conjugacy classes in \( G \). Then \( G \) has exactly \( r \) distinct irreps.

(b) Let \( V_1, \ldots, V_r \) be the distinct irreps of \( G \). Then \( \sum_{j=1}^{r} \dim(V_j)^2 = |G| \).

(c) The characters of irreps form an orthonormal basis for the space \( \text{cl}(G) \) of class functions.

(d) One has \( \mathbb{C}[G] \simeq \bigoplus_{j=1}^{r} \dim(V_j) V_j \).

Proof. Let \( A = \mathbb{C}[G] \). By Maschke’s theorem \( A \) is a semisimple ring. From the proof of Wedderburn theorem, recall that we have the isotypical component decomposition \( A = M_1 \oplus \cdots \oplus M_s \) where \( s \) is the number of distinct isomorphism types of simple \( A \) modules and this implies
\[
 A^\circ = \text{End}_A(A) = \prod_j \text{End}_A(M_j) \simeq \prod_{j=1}^{s} M_{n_j}(D_j^\circ)
\]
and hence \( A \simeq \prod_{j=1}^{s} M_{n_j}(D_j) \) where \( D_j \) are some division algebras over \( \mathbb{C} \). Since \( \mathbb{C} \) is algebraically closed, it follows that \( D_j = \mathbb{C} \) for all \( j \). Recall the unique simple \( M_{n_j}(\mathbb{C}) \) module \( V_j \simeq \mathbb{C}^{n_j} \) has dimension \( n_j \), and these already define \( s \) non-isomorphic simple \( A \) modules. So the nonisomorphic simple \( A \) modules have dimensions \( n_1, \ldots, n_s \).

Rephrasing the above, we have shown that if \( n_1, \ldots, n_s \) are the dimensions of the irreducible representations of \( G \), then
\[
 \mathbb{C}[G] \simeq \prod_{j=1}^{s} M_{n_j}(\mathbb{C}). \quad (5)
\]
It follows that \( Z(\mathbb{C}[G]) = \prod_{j=1}^{s} \mathbb{C} \). Counting dimensions of both sides as \( \mathbb{C} \) vector space, we get \( s = \dim(Z(\mathbb{C}[G])) = r \). Counting dimensions of both sides of (3.3) as \( \mathbb{C} \) vector space, we get part (b). The characters of the irreps form an orthonormal set of size \( r \) in the vector space \( \text{cl}(G) \) of dimension \( r \). Hence part (c). Finally, recall that \( M_{n_j}(\mathbb{C}) \simeq n_j V_j \) as \( M_{n_j}(\mathbb{C}) \) representation. So part (d) also follows from .

3.3.7. Since characters determine a representation, the character homomorphism \( \chi : K(G) \to \text{cl}(G) \) is an injective. The image of \( \chi \) is the free abelian group generated by the characters of irreps. So \( K(G) \) is isomorphic to a free abelian group on the irreps. So, by \( \mathbb{C} \)-linearity we can extend \( \chi : K(G) \to \text{cl}(G) \) to an injective homomorphism \( \chi : K(G) \otimes_{\mathbb{Z}} \mathbb{C} \to \text{cl}(G) \) of
\( G \)-algebras. Since the characters of irreps form a basis of \( \text{cl}(G) \) this map is onto. Thus the character defines a \( G \)-algebra isomorphism \[ \chi: K(G) \otimes \mathbb{C} \simeq \text{cl}(G) \]

3.3.8. **Character of permutation module:** Let \( X \) be a set. The vector space \( k^X \) (also sometimes written as \( k[X] \) if there is no chance of confusion with polynomial ring) has the standard basis \( \{e_x: x \in X\} \) given by \( e_x(y) = \delta_{xy} \). Let \( G \) act on a finite set \( X \). Then \( k^X \) is a \( k[G] \) module, called a permutation module. Each \( g \in G \) acts on \( k^X \) as permutation matrix (with respect to the standard basis). So
\[
\chi_{k^X}(g) = |X^g|.
\]

Below we have another proof of the theorem 3.3.6 without using the theory of semisimple rings.

3.3.9. **The regular representation:** Let \( L: k[G] \to \text{End}(k[G]) \) be the left multiplication map, given by \( x \mapsto L_x \), where \( L_x(y) = xy \). The representation \( \text{Reg} = (L, k[G]) \) is called the regular representation of \( G \). In other words, the regular representation is \( k[G] \), considered as a \( k[G] \) module. This is a permutation module. The character of the regular representation is given by \( \chi_{\text{Reg}}(1) = |G| \) and \( \chi_{\text{Reg}}(g) = 0 \) for all \( g \neq 1 \). So, if \( \rho \in \text{Irr}(G) \), then \( n_{\text{Reg}}(\rho) = \chi_{\rho}(1) = \dim(\rho) \), that is, \( k[G] \simeq \bigoplus_{\rho \in \text{Irr}(G)} \dim(\rho) \rho \). Taking norm of \( \chi_{\text{Reg}} \) and using the orthonormality of characters one gets
\[
|G| = \langle \chi_{\text{Reg}}, \chi_{\text{Reg}} \rangle = \sum_{\rho \in \text{Irr}(G)} \dim(\rho)^2.
\]

3.3.10. **The number of irreducible representations:** Let \( \text{cl}(G) \) be the space of class functions on \( G \). Let \( f \in \text{cl}(G) \) and \( (\rho,V) \in \text{Irr}(G) \). Then (3) implies that \( \rho_f = \sum_{t \in G} f(t)\rho(t) \in \text{Hom}_G(V,V) \), so by Schur’s lemma \( \rho_f = \lambda \text{id}_V \). Taking trace, one obtains,
\[
\dim(V)\lambda = \sum_{t \in G} f(t)\chi_\rho(t), \quad \text{so } \lambda = \frac{|G|}{\dim(V)} \langle f, \chi_\rho \rangle.
\]
If \( f \in \text{cl}(G) \) is orthogonal to all the characters, then the above formula implies that \( \rho_f = 0 \) for all \( \rho \in \text{Irr}(G) \). By complete irreducibility one has \( \rho_f = 0 \) for all \( \rho \in \text{Rep}(G) \). In particular applying this to the regular representation \( \text{Reg} = (L, k[G]) \), one has \( 0 = L_f e_1 = \sum_{t \in G} f(g)L_t e_1 = \sum_{t \in G} f(t)t \), so \( f = 0 \). It follows that the characters form an orthonormal basis for the space of class functions. In particular, the number of irreducible representations is equal to the number of conjugacy classes.

3.3.11. **The character homomorphism:** \( \chi: K_0(\text{Rep}_k(G)) \to \text{cl}(G) \) defined by \( \rho \mapsto \chi_\rho \) is injective since characters determine representation and surjective since the irreducible characters form a basis for the class functions. So \( \rho \mapsto \chi_\rho \) is a ring isomorphism, that identifies the Grothendieck ring of the representation category with the space of class function. Complete irreducibility implies that \( K_0(\text{Rep}_k(G)) \simeq \text{Rep}_k(G)/\simeq \), so the tensor category \( \text{Rep}_k(G) \) is determined up to equivalence by the ring \( \text{cl}(G) \).

3.3.12. **The character table:** Let \( \{e\} = C_1, \ldots, C_r \) be the conjugacy classes of \( G \). The character table of \( G \) is an \( r \times r \) table whose \((i,j)\)-th entry is \( \chi_i(C_j) \). So the rows and columns of the character table are indexed by the irreducible characters and the conjugacy classes of \( G \) respectively.
Since $\langle \chi_i, \chi_j \rangle = \delta_{ij}$, the rows of the character tables are orthogonal in the sense that

$$|G|^{-1} \sum_{k=1}^{r} |C_k| \chi_i(C_k) \bar{\chi}_j(C_k) = \delta_{ij}.$$ 

Consider the $r \times r$ matrix $A$ whose $(i,k)$-th entry is $(|C_k| / |G|)^{1/2} \chi_i(C_k)$. The row orthogonality relations above says that $AA^* = I$. Since $A$ is a square matrix, one has $A^* A = I$, that is

$$\sum_{k=1}^{r} \chi_k(C_i) \bar{\chi}_k(C_j) = \delta_{ij} |G| / |C_i|.$$ 

These orthogonality relations, together with the equation $\sum_r \chi_r(1) \chi_r = \chi_{reg}$ sometimes lets one construct the character table of a group from very little actual information about the irreducible representations.

### 3.3.13. The character table of $S_4$:

The first row indicates the size of the conjugacy classes and the row below that indicates the cycle type of each conjugacy class. Below we describe one way to construct the character table.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>6</th>
<th>8</th>
<th>3</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e$</td>
<td>(12)</td>
<td>(123)</td>
<td>(12)(34)</td>
<td>(1234)</td>
<td></td>
</tr>
<tr>
<td>$\chi_1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\chi_{sgn}$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$\chi_2$</td>
<td>2</td>
<td>0</td>
<td>-1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>$\chi_t$</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$\chi_c$</td>
<td>3</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

The first two rows are obvious, the trivial representation $\chi_1$ and the sign representation $\chi_{sgn}$. There are a total of 5 conjugacy classes, so there are five irreps. Their squared dimension must add up to $|S_4| = 24$. So the squared dimension of the remaining three representations add up to 22. There is only one possible solution to this, namely one 2 dimensional and two 3 dimensional ones. We call these $\rho_2$, $\rho_t$ and $\rho_c$ and their characters $\chi_2$, $\chi_t$ and $\chi_c$ with $\rho_2$ being the 2 dimensional one. Now we can complete the first column of the table.

If $\rho_2$ restricted to $S_3 = \text{Perm}\{1,2,3\} \subseteq S_4$ decompose as direct sum of one dimensional representations of $S_3$ and then $\chi_2((123)) = 2$ (since $\chi_1((123)) = \chi_{sgn}((123)) = 1$), but then $\sum_{g} |\chi_2(g)|^2 \geq 8.2^2 + \cdots$, contradicting $\langle \chi_2, \chi_2 \rangle = 1$. So $\rho_2$ restricted to $S_3$ must be the unique 2 dimensional representation of $S_3$. This lets us complete the first three columns of the third row. The rest of the third row can be completed using row orthogonality.

Next consider $\rho_t$ and $\rho_c$ restricted to $S_3$. As in the previous step, we can show that $\rho_t$ and $\rho_c$ restricted to $S_3$ must decompose as a direct sum of the 2 dimensional irreducible and a one dimensional representation. So $\chi_t((123)) = \chi_c((123)) = 0$ and $\chi_t((12))$, $\chi_c((12)) \in \{1, -1\}$. 

"
Now $\langle \chi_2, \chi_t \rangle = 0$ implies $\chi_t((12)(34)) = \chi_c((12)(34)) = 1$. The rest of the table can easily be completed using row and column orthogonality.

Remarks:

1. The character table implies $\rho_2 \otimes \rho_{sgn} \cong \rho_2$ and $\rho_t \otimes \rho_{sgn} \cong \rho_c$. In general, tensoring with $\rho_{sgn}$ gives an involution on the set of irreps of $S_n$.

2. Consider a cube in $\mathbb{R}^3$ with center at the origin. Label its by $1, 2, 3, 4, \bar{1}, \bar{2}, \bar{3}, \bar{4}$ where $(1, 2, 3, 4)$ are the vertices on one face in clockwise order and $1, \bar{1}$ are diagonally opposite vertices etc. The set of rotations preserving the cube permutes the 4 diagonals and gives an injective homomorphism from $S_4$ to the set of rotations preserving the unit cube, hence a 3 dimensional representation of $S_4$. This representation is visibly irreducible. The transposition $(12)$ acts by an $180^\circ$ rotation (about the line joining the midpoints of the edges $\{1, 2\}$ and $\{\bar{1}, \bar{2}\}$). So this transformation has eigenvalues $\{1, -1, -1\}$ and trace $-1$. So this representation is $\rho_c$.

3. The set of all euclidean symmetries preserving the tetrahedron (with center at the origin) permutes the 4 vertices. Thus we get a 3 dimensional irrep $\rho_T$ of $S_4$. Note that there exists $g \in S_4$ such that $\rho_T(g)$ is a reflection, so has determinant $-1$. On the other hand $\rho_c(g) = 1$ for all $g \in S_4$. So $\rho_T(g)$ and $\rho_c(g)$ cannot be conjugate. So the irreps $\rho_T$ and $\rho_c$ are distinct. So $\rho_T \not\cong \rho_t$.

4. The action of $S_4$ as rotations of the cube permutes the three lines joining the opposite faces of the cube. So we get a homomorphism from $S_4$ to $S_3$ and hence a 2 dimensional irrep of $S_4$.

3.3.14. Character table of the dihedral group of order 12. We describe a rather ad-hoc way to construct the character table of $D_6$, the dihedral group of order 12. Define $D_6$ to be the set of reflections and rotations of the regular hexagon in the complex plane whose vertices are the sixth roots of unity. Define $r, s_v, s_e \in D_6$ as follows. Let $r$ be the anticlockwise rotation of order 6. Let $s_v$ be the reflection across the $x$-axis. Let $s_e$ be a reflection across a line through the origin with slope 30 degrees. The table is given below. By definition the elements of $D_6$ are $2 \times 2$ matrices acting on the plane. This two dimensional representation is called the standard representation (row 5 of the table). The determinant of the standard representation gives a one dimensional representation called det (row 2 of table).

<table>
<thead>
<tr>
<th>sizes of conj. classes</th>
<th>1</th>
<th>1</th>
<th>2</th>
<th>2</th>
<th>3</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>conj. classes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi_1 = \chi_{trivial}$</td>
<td>1</td>
<td>-1</td>
<td>$r$</td>
<td>$r^2$</td>
<td>$s_v$</td>
<td>$s_e$</td>
</tr>
<tr>
<td>$\chi_2 = \chi_{det}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$\chi_3$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$\chi_4$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\chi_5 = \chi_{standard}$</td>
<td>2</td>
<td>-2</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\chi_6$</td>
<td>2</td>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Here is one way to construct the table. We already have rows 1, 2 and 5. The first column can be completed using the fact that the squares of the dimensions of the irreps has to add up to 12. There is a 3 dimensional permutation representation of $D_6$ where the group acts on the three mirrors of reflection joining opposite pairs of vertices of the hexagon. By counting fixed points, we compute the character of this 3 dimensional representation is $\chi_e = (3,3,0,0,1,1)$. Alternatively, this three dimensional representation is $\text{sym}^2 \chi_5$. The 3 dimensional representation contains a copy of the trivial representation (why?). Let $\chi_6 = \chi_e - \chi_{\text{trivial}}$. One verifies that $\chi_6$ has norm 1, so it is indeed character of an irrep and also that it is orthogonal to $\chi_{\text{standard}}$. So $\chi_6$ is the character of a second 2 dimensional irrep. This completes row 6. The rest of columns 2, 5 and 6 are now completed using the column orthogonality and the fact that, under an one dimensional representation, elements of order 2 can only map to $\pm 1$. Finally columns 3 and 4 can be completed using column orthogonality.

Let $C$ be the $6 \times 6$ matrix in the character table. Let $M$ be the diagonal matrix whose diagonal entries are the sizes of the conjugacy classes: $M = \text{diag}(1,1,2,2,3,3)$. Then row orthogonality says $CMC^* = 12I_6$ and column orthogonality says $C^*C = 12M^{-1}$.

### 3.3.15. Kernel of a character.

One can extract a lot of information about a finite group form its character table. Below, we describe how to detect normal subgroups. If $\chi : G \to \mathbb{C}$ is a character of $G$, define

$$\ker(\chi) = \{g \in G : \chi(g) = \chi(1)\}$$

Let $\rho : G \to \text{Aut}(V)$ be the representation with character $\chi$ and let $n = \dim(V) = \chi(1)$. Since finite group actions are uniterizable, for any $g \in G$, the eigenvalues of $\rho(g)$ are $n$ roots of unity and these can add up to $n$ if and only if $\rho(g) = I_n$. That is, $\ker(\chi) = \ker(\rho)$ is a normal subgroup of $G$.

### 3.3.16. Exercise.

(a) A subgroup $N$ of $G$ is normal if and only if there is a character $\chi : G \to \mathbb{C}$ with $\ker(\chi) = N$.

(b) Let $N$ be a normal subgroup of $G$. Let $\chi$ be a character with $\ker(\chi) = N$. If $\chi = \sum_j n_j \chi_j$ where $\chi_j$ are distinct irreducible characters and $n_j \geq 1$, then $\ker(\chi) = \cap_j \ker(\chi_j)$.

**Hint.** (a) Let $G$ acts on $\mathbb{C}[G/N]$ by $ge_{hN} = e_{ghN}$. Note that this is just the permutation representation of $G$ corresponding to the action of $G$ on $G/N$ by left translation. Verify that $g \in G$ fixes every element of $G/N$ if $g \in N$ and fixes nothing if $g \notin N$.

(b) Recall that if $\rho : G \to \text{Aut}(V)$ is the representation with character $\chi$, then $\ker(\rho) = \ker(\chi)$. \qed

It follows from the Exercise that the normal subgroups of $G$ are of the form $\cap_{j=1}^t \ker(\chi_j)$ where $\chi_1, \ldots, \chi_t$ are certain irreducible characters of $G$. The kernels of the characters of the irreps can be read off from the character table and hence, one can detect the normal subgroups of $G$ and decide when $G$ is simple.
3.4. Isotypic components.

3.4.1. Notation: For this sub-section, let $G$ be a finite group. Let $\text{Rep}(G)$ be the class of finite dimensional complex representations of $G$. Let $I$ be a set of representatives for the isomorphism classes of irreducible representations of $G$. For a sub-representation $U$ of $V$, let $p_U : V \to U$ be the $G$-linear projection map.

3.4.2. Definition. Let $V \in \text{Rep}(G)$. Let $W$ be an irrep of $G$. Let $I_W^V = I_W$ be a maximal sub representation of $V$ that can be written as $I_W \simeq \bigoplus_{j \in J} W_j$ with each $W_j \simeq W$. Let $U \subseteq V$ be an irrep such that $U \simeq W$. If $U \cap I_W = 0$, then $U \oplus I_W$ would contradict the maximality of $I_W$, so we must have $U \cap I_W \neq 0$, hence $U \subseteq I_W$ (since $U$ is an irrep). It follows that $I_W$ is the sum of all the sub-representations of $V$ that are isomorphic to $W$, so $I_W$ is uniquely defined by the isomorphism type of the irrep $W$. We say that $I_W^V = I_W$ is the isotypic component of $V$ of type $W$.

3.4.3. Lemma. Let $V \in \text{Rep}(G)$.

(a) Let $W$ be an irrep of $V$. Then all the sub-irreps of $I_W$ are isomorphic to $W$.

(b) If $W_1$ and $W_2$ are non-isomorphic irreps of $G$, then $I_{W_1} \cap I_{W_2} = 0$.

Proof. (a) Let $U$ be a sub-irrep of $I_W$. Write $I_W = \bigoplus_{j \in J} W_j$ with $W_j \simeq W$ for each $j$. Let $p_j : I_W \to W_j$ be the $G$-linear projections. Then the composites $(U \hookrightarrow I_W \xrightarrow{p_j} W_j)$ can not all be zero since $\sum_j p_j = \text{id}_{I_W}$. Sine $U$ and $W_j$ are irreps, Schur’s lemma implies that $p_j|_U : U \to W_j$ is an isomorphism for some $j$.

(b) If $I_{W_1} \cap I_{W_2} \neq 0$, then we can choose $U$ to be be any sub-irreps of $I_{W_1} \cap I_{W_2}$ and part (a) would imply that $U \simeq W_1$ and $U \simeq W_2$. \hfill $\square$

3.4.4. Theorem. Let $V \in \text{Rep}(G)$. Let $V = U_1 \oplus \cdots \oplus U_r$ be a decomposition of $V$ into irreps. Then $I_W$ is the equal to the direct sum of those $U_j$’s appearing the above decomposition that are isomorphic to $W$. In particular we have a canonical decomposition of $V$ as a direct sum of the isotypical components: $V = \bigoplus_{W \in I} I_W$.

Proof. Without loss we may assume that $U_1, \ldots, U_r$ are isomorphic to $W$ and $U_{r+1}, \ldots, U_n$ are not. Recall that $r = \text{mult}_W(V) = \dim (\text{Hom}_G(W, V))$, so any decomposition of $V$ into irreps must contain exactly $r$ copies of $W$. If $I_W$ properly contains $\bigoplus_{j=1}^r U_j$, then write $I_W = \bigoplus_{j=1}^r U_j \oplus V'$. By the lemma above each irrep appearing in $V'$ must be of type $W$. So decomposing $V'$ into irreps we obtain a decomposition of $I_W$ into irreps with more than $r$ copies of $W$, and this would yield a decomposition of $V$ into irreps that has more than $r$ copies of $W$, which is impossible. So we must have $I_W = \bigoplus_{j=1}^r U_j$. \hfill $\square$

3.4.5. Remark. The arguments of this subsection holds for any group for which for which Schur’s lemma and complete reducibility of representations hold, so the multiplicities of irreducibles in a representation is well defined.

3.4.6. Remark. Let $g$ be a $n \times n$ complex matrix of finite order. Consider the cyclic group $\langle g \rangle = G$ acting on $\mathbb{C}^n$. The distinct irreps of $G$ appearing in $\mathbb{C}^n$ correspond to the distinct eigenvalues of $g$ and the corresponding isotypic components are the eigenspaces. Thus the canonical decomposition of $\mathbb{C}^n$ into isotypic components becomes the decomposition into eigenspaces of $g$.

17
3.5. Sub-representations of the regular representation and idempotents.

3.5.1. **Notation:** Let $G$ be a finite group. Let $A = \mathbb{C}[G]$ be the complex group algebra of $G$. Recall that $A$ is the regular representation of $G$. An element $e \in A$ is an idempotent if $e^2 = e$. If $V$ is a sub-representation of $A$, let $p_V : A \rightarrow V$ be the projection.

Verify that if $U, Z$ are sub-representations of $A$, and $f : U \rightarrow Z$ is any $G$-linear map, then $f(a.u) = a.f(u)$ for all $a \in A$ and $u \in U$. The multiplication on either side is simply the multiplication in the ring $A$.

3.5.2. **Lemma.** (a) The sub-representations $V \subseteq A$ are in one to one correspondence to idempotents $e_V \in A$ via the correspondence $V = Ae_V$ and $e_V = p_V(1)$.

(b) Suppose $U$ and $Z$ are sub-representations of $A$ and $e_U, e_Z$ be the corresponding idempotents. Then $\text{Hom}_G(U, Z) = 0$ if and only if $e_U a e_Z = 0$ for all $a \in A$, that is, $e_U A e_Z = 0$.

**Proof.** (a) Given a sub-representation $V \subseteq A$, let $e_V = p_V(1)$. Then

$$p_V(a) = p_V(a.1) = ap_V(1) = ae_V.$$ 

Since $p_V$ is a projection $p_V^2 = p_V$, so $e_V = p_V^2(1) = p_V(e_V) = e_V^2$. So $e_V$ is an idempotent and $V = p_V(A) = Ae_V$. Conversely given any idempotent $e \in A$, the subspace $V = Ae$ is a $G$ sub-representation of $A$ and $a \mapsto ae$ is the projection map $p_V : A \rightarrow Ae$.

(b) Suppose $\text{Hom}_G(U, Z) = 0$. Let $a \in A$. The map $r_a : U \rightarrow A$ given by $r_a(u) = ua$ is $G$-linear. The composition $U \xrightarrow{r_a} A \xrightarrow{p_z} Z$ is $G$-linear. So $p_z \circ r_a = 0$. So

$$0 = p_z \circ r_a(e_U) = p_z(e_U a) = e_U a e_Z.$$ 

Conversely, suppose $e_U a e_Z = 0$ for all $a \in A$. Let $f : U \rightarrow Z$ be any $G$-linear map. We can write an element of $u \in U$ as $u = ae_U$ for some $a \in A$. We can write $f(e_U) = a e_Z$. Then

$$f(e_U) = f(e_U^2) = e_U f(e_U) = e_U a e_Z = 0.$$ 

Each $u \in U$ can be written as $u = be_U$. So $f(u) = bf(e_U) = 0$. \qed

3.5.3. **Idempotents in the regular representations:** Let $A = \sum_{W \in I} I_W$ be the decomposition of the regular representation into isotypical components. Let $e_W = p_I(1)$ be the idempotent corresponding to the sub-representation $I_W$. Note that $\sum_{W \in I} p_W = \text{id}_A$. So $\sum_{W \in I} e_W = 1$. Thus the decomposition of $A$ into isotypical components gives us a set of elements $\{e_W : W \in I\}$ in $A$ such that

$$e_W^2 = e_W \text{ for all } W \in I, \quad e_W A e_W = 0 \quad \text{for all } W \neq W' \quad \text{and} \quad \sum_{W \in I} e_W = 1. \quad (6)$$ 

Conversely given a set of elements in $A$ satisfy (6), let $J_W = Ae_W$. Then $A = \sum_{W \in I} J_W$. Suppose $x \in J_{W_0} \cap \left( \sum_{W \neq W_0} J_W \right)$. Then we can write $x = ae_{W_0} = \sum_{W \neq W_0} a_w e_W$. So

$$x = x e_{W_0} = \sum_{W \neq W_0} a_w e_W e_{W_0} = 0.$$ 

From the lemma above, it follows that $A$ is the direct sum of the sub-representations $J_W$ and $\text{Hom}_G(J_W, J_{W'}) = 0$ for all $W \neq W'$. It follows that $\{J_W : W \in I\}$ must be the decomposition of $A$ into isotypical components. Since the decomposition of $A$ into isotypical components is unique, a set of elements $\{e_W : W \in I\}$ satisfying (6) is also unique.
3.6. Integrality of Characters:

3.6.1. **Definition.** Let $G$ be a finite group. Let $\mathbb{Z}[G]$ denote the integral group ring of $G$, consisting of all elements of the form $\sum_{g \in G} n_g g$, with $n_g \in \mathbb{Z}$, with multiplication defined by $(\sum_{g \in G} m_g g)(\sum_{h \in G} n_h h) = \sum_{k \in G}(\sum_{gh = k} m_g n_h)k$.

The center of a ring $R$, denoted by $Z(R)$, is defined by $Z(R) = \{x \in R : xy = yx \forall y \in R\}$. The center $Z(R)$ is a commutative subring of $R$.

3.6.2. **Lemma** (Center of the group ring). Let $G$ be a finite group. Let $C_1, \ldots, C_r$ be the conjugacy classes of $G$. Define $z_j \in \mathbb{Z}[G]$ by $z_j = \sum_{g \in C_j} g$. Then

(a) $Z(\mathbb{Z}[G]) = \mathbb{Z}z_1 + \cdots + \mathbb{Z}z_r$.

(b) Let $\rho : G \to \text{Aut}(V)$ be a finite dimensional representation of $G$. Note that $\rho$ extends to a ring homomorphism $\rho : \mathbb{Z}[G] \to \text{End}_k(V)$. Then $\rho(z_j)$ is a $G$-linear map from $V$ to $V$.

(c) If $(\rho, V)$ is an irreducible representation of $G$, then $\rho(z_j) = \lambda_j \text{id}$, for some scalar $\lambda_j$.

**Proof.** Exercise. □

3.6.3. **Theorem.** Let $\rho : G \to \text{Aut}(V)$ be an $n$ dimensional representation of a finite group $G$. Let $\chi$ be the character of $V$. Let $g \in G$ and $h$ be the number of elements in the conjugacy class of $g$. Then

(a) $\chi(g)$ is an algebraic integer.

(b) Suppose $V$ is an irreducible representation. Then $h\chi(g)/n$ is an algebraic integer.

(c) Suppose $V$ is irreducible and $h$ is relatively prime to $n$. Then either $\chi(g) = 0$ or $\rho(g)$ is a scalar multiple of identity.

**Proof.** (a) Let $\zeta_1, \cdots, \zeta_n$ be the eigenvalues of $\rho(g)$. Suppose $g$ has order $m$. Let $\zeta = e^{2\pi i/n}$. Then each $\zeta_j$ is an $m$-th root of unity, that is $\zeta_j = \zeta^{m_j}$ for some $0 \leq m_j < m - 1$. In particular $\zeta_j$ are algebraic integers. So $\chi(g) = \sum_j \zeta_j$ is also an algebraic integer.

(b) Use the notation of the previous lemma. From part (c) of the previous lemma we have $\rho(z_j) = \lambda_j \text{id}$ for some. Since $Z(\mathbb{Z}[G])$ is a ring $z_i z_j = \sum_k d_{ij}^k z_k$ for some integers $d_{ij}^k$. Since $\rho : \mathbb{Z}[G] \to \text{End}(V)$ is ring homomorphism, one obtains $\lambda_i \lambda_j = \sum_k d_{ij}^k \lambda_k$. So $\mathbb{Z}[\lambda_1, \cdots, \lambda_r]$ is a finitely generated abelian group. It follows that $\lambda_j$ are algebraic integers. Let $C_j$ be the conjugacy class of $g$, then

$$h\chi(g) = \text{Tr}(\rho(z_j)) = \lambda_j \dim(V) = \lambda_j n.$$

(c) Suppose $\rho(g)$ is not a scalar multiple identity, that is, not all the $\zeta_j$ are equal. Find integers $b$ and $c$ such that $hb + nc = 1$. So

$$n^{-1}\chi(g) = b(h^{-1}\chi(g)) + c\chi(g).$$

Part (a) and (b) implies that $\chi(g)/n$ is an algebraic integer. Let $f(x) = (x - \mu_1) \cdots (x - \mu_l)$ be the monic irreducible polynomial of $\mu_1 = \chi(g)/n$. Observe that

$$|\mu_1| = n^{-1}|\chi(g)| = |n^{-1}(\zeta^{m_1} + \cdots + \zeta^{m_n})| < 1$$

since $\zeta_j$ are not all equal. The conjugates of $\mu_1$ are also algebraic integers since they satisfy the same irreducible polynomial as $\mu_1$. Since each conjugate of $\zeta$ has the form $\zeta^t$ (for some $t$), the conjugates of $\mu_1$ has the form $\mu_j = n^{-1}(\zeta^{m_1} + \cdots + \zeta^{m_n})$ (by field theory). So $|\mu_j| \leq 1$. It follows that $\prod_j |\mu_j| < 1$ but $\prod_j \mu_j \in \mathbb{Z}$ (the constant term of the $f(x)$). So $\mu_1 = 0$ (and $l = 1$). □
3.6.4. **Theorem.** Let $V$ be an irrep of a finite group $G$. Then $\dim(V)$ divides $|G|$.

**Proof.** Let $\chi$ be the character of the irrep $V$. Let $C_1, \cdots, C_r$ be the conjugacy classes of $G$. Then $\langle \chi, \chi \rangle = 1$ implies,

$$
\sum_{i=1}^r \bar{\chi}(C_i)\chi(C_i)|C_i|/\dim(V) = |G|/\dim(V).
$$

But $|C_i|\chi(C_i)/\dim(V)$ is an algebraic integer as is $\bar{\chi}(C_i)$. So $|G|/\dim(V)$ is an algebraic integer as well as a rational number. \hfill \Box

3.6.5. **Lemma.** (a) Let $G$ be a simple group of composite order. Let $\rho$ be a non-trivial irrep of $G$. If $g \in G$ is not equal to the identity then $\rho(g)$ is not a scalar operator.

(b) Suppose $G$ is a finite group of composite order that has a conjugacy class of prime power order. Then $G$ is not simple.

**Proof.** (a) Let $H = \{g \in G: \rho(g)$ is a scalar matrix$\}$. Since $G$ is simple, $\rho$ is injective. It follows that $H$ is a normal abelian subgroup of $G$. So $H = G$ or $H = \langle e \rangle$. But an abelian group of composite order is not simple, so $H \neq G$.

(b) If possible, suppose $G$ is simple. Let $1 = \rho_1, \cdots, \rho_r$ be the irreps of $G$ having dimension $n_1, \cdots, n_r$ and $\chi_1, \cdots, \chi_r$ be their characters. Let $g \in G$ be an element whose conjugacy class has $h = p^e$ elements where $p$ is prime and $e \geq 1$. Then part (a) implies $\rho_j(g)$ is not a scalar operator for $j \geq 2$. Part (c) of Lemma 3.6.3 now implies that whenever $p$ does not divide $n_j$ then $\chi_j(g) = 0$. From the character of the regular representation, we know that

$$
0 = \chi_{k[G]}(g) = 1 + \sum_{j=1}^r n_j\chi_j(g) = 1 + \sum_{j: p \text{ divides } n_j} n_j\chi_j(g) = 1 + p\beta.
$$

where $\beta = \sum_{j: p \text{ divides } n_j} (n_j/p)\chi_j(g)$. So $\beta$ is an algebraic integer. But $\beta = -1/p$ and the only rational numbers that are algebraic integers are the rational integers. This is a contradiction. \hfill \Box

3.6.6. **Theorem** (Burnside). Let $p$ and $q$ be primes. Then a group of order $p^aq^b$ is solvable.

**Proof.** Let $G$ be a group of order $p^aq^b$. We know that a group of prime power is solvable. So assume $p \neq q$ and $a, b \geq 1$. By induction on order it is enough to show that $G$ has a normal subgroup.

Recall that a group of prime power order has nontrivial center. Let $H$ be a Sylow $q$ subgroup of $G$ and $g \neq e$ be an element in the center of $H$. If $g$ is in the center of $G$ then the group generated by $g$ is a proper normal subgroup. Otherwise, let $N$ be the normalizer of $g$. Then $N \supseteq H$. So the number of elements in the conjugacy class of $g$ is equal to $|G|/|N| = p^e$ for some $e \geq 1$. Apply part (b) of the lemma above. \hfill \Box
3.7. **Induced representation:**

3.7.1. **Definition.** Let $G$ be a finite group, $H$ be a subgroup of $G$. Let $k$ be a field. All vector spaces considered today will be over $k$. Let $W$ be a representation of $H$. Define $V = \text{Fun}_H(G, W)$ to be the set of functions $\psi$ from $G$ to $W$ such that

$$\psi(xh) = h^{-1}|_W \psi(x) \quad \text{for all } x \in G, h \in H.$$  

Define a left action of $G$ on $V = \text{Fun}_H(G, W)$ by $(g\psi)(x) = \psi(g^{-1}x)$. This representation of $G$ is denoted by $\text{Ind}_H^G(W)$.

3.7.2. **Remark.** The above definition of $\text{Ind}_H^G(W)$ is simplest to state but not the most intuitive. To make it more intuitive we introduce some geometric language. The geometric picture has the added advantage that it generalizes naturally in case of continuous groups. However since we are interested in finite groups at the moment we shall not introduce any topology in our discussion. So we restrict our attention to vector bundles on finite sets.

3.7.3. **Definition.** A vector bundle of rank $n$ on a finite set $X$ is a set $E$ together with a map $E \to X$ such that $p^{-1}(x)$ is a vector space of dimension $n$ for each $x \in X$. We say that $p^{-1}(x)$ is the fiber of $E$ at $x$ and write $E_x = p^{-1}(x)$. Let $(E, p)$ and $(E', p')$ be two vector bundles on $X$. A map of vector bundles is a function $\alpha : E \to E'$ such that $p' \circ \alpha = p$ and the restriction of $\alpha$ to each fiber of $E$ is a linear transformation.

A section $s$ of a vector bundle $(E, p)$ on $X$ is a function $s : X \to E$ such that $p(s(x)) = x$ for all $x \in X$. The space of all sections form a vector space of dimension $n|X|$, which we denote by $\Gamma(X, E)$.

3.7.4. **Definition.** Let $X$ be a set with a $G$-action. A $G$-equivariant vector bundle of rank $n$ on $X$ is a vector bundle $(E, p)$ on $X$ together with an action of $G$ on $E$ such that $p(ge) = gp(e)$ for all $e \in E$ and $g|_{E_x} : E_x \to E_{gx}$ is a linear map for all $g \in G$. Let $(E, p)$ and $(E', p')$ be two $G$-equivariant vector bundles over $X$. A map of $G$-equivariant vector bundles is a map $\alpha : E \to E'$ as defined above, with the extra condition that $\alpha(ge) = g\alpha(e)$.

3.7.5. **The $G$ representation from a $G$-equivariant vector bundle:** As above, let $(E, p)$ be a $G$-equivariant vector bundle on $X$. Then the space of sections $\Gamma(X, E)$ becomes a linear representation of $G$ with the action of $G$ defined by

$$(g|_{\Gamma(X, E)}s)(y) = g|_E(s(g^{-1}|_X y)) \quad \text{for } s \in \Gamma(X, E), \ g \in G, \ y \in X.$$  

If $g \in G$ fixes some $x \in X$, then $g|_{E_x} : E_x \to E_x$ is a linear transformation. Verify that the character of the representation $\Gamma(X, E)$ is given by the following “Frobenius Character formula”:

$$\chi_{\Gamma(X, E)}(g) = \sum_{x \in X : gx = x} \text{Tr}(g|_{E_x}).$$  

3.7.6. For the rest of this section assume that $G$ acts transitively on $X$. Fix $x_0 \in X$, let $H$ be the stabilizer of $x_0$ and identify $G/H$ with $X$ by $gH \leftrightarrow gx_0$. Pick a transversal $\{t_x : x \in X\}$ for $G/H$, that is,

$$t_xx_0 = x \quad \text{and} \quad t_{x_0} = 1.$$  

Now let $E$ be a $G$-equivariant vector bundle of rank $n$ on $X$. Let $W = E_{x_0}$. Since $H$ fixes $x_0$, the vector space $W$ is a $H$ representation.
Using the transversal $t_x$ we can identify the fiber at $x$ with the fiber at $x_0$. More precisely every element of $E$ can be uniquely written as $t_xw$ for some $x \in X$ and $w \in W$. Let us temporarily write $(x, w) = t_xw$. Pick $g \in G$. Let $y = gx$. Then there exists $h \in H$ such that $gt_x = t_yh$. We have

$$g|_E(x, w) = gt_xw = t_yhw = (y, hw) = (g|_Xx, (t^{-1}_xgt_x)|_Ww).$$

Thus we see that the action of $G$ on $E$ determined once the action of $G$ on $X$ and the action of $H$ on the fiber $W = E_{x_0}$ is given. This suggests that one should be able to recover $E$ once one is given the subgroup $H$ and the representation $W$ of $H$. Indeed it is true and it is achieved by the “Borel construction”. To make this more precise, see the construction and the two lemmas below.

3.7.7. **The Borel construction:** Given a group $G$, a subgroup $H$ and a representation $W$ of $H$ of dimension $n$. Define an equivalence relation on $G \times W$ by

$$(gh, v) \sim (g, hv) \quad \text{for} \quad g \in G, h \in H, v \in W.$$ 

Define $E = (G \times W)/ \sim$. If $(g, v) \in G \times W$, then its equivalence class in $E$ is denoted by $[g, v]$. An element of $e = E$ is an equivalence class of the relation $\sim$. So $[gh, v] = [g, hv]$ for $h \in H$. There is an obvious $G$-action on $E$ given by $g[g', v] = [gg', v]$ (verify that this action is well defined). Define $p : E \rightarrow X$ by $[g, v] \mapsto gH$. Verify that this is well defined and that $(E, p)$ is a $G$-equivariant vector bundle of rank $n$ on $X$. One denotes the vector bundle $E$ by $G \times_H W$.

3.7.8. **Lemma.** Assume the setup given in 3.7.6. Then $G \times_H W$ is isomorphic to $E$ as a $G$-equivariant vector bundle.

**Proof.** Recall that $W = E_{x_0}$. Let $w \in W$ and $g \in G$. Define a map from $G \times_H W \rightarrow E$ by

$$\alpha : [g, w] \mapsto g|_E w.$$ 

This map is well defined because for each $h \in H$, $[gh, w]$ and $[g, hw]$ map to the same element of $E$. The $G$-equivariance is also visible:

$$\alpha(g'[g, w]) = \alpha([g', g, w]) = (g'|g, g)|_E w = g'|_E g|_E w = g'|_E(\alpha([g, w]))$$

Clearly $\alpha$ restricted to the fiber at $x_0$ is an isomorphism. Since $\alpha$ is $G$-equivariant, this is enough to ensure that $\alpha$ is an isomorphism on each fiber. \qed

3.7.9. **Lemma.** Given a group $G$, a subgroup $H$ and a representation $W$ of $H$ of dimension $n$. Let $X = G/H$ and $E = G \times_H W$ be the bundle defined in 3.7.7. Then the $G$-representation $\Gamma(X, E)$ is isomorphic to the $G$-representation $\text{Fun}_H(G, W)$, defined in 3.7.1.

**Proof.** Given $s \in \Gamma(X, E)$, we shall define $\psi : G \rightarrow W$. Given $g \in G$, verify that there exists unique $v \in W$ such that $s(gx_0) = [g, v]$. Define $\psi(g) = v$. In other words, the function $\psi$ is uniquely determined by the condition $[g, \psi(g)] = s(gH)$.

If $h \in H$, then observe that $[g, \psi(g)] = s(gx_0) = s(ghx_0) = [gh, \psi(gh)] = [g, hv(gh)]$.
Summing over the whole coset instead of coset representatives, we obtain,
\[ \psi(gh) = h^{-1}|W|^{-1} \psi(g), \] that is \( \psi \in \text{Fun}_H(G, W) \). Verify that the map \( s \mapsto \psi \) is an isomorphism from \( \Gamma(X, \mathcal{E}) \) to \( \text{Fun}_H(G, W) \).

3.7.10. **Definition.** The \( G \)-representation \( \Gamma(G/H, G \times_H W) \) will be denoted by \( \text{Ind}_H^G(W) \) or simply \( \text{Ind}(W) \). It is called the \( G \)-representation induced from the \( H \)-representation \( W \).

3.7.11. **Theorem** (Frobenius reciprocity). Let \( H \) be a subgroup of a group \( G \). Let \( W \in \text{Rep}(H) \) and \( U \in \text{Rep}(G) \). Then
\[ \text{Hom}_G(\text{Ind} W, U) \simeq \text{Hom}_H(W, \text{Res}(U)). \]

**Proof.** Let \( X = G/H \) and \( \mathcal{E} = G \times_H W \). Let \( x_0 \in X \) be the coset \( H \). For \( g \in G \), \( w \in W \) define a section \( \delta_{[g,w]} \in \Gamma(X, \mathcal{E}) \) by
\[ \delta_{[g,w]}(x) = [g, w] \delta_{g x_0}, \]
These are the sections supported at a single point \( x \in X \). Clearly these sections span the vector space \( \Gamma(X, \mathcal{E}) \). Verify that \( g \mapsto \delta_{[g,w]} \). Given a \( H \)-linear map \( \phi : W \to U \), we need to define a \( G \)-linear map \( \psi : \text{Ind}(W) \to U \). We shall define it on the sections supported at one point as follows:
\[ \psi(\delta_{[g,w]}) = g|_U \phi(w). \]

Note that \( \psi \) is well defined since \( \phi \) is \( H \)-linear. The \( G \)-linearity of \( \psi \) is also clear.

To get a map in the other direction we define an embedding \( i : W \hookrightarrow \Gamma(X, \mathcal{E}) \), by \( w \mapsto \delta_{[1,w]} \). Check that \( i \) is \( H \)-linear:
\[ (hi(w))(x) = (h|_{\Gamma(X, \mathcal{E})}) \delta_{[1,w]}(x) = [h, w] \delta_{x_0}, x = [1, hw] \delta_{x_0}, x = \delta_{[1,hw]}(x) = (i(hw))(x) \]
So any \( G \)-linear map \( \psi : \Gamma(X, \mathcal{E}) \to U \) gives a \( H \) linear map \( \phi : W \to U \), where \( \phi = \psi \circ i \). Verify that the maps \( \phi \mapsto \psi \) and \( \psi \mapsto \phi \) are inverses of each other.

3.7.12. **The character of induced representation:** Let \( H \) be a subgroup of a finite group \( G \), let \( W \in \text{Rep}(H) \). Let \( V = \text{Ind}(W) \) be the induced representation. Let \( X = G/H \). Choose cost representatives \( \{ t_x : x \in X \} \) for \( G/H \). Recall the Frobenius character formula (8) and the formula for the \( G \) action on sections of a bundle \( E \). From these, we obtain
\[ \chi_V(g) = \sum_{x \in X : gx = x} \text{Tr}(g|_{\mathcal{E}_x}) = \sum_{x \in X : gx = x} \chi_W(t^{-1}_x g t_x). \]

Summing over the whole coset instead of coset representatives, we obtain,
\[ \chi_V(g) = |H|^{-1} \sum_{t \in G : gt \equiv t \mod H} \chi_W(t^{-1} g t) = |H|^{-1} \sum_{h \in \text{cl}_G(g) \cap H} |N_G(g)| \chi_W(h) \]
where \( \text{cl}_G(g) \) is the conjugacy class of \( g \) in \( G \) and \( N_G(g) = \{ t \in G : tgt^{-1} = g \} \). Suppose \( C_g \cap H \) decompose into \( r \) conjugacy classes of \( H \), denoted by \( D_1, \cdots, D_r \). Then
\[ \chi_V(g) = \frac{|G|}{|H|} \sum_{h \in \text{cl}_G(g) \cap H} \frac{\chi_W(h)}{|\text{cl}_G(g)|} = \frac{|G|}{|H|} \sum_{j=1}^r \frac{|D_j|}{|\text{cl}_G(g)|} \chi_W(D_j) \]
In particular, if \( W \) is the trivial representation, then
\[ \chi_V(g) = \frac{|G|}{|H|} \frac{|\text{cl}_G(g) \cap H|}{|\text{cl}_G(g)|}. \]
4. Linear representations of Lie groups

4.1. Smooth manifolds and their tangent spaces. Disclaimer: We merely want to write out the definitions in one place. Certainly not meant to be a first introduction to manifolds. It needs to be complemented with the informal discussions and examples we saw in class.

A function between open subsets of Euclidean spaces is called $C^\infty$ (or smooth) if it has all partial derivatives of all order. Below we use the words $C^\infty$ and smooth interchangeably. A smooth (real) manifold of dimension $n$ is a space locally modelled on open subsets of $\mathbb{R}^n$ glued using smooth functions. One way to make this precise is the following.

4.1.1. Definition (Manifolds). Let $M$ be a topological space. Let $n$ be a non-negative integer. A smooth local chart $(U, \phi)$ of dimension $n$ in $M$ is an open subset $U$ of $M$ containing $p$ and a homeomorphism $\phi : U \to \phi(U) \subseteq \text{open} \mathbb{R}^n$. A local chart $(V, \psi)$ is a restriction of a local chart $(U, \phi)$ if $V \subseteq U$ and $\psi = \phi|_V$.

A smooth atlas of dimension $n$ for $M$ is a collection $\mathcal{A} = \{(U_i, \phi_i) : i \in I\}$ of dimension $n$ local charts such that $\bigcup_{i \in I} U_i = M$ and such that for all $i, j$ the functions

$$\phi_j \circ \phi_i^{-1}|_{\phi_i(U_i \cap U_j)} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$$

are smooth (i.e. $C^\infty$) isomorphisms (draw a picture). The functions $\{\phi_j \circ \phi_i^{-1}|_{\phi_i(U_i \cap U_j)} : i, j\}$ are called the transition functions of the smooth atlas $\mathcal{A}$.

Let $\mathcal{A}$ and $\mathcal{A}'$ are two smooth atlases on $M$. Say that two atlases are compatible if their union is still an atlas. This defines an equivalence relation on the collection of smooth atlases on $M$. A smooth (real) manifold of dimension $n$ is a second countable Hausdorff space with an equivalence class smooth atlases of dimension $n$. One way to describe a manifold is to specify the topological space $M$ and describe a smooth atlas on $M$.

Let $M$ and $N$ be smooth manifolds. A continuous function $f : M \to N$ is called a smooth if for each $p \in M$, there exists local charts $(U, \phi)$ in $M$ and $(V, \psi)$ in $N$ with $p \in M$ and $f(U) \subseteq V$ and $\psi \circ f \circ \phi^{-1} : \phi(U) \to \psi(V)$ is a $C^\infty$ function. One verifies that the smoothness of a function $f$ does not depend on the choice of atlases or charts. Now we have a category of smooth manifolds and smooth maps. The first examples are $\mathbb{R}^n$ (with all local $C^\infty$ isomorphisms defining local charts). An open subset $N$ of a manifold $M$ is again a manifold with the manifold structure inherited from $M$ (just intersect the domains of the local charts with the open set $N$).

One defines topological manifolds, differentiable manifolds, algebraic manifolds, real analytic manifolds by changing $C^\infty$ functions to appropriate class of functions. One defines complex manifolds by changing $\mathbb{R}$ with $\mathbb{C}$ and by taking analytic transition functions. In the discussion below, we shall stick to smooth manifolds but most of it goes through in any of the other categories with obvious modifications.

4.1.2. Definition (Tangent space). Let $M$ be a smooth manifold and $p \in M$. A smooth map $\gamma : (-\epsilon, \epsilon) \to M$ (for some $\epsilon > 0$) such that $\gamma(0) = p$ is called a smooth curve in $M$ through $p$. Let $\gamma_1 : (-\epsilon_1, \epsilon_1) \to M$ and $\gamma_2 : (-\epsilon_2, \epsilon_2) \to M$ be two smooth curves in $M$ through $p$. Say that $\gamma_1 \sim_p \gamma_2$ if there is an open chart $(U, \phi)$ around $p$ such that $(\phi \circ \gamma_1)'(0) = (\phi \circ \gamma_2)'(0)$. Verify that the definition of $\sim_p$ does not depend on the choice of the local chart around $p$. Verify that $\sim_p$ is an equivalence relation on $T_p$. A tangent vector to $M$ at $p$ can be defined as an equivalence class. In this description, we describe a tangent vector by specifying a
smooth curve $\gamma : (-\epsilon, \epsilon) \to M$ such that $\gamma(0) = p$. We denote corresponding tangent vector by $[\gamma]$. Let $T_p(M)$ be the set of tangent vectors to $M$ at $p$. A smooth map $f : M \to N$ of manifolds determines a function $df_p : T_p(M) \to T_{f(p)}N$ defined by $df_p([\gamma]) = [f \circ \gamma]$. Verify that this map is well defined.

4.1.3. Exercise (a) Show that if $I : M \to M$ is the identity map, then $dI_p : T_p(M) \to T_p(M)$ is the identity map for all $p \in M$.

(b) If $M' \xrightarrow{g} M \xrightarrow{f} M''$ are smooth maps of manifolds, then $d(f \circ g)_p = df_{g(p)} \circ dg_p$ for all $p \in M$. One often abbreviates this equation as $d(f \circ g) = df \circ dg$.

(c) Show that there is a natural identification $T_p(\mathbb{R}^n) = \mathbb{R}^n$ for all $p \in \mathbb{R}^n$.

4.1.4. Remark.

- If $M', M, M''$ are open subsets of $\mathbb{R}^n$, then part (b) of the exercise is just chain rule from multivariable calculus. The general case follows from this.
- From part (a) and (b) it follows that if $f : M \to N$ is a smooth isomorphism, then so is $df_p$ for all $p \in M$.
- Let $M$ be a smooth manifold of dimension $n$ and $p \in M$. Let $(U, \phi)$ be a local chart around $p$. Then $d\phi_p : T_p(M) \to T_{\phi(p)}(\mathbb{R}^n) = \mathbb{R}^n$ is an isomorphism. Using this isomorphism we can turn $T_p(M)$ into a $n$ dimensional vector space. Verify that this vector space structure does not depend on the choice of the local chart. The $n$ dimensional vector space $T_p(M)$ is called the tangent space of $M$ at $p$.

- Verify that if $f : M \to N$ is a smooth map, then $d\phi_p : T_p M \to T_{\phi(p)}N$ is a linear map for all $p \in M$. If $M$ and $N$ are open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively then the matrix for the linear map $d\phi_p$ with respect to the usual basis of $\mathbb{R}^m$ and $\mathbb{R}^n$ is just the $n \times m$ matrix of partial derivatives of $f$ at $p$ (i.e. the Jacobian matrix).

As we already remarked, an open subset of a manifold is a manifold in obvious manner. But there is some subtleties involved in defining closed submanifolds of a smooth manifold. This is addressed in the definition below.

4.1.5. Definition (Submanifolds). An immersion $f : N \to M$ of smooth manifolds is a smooth map such that $df_p$ is an injective map for all $p \in N$. If $f : N \to M$ is an immersion, we say that $f(N)$ is an immersed submanifold of $M$. But globally, $f$ need not be injective. Think of a smooth non-simple closed curve in the plane, for example, a figure 8.

Let $f : N \to M$ is an injective immersion. One can “transport the manifold structure from $N$ to $f(N)$” to define a manifold structure on $f(N)$ inherited from $M$ such that $f : N \to f(N)$ is a smooth isomorphism. But note that the subspace topology on $f(N)$ need not agree with the topology $f(N)$ inherits from its manifold structure. Think of the image of an irrational rotation $f : \mathbb{R} \to T$ on the torus $T = \mathbb{R}^2/\mathbb{Z}^2$. The real line injectively and smoothly maps into $T$ with dense image so the topology on $f(\mathbb{R})$ as a subspace of $T$ is certainly different from the topology it inherits from $\mathbb{R}$. The next definition rules out these kind of possibilities.

A smooth map $f : N \to M$ is an embedding if $f$ is an injective immersion and also is a closed map (i.e. takes closed sets to closed sets). The image $f(N)$ is sometimes called a closed submanifold of $M$. Again, $f(N)$ inherits a manifold structure from $N$ and $f : N \to f(N)$ is a smooth isomorphism. The topology on $f(N)$ agrees with the subspace topology inherited from $M$.

Unless otherwise stated, when we say $N \subseteq M$ is a closed submanifold of a manifold $M$, we usually mean that $N$ is a manifold and the inclusion map $i : N \to M$ is an embedding.
If \( N \) is a submanifold of \( M \), in fact the manifold structure on \( N \) is inherited from \( M \). We explain this below in case of submanifolds of \( \mathbb{R}^n \). The general statement is similar.

4.1.6. **Submanifolds of Euclidean space:** We want to note two important facts that we shall not prove. These allow us to work with manifolds in very concrete terms.

- Any smooth manifold can be embedded in Euclidean space. More precisely, Whitney proved that if \( M \) is a smooth manifold of dimension \( n \) then there is an embedding \( f : M \to \mathbb{R}^{2n} \). So while studying smooth manifolds, we lose no generality if we only look at closed submanifolds of Euclidean space.

- Let \( M \) be a smooth \( n \) dimensional submanifold of \( \mathbb{R}^{n+r} \). Identify \( \mathbb{R}^n \) inside \( \mathbb{R}^{n+r} \) in the standard manner (by setting the last \( r \) coordinates equal to 0). Then for each \( p \in M \), there exists an open subset \( U \subseteq \mathbb{R}^{n+r} \) and a \( C^\infty \) isomorphism \( \phi : U \to \phi(U) \subseteq \mathbb{R}^{n+r} \) such that

\[
\phi(M \cap U) = \mathbb{R}^n \cap \phi(U)
\]

In other words, locally \( M \subseteq \mathbb{R}^{n+r} \) just looks like a \( n \) dimensional subspace \( \mathbb{R}^n \subseteq \mathbb{R}^{n+r} \). If we write \( \phi(x) = (\phi_1(x), \ldots, \phi_{n+r}(x)) \) and think of these as a set of (curvelinear) coordinates no \( \mathbb{R}^{n+r} \), then \( M \) is locally described by setting some of these coordinates equal to 0. Further, \( \phi|_{M \cap U} : M \cap U \to \phi(M \cap U) \) is a smooth isomorphism where the manifold structure on \( \phi(M \cap U) \) is as an open subset of \( \mathbb{R}^n \). In this sense, the manifold structure on \( M \) is induced from the ambient Euclidean space \( \mathbb{R}^{n+r} \).

Often one defines submanifolds of \( \mathbb{R}^n \) using implicit function theorem. In these cases, the above statement is automatic by looking at the proof of implicit function theorem (from inverse function theorem). We illustrate this with an example.

4.1.7. **Tangent spaces of submanifolds of \( \mathbb{R}^n \).** Let \( M \) be a closed \( n \) dimensional submanifold of \( \mathbb{R}^d \). Let \( p \in M \). Verify that the tangent space \( T_p(M) \) of \( M \) at \( p \) can be identified with the \( n \) dimensional subspace of \( \mathbb{R}^d \) consisting of all \( v \in \mathbb{R}^d \) such that \( (p + v) \) is tangent to \( M \) at \( p \), that is, there exists smooth \( \gamma : (-\epsilon, \epsilon) \to M \) such that \( \gamma(0) = p \) and \( \gamma'(0) = v \).

Recall that a smooth map \( f : M \to N \) of manifolds defines linear maps \( df_p : T_p(M) \to T_{f(p)}(N) \) between tangent spaces and one has the chain rule \( d(f \circ g) = df \circ dg \). Let \( v \) be a tangent vector to \( M \) at \( p \). To compute \( df_p(v) \), choose a curve \( \gamma : (-\epsilon, \epsilon) \to M \) such that \( \gamma(0) = p \) and \( \gamma'(0) = v \). Then

\[
df_p(v) = (f \circ \gamma)'(0).
\]

Assume \( M \subseteq \mathbb{R}^m \) and \( N \subseteq \mathbb{R}^n \) are submanifolds in Euclidean space. Let \( f : M \to N \) be a smooth map. Let \( p \in M \). From our discussion above, it follows that there exists some open neighborhood \( U \) of \( p \) and a smooth function \( F : \mathbb{R}^m \to \mathbb{R}^n \) such that \( F|_{U \cap M} = f|_{U \cap M} \). Let \( dF_p \) be the Jacobian matrix of \( F \) at \( p \) (i.e. the matrix of partial derivatives). Let \( v \) be a tangent vector at \( p \). Then \( df_p(v) := dF_p(v) \). These discussions go through unchanged for real analytic or complex manifold. (In the complex case one consider complex analytic curves (i.e. analytic maps from a small ball in \( \mathbb{C} \) to our manifold and one takes holomorphic derivatives).
4.2. Lie groups and their Lie algebras.

4.2.1. Definition. Let $k$ be a field. A Lie algebra over $k$ is a $k$-vector space $L$ with a $k$-bilinear antisymmetric product $[\ ,\ ] : L \times L \to L$, called the Lie bracket satisfying the Jacobi identity. A Lie algebra homomorphism is a linear map that preserves the Lie bracket. Exercise: Define subalgebra, ideal, quotient.

4.2.2. Example.
- Any associative algebra is a Lie algebra under the commutator bracket: $[x, y] = xy - yx$.
- In particular, if $V \simeq k^n$ is a finite dimensional $k$ vector space, then the Lie algebra corresponding to $\text{End}(V) \simeq M_n(k)$ is denoted by $\mathfrak{gl}(V) \simeq \mathfrak{gl}_n(k)$.
- Let $R$ be an associative $k$ algebra. A derivation of $R$ is a $k$-linear map $\delta : R \to R$ satisfying the Liebnitz rule: $\delta(fg) = \delta(f)g + f\delta(g)$. Verify that the set of all derivations $\text{der}_k(R)$ is a Lie algebra under commutator bracket.

4.2.3. Lemma. Let $k$ be a field. $L$ be a $k$-vector space with a $k$ anti-symmetric bilinear product $[\ ,\ ] : L \times L \to L$. For $x \in L$ define $\text{ad}(x) : L \to L$ by $\text{ad}(x)(y) = [x, y]$. TFAE
(a) $\text{ad}$ is a derivation of the product $[\ ,\ ]$.
(b) $\text{ad} : L \to \mathfrak{gl}(L)$ is a Lie algebra homomorphism.
(c) $[\ ,\ ]$ satisfies the Jacobi identity.

4.2.4. Definition. A real (or complex) Lie group $G$ is a smooth manifold as well as a group such that the map $G \times G \to G$ taking $(x, y) \to xy^{-1}$ is a smooth (or complex analytic) map. A Lie group homomorphism is a group homomorphism which is also a smooth (or analytic) map.

We shall soon see that the group multiplication of a Lie group $G$ induces a Lie algebra structure on the tangent space $T_e(G)$ of $G$ at identity. This is called the Lie algebra of $G$ and is usually denoted by $\mathfrak{g}$. Since $\text{GL}_n(k)$ is an open subset of $M_n(k)$, its Lie algebra (as a set) is simply the set of all $n \times n$ matrices. We will soon verify that the Lie algebra structure induced on $M_n(k)$ is the commutator bracket, that is, the Lie algebra of $\text{GL}_n(k)$ is $\mathfrak{gl}_n(k)$.

Our primary object of interest are the real and complex Lie groups and their finite dimensional representations. But we will spend most of our time studying the Lie algebras of these Lie groups and their representations. Let us first justify this strategy, before describing the definition of the bracket on $\mathfrak{g}$.

4.2.5.
- A Lie group homomorphism $f : G \to H$ determines a Lie algebra homomorphism $f_* : \mathfrak{g} \to \mathfrak{h}$.
- If $G$ is connected, then $f$ is determined by $f_*$.
- If $G$ is connected and simply connected, then any Lie algebra homomorphism is the differential of a group homomorphism.

4.2.6. Definition of the bracket on $\mathfrak{g}$: Let $G$ be a (smooth real) Lie group. Write $\mathfrak{g} = T_eG$. Take $g \in G$. Let $c_g : G \to \text{Aut}(G)$ be the conjugation by $g$ map:
$$c_g(h) = ghg^{-1}$$
“Differentiating $c_g$ at identity” we get the map $Ad(g) := (c_g)_* : \mathfrak{g} \to \mathfrak{g}$. One verifies that this yields a smooth map of Lie groups

$$Ad : G \to GL(\mathfrak{g}).$$

“Differentiating $Ad$ at identity, we get a linear map

$$ad := Ad_* : \mathfrak{g} \to \text{End}_k(\mathfrak{g}).$$

This defines the bilinear map $[,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ by

$$[x, y] = ad(x)(y).$$

Note that if $G$ is abelian, then $Ad$ is the constant map $id_\mathfrak{g}$, so $ad = 0$, that is, $[x, y] = 0$ for all $x, y \in \mathfrak{g}$.

4.2.7. Lemma. Let $\rho : G \to H$ be a Lie group homomorphism. Then the induced map

$$\rho_* = d\rho_* : \mathfrak{g} \to \mathfrak{h}$$

preserves the bracket we just defined.

Proof. Let $g \in G$. Then verify that

$$\rho \circ c_g = c_{\rho(g)} \circ \rho.$$ (draw the commutative diagram). Differentiating both sides at identity and applying the chain rule, we find that

$$\rho_* \circ Ad(g) = Ad(\rho(g)) \circ \rho_*.$$ In other words, the two maps $G \to \text{Hom}_k(\mathfrak{g}, \mathfrak{h})$ defined by $\phi(g) = \rho_* \circ Ad(g)$ and $\phi'(g) = (Ad \circ \rho(g)) \circ \rho_*$ are equal. Take $x \in \mathfrak{g}$. Using the chain rule again, we find

$$\phi_* (x) = \rho_* \circ ad(x) \quad \text{and} \quad \phi'_* (x) = (dAd_{\rho_*} \circ d\rho_*)(x) \circ \rho_* = (ad(\rho_*) (x)) \circ \rho_* = ad(\rho_*) (x) \circ \rho_*$$

where we have to remember that $\rho_*$ is linear so, so its differential is itself. So for all $x, y \in \mathfrak{g}$, we have $\rho_* \circ ad(x)(y) = (ad(\rho_*) (x)) \circ \rho_*(y) = ad(\rho_*) (y)$ (draw the commutative diagram). It follows that

$$\rho_* [x, y] = \rho_* \circ ad(x)(y) = (ad(\rho_*) (x)) \circ \rho_*(y) = ad(\rho_*) (\rho_*(y)) = [\rho_*(x), \rho_*(y)].$$

□

From the lemma, it follows that if $j : H \to G$ is a local immersion near identity of Lie groups, then since the induced map $j_*$ is injective, once we identify $\mathfrak{h}$ inside $\mathfrak{g}$ using $j_*$, the bracket on $\mathfrak{h}$ is induced from the bracket on $\mathfrak{g}$. In particular, this is true if $H$ is a closed Lie subgroup of $G$.

4.2.8. the bracket on the Lie algebra of $GL_n(\mathbb{R})$. Let $G = GL_n(\mathbb{R})$ and let $\mathfrak{g}$ be its Lie algebra. Since $G$ is an open subset of the vector space $M_n(\mathbb{R})$, the tangent space identity is $\mathfrak{g} = M_n(\mathbb{R})$. One computes $[x, y] = xy - yx$ for $x, y \in \mathfrak{g}$. (We did this in class. Need to fill in the details here). Let $H$ be a closed Lie subgroup of $G$. Then as we already commented, the bracket on $\mathfrak{h}$ is induced from $\mathfrak{g}$ so it is again just the commutator bracket and so it satisfies skew symmetry and Jacobi identity.

So at least for closed Lie subgroups $H$ of $GL_n(\mathbb{R})$ (or $GL_n(\mathbb{C})$), the bracket we defined on $\mathfrak{h} = T_e(H)$ turns it into a Lie sub-algebra of $\mathfrak{gl}_n(\mathbb{R})$ (or $\mathfrak{gl}_n(\mathbb{C})$). Lie algebra of $\mathfrak{h}$ is simply the set of all matrices tangent to the subgroup $H$ at identity with the commutator bracket.
4.2.9. Lie algebras, vector fields and derivations. We only sketch the connections between these concepts. You should read the details in a textbook (like Warner or Helgason) at some point. A smooth vector field $X$ on a smooth manifold $M$ is a choice of tangent vector $X(p)$ for each $p \in M$ such that $p \mapsto X(p)$ “varies smoothly with $p$”. One way to make this precise is to think of $M$ as a closed submanifold of Euclidean space so that the tangent spaces are all subspaces of the ambient Euclidean space. As usual, then one has to verify that this notion of smoothness does not depend on the choice of embedding. Let $C^\infty(M)$ be the ring of smooth (real valued) functions on $M$. If $v$ is a tangent vector at $p \in M$ and $f \in C^\infty(M)$, then define

$$v(f) = df_p v.$$ 

Verify that

$v : C^\infty(M) \to \mathbb{R}$

satisfies the Liebniz rule, that is

$$v(f_1 f_2) = v(f_1) f_2(p) + f_1(p) v(f_2).$$

In fact, another intrinsic way to define the tangent space of a manifold $M$ at a point $p$ is to define a tangent vector as a real valued $\mathbb{R}$ linear map on the vector space of “smooth local functions defined near $p$” satisfying Liebniz rule.

Now if $p \mapsto X(p)$ is a smooth vector field on $M$, and $f : M \to \mathbb{R}$ is a smooth function, then a new function $Xf : M \to \mathbb{R}$ by $(Xf)(p) = X(p)f$. Since $X$ is smooth, it is easy to convince oneself that $Xf$ is also smooth. It follows that $f \mapsto Xf$ is a derivation of $C^\infty(M)$. In fact, the set $\text{Der}(C^\infty(G))$ of “smooth derivations” of $C^\infty(M)$ can be identified with the set of smooth vector fields on $M$ and we will denote both by $\text{Der}(C^\infty(M))$. We know the set of derivations of an associative ring is a Lie algebra under commutator bracket. So space of smooth vector fields on $M$ is a Lie algebra.

Now let $G$ be a Lie group. A vector field $v$ on a Lie group $G$ is left invariant if for all $g, h \in G$, one has $(dm_g)_h v(h) = v(gh)$ where $m_g : G \to G$ is the map $m_g(h) = gh$. Verify that this is equivalent to requiring $(dm_g)_e v(e) = v(g)$ for all $g \in G$. Clearly a left invariant vector field on a Lie group $G$ is determined by its value at identity. Conversely, given a tangent vector $x \in T_e(G) = g$ we get a Left invariant vector field by translating it using left multiplication: $v_x(g) : = (dm_g)_e(x)$. In this way $g$ is identified with the set of left invariant vector fields on $G$ and hence with a subset of of $\text{Der}(C^\infty(G))$. One can show that the Lie bracket we defined on $g$ in 4.2.6 is the same as it inherits from $\text{Der}(C^\infty(G))$. If we made all this precise, we would get a proof that the bracket on $g$ satisfies antisymmetry and Jacobi identity in general. Another proof will be sketched below after we define the exponential map.

Our next objective is to define the exponential map, which is the main tool for going back to the Lie groups from Lie algebras. Again the treatment is terse. Please see a textbook for a more detailed exposition, for example, Fulton and Harris Chapter 8.

4.2.10. Definition (One parameter subgroups and the exponential map.). Let $G$ be a Lie group and let $g$ be its Lie algebra. Choose $x \in g$. As we saw, this defines a left invariant vector field $g \mapsto v_x(g)$ defined by $v_x(g) = (dm_g)_e(x)$. By existence and uniqueness theorem of first order ordinary differential equations, there exists a unique curve $\phi = \phi_x : (-\epsilon, \epsilon) \to G$ for some $\epsilon > 0$ such that $\phi(0) = e$ and $\phi'(t) = v(\phi(t))$ for $t \in (-\epsilon, \epsilon)$. Since $v_x$ is left invariant, it follows quickly that $\phi(s) \phi(t) = \phi(s + t)$ for all $s$ and $t$ near 0 (whenever both sides are
defined). Now one can extend $\phi$ uniquely to a smooth group homomorphism $\phi : \mathbb{R} \to G$. A smooth group homomorphism $\phi : \mathbb{R} \to G$ is called a one parameter subgroup of $G$. In this way, each element $x$ of Lie algebra $\mathfrak{g}$ gives a left invariant vector field $v_x$, which “integrates to” an one parameter subgroup $\phi_x$ of $G$ that passes through identity and follows this vector field $v_x$. Exercise: verify that $\phi_x$ is the unique Lie group homomorphism $\mathbb{R} \to G$ such that $\phi_x'(0) = x$.

Define

$$\exp : \mathfrak{g} \to G \text{ by } \exp(x) = \phi_x(1).$$

By the uniqueness property of $\phi_x$ it follows that if $\lambda \in \mathbb{R}$, then

$$\phi_{\lambda x}(t) = \phi_x(\lambda t)$$

It follows that $\phi_x(\lambda) = \phi_{\lambda x}(1) = \exp(\lambda x)$. In other words, $\lambda \mapsto \exp(\lambda x)$ is the one parameter subgroup whose tangent vector at identity is $x$. It follows that the exponential map is the unique map $\mathfrak{g} \to G$ whose differential at $e$ is the identity map $\mathfrak{g} \to \mathfrak{g}$ and whose restriction to the lines the origin are the one parameter subgroups.

Let $\rho : G \to H$ be any Lie group homomorphism and let $\rho_* : \mathfrak{g} \to \mathfrak{h}$ be the induced map of Lie algebras. We have the following consequences:

- Take $x \in \mathfrak{g}$. Then verify that $t \mapsto \rho(\exp(tx))$ is a one parameter subgroup in $H$ with tangent vector $\rho_*(x)$. By uniqueness property of the exponential map, we must have $\rho(\exp(tx)) = \exp(tp_*(x))$. Thus the exponential map is natural, in the sense that

$$\rho \circ \exp = \exp \circ \rho_*$$

- Since the differential of the exponential is identity, by inverse function theorem, $\exp$ is local diffeomorphism from a neighborhood of $0$ in $\mathfrak{g}$ to a neighborhood of $e$ in $G$. So the naturality $\rho \circ \exp = \exp \circ \rho_*$ implies that $\rho$ is determined in a neighborhood of identity by $\rho_*$. If $G$ is connected, then $G$ is generated as a group by any neighborhood of identity (Why?). So $\rho$ is determined by its restriction to any neighborhood of identity and hence, $\rho$ is determined by its differential $\rho_*$. 

For $k = \mathbb{R}$ or $\mathbb{C}$, the exponential map $M_n(k) \to \text{GL}_n(k)$ is given by the matrix exponential. By naturality, the same holds for any closed subgroup of $\text{GL}_n(k)$ and its Lie algebra. In this case one can verify directly, the Campbell-Baker-Hausdorff (or CBH) formula:

$$\exp(x) \exp(y) = \exp(x + y + \frac{1}{2}[x, y] + \frac{1}{12}[x, [x, y]] + \frac{1}{12}[y, [y, x]] - \frac{1}{24}[y, [x, [x, y]]] + \cdots).$$

The CBH formula holds for any Lie group but we will not prove this.

4.2.11. Exercise. Lemma 4.2.7 implies that $\text{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ preserves the bracket for all $g \in G$. The map $\text{Ad} : G \to \text{GL}(\mathfrak{g})$ is a representation of $G$ on its $\mathfrak{g}$, called the Adjoint representation of $G$. We also find that $\text{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ preserves the bracket, that is,

$$\text{ad}[x, y] = [\text{ad}(x), \text{ad}(y)],$$

where the bracket on the right hand side is the commutator bracket. This is equivalent to saying that $[\ , \ ]$ satisfies the Jacobi identity. Use the exponential map to show that the bracket is antisymmetric as well. (Hint: Take $x \in \mathfrak{g}$. Then $x$ belongs to the lie algebra of the one parameter subgroup $H = \{\exp(tx) : t \in \mathbb{R}\} \subseteq G$. Since the one parameter subgroup is abelian its Lie bracket is 0, so $[x, x] = 0$). Now we have a general proof that $T_e(G) = \mathfrak{g}$ is indeed a Lie algebra for any Lie group $G$. 

30
4.2.12. Examples.

- Recall that since $GL_n(\mathbb{R})$ is an open subset of $M_n(\mathbb{R})$, its Lie algebra is simply $\mathfrak{gl}_n(\mathbb{R}) = M_n(\mathbb{R})$. We have already seen that the Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ of $SL_n(\mathbb{R})$ consists of all trace 0 matrices.

$$\mathfrak{sl}_n(\mathbb{R}) = \{ x \in \mathfrak{gl}_n(\mathbb{R}) : \text{tr}(x) = 0 \}.$$

- Let $G$ be a closed Lie subgroup of $GL_n(\mathbb{R})$ with Lie algebra $\mathfrak{g}$. Note that $x \in \mathfrak{g}$ if and only if $\exp(tx) \in G$ for $t$ sufficiently small, hence for all $t$. For example, let $G = O_n(\mathbb{R})$. We write $\mathfrak{g} = \mathfrak{o}_n(\mathbb{R})$. If $x \in \mathfrak{o}_n(\mathbb{R})$, then $\exp(tx) \in O_n(\mathbb{R})$, so $\exp(tx)^t \exp(tx) = I$. Differentiating at $t = 0$, one gets $x^t + x = 0$. Conversely, if $x^t = -x$, then $\exp(tx)^t \exp(tx) = \exp(-tx) \exp(tx) = I$. So

$$\mathfrak{o}_n(\mathbb{R}) = \{ x \in \mathfrak{gl}_n(\mathbb{R}) : x^t + x = 0 \}.$$

- Since $SO_n(\mathbb{R}) = SL_n(\mathbb{R}) \cap O_n(\mathbb{R})$, its Lie algebra is

$$\mathfrak{so}_n(\mathbb{R}) = \{ x \in \mathfrak{gl}_n(\mathbb{R}) : x^t + x = 0, \text{tr}(x) = 0 \}.$$

- Let $J$ be the $2n \times 2n$ skew symmetric matrix $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$. Recall we defined the symplectic group $Sp_{2n}(\mathbb{R})$ as the set of all matrices $g \in GL_{2n}(\mathbb{R})$ such that $g^t J g = J$. Its Lie algebra is

$$\mathfrak{sp}_{2n}(\mathbb{R}) = \{ x \in \mathfrak{gl}_{2n}(\mathbb{R}) : x^t J + J x = 0 \}.$$

The Lie algebras of the corresponding complex Lie groups $GL_n(\mathbb{C})$, $SL_n(\mathbb{C})$, $O_n(\mathbb{C})$, $SO_n(\mathbb{C})$, $Sp_{2n}(\mathbb{C})$ have exactly the same description (just replace $\mathbb{R}$ by $\mathbb{C}$). Indeed, the Lie groups above can be defined over any commutative ring $R$ (and in particular for any field $k$), as Linear algebraic groups, meaning that they are subgroups of $GL_n(R)$ defined as the set of solution of a bunch of polynomial equations (in the matrix entries). In this context, one can define the Lie algebra of a Linear algebraic group $G(k) \subseteq GL_n(k)$ as the set of all $x \in M_n(k)$ such that such that $1 + \epsilon x \in G(k[\epsilon]/(\epsilon^2))$. This definition matches with our Lie theory definition for $k = \mathbb{R}$ or $\mathbb{C}$. For example, $x \in \mathfrak{sp}_{2n}(\mathbb{C})$ if and only if

$$(1 + \epsilon x)^t J (1 + \epsilon x) = J$$

where $\epsilon$ is such that $\epsilon^2 = 0$. This is clearly equivalent to $x^t J + J x^t = 0$. Of course, note that if $\epsilon^2 = 0$, then $(1 + \epsilon x) = \exp(\epsilon x)$. So $(1 + \epsilon x)^t J (1 + \epsilon x) = J$ is the formal version of the equality $\exp(\epsilon x)^t J \exp(\epsilon x) = J$.

4.2.13. Exercise (More examples). (a) Show that the Lie algebra of the Unitary group $U(n)$ is

$$\mathfrak{u}_n = \{ g \in GL_n(\mathbb{C}) : x^* + x = 0 \}.$$

Here $x^*$ denotes the conjugate-transpose of $x$. Note that $U_n$ is a real Lie subgroup of $GL_n(\mathbb{C})$; it is not a complex Lie group.

(b) Let $k = \mathbb{R}$ or $\mathbb{C}$. Let $B_n(k)$ (resp. $N_n(k)$) be the set of upper triangular (resp. strictly upper triangular) matrices in $GL_n(k)$. Show that their Lie algebras $\mathfrak{b}_n(k)$ and $\mathfrak{n}_n(k)$ consist of all upper triangular (resp. strictly upper triangular) matrices.

4.2.14. Representations of Lie groups and their Lie algebras. A (finite dimensional) representation of a real (resp. complex) Lie group $G$ is a Lie group homomorphism $\rho : G \rightarrow$
GL(V) where V is a (finite dimensional) real (resp. complex vector space). If g is the Lie algebra of G, then ρ determines a Lie algebra homomorphism

\[ \rho_* : g \to \mathfrak{gl}(V) \]

which we call the associated representation of the Lie algebra g. As usual we say that G (resp. g) acts on V. write \( \rho(g)v = g|_V v = gv \) and \( \rho_* (x)v = x|_V v = xv \) for \( g \in G, x \in g, v \in V \). Each \( g|_V \) (or \( x|_V \)) are linear maps on V that vary smoothly as we vary \( g \) (or \( x \)). Saying that \( \rho \) is a group homomorphism implies

\[ (g_1 g_2)v = g_1(g_2v) \text{ for } g_1, g_2 \in G \text{ and } v \in V. \]

Saying that \( \rho_* \) is a Lie algebra homomorphism implies

\[ [x_1, x_2]v = x_1(x_2v) - x_2(x_1v) \text{ for } x_1, x_2 \in g \text{ and } v \in V. \]

4.2.15. New representations from old ones. Let G be a Lie group and g its Lie algebra. Suppose \( G \) acts on vector spaces \( V \) and \( W \). Then \( G \) acts on the vector spaces \( V \otimes W, V^*, \text{Hom}(V, W), \otimes^k V, \text{sym}^k(V), \text{alt}^k(V) \) in the usual manner (as in the case of finite groups). We can work out the action of the Lie algebra g of G on these vector spaces by formally “differentiating the action of G”. As an example, let us work out the action of g on \( V \otimes W \). If \( g \in G, v \in V, w \in W \), then

\[ g(v \otimes w) = gv \otimes gw. \]

If \( x \in g \), then formally \( (1 + \epsilon x) \in G \) where \( \epsilon^2 = 0 \). So

\[ (1 + \epsilon x)(v \otimes w) = ((1 + \epsilon x)v) \otimes ((1 + \epsilon x)w) . \]

Equating the coefficients of \( \epsilon \) on both sides we obtain

\[ x(v \otimes w) = xv \otimes w + v \otimes xw. \]

This is the Liebnitz rule. The actions of the Lie algebra on \( \otimes^k V, \text{sym}^k(V), \text{alt}^k(V) \) are also obtained by applying Liebnitz rule.

The above discussion leads us to the following purely algebraic definitio n of representations of Lie algebras in general.

4.2.16. representations of Lie algebras. Let \( k \) be a field and let L be a Lie algebra define over \( k \). A linear representation \( (V, \rho) \) of L is a vector space \( V \) together with a Lie algebra homomorphism

\[ \rho : L \to \mathfrak{gl}(V). \]

Let \( \phi \in V^*, v \in V \) and \( x \in L \). Then we have the natural pairing \( (, ) : V^* \times V \to k \). Think of \( k \) as the trivial represenation of g. Formally applying “Liebnitz rule”, we find that we should have:

\[ 0 = x(\phi, v) = (x\phi, v) + (\phi, xv). \]

So we define

\[ (x|_V, \phi)(v) = -\phi(xv). \]

Let V and W be representations of L. Define the action of L on \( V \otimes W \) by

\[ x|_{V \otimes W}(v \otimes w) = xv \otimes w + v \otimes xw. \]
Similarly define \( x|_{alt^3 V} \) by

\[
x|_{alt^3 V}(v_1 \wedge v_2 \wedge v_3) = xv_1 \wedge v_2 \wedge v_3 + v_1 \wedge xv_2 \wedge v_3 + v_1 \wedge v_2 \wedge xv_3
\]

and so on. One verifies easily that \( x|_V, x|_{V \otimes W}, x|_{alt^3 V} \) etc define representations of the Lie algebra \( L \).

4.2.17. **Remark.** There are interesting examples of infinite dimensional Lie algebra representations where at least a priori we do not have a Lie group in the picture. Some examples of important infinite dimensional Lie algebras are the Heisenberg algebra, the current algebras, the Witt algebra of vector fields on circle and its central extension the Virasoro algebra. The Heisenberg algebra is defined as \( H = \mathbb{C}[t^\pm] \oplus \mathbb{C}K \) with the bracket defined by

\[
[b_n, b_m] = n\delta_{n,-m}K
\]

where \( b_n = t^n \) and where \( K \) is central. Show that this Lie algebra has a representation on \( V = \mathbb{C}[b_{-1}, b_{-2}, \ldots] \) where such that \( b_n \) acts by multiplication for \( n < 0 \), \( K \) acts as 1 and \( b_n \) acts by \( n\partial/\partial b_{-n} \) for \( n \geq 0 \). This is the famous Fock space representation. It is determined by the following conditions: \( b_n \) acts as multipication for \( n < 0 \), \( K \) acts as identity and \( b_n \cdot 1 = 0 \) for \( n \geq 0 \). We say that \( b_n \)'s act as the creation operators for \( n < 0 \) and annihilation operators for \( n \geq 0 \). These names come from the analysis of quantization of a harmonic oscillator where these operators naturally arise. Each element \( a \) of the Fock space \( V \) determines a “field” \( Y(a, z) \in \text{End}(V)[[z^\pm]] \) as follows: \( Y(b_{-1}, z) = b(z) = \sum_{n \in \mathbb{Z}} b_n z^{-n-1} \) and

\[
Y(b_j b_j \cdots b_j, z) = \frac{1}{(-j_1 - 1)! \cdots (-j_k - 1)!} : \partial_z^{-j_1-1}b(z) \cdots \partial_z^{-j_k-1}b(z) :
\]

where \( :a_1(z) \cdots a_k(z) : \) denotes the “normally ordered product” where all the annihilation operators are moved to the right of all the creation operators. The vector space \( V \) with the fields \( Y \) form an example of a vertex operator algebra that have important applications in two dimensional conformal field theory as well as in geometric representation theory. One important feature of the Fock space is that it contains a Virasoro field, i.e. a representation of the Virasoro algebra. The Virasoro algebra is a central extension of the Witt algebra. The Witt algebra is \( W = \mathbb{C}[t^\pm] \partial_t \) with the bracket

\[
[L_m, L_n] = (m-n)L_{m+n}
\]

where \( L_m = -t^{n+1}\partial_t \). The Virasoro algebra with central charge \( c \) is \( \text{Vir}_c = W \oplus \mathbb{C}C \) with \( C \) central and bracket

\[
[L_m, L_n] = (m-n)L_{m+n} + c\frac{m^3 - m}{12}\delta_{m,-n}
\]
4.3. Solvability and Semisimplicity. Lie, Engel and Weyl’s theorems.
Let $Z(G)$ denote the center of a group $G$.

4.3.1. Exercise. Let $G$ be a connected Lie group. Show that a discrete normal subgroup $G$ must be in $Z(G)$. If $Z(G)$ is discrete, then show that $G/Z(G)$ has trivial center.

4.3.2. Simply connected and adjoint forms of a Lie group. Of course, the Lie algebra of a Lie group only knows about the connected component of identity. But there can also be more than one connected Lie groups with the same Lie algebra. These can be described roughly as follows. The proofs are omitted.

If $G$ is a connected Lie group and $Z$ is a discrete subgroup of its center, then $G/Z$ is a Lie group and the projection $p : G \to G/Z$ is a map of covering spaces. Of course the Lie groups $G$ and $G/Z$ have the same Lie algebra. We say that such Lie groups are isogenous. For example, let $\Lambda \simeq \mathbb{Z}^2$ be a lattice in $\mathbb{C}$ and let $\Lambda'$ be a finite index sublattice. Then $\mathbb{C}/\Lambda' \to \mathbb{C}/\Lambda$ is an isogeny.

Let $G$ be a connected Lie group with Lie algebra $\mathfrak{g}$. Among all Lie groups that are isogenous to $G$, there is a “largest one” (constructed by taking the universal cover $\tilde{G}$ of $G$ and lifting the group structure to $\tilde{G}$). This is called the simply connected form of $G$. If the center $Z(\tilde{G})$ of $\tilde{G}$ is discrete (as is the case for all semisimple Lie groups), then these is a “smallest one” among all groups isogenous to $G$, namely $\tilde{G}/Z(\tilde{G})$. This is called the centerless or adjoint form of $G$. The adjoint form of $G$ is isomorphic to the the image of the map $\text{Ad} : G \to \text{GL}(\mathfrak{g})$ and for semisimple groups the image is simply the connected component of identity in $\text{Aut}(\mathfrak{g})$ (which means linear maps $\alpha : \mathfrak{g} \to \mathfrak{g}$ such that $[\alpha x, \alpha y] = [x, y]$). The rest of the isogenous Lie groups are obtained by quotienting $\tilde{G}$ by subgroups of $Z(\tilde{G})$. These are all the connected Lie groups with Lie algebra $\mathfrak{g}$. So by classifying complex semisimple Lie algebras, we shall be classifying connected complex semisimple Lie groups up to isogeny.

4.3.3. Example. Identify the quaternions with $\mathbb{C}^2$. The unit norm real quaternions act on $\mathbb{C}^2$ by left multiplication preserving the norm. This identifies the unit quaternions with $SU(2)$ and shows $SU(2)$ is homeomorphic to the 3-sphere and hence is simply connected. The action of $SU(2)$ on quaternions fix the element 1 and hence acts on its orthogonal complement, the 3 dimensional real subspace of imaginary quaternions, also preserving the norm. This gives us a map $SU(2) \to SO(3)$ which is a 2:1 covering map. In fact the center of $SU(2)$ is a $\mathbb{Z}/2$ and $SO(3)$ is the quotient. This is an example of nontrivial isogeny. Here $SU(2)$ is a simply connected form and $SO(3)$ is the adjoint form.

4.3.4. Definition. Let $\mathfrak{g}$ be a Lie algebra over a field $k$ of characteristic zero. The center of $\mathfrak{g}$, denoted $Z(\mathfrak{g})$, is the set of all $x \in \mathfrak{g}$ such that $[x, y] = 0$ for all $y \in \mathfrak{g}$. We say that $\mathfrak{g}$ is abelian if $Z(\mathfrak{g}) = \mathfrak{g}$. Find the center of $\mathfrak{gl}_n(k)$ and the center of $\mathfrak{sl}_n(k)$ and the Lie algebra of all upper triangular matrices in $\mathfrak{gl}_n(k)$.

A Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is called an ideal in $\mathfrak{g}$ if $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$. If the Lie algebra of a connected Lie group, then the ideals in the Lie algebra correspond to normal subgroups of the Lie group. More precisely, if $H$ is a connected Lie subgroup of a connected Lie group $G$ with Lie algebras $\mathfrak{h}$ and $\mathfrak{g}$, then $H$ is normal in $G$ if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. The correspondence between normal subgroups and ideals motivate much of the nomenclature introduced below. For example, a Lie algebra $\mathfrak{g}$ is called simple if $\dim(\mathfrak{g}) > 1$ and it has no nonzero proper ideal.
If \( I \) and \( J \) are ideals in \( \mathfrak{g} \), then verify that \((I + J), I \cap J \) and \([I, J]\) are ideals as well. The third one uses Jacobi identity. Define the lower central series of \( \mathfrak{g} \)
\[
\mathfrak{g} = D_0(\mathfrak{g}) \supseteq D_1(\mathfrak{g}) \supseteq D_2(\mathfrak{g}) \supseteq \cdots
\]
and the derived series
\[
\mathfrak{g} = D^1(\mathfrak{g}) \supseteq D^1(\mathfrak{g}) \supseteq D^2(\mathfrak{g}) \supseteq \cdots
\]
by
\[
D_k(\mathfrak{g}) = [\mathfrak{g}, D_{k-1}(\mathfrak{g})] \quad \text{and} \quad D^k(\mathfrak{g}) = [D^{k-1}(\mathfrak{g}), D^{k-1}(\mathfrak{g})].
\]
Clearly \( D_k(\mathfrak{g}) \subseteq D^k(\mathfrak{g}) \). Verify that \( D^k(\mathfrak{g}) \) and \( D_k(\mathfrak{g}) \) are all ideals in \( \mathfrak{g} \). The ideal \([\mathfrak{g}, \mathfrak{g}] = D^1(\mathfrak{g}) = D_1(\mathfrak{g})\) is called the commutator subalgebra of \( \mathfrak{g} \). One says that \( \mathfrak{g} \) is nilpotent if \( D_k(\mathfrak{g}) = 0 \) for some \( k = 0 \), in other words bracket of any \( k \) elements is equal to 0. We say that \( \mathfrak{g} \) is solvable if \( D^k(\mathfrak{g}) = 0 \) for some \( k \). We say that \( \mathfrak{g} \) is semisimple if it has no nonzero solvable ideal. We shall see that semisimple Lie algebras are a direct sum of simple ideals. Finally note that simple Lie algebras are the diametrical opposite of solvable ones: for a simple Lie algebra \( \mathfrak{g} \), one has \( \mathfrak{g} = D^1(\mathfrak{g}) = D^2(\mathfrak{g}) = \cdots \).

**4.3.5. Quotients and isomorphism theorems** Let \( \mathfrak{h} \) be an ideal in a Lie algebra \( \mathfrak{g} \). Then verify that the bracket of \( \mathfrak{g} \) induces a bracket on the quotient vector space \( \mathfrak{g}/\mathfrak{h} \) satisfying
\[
[x + \mathfrak{h}, y + \mathfrak{h}] = [x, y] + \mathfrak{h}
\]
for all \( x, y \in \mathfrak{g} \). With this bracket \( \mathfrak{g}/\mathfrak{h} \) becomes Lie algebra and we have the obvious short exact sequence of Lie algebra homomorphisms
\[
0 \rightarrow \mathfrak{h} \rightarrow \mathfrak{g} \rightarrow \mathfrak{g}/\mathfrak{h} \rightarrow 0.
\]
Verify that the kernel of every Lie algebra homomorphism is an ideal. Conversely, every ideal \( \mathfrak{h} \) in \( \mathfrak{g} \) arises as a kernel of a Lie algebra homomorphism, for example, the natural projection from \( \mathfrak{g} \) to \( \mathfrak{g}/\mathfrak{h} \). **Exercise:** State and prove the usual isomorphism theorems for Lie algebras and ideals.

**4.3.6. Exercise on solvable algebras.** (a) Let \( \mathfrak{b}_n(k) \) (resp. \( \mathfrak{n}_n(k) \)) be the subalgebra of \( \mathfrak{gl}_n(k) \) consisting of all upper triangular (resp. strictly upper triangular) matrices. Compute the derived series and upper central series of these Lie algebras and verify that \( \mathfrak{b}_n(k) \) is solvable and \( \mathfrak{n}_n(k) \) is nilpotent. These are the standard examples of solvable and nilpotent Lie algebras.

(b) Subalgebras and homomorphic images of a solvable Lie algebra is solvable.

(c) Let \( 0 \rightarrow \mathfrak{g}' \rightarrow \mathfrak{g} \rightarrow \mathfrak{g}'' \rightarrow 0 \) be a short exact sequence of Lie algebra homomorphism. Show that \( \mathfrak{g} \) is solvable if and only if \( \mathfrak{g}' \) and \( \mathfrak{g}'' \) are solvable.

(d) Show that if \( I \) and \( J \) are solvable ideals, then so is \( I + J \). (Hint: By second isomorphism theorem, one has the short exact sequence \( 0 \rightarrow J \rightarrow (I + J) \rightarrow I/(I \cap J) \rightarrow 0 \)).

**4.3.7. Radical and semisimplicity.** Let \( \mathfrak{g} \) be a Lie algebra. The exercise above implies that the sum of all the solvable ideals in \( \mathfrak{g} \), called the radical of \( \mathfrak{g} \) and denoted by \( \text{rad}(\mathfrak{g}) \), is the largest solvable ideal of \( \mathfrak{g} \) and \( \mathfrak{g}/\text{rad}(\mathfrak{g}) \) is semisimple. In particular, \( \mathfrak{g} \) is semisimple, if and only if \( \text{rad}(\mathfrak{g}) = 0 \). At least over \( \mathbb{C} \) (probably over any algebraic closed field of characteristic zero?) the exact sequence \( 0 \rightarrow \text{rad}(\mathfrak{g}) \rightarrow \mathfrak{g} \rightarrow \mathfrak{g}/\text{rad}(\mathfrak{g}) \rightarrow 0 \) in fact splits (called the Levi decomposition) and this is a consequence of Ado’s theorem, which states that \( \mathfrak{g} \) has a faithful finite dimensional representation, that is, it can be realized as a Lie subalgebra of \( \mathfrak{gl}(V) \) for some finite dimensional vector space \( V \). But we wont prove these facts.
Next we state Lie and Engel’s theorem that explain the structure of nilpotent and solvable Lie algebras and their representations.

4.3.8. Theorem (Engel’s theorem). Let $V$ be a finite dimensional vector space over a field $k$. Let $g$ be a Lie subalgebra of $gl(V)$ such that each element of $g$ acts via nilpotent endomorphism on $V$. Then there exists $v \in V - \{0\}$ such that $g.v = 0$.

**Sketch of proof.** Induct on $\dim(g)$. **Claim 1:** $g$ contains a codimension 1 ideal. Let $h$ be a proper subalgebra of $g$ of maximal possible dimension. If $x \in h$, then $\text{ad}(x)$ takes $h$ to itself, and acts on $g/h$ as a nilpotent endomorphism (by homework exercise: nilpotent implies ad-nilpotent). So by induction, there exists $y \in g/h - \{0\}$ such that $\text{ad}(x)(y) = 0$ for all $x \in h$. In other words, $[h, y] \subseteq h$. This implies $(h + ky)$ is a subalgebra, forcing $h$ to be in fact an ideal in $g$ of codimension 1. Proving claim 1.

Let $W = \{v \in V: hv = 0\}$. By induction hypothesis, $W \neq 0$. Verify that $[y, h] \subseteq h$ implies $yW \subseteq W$. But then $y|_W$, being a nonzero nilpotent endomorphism, kills a nonzero vector $v$. So $gv = 0$. □

4.3.9. Theorem (Lie’s theorem). Let $V$ be a finite dimensional complex vector space. Let $g$ be a Lie subalgebra of $gl(V)$ which is solvable. Then there exists $v \in V - \{0\}$ such that $g.v = C.v$.

Before the proof, we need a lemma.

4.3.10. Lemma. Let $V$ be a representation of a Lie algebra $g$. Let $h$ be an ideal in $g$. Let $\lambda: h \to C$ be a linear map. Let $W = \{v \in V: xv = \lambda(x)v \text{ for all } x \in h\}$. Then $gW \subseteq W$.

**Sketch of proof.** If $W = 0$, we have nothing to prove. Let $w \in W - \{0\}$, $y \in g$. Define $U = \text{span}\{w, yw, y^2w, \cdots\}$. We shall show that the hypothesis $W \neq 0$, actually implies $\lambda([g, h]) = 0$. Choose a basis $B$ of $U$ such that $B \subseteq \{w, yw, y^2w, \cdots\}$. For any $x \in h$, note that

$$x(yw) = y(xw) + [x, y]w = \lambda(x)yw + \lambda([x, y])w.$$ 

and more generally,

$$x(y^kw) = yx(y^{k-1}w) + [x, y](y^{k-1}w).$$

By induction on $k$, it follows that for all $x \in h$, one has

$$x(y^kw) = \lambda(x)y^kw + \text{ a linear combination of } y^{k-1}w, y^{k-2}w, \cdots, w.$$ 

In other words, if $x \in h$, then $xU \subseteq U$ and the matrix of $x|U$ with respect to the basis $B$ is upper triangular with diagonal entries $\lambda(x)$. So $\text{tr}(x|_U) = \lambda(x)\dim(U)$. Now if $x \in h$, then $[x, y] \in h$ is a commutator of two endomorphisms of $U$, hence has trace 0. So $\lambda([x, y]) = 0$. It follows that $x(yw) = \lambda(x)yw + \lambda([x, y])w = \lambda(x)yw$. So $yv \in W$. □

Now we are ready to prove Lie’s theorem.

**Proof.** Induct on $\dim(g)$. Note that the dimension 1 case uses the fact that $C$ is algebraically closed, so any endomorphism as an eigenvector. Note that the preimage $h$ of a codimension 1 subspace in the abelian algebra $g/[g, g]$ is a codimension 1 ideal in $g$. By induction, $h$ acting on $V$ has a common eigenvector $v_0$. Define $\lambda: h \to C$ by $xv_0 = \lambda(x)v_0$ for all $x \in h$. Clearly $\lambda$ is a linear function on $h$. Let $W = \{v \in V: xv = \lambda(x)v \text{ for all } x \in h\}$. Choose $y \in g - h$. By lemma, $yW \subseteq W$. Choose an eigenvector $v$ for $y|_W$. Since $h + Cy = g$, this $v$ is a common eigenvector for $g$. □
Note that the conclusions of Engel and Lie’s theorems says that the elements of $\mathfrak{g}$ have a common eigenvector in $V$. Now we can identify $\mathfrak{g} \subseteq \mathfrak{gl}(V/v \mathbb{C})$ and again find a common eigenvector. By iterating this process, we find the following consequences of Engel’s (resp. Lie’s) theorems: there is a basis of $V$ with respect to which all matrices of all the elements of $\mathfrak{g}$ are strictly upper triangular (resp. upper triangular). This is why we remarked that $\mathfrak{n}_n(k)$ (resp. $\mathfrak{b}_n(k)$) are the “standard” examples of nilpotent (resp. solvable) Lie algebras. Another consequence of Lie’s theorem is that any irrep of a solvable Lie algebra is one dimensional.

4.3.11. Theorem. Let $\mathfrak{g}$ be a complex Lie algebra. Then irrep of $\mathfrak{g}$ has the form $V_0 \otimes L$ where $V_0$ is an irrep of the semisimple algebra $\mathfrak{g}/\text{rad}(\mathfrak{g})$ and $L$ is a one dimensional representation. If $\mathfrak{g}'$ is a subalgebra of $\mathfrak{g}$ that maps isomorphically onto $\mathfrak{g}/\text{rad}(\mathfrak{g})$, then any irrep of $\mathfrak{g}$ restricts to an irrep of $\mathfrak{g}'$ and any irrep of $\mathfrak{g}'$ extends to an irrep of $\mathfrak{g}$.

Because of this, understanding representation theory of complex reduces to understanding representations of the semisimple ones. The semisimple algebras have lots of interesting representations. First we have the following:

4.3.12. Theorem (Weyl’s theorem). Every finite dimensional complex representation of a complex semisimple Lie algebra $\mathfrak{g}$ is a direct sum of irreps.

So the category $\text{Rep-}\mathfrak{g}$ of finite dimensional complex representations of $\mathfrak{g}$ is indeed semisimple.

4.3.13. Jordan decomposition. Let $V$ be a finite dimensional complex vector space. Let $x \in \text{End}(V)$ be a linear map. Then one has the Jordan decomposition $x = x_s + x_n$ where $x_s$ is semisimple (i.e. diagonalizable) and $x_n$ is nilpotent, $x_s$ and $x_n$ commute and in fact are polynomials in $x$.

Let $\mathfrak{g}$ be a Lie algebra. Let $x \in \mathfrak{g}$ Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation. In general we cannot say anything about the diagonalizability of $\rho(x)$.

Exercise: Let $\mathfrak{g} = \mathbb{C}$ and $x \in \mathfrak{g} - \{0\}$. Find three representations $\rho_1, \rho_2, \rho_3 : \mathfrak{g} \to \mathfrak{gl}(\mathbb{C}^2)$ such that $\rho_1(x)$ is diagonalizable, $\rho_2(x)$ is nilpotent and $\rho_3(x)$ is neither.

However, if $\mathfrak{g}$ is semisimple and if $x \in \mathfrak{g}$ acts as a semisimple transformation in one faithful representation, then it acts as semisimple transformation in all representations. More precisely, we have the following.

4.3.14. Theorem (Preservation of Jordan decomposition). Let $\mathfrak{g}$ be a finite dimensional complex semisimple Lie algebra. Let $x \in \mathfrak{g}$. Then there exists $x_s, x_n \in \mathfrak{g}$ such that $x = x_s + x_n$ and for any finite dimensional representations $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ one has $\rho(x_s) = \rho(x)_s$ and $\rho(x_n) = \rho(x)_n$.

For now we shall omit the proof of Weyl’s theorem and preservation of Jordan decomposition. See the appendix in Fulton and Harris or Humphrey’s book.

4.3.15 (Exercise). (a) Let $\mathfrak{h}$ be an ideal in a Lie algebra $\mathfrak{g}$. Show that $\mathfrak{g}$ is semisimple if and only if $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are semisimple.

(b) Show that every semisimple Lie algebra is a direct sum of simple ideals.
4.4. Representations of \( \mathfrak{sl}_2(\mathbb{C}) \). For this section, let \( \mathfrak{g} = \mathfrak{sl}_2(\mathbb{C}) \). Let
\[
E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

4.4.1. Exercise. Verify that \( \mathfrak{g} \) can be described as a Lie algebra generated by \( E, F, H \) subject to the commutator relations:
\[
\]
Using these relations, verify directly that \( \mathfrak{g} \) is a simple Lie algebra.

4.4.2. Exercise. Consider the adjoint representation of \( \mathfrak{g} \) and the representation \( \text{sym}^3(\mathbb{C}^2) \) of \( \mathfrak{g} \). Write down the matrices for \( E, F, H \) with respect to suitable basis of these representations. Verify that \( H \) acts a diagonalizable matrix on both representations.

For now, we shall assume the facts about representations of semisimple Lie algebras. In particular, that,
\[
\circ \quad \text{any finite dimensional representation of } V \text{ decomposes as direct sum of irreps}
\circ \quad \text{The element } H \text{ acts as a diagonalizable transformation on any representation (since it is diagonalizable in the standard defining representation } \mathbb{C}^2.
\]

We want to describe the (finite dimensional) irreps of \( \mathfrak{g} \). Let \( V \) be an irrep of \( \mathfrak{g} \). Note that \( \mathfrak{h} = H \mathbb{C} \) is a maximal abelian subalgebra of \( \mathfrak{g} \) consisting of diagonalizable elements. We first consider the action of this subalgebra \( H \mathbb{C} \) on \( V \) and decompose \( V \) into eigenspaces for the action of \( H \). So we write
\[
V = \bigoplus_{\alpha} V_{\alpha} \quad \text{where } Hv = \alpha v \text{ for all } v \in V_{\alpha}.
\]
Here \( \alpha \) varies over complex numbers. Of course, since \( V \) is finite dimensional, \( V_{\alpha} \) is nonzero only for finitely many values of \( \alpha \). These are the eigenvalues of \( H|_V \) and are called the weights of the representation \( V \). The eigenspaces \( V_{\alpha} \) are called the (nonzero) weight spaces.

Next, we work out how \( E \) and \( F \) acts on these eigenspaces:

4.4.3. Lemma. One has \( E(V_\alpha) \subseteq V_{\alpha+2} \) and \( F(V_\alpha) \subseteq V_{\alpha-2} \).

Choose any weight \( \alpha_0 \in w(V) \). The lemma above implies that
\[
\bigoplus_{k \in \mathbb{Z}} V_{\alpha_0+2k}
\]
is stable under the action of \( E, F, H \). Since \( V \) is an irrep we must have
\[
V = \bigoplus_{k \in \mathbb{Z}} V_{\alpha_0+2k}
\]
Because \( V \) is finite dimensional, there is a largest integer \( m \) for which \( V_{\alpha_0+2m} \neq 0 \). Write \( n = \alpha_0 + 2m \). Observe that \( EV_n = 0 \). Choose a nonzero vector \( v \) in \( V_n \). Such a vector is called a highest weight vector.

4.4.4. Lemma. One has \( E(F^m v) = m(n - m + 1)F^{m-1}v \) for all \( m \geq 0 \).

Proof. Induct on \( n \). It is a direct computation using the commutator relations. \( \square \)

It follows that \( \text{span}\{v, Fv, F^2v, \cdots\} \) is stable under \( E, F, H \). So
\[
V = \text{span}\{v, Fv, F^2v, \cdots\}.
\]
Note that \( F^k v \in V_{n-2k} \). Since \( V \) is finite dimensional, \( V_{n-2k} \) must be 0 for sufficiently large \( k \), so \( F^k v = 0 \) for sufficiently large \( k \). Choose the smallest positive integer \( m \) such that
\( F^m v = 0 \). Now 4.4.4 implies \((n - m + 1) = 0\). So infact \( n = m - 1 \) is a non-negative integer! It follows that \( V = \mathbb{C}v + \mathbb{C}Fv + \cdots + \mathbb{C}F^m v \). Note that \( F^k v \in V_{n-2k} \) (i.e. these are eigenvectors of \( H \) with distinct eigenvalues), and hence the sum in the previous sentence must be a direct sum. So the direct sum decomposition \( V = \bigoplus \alpha V_{\alpha} \) takes the form:

\[
V = V_n \oplus V_{n-2} \oplus V_{n-4} \oplus \cdots \oplus V_{-n}
\]

Each weight space \( V_{n-2k} \) is one dimensional and is spanned by \( F^k v \). Of course the action of \( H \) preserves each weight space. The matrix \( E \) (resp \( F \)) acts as “raising operator” (resp. “lowering operator”) meaning that it increases (resp. decreases) weight of a vector by 2.

As for as existence of irreps, note that \( V_{(n)} = \text{Sym}^n (\mathbb{C}^2) \) is an irrep with exactly this weight space decomposition. In conclusion:

\textit{The irreps of} \( \mathfrak{sl}_2(\mathbb{C}) \) \textit{are} \( \text{sym}^n (\mathbb{C}^2) \) \textit{for} \( n \geq 2 \).

Recall that we have already described these irreps in the introduction. Before we conclude, we want to extract a couple of observations from our discussion above:

\textit{Each irrep has a highest weight vector (killed by} \( E \) \textit{) and the irrep is generated by successively applying the lowering operator} \( F \). \textit{Each irrep is determined completely by its set of weights, in fact by the highest weight} \( n \).

All these facts generalize for all complex simple Lie algebras.
4.5. **Representations of \( \mathfrak{sl}_3(\mathbb{C}) \).** Our next job is to describe the irreps of \( \mathfrak{sl}_3(\mathbb{C}) \). The basic plan follows the roadmap laid out by our analysis of \( \mathfrak{sl}_2(\mathbb{C}) \), but some new ideas are needed to carry it through. Fortunately, the discussion for \( \mathfrak{sl}_3(\mathbb{C}) \), generalizes in obvious manner to \( \mathfrak{sl}_n(\mathbb{C}) \). For this section let \( \mathfrak{g} = \mathfrak{sl}_n(\mathbb{C}) \) and \( n = 3 \).

4.5.1. **Exercise.** (Multiplying elementary matrices). Let \( E_{ij} \) be the \( n \times n \) matrix whose \((i, j)\)-th entry is 1 and all other entries are equal to 0. So the \( i, j\)-th entry of \( E_{ij} \) is \((E_{ij})_{k,l} = \delta_{ik}\delta_{jl}\). The \( E_{ij} \)'s form the standard basis for \( \mathfrak{gl}_n(k) \) (for any \( k \)) and they satisfy the relations:

\[
E_{ij}E_{kl} = \delta_{jk}E_{il}.
\]

Using these, verify that if \( h = \sum_i a_i E_{ii} \) is a diagonal matrix, then

\[
[h, E_{ij}] = (a_i - a_j)E_{ij}.
\]

4.5.2. **The weight space decomposition.** The starting point of our analysis of a representation \( V \) of \( \mathfrak{sl}_2(\mathbb{C}) \) was to consider the action of the diagonal matrix \( H \) on \( V \). While \( \mathfrak{sl}_2(\mathbb{C}) \) has upto scale a unique diagonal matrix, for \( n \geq 3 \), there is a \((n-1)\) dimension worth of them. Let \( \mathfrak{h} \) be the set of diagonal matrices in \( \mathfrak{g} \). Our starting point will be to replace \( H \mathbb{C} \) with \( \mathfrak{h} \) and make appropriate modifications (in both cases, these are maximal abelian subalgebras consisting of diagonalizable elements).

Let

\[
H_1 = E_{11} - E_{22}, \quad H_2 = E_{22} - E_{33}.
\]

These two matrices form a basis of \( \mathfrak{h} \). By invariance of Jordan decomposition, \( H_1|_V \) and \( H_2|_V \) are diagonalizable and these matrices commute. So \( H_1|_V \) and \( H_2|_V \) are simultaneously diagonalizable. Since \( H_1, H_2 \) form a basis of \( \mathfrak{h} \), it follows that:

*We can choose a basis of \( V \) consisting of simultaneous eigenvectors of all elements of \( \mathfrak{h} \).*

Let \( U \) be a nonzero eigenspace. So if \( h \in H \), then \( hv \) is a scalar multiple of \( v \) for all \( v \in U \). This scalar depends on \( h \), so call it \( \alpha(h) \). In other words, there exists, \( \alpha : \mathfrak{h} \rightarrow \mathbb{C} \) such that \( hv = \alpha(h)v \) for all \( h \in H \) and all \( v \in U \). Verify that \( \alpha \) is a linear map, that is, \( \alpha \in \mathfrak{h}^* \). We call this \( \alpha \) an eigenvalue for \( \mathfrak{h} \) corresponding to eigenspace \( U \). Write \( U = V_\alpha \). In summary, we find that \( V \) can be decomposed into eigenspaces:

\[
V = \bigoplus_{\alpha \in \mathfrak{h}^*} V_\alpha
\]

where

\[
hv = \alpha(h)v \quad \text{for all} \quad v \in V, h \in \mathfrak{h}.
\]

This is called the weight space decomposition of \( V \). Since \( V \) is finite dimensional, \( V_\alpha \neq 0 \) only for finitely many \( \alpha \)'s. These are called the weights of \( V \) and corresponding \( V_\alpha \)'s are the (non-zero) weight spaces.

4.5.3. **The root space decomposition** Let us apply the weight space decomposition to the adjoint representation of \( \mathfrak{g} \) on \( \mathfrak{g} \). Define \( L_j : \mathfrak{h} \rightarrow \mathbb{C} \) are the linear maps defined by

\[
L_j(\text{diag}(a_1, \cdots, a_n)) = a_j.
\]

Note that \( L_1, \cdots, L_n \) span \( \mathfrak{h}^* \) subject to \( L_1 + \cdots + L_n = 0 \). Thus

\[
\mathfrak{h}^* \simeq (\mathbb{C}l_1 \oplus \cdots \oplus \mathbb{C}l_n)/(l_1 + \cdots + l_n = 0).
\]

via the map \( l_j \rightarrow L_j \). From the commutator relation in the exercise above, we have

\[
ad(h)E_{ij} = (L_i - L_j)(h)E_{ij}.
\]
So each $E_{ij} \mathbb{C}$ is a one dimensional weight space for the adjoint representation with weight $(L_i - L_j)$ and $\mathfrak{h}$ is the zero-th weight space (since $\mathfrak{h}$ is commutative). The nonzero weights of the adjoint representation are called the \textit{roots} of $\mathfrak{g}$. So the roots of $\mathfrak{g}$ are:

$$R = \{(L_i - L_j) : i \neq j\}.$$ 

The corresponding weight spaces are called the root spaces. The root spaces are

$$\mathfrak{g}_{L_i - L_j} = \mathbb{C} E_{ij}$$

and the zero-th weight space is:

$$\mathfrak{g}_0 = \mathfrak{h}.$$

The weight space decomposition of $\mathfrak{g}$ (as adjoint representation) is called the \textit{root space decomposition}:

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \left( \oplus_{r \in R} \mathfrak{g}_r \right).$$

\textbf{4.5.4. The weight diagrams.} Let $\mathfrak{g}_R^*$ be the two dimensional real vector subspace of $\mathfrak{h}$ spanned by $L_1, L_2, L_3$. We represent the vector $L_1$ in this vector space as the unit vector in the $x$-axis direction and the vectors $L_2$ and $L_3$ as the vectors obtained by rotating $L_1$ anticlockwise by 120 and 240 degrees. We shall see later that there is a natural inner product in the inner product on $\mathfrak{h}_\mathbb{R}$ with respect to which $L_1, L_2, L_3$ has norm 1. We will soon find that the weights of a representation of $\mathfrak{g}$ lie in the lattice

$$\Lambda_W = \mathbb{Z}\langle L_j \rangle.$$ 

This lattice is called the \textit{weight lattice}. A finite dimensional representation of $\mathfrak{g}$ is thus represented by a 2 dimensional “weight diagram” which consists of a finite set of dots (with multiplicities indicating the dimension of the corresponding weight space) in the plane $\mathfrak{g}_R^*$, with each dot actually lying in the lattice $\Lambda_W$.

\textbf{Exercise.} Draw the weight diagrams of the standard representation $\mathbb{C}^3$, its dual and of the adjoint representation $\mathfrak{g}$.

Let us now get back to a finite dimensional representation $V$ of $\mathfrak{g}$ and work out how the root spaces act on the weight spaces of $V$.

\textbf{4.5.5. Lemma.} Let $V$ be a finite dimensional representation of $\mathfrak{g}$ with weight space decomposition $V = \oplus_{\alpha \in \mathfrak{h}} V_\alpha$. Then

$$[\mathfrak{g}_\alpha, V_\beta] \subseteq V_{\alpha + \beta}.$$ 

In particular

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha + \beta}.$$ 

\textbf{Proof.} Exercise. \qed

\textbf{4.5.6. Positivity and highest weight.} Define the root lattice $\Lambda_R$ to be the $\mathbb{Z}$-span of the roots of $\mathfrak{g}$. Now assume that the representation $V$ is irreducible. Choose a nonzero weight space $V_{\alpha_0}$ of $V$. The lemma above implies that $\oplus_{\lambda \in \Lambda_R} V_{\alpha_0 + \lambda}$ is stable under each root space (and of course, also stable under $\mathfrak{h}$). So

$$V = \oplus_{\lambda \in \Lambda_R} V_{\alpha_0 + \lambda}.$$
In other words, the weight of an irrep all lie in a translate of the root lattice \( \Lambda_R \). Next, we want to define a notion of the highest weight that is killed by half of the root spaces (to be called the positive root spaces). For this, choose a linear functional
\[
l : \mathfrak{h}^*_R \to \mathbb{R}.
\]
and extend it linearly to \( \mathfrak{h} \). We choose the linear functional generic enough so that \( l \) does not vanish on \( \Lambda_R \). Since any two root of \( V \) differ by an element of \( \Lambda_R \), it follows that the real part of functional \( l \) takes distinct values on all the weights of \( V \). Since \( V \) is finite dimensional, there exists a unique weight \( \alpha \) of \( V \) for which \( \text{Re}(l(\alpha)) \) is maximal. This is called the highest weight of \( V \) and the corresponding weight space \( V_\alpha \) is called the highest weight space. Choose \( v \in V_\alpha - \{0\} \). This is called a highest weight vector.

To be specific, let us choose \( l \) as follows. Let \( a > b > c \) be generic real numbers such that \( a + b + c + 0 \) and let
\[
l(a_1 L_1 + a_2 L_2 + a_3 L_3) = a a_1 + b a_2 + c a_3.
\]
All the discussion below is contingent on this choice of \( L \). Let \( R_+ \) (resp \( R_- \)) be the set of roots \( \alpha \) for which \( \text{Re}(l(\alpha)) > 0 \) (resp. \( < 0 \)). These are called the positive (resp. negative) roots. One has
\[
R_+ = \{ L_1 - L_2, L_2 - L_3, L_1 - L_3 \}, \quad \text{and} \quad R_- = -R_+ = R - R_+.
\]
Define the positive and negative part of the Lie algebra by
\[
\mathfrak{g}_+ = \bigoplus_{\alpha \in R_+} \mathfrak{g}_\alpha \quad \text{and} \quad \mathfrak{g}_- = \bigoplus_{\alpha \in R_-} \mathfrak{g}_\alpha.
\]
Observe that \([\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha + \beta}\) implies that \( \mathfrak{g}_+ \) and \( \mathfrak{g}_- \) are subalgebras. One has the triangular decomposition of the Lie algebra:
\[
\mathfrak{g} = \mathfrak{g}_+ \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_-.
\]
The root spaces in \( \mathfrak{g}_+ \) (resp. \( \mathfrak{g}_- \)) are \( \mathbb{C}E_{12}, \mathbb{C}E_{23} \) and \( \mathbb{C}E_{13} \) (resp. \( \mathbb{C}E_{21}, \mathbb{C}E_{32} \) and \( \mathbb{C}E_{31} \)). These are called the positive (resp. negative) root spaces. We shall call the vectors \( E_{12}, E_{23}, E_{13} \) raising operators and the other three lowering operators.

Now we have the following lemma which corresponds to the statement for \( \mathfrak{sl}_2(\mathbb{C}) \) that the highest weight vector of an irrep is killed by the raising operator \( E \) and the irrep is spanned by taking the highest weight vector and applying the lowering operator repeatedly.

4.5.7. Lemma. Let \( V \) be a finite dimensional irrep of \( \mathfrak{g} \). If \( v \) is a highest weight vector of \( V \), then \( \mathfrak{g}_+ v = 0 \). The irrep \( V \) is generated by the images of \( V \) under the successive application of the lowering operators.

Sketch of proof. Choose the weight \( \alpha \) of \( V \) for which \( \text{Re}(l(\alpha)) \) is maximal and choose \( v \in V_\alpha - \{0\} \). Let \( E \in \mathfrak{g}_\beta \) for some \( \beta \in R_+ \). Then \( E v \in V_{\alpha + \beta} \) and \( \text{Re}(l(\alpha + \beta)) > \text{Re}(l(\beta)) \), so we must have \( V_{\alpha + \beta} = 0 \) and hence \( E v = 0 \). So \( \mathfrak{g}_+ v = 0 \).

Let \( V' \) be the subspace of \( V \) spanned by the vectors obtained by repeatedly applying lowering operators to \( v \). It suffices to show that \( V' \) is stable under the raising operators. First, note that it is enough to consider the raising (resp lowering) operators \( E_{12}, E_{23} \) (resp. \( E_{21}, E_{32} \)) since the third one is a commutator of these two.

\[\text{To generalize this part for other Lie algebras, we shall need the notion of simple roots.}\]
We do this inductively as follows. Let $W_n$ be the subspace of $V$ spanned by vectors of the form $E_1E_2\cdots E_kv$ where $k \leq n$ and each $E_j \in \{E_{21}, E_{32}\}$. Using the commutator relations $[g_\alpha, g_\beta] \subseteq g_{\alpha+\beta}$ one verifies that $E_{12}$ and $E_{23}$ takes $W_n$ into $W_{n-1}$.

In fact the proof of the lemma above in fact proves the following stronger statement:

4.5.8. **Lemma.** If $V$ is any representation of $\mathfrak{g}$ and $v$ is a highest weight vector in $V$, then the subspace $W$ of $V$ generated by repeatedly applying the lowering operators is irreducible.

**Proof.** Let $W$ be Let $\alpha = \text{weight}(v)$. Applying any lowering operator to $v$ gives a vector of lower weight, so $W_\alpha$ is one dimensional. If $W = W' \oplus W''$ for some subrepresentation $W',W''$, then verify that $W_\alpha = W_\alpha \oplus W_\alpha''$ since the projections to $W'$ and $W''$ commute with the action of $\mathfrak{h}$. Since $W_\alpha$ is one dimensional, we must have $W'_\alpha = 0$ or $W''_\alpha = 0$. So $v \in W'_\alpha$ or $W''_\alpha$ (and the other one is zero), which implies $W = W'_\alpha$ or $W''_\alpha$.

4.5.9. **The root $\mathfrak{sl}_2$'s.** Next we want to understand the set of weights that appear in an irrep $V$ of $\mathfrak{g}$. The basic tool for this is to consider $V$ as representation of some subalgebras of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$, called the “root $\mathfrak{sl}_2$'s”. For each positive root $\alpha \in R$, let

$$\mathfrak{sl}_2,\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$$

For each root $\alpha \in R$, we shall choose $E_{\pm \alpha} \in \mathfrak{g}_{\pm \alpha} - \{0\}$ and let $H_\alpha$ be the unique element of $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ such that $\alpha(H_{ij}) = 2$. Explicitly, for the root $\alpha = L_i - L_j$, we choose

$$E_\alpha = E_{ij}, \ E_{-\alpha} = E_{ji}, \text{ and } H_\alpha = [E_\alpha, E_{-\alpha}] = E_{ii} - E_{jj}.$$ 

Note $H_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{h}$. One has

$$\mathfrak{sl}_2,\alpha = C E_\alpha \oplus C E_{-\alpha} \oplus C H_\alpha$$

Verify that for each $\alpha \in R$, the subspace $\mathfrak{sl}_2,\alpha$ is a subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. These are called the root $\mathfrak{sl}_2$'s.

4.5.10. **Restricting $\mathfrak{sl}_3(\mathbb{C})$ irrep to root $\mathfrak{sl}_2$.** Let $V$ be an irrep of $\mathfrak{g}$. Fix a positive root $\alpha \in R$ and consider $V$ as a representation of $\mathfrak{sl}_2,\alpha$. Let $\beta$ be a weight of $V$. Observe that

$$W = \oplus_{n \in \mathbb{Z}} V_{\beta+n\alpha}$$

is a (not necessarily irreducible) representation of $\mathfrak{sl}_2,\alpha$. Choose one of the weights $\gamma = \beta+n\alpha$. We ask: How does $H_\alpha$ act on $V_\gamma$.

Answer: $H_\alpha$ acts as multiplication by $\gamma(H_\alpha) = \beta(H_\alpha) + 2n$ since $\alpha(H_\alpha) = 2$.

In particular, if $\gamma \neq \gamma'$, then $\gamma(H_\alpha) \neq \gamma'(H_\alpha)$. So:

$V_{\beta+n\alpha}$ are the weight spaces of the $\mathfrak{sl}_2,\alpha$ representation $W$ with weight $(\beta+n\alpha)(H_\alpha)$.

Since we completely understand what the weight spaces of a finite dimensional $\mathfrak{sl}_2(\mathbb{C})$ representation look like, this has two important consequences:

- $\gamma(H_\alpha) \in \mathbb{Z}$ for all $\alpha \in R$.
- If $\gamma' = \gamma + n\alpha$ such that $\gamma(H_\alpha) = -\gamma'(H_\alpha)$, then dim $V_\gamma = \dim V_{\gamma'}$.
- The weights of the form $\{\beta + n\alpha: n \in \mathbb{Z}\}$ form an “unbroken string”.

The first two of these in turn respectively imply the “integrality of weights” and “weyl group symmetry” of the weights of $\mathfrak{g}$-irreps.
4.5.11. **Integrality of weights.** Let $\gamma$ be any weight of a $g$-irrep $V$. Then we have $\gamma(H_\alpha) \in \mathbb{Z}$ for each root $\alpha \in R$. Write $\gamma = aL_1 + bL_2$ with $a, b \in \mathbb{C}$. Then $\gamma(H_{L_1-L_2}) = a$ and $\gamma(H_{L_2-L_3}) = b$. So $a, b \in \mathbb{Z}$. In other words, the weights of $V$ all lie in the Weight lattice $\Lambda_W = \mathbb{Z}L_1 + \mathbb{Z}L_2$ (and as we already know, any two of them differ by the root lattice $\Lambda_R$).

4.5.12. **Weyl group symmetry.** In any finite dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$ the weights $n$ and $-n$ appear with same multiplicity (where multiplicity of a weight means the dimension of the weight space). This gave us that if $\gamma$ and $\gamma' = \gamma + n\alpha$ are two weights of an irrep $V$ of $g$ such that $\gamma(H_\alpha) = -\gamma'(H_\alpha)$, then $\dim(V_\gamma) = \dim(V_{\gamma'})$.

Note that

$$(\gamma + n\alpha)(H_\alpha) = \gamma(H_\alpha) + 2n.$$ 

So two weights $\gamma$ and $\gamma' = \gamma + n\alpha$ satisfy $\gamma'(H_\alpha) = -\gamma(H_\alpha)$ if and only if $\gamma(H_\alpha) + 2n = -\gamma(H_\alpha)$, that is,

$$\gamma' = \gamma - \gamma(H_\alpha)\alpha.$$ 

Define $R_\alpha : g^* \rightarrow g^*$ by

$$R_\alpha(\gamma) = \gamma - \gamma(H_\alpha)\alpha.$$ 

Thus we have $\dim(V_\gamma) = \dim(V_{R_\alpha(\gamma)})$ for all $\alpha \in R_+$. Note that $R_\alpha$ fixes the hyperplane $H_\alpha^+ := \{ \gamma \in h^* : \gamma(H_\alpha) = 0 \}$ and takes $\alpha$ to $-\alpha$. The linear map $R_\alpha$ is called the reflection in $\alpha$ (more precisely, at the moment, it should be called the pseudoreflection in $(\alpha, H_\alpha)$). Observe that $R_\alpha$ preserves the root lattice, the weight lattice and the real vector space $h_g$. On our weight diagrams, $R_\alpha$ indeed acts as orthogonal reflection across the hyperplane orthogonal to $\alpha$. The group $W$ generated by these reflections is called the Weyl group of $g$. The We thus have

$$\dim(V_\gamma) = \dim(V_{w(\gamma)}) \text{ for all } w \in W.$$ 

All the discussion so far generalizes for any simple Lie algebra with some extra work and we shall work through it after filling in the general background material on complete reducibility, killing form etc.

Finally it is an easy exercise to verify that the highest weight $\alpha$ of an irrep lies in the cone $C$ bounded by the hyperplanes $H_{L_1}^+$ and $H_{L_2-L_3}^+$ in the half plane $\{ y \in h_g^* : l(y) > 0 \}$. This cone is called the Weyl chamber. In other words, the highest weight of an irrep has the form $aL_1 - bL_3$ for $a, b \in \mathbb{Z}_{\geq 0}$. The above discussion proves much of the following theorem that describes the finite dimensional irreps of $\mathfrak{sl}_3(\mathbb{C})$.

4.5.13. **Theorem.** For each $a, b \in \mathbb{Z}_{\geq 0}$, there exists a unique irreducible representation $\Gamma_{aL_1-bL_3}$ of $g = \mathfrak{sl}_3(\mathbb{C})$ with highest weight $aL_1 - bL_3$ and these are all the finite dimensional irreps of $g$. Write $\alpha = aL_1 - bL_3$. The set of weights of $\Gamma_\alpha$ is the subset of $\Lambda_W$ described by

$$\text{convex hull}\{ w\alpha : w \in W \} \cap (\alpha + \Lambda_R).$$

Note that the convex hull is a hexagon and the weights are the points in this hexagon that lie in the weight lattice and differ from the highest weight by an element of the root lattice.

**Sketch of completion of the proof.** We already know that each irrep has a highest weight, which lies in $C \cap \Lambda_W$, that the irrep is determined by its highest weight and that the set of weights of an irrep with highest weight $\alpha$ is as described in the theorem. It remains to actually construct the irreps.

44
Let $a, c \in \mathbb{Z}_{\geq 0}$, $\alpha = aL_1 - bL_3$. We want to construct an irrep $\Gamma_\alpha$ with highest weight $\alpha$. Write $\Gamma_\alpha = \Gamma_{a,b}$. Let $V = \mathbb{C}^3$ be the standard 3-dimensional representation of $\mathfrak{g}$. One verifies that $V = \Gamma_{1,0}$ and $V^* = \Gamma_{0,1}$. Let $u \in V$ and $w \in V^*$ be highest weight vectors with weights $L_1$ and $-L_2$ respectively. Let $U_{a,b} = \text{sym}^a(V) \otimes \text{sym}^b(V^*)$. Verify that $v = u^a \otimes w^b$ is the highest weight vector of with weight $aL_1 - bL_3$. So applying the lowering opeartors repeatedly to $v$, one obtains a subrepresentation of $U_{a,b}$ that is an irrep of highest weight $aL_1 - bL_3$. \hfill $\square$

4.5.14. **Remark.** One can actually make more precise statements. First of all verify that the set of weights of $\text{sym}^a(V)$ (resp. $\text{sym}^a(V^*)$) are the elements of $aL_1 + \Lambda_W$ that lie in the equilateral triangle with vertices $aL_1, aL_2, aL_3$ (resp. $-aL_1, -aL_2, -aL_3$) and that these weights all occur with multiplicity 1. These are two cases when the hexagon becomes a triangle. Since an irrep with highest weight $aL_1$ (resp. $-aL_3$) must contain all the weights in these triangles, a consequence of multiplicity 1 is that $\text{sym}^a(V)$ and $\text{sym}^a(V^*)$ are actually irreducible. In other words

$$\Gamma_{a,0} = \text{Sym}^a(V) \text{ and } \Gamma_{0,a} = \text{sym}^a(V^*).$$

It remains to identify the irreps $\Gamma_{a,b}$ inside $U_{a,b} = \Gamma_{a,0} \otimes \Gamma_{0,b}$ when $a, b$ are both strictly positive. One has a natural contraction map

$$i_{a,b} : U_{a,b} \rightarrow U_{a-1,b-1}$$

given by

$$i_{a,b} : (v_1 \cdots v_a) \otimes (v_1^* \cdots v_b^*) \mapsto \sum_{i,j} \langle v_j^*, v_i \rangle (v_1 \cdots \hat{v}_i \cdots v_a) \otimes (v_1^* \cdots \hat{v}_j^* \cdots v_b^*)$$

where $\langle v_j^*, v_i \rangle$ is the natural pairing between a vector space and its dual and a hat over an element means we omit it. Verify that $i_{a,b}$ is a map of $\mathfrak{g}$-representations (also called $\mathfrak{g}$-equivariant), so $\ker(i_{a,b})$ is a subrepresentation of $U_{a,b}$. On the other hands, $aL_1 - bL_3$ is not a weight of $U_{a-1,b-1}$ since it has highest weight $(a - 1)L_1 - (b - 1)L_3$. So the highest weight vector of $U_{a,b}$ must be in the kernel of $i_{a,b}$ which implies $\Gamma_{a,b} \subseteq \ker i_{a,b}$. In fact

$$\Gamma_{a,b} = \ker(i_{a,b}).$$

We shall not prove this in general, but it will be a good exercise to verify this by hand for the example $a = 2, b = 1$. In other words, compute the weights of $U_{2,1} = \text{sym}^2(V) \otimes V^*$ and consider the contraction $i_{2,1} : U_{2,1} \rightarrow V$ to verify directly that $\ker(i_{2,1})$ is an irrep. Thus the decomposition of $U_{2,1}$ into irreps is given by

$$U_{2,1} \simeq \Gamma_{2,1} \oplus V.$$  

More generally, one actually has for $a \geq b$,

$$U_{a,b} \simeq \bigoplus_{i=0}^{b} \Gamma_{a-i,b-i}.$$
4.6. **Killing form, Jordan decomposition, complete reducibility.** We need some preparation. First we recall the Jordan decomposition from Linear algebra.

4.6.1. **Theorem** (Jordan normal form). Let $V$ be a finite dimensional complex vector space. Any $x \in \text{End}(V)$ can be uniquely written in the form $x = x_s + x_n$ such that $x_s$ and $x_n$ are commuting elements of $\text{End}(V)$ and such that $x_s$ is diagonalizable, $x_n$ is nilpotent. In fact $x_s$ and $x_n$ are polynomials in $x$.

**Proof.** Choose a basis of $V$ so that the matrix of $x$ is in the Jordan form and identify linear maps with matrices using this basis. Let $x_s$ be the diagonal part of $x$ and $x_n$ be the upper triangular part. To show that $x_s$ is a polynomial in $x$, let $\prod_i (x - \lambda_i)^{a_i}$ be the characteristic polynomial of $x$. Let $V_i = \ker(x - \lambda_i)^{a_i}$ be the generalized eigenspace corresponding to the eigenvalue $\lambda_i$. (In fact if $x$ is in the Jordan form, then $V_i$ can be chosen to be the span of the basis elements corresponding to the Jordan blocks of type $\lambda_i$). One has $V = \bigoplus_i V_i$. By chinese remainder theorem, choose $p(t) \in \mathbb{C}[t]$ such that $p(t) \equiv \lambda_i \mod (t - \lambda_i)^{a_i}$ and $p(t) \equiv 0 \mod t$ (this is redundant if some $\lambda_j = 0$). Now, note that $p(x)|_{V_i} = \lambda_i \text{id}_{V_i}$ for each $i$. So $p(x) = x_s$. This proves that there exists $x_s$ and $x_n$ as required and that they are polynomials in $x$.

To prove the uniqueness, let $x = s + n$ be a decomposition of $x$ into commuting endomorphisms where $s$ is diagonalizable and $n$ is nilpotent. Then $s$ and $n$ commute with $x$ and hence they commute with $x_s$ and $x_n$. The sum of two commuting diagonalizable (resp. nilpotent) matrices is again diagonalizable (resp. nilpotent). So $(x_s - s) = (n - x_n)$ is diagonalizable and nilpotent at the same time, hence must be equal to 0. \qed

4.6.2. **Lemma** (A corollary to Engel’s theorem). A finite dimensional Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\text{ad}(x)$ is nilpotent for all $x \in \mathfrak{g}$.

**Proof.** Exercise. \qed

4.6.3. **Exercise.** Let $V$ be a finite dimensional complex vector space.

(a) Show that the restriction of a diagonalizable endomorphism $x \in \text{End}(V)$ to a $x$-stable subspace $W$ (which means $xW \subseteq W$) is again diagonalizable.

(b) If $a \in \mathfrak{gl}(V)$ is a diagonalizable, then $[a, \cdot] \in \text{End}(\mathfrak{gl}(V))$ is diagonalizable.

(c) Let $\mathfrak{g}$ be a Lie subalgebra of $\mathfrak{gl}(V)$. Show that $\text{ad}(x_s) = \text{ad}(x_s)$ and $\text{ad}(x_n) = \text{ad}(x_n)$ for all $x \in \mathfrak{g}$.

**Sketch of proof.** (a) By induction on $n$, one shows that if $v_1 + \cdots + v_n \in W$ where $v_j$’s are eigenvectors of $x$ corresponding to distinct eigenvalues, then each $v_j \in W$. It follows that $W$ is spanned by the set of eigenvectors of $x$ that lie in $W$. Hence $W$ has a basis consisting of eigenvectors of $x$.

(b) Identify $E = \mathfrak{gl}(V)$ with $n \times n$ matrices and just compute $\text{ad}(a)(E_{ij})$ where $E_{ij}$ are the elementary matrices.

(c) Let $x = x_s + x_n$ be the Jordan decomposition of some $x \in \mathfrak{g}$. By (b), $T = [x_s, \cdot] \in \text{End}(\mathfrak{gl}(V))$ is diagonalizable and hence $\text{ad}(x_s) = T|_{\mathfrak{g}}$ is diagonalizable by (a). By a previous exercise, $\text{ad}(x_n)$ is nilpotent. Since $x_s$ and $x_n$ commute, so does $\text{ad}(x_s)$ and $\text{ad}(x_n)$. Now use the uniqueness of the Jordan decomposition. \qed

4.6.4. **Definition.** In this section $\mathfrak{g}$ will be a finite dimensional complex Lie algebra. Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite dimensional complex representation of $\mathfrak{g}$. Let $B_\rho$ (or $B_V$) be the
symmetric bilinear form on \( \mathfrak{g} \) defined by
\[
B_\rho(x, y) = \text{tr}(\rho(x) \circ \rho(y)).
\]
The bilinear form \( B(x, y) = B_{ad}(x, y) = \text{tr}(\text{ad}(x) \circ \text{ad}(y)) \) is called the killing form of \( \mathfrak{g} \).

4.6.5. **Exercise.** Verify that
\[
B_\rho([x, y], z) = B_\rho(x, [y, z])
\]
and if \( \mathfrak{a} \) is an ideal in \( \mathfrak{g} \), then so is \( \mathfrak{a}^\perp \) (the orthogonal complement of \( \mathfrak{a} \) respect to \( B_\rho \)).

The Killing form can detect semisimplicity and solvability of a Lie algebra. Recall the commutator subalgebra \( D(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}] \) of a Lie algebra \( \mathfrak{g} \). This is an ideal in \( \mathfrak{g} \) and \( \mathfrak{g}/D(\mathfrak{g}) \) is abelian. So \( \mathfrak{g} \) is solvable, if and only if \( D(\mathfrak{g}) \) is. Also note that an ideal \( \mathfrak{h} \) in \( \mathfrak{g} \) is solvable if and only if \( \text{ad}(\mathfrak{h}) \) is, since the kernel of the adjoint representation is abelian, hence solvable.

4.6.6. **Theorem.** TFAE: (a) \( \mathfrak{g} \) is solvable. (b) \( B(\mathfrak{g}, D(\mathfrak{g})) = 0 \). (c) \( B(D(\mathfrak{g}), D(\mathfrak{g})) = 0 \).

**Proof.** Suppose \( \mathfrak{g} \) is solvable. Then from Lie’s theorem, we know that one can choose a basis of \( \mathfrak{g} \) such that \( \text{ad}(x) \) is upper triangular for all \( x \in \mathfrak{g} \). Note that if \( x \in D(\mathfrak{g}) \) and \( y \in \mathfrak{g} \), then \( \text{ad}(x) \circ \text{ad}(y) \) is strictly upper triangular since \( \text{ad}(y) \) is upper triangular and \( \text{ad}(x) \) is strictly upper triangular. It follows that \( \text{tr}(\text{ad}(x) \circ \text{ad}(y)) = 0 \), that is \( B(\mathfrak{g}, D(\mathfrak{g})) = 0 \). This proves (a) implies (b). Clearly (b) implies (c). Assume (c). The Cartan’s criteria proved below, applied to the Lie algebra \( \text{ad}(D(\mathfrak{g})) \subseteq \mathfrak{gl}(\mathfrak{g}) \) shows that \( \text{ad}(D(\mathfrak{g})) \) is abelian, so \( D(\mathfrak{g}) \) is solvable, as is \( \mathfrak{g} \).

4.6.7. **Theorem** (Cartan’s criteria). If \( \mathfrak{g} \) is a subalgebra of \( \mathfrak{gl}(V) \) such that \( B_V(x, y) = \text{tr}(xy) = 0 \) for all \( x, y \in \mathfrak{g} \), then \( \mathfrak{g} \) is solvable.

**Proof.** Since \( \mathfrak{g}/D(\mathfrak{g}) \) is abelian (hence solvable), it suffices to show that \( D(\mathfrak{g}) \) is solvable. We shall show something stronger, that \( D(\mathfrak{g}) \) is a nilpotent ideal. By the corollary to Engel’s theorem, it is enough to show that \( \text{ad}_{D(\mathfrak{g})}(x) \) is nilpotent for all \( x \in D(\mathfrak{g}) \).

Choose \( x \in D(\mathfrak{g}) \) so \( x = \sum_j [y_j, z_j] \) for some \( y_j, z_j \in \mathfrak{g} \). Work in a basis of \( V \) in which the matrix of \( x \) is in Jordan form \( x = x_s + x_n \) where \( D = x_s = \text{diag}(\lambda_1, \ldots, \lambda_n) \) and \( x_n \) is strictly upper triangular. It suffices to show that \( D = 0 \), because that would imply \( x \) is nilpotent, hence \( \text{ad}(x) \) is nilpotent. Now \( D = 0 \) if and only if \( \sum_j \bar{\lambda}_j \lambda_j = 0 \) and
\[
\sum_j \bar{\lambda}_j \lambda_j = \text{tr}(D \cdot x_s) = \text{tr}(D \cdot x) = \sum_j \text{tr}(D \circ [y_j, z_j]) = \sum_j \text{tr}([D, y_j] \circ z_j)
\]
For the first equality, note that \( x_n \) is strictly upper triangular and \( D \) is diagonal. So it would be enough to prove the claim that \( \text{ad}(D)(\mathfrak{g}) \subseteq \mathfrak{g} \) since then each trace in the last sum is equal to zero by the hypothesis of the theorem.

Observe that \( \text{ad}(\bar{D}) \) and \( \text{ad}(D) \) act on \( \mathfrak{gl}(V) \) via complex conjugate diagonal matrices with respect to the standard basis of \( \mathfrak{gl}(V) \). So there is a polynomial \( p \) such that \( \text{ad}(\bar{D}) = p(\text{ad}(D)) = p(\text{ad}(x_s)) \) (the equality holds in \( \mathfrak{gl}(V) \) hence in \( \mathfrak{gl}(\mathfrak{g}) \)). Finally \( \text{ad}(x_s) = \text{ad}(x)_s \) is a polynomial in \( \text{ad}(x) \). Since \( \text{ad}(x) \) takes \( \mathfrak{g} \) to \( \mathfrak{g} \), we have proved the claim. \( \square \)

If \( b : V \times V \to \mathbb{C} \) is a symmetric bilinear form, define \( \text{rad}(b) = \{ x \in V : b(x, V) = 0 \} \). One says that \( b \) is nondegenerate if \( \text{rad}(b) = 0 \).

4.6.8. **Theorem.** A Lie algebra \( \mathfrak{g} \) is semisimple if and only if the killing form \( B \) is nondegenerate.
sketch of proof. Assume \( \mathfrak{g} \) is semisimple. Verify that \( \text{rad}(B) \) is an ideal in \( B \). By Cartan’s criteria, \( \text{ad}(\text{rad}(B)) \) is solvable, so \( \text{rad}(B) \) is a solvable ideal. So if \( \mathfrak{g} \) is semisimple, then \( \text{rad}(B) = 0 \) and hence \( B \) is nondegenerate.

Conversely suppose \( B \) is nondegenerate. Let \( \mathfrak{a} \) be any abelian ideal in \( \mathfrak{g} \). If \( x \in \mathfrak{a} \) and \( y \in \mathfrak{g} \), then verify that \( A = \text{ad}(x) \circ \text{ad}(y) \) maps \( \mathfrak{g} \) into \( \mathfrak{a} \) and \( \mathfrak{a} \) to 0, which implies \( \text{tr}(A) = 0 \). This implies \( \mathfrak{a} \subseteq \text{rad}(B) \), so \( \mathfrak{a} = 0 \). Since \( \mathfrak{g} \) has no nonzero abelian ideal, \( \mathfrak{g} \) is semisimple. \( \square \)

4.6.9. Corollary. Let \( \mathfrak{g} \) be a finite dimensional complex semisimple Lie algebra. Then \( \mathfrak{g} \) is a direct sum of simple ideals. Ideals and homomorphic images of \( \mathfrak{g} \) are semisimple and \( \mathfrak{g} = D(\mathfrak{g}) \). The adjoint representation \( \text{ad}: \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \) is an isomorphism from \( \mathfrak{g} \) onto \( \text{Der}(\mathfrak{g}) \). Any one dimensional representation of \( \mathfrak{g} \) is trivial (since \( \mathfrak{g} = D(\mathfrak{g}) \)).

**Proof.** Exercise.

4.6.10. Corollary. Let \( \mathfrak{g} \) be a Lie subalgebra of \( \mathfrak{gl}(V) \). If \( \mathfrak{g} \) is semisimple, then the form \( B_V : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C} \) defined by \( B_V(x,y) = \text{tr}(xy) \) is nondegenerate.

**Proof.** Note that \( \mathfrak{r} = \text{rad}(B_V) \) is the orthogonal complement of \( \mathfrak{g} \) with respect to \( B_V \), hence is an ideal in \( \mathfrak{g} \). The ideal \( \mathfrak{r} \) is solvable by Cartan’s criteria, hence must be 0. \( \square \)

Next we want to sketch an algebraic proof of Weyl’s theorem on complete reducibility of finite dimensional complex representations of a complex semisimple Lie algebra \( \mathfrak{g} \).

4.6.11. Exercise. For this exercise only, \( \mathfrak{g} \) is any finite dimensional Lie algebra and \( V,W \) are finite dimensional representations of \( \mathfrak{g} \).

(a) Use the identification \( \text{Hom}(V,W) = V^* \otimes W \) to show that if \( V,W \) are representations of \( \mathfrak{g} \), then the action of \( \mathfrak{g} \) on \( \text{Hom}(V,W) \) is given by

\[
(xf)(v) = x(f(v)) - f(xv).
\]

If \( \mathfrak{g} \) happens to be the Lie algebra of a Lie group \( G \) and \( V,W \) are representations of \( G \), then this can also be checked by differentiation the action of \( G \) on \( \text{Hom}(V,W) \).

(b) Conclude that \( \text{Hom}_\mathbb{C}(V,W)^\mathfrak{g} = \text{Hom}_\mathbb{C}(V,W) \). Here, as usual, for a \( \mathfrak{g} \)-representation \( M \), the notation \( M^\mathfrak{g} \) means all \( v \in M \) such that \( \mathfrak{g}v = 0 \).

(c) Let \( 0 \to W \to V \) be maps of \( \mathfrak{g} \)-representations. Then restriction of homomorphisms to \( W \) yields \( \rho : \text{Hom}_\mathbb{C}(V,W) \to \text{Hom}_\mathbb{C}(W,W) \to 0 \). The sequence \( 0 \to W \to V \) splits if and only if \( \text{Hom}_\mathbb{C}(V,W) \) contains a copy of the trivial representation that maps under \( \rho \) isomorphically onto \( \mathbb{C} \text{id}_W \).

The algebraic proof of Weyl’s theorem uses the Casimir element, a canonical element in the center of the universal enveloping algebra that is closely related to the Laplacian. The Casimir helps us by “detecting the trivial representation” among all irreps and plays the same role as the operator \( (1 - |G|^{-1} \sum_{g \in G} g) \) in finite group representation.

4.6.12. The Casimir element. Let \( \mathfrak{g} \) be a subalgebra of \( \mathfrak{gl}(V) \). As a corollary of Cartan Criteria, we saw that \( B_V : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C} \) defined by \( B_V(x,y) = \text{tr}(xy) \) is a nondegenerate bilinear form on \( \mathfrak{g} \). So \( B_V \) defines an isomorphism \( \mathfrak{g}^* \to \mathfrak{g} \) and hence an isomorphism \( \text{Hom}_\mathbb{C}(\mathfrak{g}, \mathfrak{g}) \simeq \mathfrak{g}^* \otimes \mathfrak{g} \simeq \mathfrak{g} \otimes \mathfrak{g} \). The image of \( \text{id}_\mathfrak{g} \) under this isomorphism gives a canonical element of \( \mathfrak{g} \otimes \mathfrak{g} \) and this lets us define a canonical operator on \( V \) called the casimir element. Explicitly, let \( \{u_i\} \) and \( \{u'_i\} \) be dual basis of \( \mathfrak{g} \) with respect to the form \( B_V \). Then define

\[
c = c_V = \sum_i u_i u'_i.
\]
4.6.13. **Lemma.** One has $c_V \in \text{End}_\mathfrak{g}(V)$ and $\text{tr}(c_V) = \dim(\mathfrak{g})$.

**Proof.** Take $x \in \mathfrak{g}$. In $\text{End}_\mathbb{C}(V)$, we compute

$$c_V x - xc_V = \sum_i (u_i u'_i x - xu_i u'_i) = \sum_i (u_i[u'_i, x] + [u_i, x] u'_i).$$

Now

$$\sum_i [u_i, x] u'_i = \sum_{ij} u_j B_V([u_i, x], u'_j) u'_i = -\sum_{ji} u_j B_V([u'_j, x], u_i) u'_i = -\sum_j u_j [u'_j, x]$$

where the second equality follows from since $B_V$ is $\mathfrak{g}$ equivariant and symmetric and the Lie bracket is antisymmetric and the first and the third equality follows by writing $[u_i, x]$ and $[u'_j, x]$ with respect to the dual bases. So $c_V x = xc_V$ for all $x \in \mathfrak{g}$, that is, $c_V \in \text{End}_\mathfrak{g}(V)$. Finally, $\text{tr}(c_V) = \sum_i \text{tr}(u_i u'_i) = \sum_i B_V(u_i, u'_i) = \dim(\mathfrak{g})$. \(\square\)

Now we are ready to revisit Weyl’s theorem: a finite dimensional representation of a finite dimensional complex semisimple Lie algebra $\mathfrak{g}$ is a direct sum of irreps. It suffices to prove the following.

4.6.14. **Theorem.** Let $V$ be a finite dimensional complex representation of $\mathfrak{g}$ and let $W$ be a subrepresentation of $V$. Then $W$ is a direct summand of $V$.

**Sketch of Proof.** By induction on $\dim(V)$, assume the theorem holds for all representations of dimension less that $\dim(V)$.

Step 1: First assume that $W$ is an irrepsuch that $\dim(V/W) = 1$. Since $V/W$ is one dimensional representation of $\mathfrak{g}$ it is trivial, so $c = c_V$ acts trivially on it, in other words $c(V) \subseteq W$. Since $c$ is a map of $\mathfrak{g}$-representations, by Schur’s lemma $c|_W = \lambda \text{id}_W$ for some $\lambda \in \mathbb{C}$ and $\lambda \neq 0$ because otherwise, we would get $\text{tr}(c) = 0$, contradicting $\text{tr}(c) = \dim(\mathfrak{g})$. It follows that $V = \ker(c) \oplus W$. This proves the theorem in this special case.

Step 2: Next assume $W$ is any subrepresentation of $V$ such that $\dim(V/W) = 1$. If $W$ is not an irrep, then let $Z$ be a proper nonzero subrep. By induction hypothesis, the subrep $W/Z \subseteq V/Z$ has a complement, i.e, there exists some subrep $Y$ containing $Z$ such that $V/Z = W/Z \oplus Y/Z$. Again by induction, write $Y = Z \oplus U$ for some subrepresentation $U$. Then verify that $V = W \oplus U$.

Step 3: Finally, assume that $W$ is any subrepresentation. We need to show that the exact sequence of $\mathfrak{g}$ modules $0 \to W \to V$ splits, that is there exists $f \in \text{Hom}_\mathfrak{g}(V, W)$ such that $f|_W = \text{id}$. The inclusion $0 \to W \to V$ gives us a surjection

$$\text{Hom}_\mathbb{C}(V, W) \xrightarrow{\text{res}} \text{Hom}_\mathbb{C}(W, W) \to 0$$

of $\mathfrak{g}$ representaitions where $\text{res}(\alpha) = \alpha|_W$. Now $\text{Hom}_\mathbb{C}(W, W)$ contains a one dimensional trivial representation $\mathbb{C} \text{id}_W$ and the above surjection, yields a surjection of $\mathfrak{g}$-representations

$$\text{res}^{-1}(\mathbb{C} \text{id}_W) \xrightarrow{\text{res}} \mathbb{C} \text{id}_W \to 0. \quad (10)$$

By step 2, this exact sequence splits since that would yield a one dimensional $U \subseteq \text{res}^{-1}(\mathbb{C} \text{id}_W) \subseteq \text{Hom}_\mathbb{C}(V, W)$ that maps isomorphically onto $\mathbb{C} \text{id}_W$ under $\text{res}$. Choose $f \in U$ such that $\text{res}(f) = f|_W = \text{id}$. Since $U$ is one dimensional representation of $\mathfrak{g}$, it is trivial, so $f \in \text{Hom}(V, W)$ is $\mathbb{C} \text{id}_V$. This $f : V \to W$ gives a splitting of the sequence $0 \to W \to V$. In other words, $V = W \oplus \ker(f)$. \(\square\)
4.7. **Representations of complex semisimple Lie algebras.** Let \( g \) be a finite dimensional complex semisimple Lie algebra. We are now ready to classify the finite dimensional complex representations of \( g \). The general story closely parallels the story for the example of \( \mathfrak{sl}_3(\mathbb{C}) \). The proofs of some of the results below are identical to the proofs we already did for \( \mathfrak{sl}_3(\mathbb{C}) \). If so, we omit the proofs.

4.7.1. **Cartan subalgebras.** First, we choose and fix a maximal abelian subalgebra of \( g \) consisting of diagonalizable elements, call it \( h \). Such an \( h \) always exists (just take one of maximal possible dimension). Such a subalgebra is called a *Cartan subalgebra* of \( g \). We need two facts about \( h \):

1. *All Cartan subalgebras in \( g \) are conjugate under \( \text{Aut}(g) \).*
2. *\( h \) is self normalizing, that is, \( h = \{ x \in g : [h, x] = 0 \} \).*

Which \( h \) we choose does not matter because of fact 1. The first step is to study the action of \( h \) on an irrep of \( g \).

4.7.2. **The weight space decomposition.** Let \( V \) be a finite dimensional complex representation of \( g \). By the invariance of Jordan decomposition, for each \( h \in h \), the linear map \( h|_V \) is diagonalizable. Since \( h \) is abelian and commuting diagonalizable matrices are simultaneously diagonalizable, we can decompose \( V \) into a direct sum of simultaneous eigenspaces of each element of \( h \). Let \( U \) be one of these eigenspaces. If \( h \in h \), then \( h|_U \) is a scalar multiple of identity. Let \( \alpha(h) \) be this scalar. One verifies that \( h \mapsto \alpha(h) \) is a linear function on \( h \). In other words, there exists some \( \alpha \in h^* \) such that \( hu = \alpha(h)u \) for all \( h \in h \) and all \( u \in U \). Write \( U = V_\alpha \). Such an \( \alpha \) is called a an *weight* of the representation \( V \) and the corresponding eigenspace \( V_\alpha \) is called a (nonzero) weight space of \( V \). Thus we obtain the weight space decomposition of \( V \):

\[
V = \bigoplus_{\alpha \in h^*} V_\alpha
\]

where

\[
hv = \alpha(h)v \quad \text{for all } h \in h \text{ and for all } v \in V_\alpha.
\]

Of course, since \( V \) is finite dimensional, \( V_\alpha \) is nonzero only for finitely many \( \alpha \). We denote this finite set of \( \alpha \)'s by \( \text{wt}(V) \). So in fact \( V = \bigoplus_{\alpha \in \text{wt}(V)} V_\alpha \). The dimension of \( V_\alpha \) is called the multiplicity of the weight \( \alpha \) in \( V \).

4.7.3. **The root space decomposition.** Apply the above discussion to the adjoint representation \( \text{ad} : g \to \mathfrak{gl}(g) \). Weight space decomposition of the adjoint representation \( g \) is called the *root space decomposition* of \( g \). The nonzero weights in this case are called the *roots* of \( g \). We denote the set of roots of \( g \) by \( R = R(g) \). We shall see later that the set of \( \mathbb{Z} \)-linear combinations of the roots, denoted

\[
\Lambda_R = \mathbb{Z}\langle R \rangle.
\]

forms a lattice in \( h^* \). It is called the root lattice. We write

\[
h^*_R = \mathbb{R}\langle R \rangle.
\]

So we have

\[
g = \bigoplus_{\alpha \in h^*} g_\alpha = g_0 \oplus \left( \bigoplus_{\alpha \in R} g_\alpha \right)
\]

where

\[
\text{ad}(h)x = [h, x] = \alpha(h)x \quad \text{for all } h \in h \text{ and for all } x \in g_\alpha.
\]

Fact 2 about the Cartan subalgebra \( h \) quoted above implies that \( g_0 = h \).
As before, the lemma below is crucial and tells us how the root spaces in the Lie algebra $\mathfrak{h}$ acts on the weight spaces of any representation $V$.

4.7.4. **Lemma.** One has $\mathfrak{g}_\alpha V_\beta \subseteq V_{\alpha+\beta}$ for all $\alpha, \beta \in \mathfrak{h}^*$. In particular, $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.

Before proceeding further, we quote without proof some facts about roots and root spaces which will be proved in the next section. The basic ingredient of the proof is the non-degeneracy of the Killing form.

4.7.5. **Lemma.**

1. Each root space $\mathfrak{g}_\alpha$ is one dimensional.
2. The root lattice $\Lambda_R$ has rank equal to $\dim_{\mathbb{C}}(\mathfrak{h})$. So $\mathfrak{h}^*_R \otimes \mathbb{C} = \mathfrak{h}$.
3. Let $\alpha \in R$. Then $n\alpha \in R$ for $n \in \mathbb{Z}$ if and only if $n \in \{\pm 1\}$.
4. $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \neq 0$.
5. $[\mathfrak{g}_\alpha, [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]] \neq 0$.

Now assume $V$ is an irrep of $\mathfrak{g}$. Let $\beta \in \text{wt}(V)$. Then $\bigoplus_{\alpha \in \Lambda_R} V_{\beta+\alpha}$ is stable under each root space by lemma 4.7.4 and also stable under $\mathfrak{h}$, and hence is a subrepresentation of $V$, so

$$V = \bigoplus_{\alpha \in \Lambda_R} V_{\beta+\alpha}$$

In other words, the weights of an irrep $V$ differ from each other by elements of the root lattice.

4.7.6. **Positive roots, highest weight vector.** Choose a linear functional $l : \mathfrak{h}^*_R \to \mathbb{R}$.

and extend it linearly to $\mathfrak{h}$. We choose the linear functional generic enough so that $l$ does not vanish on $\Lambda_R$. Since any two root of $V$ differ by an element of $\Lambda_R$, it follows that the real part of functional $l$ takes distinct values on all the weights of $V$. Since $V$ is finite dimensional, there exists a unique weight $\alpha$ of $V$ for which $\text{Re}(l(\alpha))$ is maximal. This is called the highest weight of $V$ and the corresponding weight space $V_\alpha$ is called the highest weight space. Choose $v \in V_\alpha - \{0\}$. This is called a highest weight vector.

Let $R_+$ (resp. $R_-$) be the set of roots $\alpha$ for which $\text{Re}(l(\alpha)) > 0$ (resp. $< 0$). These are called the positive (resp. negative) roots. Define the positive and negative part of the Lie algebra by

$$\mathfrak{g}_+ = \bigoplus_{\alpha \in R_+} \mathfrak{g}_\alpha \quad \text{and} \quad \mathfrak{g}_- = \bigoplus_{\alpha \in R_-} \mathfrak{g}_\alpha.$$ 

Observe that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$ implies that $\mathfrak{g}_+$ and $\mathfrak{g}_-$ are subalgebras. One has the triangular decomposition of the Lie algebra:

$$\mathfrak{g} = \mathfrak{g}_+ \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_-.$$ 

If $\alpha \in R_+$ (resp. $\alpha \in R_-$), then the root spaces $\mathfrak{g}_\alpha$ is called a positive root space (resp. negative root space). Elements of the positive root space are called the raising operators and the elements of negative root space are called the lowering operators. As before, the highest weight vector is killed by all the raising operators.

To get more detailed information about the weights of $V$, we need to introduce the root $\mathfrak{sl}_2$'s.
4.7.7. The root \( \mathfrak{sl}_2 \)'s. Let \( \alpha \) be any positive root of \( \mathfrak{g} \). Define
\[
\mathfrak{sl}_\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}].
\]
By lemma 4.7.5, \( \mathfrak{sl}_\alpha \) is 3 dimensional. Verify that 4.7.5(5) implies the existence of a unique \( H_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \) such that
\[
\alpha(H_\alpha) = 2.
\]
Choose \( E_\alpha \in \mathfrak{g}_\alpha \setminus \{0\} \) and \( F_\alpha \in \mathfrak{g}_{-\alpha} \) such that
\[
[E_\alpha, F_\alpha] = H_\alpha.
\]
Verify that \( \{E_\alpha, F_\alpha, H_\alpha\} \) satisfy the commutator relations of the standard generators of \( \mathfrak{sl}_2(\mathbb{C}) \). In particular \( \mathfrak{sl}_\alpha \simeq \mathfrak{sl}_2(\mathbb{C}) \).

4.7.8. Integrality and Weyl group Symmetry. Define the hyperplane
\[
\Omega_\alpha = \{ \beta \in \mathfrak{h}^* : (H_\alpha, \beta) = 0 \}.
\]
For each \( \alpha \in \mathbb{R} \), define the involutions \( R_\alpha : \mathfrak{h}^* \rightarrow \mathfrak{h}^* \) by
\[
R_\alpha(\beta) = \beta - \beta(H_\alpha)\alpha.
\]
This is an involution that fixes \( \Omega_\alpha \) pointwise and takes \( \alpha \) to \( -\alpha \). It is called the reflection in \( \alpha \). The group \( W \) generated by these involutions is called the Weyl group of \( \mathfrak{g} \).

Given \( \beta \in \text{wt}(V) \), consider the subspace
\[
U = \oplus_{n \in \mathbb{Z}} V_{\beta + n\alpha}.
\]
Note that \( U \) is an \( \mathfrak{sl}_\alpha \) representation. Since \( \alpha(H_\alpha) = 2 \), we find that \( H_\alpha \) acts on \( V_{\beta + n\alpha} \) with weight \( \beta(H_\alpha) + 2n \). So we find
\[
V_{\beta + n\alpha} \text{ is the weight space of the } \mathfrak{sl}_\alpha \text{ representation } U \text{ with weight } \alpha.
\]
The weights of finite dimensional \( \mathfrak{sl}_2(\mathbb{C}) \) representations are integers. So we find the following integrality condition for weights:
\[
\beta(H_\alpha) \in \mathbb{Z} \text{ for all } \alpha \in \mathbb{R} \text{ and for all } \beta \in \text{wt}(V).
\]
Define
\[
\Lambda_W = \{ \beta \in \mathfrak{h}^* : \beta(H_\alpha) \in \mathbb{Z} \text{ for all } \alpha \in \mathbb{R} \}.
\]
So the weights of any irrep lie in \( \Lambda_W \). We shall see in the next section that
\textbf{Fact:} \( \Lambda_W \) is a lattice in \( \mathfrak{h}_R^* \) containing the root lattice \( \Lambda_R \).

The lattice \( \Lambda_W \) is called the weight lattice of \( \mathfrak{g} \).

Next, The weights (i.e. the eigenvalues of \( H_\alpha \)) \( k \) and \( -k \) occur with the same multiplicity in the \( \mathfrak{sl}_\alpha \) representation \( U \). As in the case of \( \mathfrak{sl}_3(\mathbb{C}) \), this implies
\[
\dim(V_\beta) = \dim(V_{R_\alpha(\beta)}) \text{ for all } \alpha \in \mathbb{R}, \text{ and for all } \beta \in \text{wt}(V).
\]
In particular, note that the weights all lie in the real vector space \( \mathfrak{h}_R^* \), so
\[
\text{Re}l(\alpha) = l(\alpha).
\]
To complete the description of the weights of an irrep we need the next proposition already seen in the case of \( \mathfrak{sl}_3 \). The proof requires some facts about simple roots. So we should skip it now and come back and do it as an exercise copying the proof of the \( \mathfrak{sl}_3 \) case after learning about simple roots.
4.7.9. **Proposition.** Any finite dimensional irrep of \( g \) has a highest weight vector \( v \) that is unique up to scalar. One has \( g_+v = 0 \). The representation \( V \) is generated by the images of \( v \) under repeated application of the lowering operators. In other words \( U(g_-)v = V \).

**Proof.** Exercise, to be done after learning about simple roots. \( \square \)

It is easy exercise to show that the highest weight of any irrep lies in one connected component \( C \) of \((\mathfrak{h}^* - \cup_{\alpha \in \Lambda_R} \Omega_\alpha)\), called the Weyl chamber. As we shall explain below, given any \( \alpha \in C \cap \Lambda_W \), there is an irrep \( V(\alpha) \) with highest weight \( \alpha \). Summarizing, we have arrived at the following theorem:

4.7.10. **Theorem.** A finite dimensional representation of \( g \) is determined by its highest weight \( \alpha \) that lies in \( C \cap \Lambda_W \). Take any \( \alpha \in C \cap \Lambda_W \). There exist an unique irrep \( V(\alpha) \) with highest weight \( \alpha \). The set of weights of \( V(\alpha) \) are precisely \( (\alpha + \Lambda) \cap \text{conv}(W\alpha) \).

4.7.11. **Constructing the irreps with a given highest weight.** On the Lie algebra level, write the triangular decomposition \( g = g_- \oplus g_0 \oplus g_+ \). Let \( b = g_0 \oplus g_+ \). Let \( \lambda \) be a dominant weight. Then \( \lambda : g_0 \to \mathbb{C} \) is a one dimensional representation of the abelian Lie algebra \( g_0 \).

Extend this to a one dimensional representation of \( b \) by letting \( g_+ \) act as 0. Call this one dimensional representation \( \mathbb{C}|\lambda\rangle \). It is generated by a vector \( |\lambda\rangle \) such that \( h|\lambda\rangle = \lambda(h) \) for \( h \in \mathfrak{h} \) and \( x|\lambda\rangle = 0 \) for \( x \in g_- \). Define the Verma module of highest weight \( \lambda \) as the induced module

\[ V_\lambda = U(g) \otimes_{U(b)} \mathbb{C}|\lambda\rangle. \]

One can show that all sum of the proper submodules of \( V_\lambda \) is again a proper submodule (since the weight \( \lambda \) does not occur in any proper submodule since the there is only one dimensional weight space of weight \( \lambda \) spanned by \( |\lambda\rangle \) and this generates the whole Verma module. So the sum of all the proper submodules of \( V_\lambda \) is the unique maximal proper submodule \( P_\lambda \) and the quotient \( V_\lambda/P_\lambda \) happens to be the irrep of highest weight \( \lambda \).

On the Lie group level, let \( G \) be a complex semisimple Lie group. Let \( B \) be a maximal solvable subalgebra (called a Borel subalgebra). The Borel subalgebra is obtained by exponentiating the Lie subalgebra \( b \). If \( G = \text{SL}_n(\mathbb{C}) \), we can choose \( B \) to be the set of upper triangular matrices in \( G \). Let \( H \subseteq B \) be the Cartan subgroup (obtained by exponentiating the Cartan subalgebra) and let \( N \) be the exponentiation of \( g_- \). Then \( B/N \simeq H \). Let \( \lambda \) be a dominant weight. Then we have a one dimensional representation \( \mathbb{C}|\lambda\rangle \) of \( H \) with weight \( \lambda \) defined by

\[ \exp(h)|\lambda\rangle = \exp(\lambda(h))|\lambda\rangle. \]

We can consider \( \mathbb{C}|\lambda\rangle \) as a one dimensional representation of the Borel subalgebra since \( B/N \simeq H \). Consider the manifold \( G/B \) (called the Flag manifold). On this manifold we have the principal \( B \)-bundle \( G \to G/B \). The one dimensional representation \( \mathbb{C}|\lambda\rangle \) allows us to replace the fibers \( B \) with \( \mathbb{C}|\lambda\rangle \) to obtain a Line bundle \( L_\lambda \) on the flag manifold \( G/B \). Explicitly

\[ L_\lambda = G \times_B \mathbb{C}|\lambda\rangle. \]

The group \( G \) acts on the space of global sections of this Line bundle and this is irrep \( L_\lambda \) of \( G \) of highest weight \( \lambda \):

\[ L_\lambda = H^0(G/B, \mathcal{L}_\lambda). \]

This is called the Atiyah-Borel-Weil theorem. Notice that this construction is identical to out construction of induced represenations for finite groups as sections of vector bundles.
4.8. **On roots and root spaces of semisimple Lie algebras.** Let \( \mathfrak{g} \) be a finite dimensional complex semisimple Lie algebra. We want to prove a sequence of small lemmas that give us information regarding the roots and root spaces of \( \mathfrak{g} \). We shall freely use the notations and definitions introduced in the previous section: In particular, a Cartan algebra \( \mathfrak{h} \) is fixed. We have the root space decomposition of \( \mathfrak{g} \):

\[
\mathfrak{g} = \mathfrak{g}_0 \oplus \left( \oplus_{\alpha \in R} \mathfrak{g}_\alpha \right)
\]

where \( \mathfrak{g}_0 = \mathfrak{h} \) and \([h, x] = \alpha(h)x \) for all \( h \in \mathfrak{h} \) and all \( x \in \mathfrak{g}_\alpha \). Let \( B(x, y) = \text{tr}(\text{ad}(x) \circ \text{ad}(y)) \) be the Killing form. Recall that semisimplicity of \( \mathfrak{g} \) is equivalent to nondegeneracy of \( B \). The adjoint action of one root space on another together with the \( \mathfrak{g} \)-invariance of the killing form, lets us extract more detailed information about orthogonality of various root spaces with respect to \( B \) and these in turn imposes strong restrictions on the set of roots \( R \). The other main ingredient in the proof, as always is the understanding of representations of \( \mathfrak{sl}_2(\mathbb{C}) \).

4.8.1. **Lemma.** (a) Suppose \( \alpha, \beta \in R \) and \( \alpha + \beta \neq 0 \). Then \( B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0 \).

(b) If \( \alpha \in R \), then \( -\alpha \in R \) and the pairing \( B : \mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha} \to \mathbb{C} \) is nondegenerate.

(c) The pairing \( B : \mathfrak{h} \times \mathfrak{h} \to \mathbb{C} \) is nondegenerate.

**Proof.** (a) Let \( x \in \mathfrak{g}_\alpha \) and \( y \in \mathfrak{g}_\beta \). Then \( \text{ad}(x) \circ \text{ad}(y) \) takes \( \mathfrak{g}_\gamma \) to \( \mathfrak{g}_{\alpha + \beta + \gamma} \). This implies that the matrix of \( \text{ad}(x) \circ \text{ad}(y) \) with respect to a suitable basis of \( \mathfrak{g} \) has all zeros on diagonals.

(b) Let \( \alpha \in R \). If \( x \in \mathfrak{g}_\alpha - \{0\} \) is orthogonal to all elements of \( \mathfrak{g}_{-\alpha} \), then, by part (a), \( x \) is orthogonal to all the root spaces, hence orthogonal to \( \mathfrak{g} \), contradicting the nondegeneracy of \( B \). So \( \mathfrak{g}_{-\alpha} \) must be nondegenerate.

Proof of part (c) is similar to part (b). \( \square \)

4.8.2. **Lemma.** The roots span \( \mathfrak{h}^* \).

**Proof.** Suppose \( x \in \mathfrak{h} \) such that \( \alpha(x) = 0 \) for all \( \alpha \in R \). This implies \( x \in Z(\mathfrak{g}) = \{0\} \). \( \square \)

4.8.3. **Lemma.** Let \( \alpha \in R \), \( x \in \mathfrak{g}_\alpha \), \( y \in \mathfrak{g}_{-\alpha} \) and \( h \in \mathfrak{h} \). Then

\[
B(h, [x, y]) = B([h, x], y) = \alpha(h)B(x, y).
\]

4.8.4. **Lemma.** If \( \alpha \in R \), then \( [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \neq 0 \).

**Proof.** Since \( B : \mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha} \to \mathbb{C} \) is non-degenerate, there exists \( x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha} \) such that \( B(x, y) \neq 0 \). Since the roots span \( \mathfrak{h}^* \), There exists \( h \in \mathfrak{h} \) such that \( \alpha(h) \neq 0 \). So \( B(h, [x, y]) \neq 0 \), in particular \([x, y] \neq 0 \). \( \square \)

4.8.5. **The bilinear form on roots.** The restriction of \( B \) to \( \mathfrak{h} \times \mathfrak{h} \) is nondegenerate. So the form \( B|_{\mathfrak{h} \times \mathfrak{h}} \) determines an isomorphism \( \tilde{B} : \mathfrak{h} \to \mathfrak{h}^* \) given by \( \tilde{B}(x) = B(x, \cdot) \). Using this isomorphism, we can transport the bilinear form \( B \) from \( \mathfrak{h} \) to \( \mathfrak{h}^* \) and get a non-degenerate bilinear form \( B^* : \mathfrak{h}^* \times \mathfrak{h}^* \to \mathbb{C} \), also called the killing form. For \( \alpha \in \mathfrak{h}^* \), let \( T_{\alpha} \) be the element of \( \mathfrak{h} \) dual to \( \alpha \) with respect to \( B \), that is, \( T_{\alpha} = \tilde{B}^{-1}(\alpha) \), or in other words

\[
B(T_{\alpha}, \cdot) = \alpha(\cdot).
\]

Then, One has

\[
B^*(\alpha, \beta) = B(T_{\alpha}, T_{\beta}) \quad \text{for all} \quad \alpha, \beta \in \mathfrak{h}^*.
\]

Let \( \alpha \in \mathfrak{h}^* \) and \( h \in \mathfrak{h} \). Then

\[
B^*(\alpha, B(h, \cdot)) = B(T_{\alpha}, h) = \alpha(h).
\]
Now we are ready for the lemma that gives us enough information about the root spaces to construct the root \( \mathfrak{sl}_2 \)'s.

4.8.6. Lemma. Let \( \alpha \) be a root.

(a) If \( x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha} \), then \([x, y] = B(x, y)T_\alpha\).

(b) \( B^*(\alpha, \alpha) = \alpha(T_\alpha) \neq 0 \).

(c) \([ [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}], \mathfrak{g}_\alpha] \neq 0 \).

(d) \( \mathfrak{g}_\alpha \) is one dimensional.

(e) \( k\alpha \) is a root for some \( k \in \mathbb{C} \) if and only if \( k \in \{ \pm 1 \} \).

Proof. (a) Since \( B|_{\mathfrak{h} \times \mathfrak{h}} \) is non-degenerate, it suffices to show that pairing either side of the required equality with an arbitrary \( h \in \mathfrak{h} \) gives the same number. We compute

\[
B(h, B(x, y)T_\alpha) = B(x, y)B(h, T_\alpha) = \alpha(h)B(x, y) = B(h, [x, y]).
\]

(b) If possible, suppose \( \alpha(T_\alpha) = 0 \). Choose \( x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha} \) such that \( c = B(x, y) \neq 0 \). Then \([x, y] = cT_\alpha \). So \( \mathfrak{s} = \mathbb{C}x + \mathbb{C}y + \mathbb{C}T_\alpha \) is a subalgebra of \( \mathfrak{g} \). Verify that \( \alpha(T_\alpha) = 0 \) implies that \( \mathfrak{s} \) is solvable. So by Lie's theorem, \( \text{ad}\_\mathbb{C}(x) \) and \( \text{ad}\_\mathbb{C}(y) \) are upper triangularizable. So \( \text{ad}\_\mathbb{C}(x, y) \) is nilpotent. So \( \text{ad}\_\mathbb{C}(T_\alpha) \) is nilpotent. But \( T_\alpha \in \mathfrak{h} \) implies \( \text{ad}\_\mathbb{C}(T_\alpha) \) is diagonalizable. So \( \text{ad}\_\mathbb{C}(T_\alpha) = 0 \), so \( T_\alpha = 0 \), and hence \( \alpha = 0 \) which is a contradiction.

(c) Follows from (b). If \( x \in \mathfrak{g}_\alpha \setminus \{0\} \), then \( 0 \neq \alpha(T_\alpha)x = [T_\alpha, x] \in [[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}], \mathfrak{g}_\alpha] \).

We prove (d) and (e) together. From (a), we know that \([ \mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha} ] = \mathbb{C}T_\alpha \). By (c), there exists unique \( H_\alpha \in [ \mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha} ] \) such that \( \alpha(H_\alpha) = 2 \). In fact

\[
H_\alpha = \frac{2T_\alpha}{\alpha(T_\alpha)} = \frac{2T_\alpha}{B^*(\alpha, \alpha)}.
\]

Choose \( x \in \mathfrak{g}_\alpha \neq 0 \). There exists \( y \in \mathfrak{g}_{-\alpha} \) such that \([x, y] = H_\alpha \). Let \( \mathfrak{s} = \text{span}\{x, y, H_\alpha\} \). As before, verify that \( \mathfrak{s} \simeq \mathfrak{sl}_2(\mathbb{C}) \). Consider the adjoint action of \( \mathfrak{s} \) on

\[
V = \mathfrak{h} \oplus \left( \bigoplus_{k \in \mathbb{C}} \mathfrak{g}_{k\alpha} \right).
\]

Note that \( \alpha(H_\alpha) = 2 \) implies that \( \mathfrak{g}_{k\alpha} \) is the weight space of the \( \mathfrak{sl}_2(\mathbb{C}) \) representation \( V \) having weight \( 2k \). So \( \mathfrak{h} \) is the zero weight space of \( V \). So from \( \mathfrak{sl}_2(\mathbb{C}) \) representation theory, we know that the only \( k \)'s appearing in the above direct sum are in \( \frac{1}{2}\mathbb{Z} \).

Note that \( \ker(\alpha) \oplus \mathfrak{s} \) is a subrepresentation of \( V \) and \( \mathfrak{s} \) acts trivially on \( \ker(\alpha) \) and irreducibly on \( \mathfrak{s} \). Together, \( \ker(\alpha) \oplus \mathfrak{s} \) exhaust the 0 weight space of \( V \). So if \( V = \ker(\alpha) \oplus \mathfrak{s} \oplus U \), then \( U \) has no 0 weight space. The representation theory of \( \mathfrak{sl}_2 \) implies that \( U \) has no even weight space. So the only nonzero even weight spaces in \( V \) are those in \( \mathfrak{s} \), namely \( \mathbb{C}x \subseteq \mathfrak{g}_\alpha \) and \( \mathbb{C}y \subseteq \mathfrak{g}_{-\alpha} \) with weights 2 and \(-2\). In particular, this proves \( \mathfrak{g}_{2\alpha} = 0 \), so \( 2\alpha \notin \mathbb{R} \). Thus, we have proved that \( \text{twice a root is not a root} \). But then, half of a root is not a root either. In particular, \( \mathfrak{g}_{\alpha/2} = 0 \). So 1 is not a weight of \( V \), in particular 1 is not a weight of \( U \). But in any irrep of \( \mathfrak{sl}_2(\mathbb{C}) \) either weight 0 or 1 occurs. This forces \( U = 0 \) and hence \( V = \ker(\alpha) \oplus \mathfrak{s} = \mathfrak{h} \oplus \mathbb{C}x \oplus \mathbb{C}y \). In particular this proves \( \mathfrak{g}_\alpha = \mathbb{C}x, \mathfrak{g}_{-\alpha} = \mathbb{C}y \) and \( \mathfrak{g}_{k\alpha} \neq 0 \) for some nonzero complex number \( \mathbb{C} \) if and only if \( k = \pm 1 \). \( \square \)

4.8.7. Lemma. Let \( \mathfrak{h}_\mathbb{R} \) be the real span of \( H_\alpha \)'s as \( \alpha \) varies in \( R \). Then \( B|_{\mathfrak{h}_\mathbb{R} \times \mathfrak{h}_\mathbb{R}} \) is positive definite.
Proof. Let \( x, y \in \mathfrak{h} \). If \( z \in \mathfrak{h} \), then \( \text{ad}(x) \circ \text{ad}(y)z = 0 \). If \( z \in \mathfrak{g}_\alpha \), then \( \text{ad}(x) \circ \text{ad}(y)z = \alpha(x)\alpha(y)z \). Since each root space is 1 dimensional, it follows that

\[
B(x, y) = \sum_{\alpha \in R} \alpha(x)\alpha(y).
\]

By integrality of weights, we know that \( \alpha(H_\beta) \in \mathbb{Z} \subseteq \mathbb{R} \) for all \( \alpha, \beta \in R \). So \( B|_{\mathfrak{h}_\mathbb{R} \times \mathfrak{h}_\mathbb{R}} \) is real valued. If \( h \in \mathfrak{h}_\mathbb{R} \), then \( B(h, h) = \sum_{\alpha \in R} \alpha(h)^2 \geq 0 \). If \( B(h, h) = 0 \), then \( \alpha(h) = 0 \) for all \( \alpha \in R \), which implies \( h = 0 \) since the roots span \( \mathfrak{h}^* \).

Write

\[
c_\alpha = \alpha(T_\alpha) = B(T_\alpha, T_\alpha) = B^*(\alpha, \alpha).
\]

By definition of \( H_\alpha \), we have

\[
H_\alpha = 2T_\alpha/c_\alpha.
\]

It follows that

\[
B(H_\alpha, H_\alpha)B^*(\alpha, \alpha) = 4.
\]

Recall that under the isomorphism \( \tilde{B}^{-1} : \mathfrak{h}^* \to \mathfrak{h} \), the root \( \alpha \) corresponds to \( T_\alpha \). In other words

\[
\alpha \leftrightarrow \frac{2H_\alpha}{B(H_\alpha, H_\alpha)} \quad \text{and} \quad \frac{2\alpha}{B^*(\alpha, \alpha)} \leftrightarrow H_\alpha.
\]

It follows that, under the isomorphism \( \tilde{B} \), the real vector space \( \mathfrak{h}_\mathbb{R} \) corresponds to the real vector space \( \mathfrak{h}^*_\mathbb{R} = \mathbb{R}\langle R \rangle \) and hence \( B^* \) defines a positive definite bilinear form on \( \mathfrak{h}^*_\mathbb{R} \). We let \( \mathbb{E} \) denote the real vector space \( \mathfrak{h}^*_\mathbb{R} \) with the positive definite bilinear form \( B^* \). We may identify \( \mathbb{E} \) with an Euclidean space. For simplicity, we shall just write \( B^*(\alpha, \beta) = (\alpha, \beta) \) for \( \alpha, \beta \in \mathbb{E} \). Note that

\[
\beta(H_\alpha) = 2\beta(T_\alpha)/c_\alpha = 2(\beta, B(T_\alpha, \cdot))/(\alpha, \alpha) = 2(\beta, \alpha)/(\alpha, \alpha).
\]

So the integrality condition translates into

\[
2(\beta, \alpha)/(\alpha, \alpha) \in \mathbb{Z} \quad \text{for all} \quad \alpha, \beta \in R.
\]

4.8.8. Lemma. For each \( \alpha \in R \), the map \( R_\alpha \) (restricted to \( \mathbb{E} \)) is an orthogonal reflection in \( \mathbb{E} \) with respect to the mirror \( \alpha^\perp \). So the Weyl group \( W \subseteq O(\mathbb{E}) \).

Proof. One has

\[
R_\alpha(\beta) = \beta - \beta(H_\alpha)\alpha = \beta - 2\frac{(\beta, \alpha)}{(\alpha, \alpha)}\alpha
\]

which is just the euclidean reflection across the hyperplane \( \alpha^\perp \). \( \square \)

4.9.1. Definition. Starting from a complex semisimple Lie algebra \( g \), after fixing a Cartan \( h \) (of complex dimension \( n \)), we have now arrived at a finite set of vectors \( R \) (called roots) in a Euclidean vector space \( \mathbb{E} \) (of real dimension \( n \)) satisfying:

1. \( R \) spans \( \mathbb{E} \).
2. \( \alpha \in R \) implies \( \alpha \mathbb{R} \cap R = \{-\alpha, \alpha\} \).
3. If \( \alpha \in R \), then \( R_\alpha(R) \subseteq R \).
4. If \( \alpha, \beta \in R \), then \( n_{\beta, \alpha} = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z} \).

A set of vectors \( R \) in \( \mathbb{R}^n \) satisfying the above four conditions is called a (reduced) root system of rank \( n \).

Of course, if \( R \) is a root system in \( \mathbb{R}^n \), then scaling all the vectors of \( R \) we again get a root system which we shall identify with \( R \). If all the roots of \( R \) have the same length, then we shall use the convention that each vector has squared length (or norm) 2. In this case condition (4) just reads \( (\beta, \alpha) \in \mathbb{Z} \) for all \( \alpha, \beta \in R \).

The group \( W \) generated by \( R_\alpha \)'s is called the Weyl group of the root system \( R \). Note that condition (4) implies that \( W \) takes \( \Lambda_R = \mathbb{Z}(R) \) to itself. A root system \( R \) is reducible, if \( R \) can be written as disjoint union of two root systems \( R_1 \) and \( R_2 \) where each vector of \( R_1 \) is orthogonal to each vector in \( R_2 \). We shall write \( R = R_1 \times R_2 \). A root system is irreducible if it is not reducible. There is only one root system of rank 1 consisting of 2 vectors. It is called \( A_1 \). So \( A_1 \times A_1 \) is a reducible rank 2 root system. The root system of \( sl_3(\mathbb{C}) \) is irreducible of rank 2. It is called \( A_2 \). Let \( B_2 = \{ u + v : u, v \in A_1 \times A_1 \} \) and \( G_2 = \{ u + v : u, v \in A_2 \} \).

4.9.2. Exercise. (a) If \( \alpha, \beta \) are two non-proportional roots, then \( (\mathbb{R} \alpha \oplus \mathbb{R} \beta) \cap R \) is a rank 2 root system. (Hint: Just observe that the four axioms are obviously satisfied).

(b) Show that there are four root systems of rank 2, namely \( A_1 \times A_1, A_2, B_2 \) and \( G_2 \). Draw these. The first one is reducible, while the other three are irreducible. (Hint: Show that the integrality condition (4) severely restricts the possible angles between two roots).

Because of the above exercise we know what configuration of two roots may look like. In fact we have complete information on two dimensional slices of any root systems. Part (a), (b), (c) of the next lemma follows from this by inspection of the four rank 2 examples.

4.9.3. Lemma. If \( \alpha, \beta \) are two non-proportional roots,

(a) Then the roots of the form \( \beta - p\alpha, \beta - (p - 1)\alpha, \cdots, \beta + q\alpha, \)
(also called the \( \alpha \)-string through \( \beta \)) has at most 4 elements, that is, \( p + q \leq 3 \). Further, \( p - q = n_{\beta, \alpha} \).

(b) If \( (\beta, \alpha) > 0 \), then \( \alpha - \beta \in R \).

(c) If \( \alpha, \beta \) are two non-orthogonal roots and \( (\alpha, \alpha) \geq (\beta, \beta) \), then \( (\alpha, \alpha)/(\beta, \beta) \in \{1, 2, 3\} \).

(d) If \( (\beta, \alpha) < 0 \), then \( \alpha + \beta \in R \).

(e) If \( (\alpha, \beta) = 0 \), then \( \alpha - \beta \) and \( \alpha + \beta \) are both roots or non-roots.

Proof. Part (d) follows from part (b) since negative of a root is a root. Part (e) follows since \( W_\alpha(\beta + \alpha) = \beta - \alpha \) if \( \alpha \perp \beta \). □

Choose a linear functional \( l : \mathbb{E} \to \mathbb{R} \) such that \( l(\alpha) \neq 0 \) for all \( \alpha \in R \). Then the hyperplane \( \{ x \in \mathbb{E} : l(x) = 0 \} \) splits the roots into two parts \( R_+ \) and \( R_- \), called the positive
and negative roots. Here positive roots means roots \( r \) such that \( l(r) > 0 \). Now it is time to introduce the key remaining notion.

4.9.4. Definition. A positive root is called simple if it is not the sum of two other positive roots.

Let \( \Delta \) be a set of simple roots.

4.9.5. Lemma. Difference of two simple roots is not a root.

Proof. Let \( \alpha, \beta \) be two non-proportional simple roots. If \( (\alpha - \beta) \) is a positive root, then \( \alpha = (\alpha - \beta) + \beta \) would not be simple. So \( (\alpha - \beta) \) is not a positive root. For the same reason, \( (\beta - \alpha) \) is not a root. \( \square \)

4.9.6. Lemma. If \( \alpha, \beta \in \Delta \), then \( \langle \alpha, \beta \rangle \leq 0 \), that is the angle between two simple roots cannot be acute.

Proof. If \( \langle \alpha, \beta \rangle > 0 \), then \( (\alpha - \beta) \) would be a root, contradicting the previous lemma. \( \square \)

4.9.7. Lemma. Let \( S \) be a set of vectors in Euclidean space lying on one side of a hyperplane \( H \) through the origin such that \( \langle s, s' \rangle \leq 0 \) for all \( s, s' \in S \) and \( s \neq s' \). Then \( S \) is linearly independent. In particular, the simple roots are linearly independent.

Proof. Choose a linear functional \( l \) vanishing on \( H \) such that \( l(s) > 0 \) for all \( s \in S \). If \( S \) is not independent, then there is a relation of the form

\[
v = \sum_{i \leq k} c_i s_i = \sum_{j > k} c_j s_j
\]

where \( \{s_1, \ldots, s_n\} = S \) and \( c_j \)'s are all non-negative real numbers and at least some \( c_i \) is nonzero. Then \( v^2 = \sum_{i,j} c_i c_j \langle s_i, s_j \rangle \leq 0 \), so \( v^2 = 0 \), hence \( v = 0 \). But then \( 0 = l(v) = \sum_{i \leq k} c_i l(s_i) \) implies \( c_i = 0 \) for all \( i \leq k \) since each \( l(s_i) > 0 \). Similarly \( c_i = 0 \) for all \( i > k \). This is a contradiction. \( \square \)

4.9.8. Lemma. There are precisely \( n \) simple roots. Each positive root can be written uniquely as a non-negative integer linear combination of the simple roots.

Proof. The uniqueness follows from linear independence of simple roots. Argue existence by contradiction. If possible, choose a positive root \( \alpha \) with least possible \( l(\alpha) \) that cannot be written as required, that is, as non-negative integer linear combination of the simple roots. Then, in particular, \( \alpha \) is not simple. So \( \alpha = \beta + \gamma \) for some positive roots \( \beta \) and \( \gamma \). But then \( l(\beta), l(\gamma) \) are strictly less than \( l(\alpha) \) and hence \( \beta, \gamma \) can be written as required, which implies \( \alpha \) can be written that way too; a contradiction. \( \square \)

Combining the above facts, we have proved the following theorem.

4.9.9. Theorem. A root system \( R \subseteq \mathbb{R}^n \) has exactly \( n \) simple roots and these form a basis for \( \mathbb{R}^n \). Every root is a integer linear combination of simple roots with all non-negative coefficients or all non-positive coefficients.

It follows that \( \Lambda_R = \mathbb{Z}\langle R \rangle \) is indeed a lattice in \( \mathbb{R}^n \). This is the root lattice. It also follows that the Weyl group \( W \) is discrete since it preserves the root lattice. Since \( W \), being subset of \( O(E) \) is also compact, it follows that the Weyl group \( W \) is finite. Finally it follows that \( \Lambda_W = \{ \beta \in E: 2(\beta, \alpha)/(\alpha, \alpha) \in \mathbb{Z} \text{ for all } \alpha \in R \} \) is also a lattice in \( E \) containing \( \Lambda_R \) as a finite index sublattice.
4.9.10 (Exercise.) Let $R$ be a reflection in a non-zero vector $s$ in an Euclidean space $\mathbb{E}$. If $V$ is a subspace of $\mathbb{E}$ such that $RV \subseteq V$, then either $s \in V$ or $V \subseteq s^\perp$.

*Proof.* Suppose $V \notin s^\perp$. Choose $v \in V - s^\perp$. Then $v - Rv = 2(v, s)s/(s, s) \in V$ and hence $s \in V$. □

4.9.11. **Lemma.** Assume the root system $R$ is irreducible. Then $\mathbb{E}$ is an irrep of the Weyl group $W$.

*Proof.* Suppose $V$ is a $W$-stable subspace of $\mathbb{E}$. Then $V^\perp$ is $W$-stable as well since $W \subseteq O(\mathbb{E})$. If $s$ is any root, then the reflection $R_s$ preserves $V$ and $V^\perp$. So either $s \in V$ or $V \subseteq s^\perp$. In other words, either $s \in V$ or $s \in V^\perp$. In other words, every root is either in $V$ or $V^\perp$, contradicting the irreducibility of $R$. □

4.9.12. **Exercise.** In an irreducible root system, there are at most two distinct root lengths.

*Proof.* Let $\alpha, \beta$ be two roots. By the previous lemma, $W\beta$ spans $\mathbb{E}$. So there exists $w \in W$ such that $(w\beta, \alpha) \neq 0$. So the ratio of $(\alpha, \alpha)$ and $(\beta, \beta)$ is among $3, 2, 1, 1/2, 1/3$. Since this must hold for any two roots, only two distinct length ratios are possible. □
4.10. Dynkin diagrams and the classification of root systems. We want to classify the root systems. It is easy to see that any root system is an orthogonal direct sum of the irreducible root systems. So it is enough to classify the irreducibles. A root system is called simply laced if all the lattice vectors have the same length. In this case, we choose all the roots to have norm 2. (Norm means squared length). We saw that an irreducible root system has only two possible root lengths. So a non-simply laced irreducible root system has long roots and short roots and the long roots are either $\sqrt{2}$ or $\sqrt{3}$ times longer than the short roots. We will sketch the classification of the simply laced irreducible root systems. The general classification is similar but needs some more details.

4.10.1. Definition. Let $\Phi$ be a root system. Choose a set of positive roots and this determines a set of simple roots. Let $s$ and $t$ are two non-proportional simple roots and $(t, t)/(s, s) \geq (s, s)$. Then verify that there are the following possibilities:

1. $s$ and $t$ are at 120 degrees and their $(t, t)/(s, s) = 1$.
2. $s$ and $t$ are at 135 degrees and their $(t, t)/(s, s) = \sqrt{2}$.
3. $s$ and $t$ are at 150 degrees and their $(t, t)/(s, s) = \sqrt{3}$.
4. $s$ and $t$ are at 120 degrees and their $(t, t)/(s, s) = 1$.

The Dynkin diagram of $\Phi$ is a graph with vertex set equal to a set of simple roots and edges determined by the inner products between the simple roots. If $s$ and $t$ are two orthogonal simple roots then there is no edge between $s$ and $t$. Otherwise, we draw 1 or 2 or 3 edges joining $s$ and $t$ in cases 1, 2 and 3 above respectively. When $t$ is longer than $s$ there is an arrow on the double or the triple edge pointing towards the shorter root. One can prove that the Weyl group acts transitively on the set of positive systems, so the Dynkin diagram is independent of the choice of the positive system. One verifies that a root system is irreducible if and only if its Dynkin diagram is connected. An irreducible root system is simply laced if and only if the Dynkin diagram is simply laced i.e. has no multiple edges.

4.10.2. Exercise (a) Let $L \cong \mathbb{Z}^n$ be a free $\mathbb{Z}$-module of rank $n$ and let $(\ , \ ) : L \times L \to \mathbb{Q}$ be a bilinear form. We say $L$ is a lattice of rank $n$. If the bilinear form takes values in $\mathbb{Z}$ we say $L$ is an integral lattice. If all the norms of lattice vectors are even integers, then $L$ is called an even lattice. Verify that even lattices are integral. Verify that every integral lattice is either even or has an even lattice of index 2. If $L$ is an even lattice show that the norm 2 vectors of $L$ form a root system. An even lattice $L$ is a root lattice if $L$ is spanned by its norm 2 vectors.

(b) Define the following lattices:

- $A_n$ lattice: $\{(x_0, \ldots, x_n) \in \mathbb{Z}^{n+1}: \sum_i x_i = 0\}, n \geq 1$
- $D_n$ lattice: $\{(x_1, \ldots, x_n) \in \mathbb{Z}^n: \sum_i x_i \equiv 0 \mod 2\}, n \geq 4$
- $E_8$ lattice: $\{(x_1, \ldots, x_8) \in \mathbb{Z}^8 \cup (\frac{1}{2} + \mathbb{Z})^8: \sum_i x_i \equiv 0 \mod 2\}$.

The orthogonal complement of a norm 2 vector in the $E_8$ lattice is called the $E_7$ lattice. The orthogonal complement of (the span of) two non-orthogonal norm 2 vectors in the $E_8$ lattice is called the $E_6$ lattice. Verify that $A_n$, $D_n$, $E_6$, $E_7$, $E_8$ are all root lattices. It follows the
norm 2 vectors of these lattices form simply laced root systems, called root systems of type $A_n, D_n, E_6, E_7, E_8$.

4.10.3 (Exercise). (a) Choose the positive roots in $A_n$ using the functional $l : \mathbb{E} \to \mathbb{R}$ defined by $l(x_0, \ldots, x_n) = \sum_i a_i x_i$ where $a_0 > a_1 > \cdots > a_n$ are generic real numbers such that $\sum_i a_i = 0$. Verify that the positive roots of $A_n$ are the vectors of the form $(e_i - e_j)$ for $i < j$.

(b) Choose the positive roots in $D_n$ using the functional $l : \mathbb{E} \to \mathbb{R}$ defined by $l(x_1, \ldots, x_n) = \sum_i a_i x_i$ where $a_1 > \cdots > a_n > 0$ are generic real numbers. Verify that the positive roots of $A_n$ are the vectors of the form $(e_i \pm e_j)$ for $i < j$.

(c) Choose the positive roots in $E_8$ using the functional used for $D_8$. Describe the positive roots. Also work out the positive roots for $E_7$ and $E_6$.

(d) Find a set of simple roots for the root systems above and draw the Dynkin diagrams.

4.10.4. Definition. Let $\Phi$ be a root system with a fixed positive system $\Phi^+$ and corresponding simple system $\Delta$. The half sum of the set of positive roots is called a Weyl vector $\rho$. We shall say that $(\rho, r)$ is the height of the root $r$. Verify that if $\Phi$ is an irreducible simply laced diagram, then the simple roots are precisely the roots of height 1 and all other roots have strictly higher height. The root with highest value of $(\rho, r)$ is called the highest root and its negative is called the lowest root. Adding the lowest root to the set of simple roots and drawing the edges using the same rules used for drawing Dynkin diagrams one obtains the affine diagrams.

4.10.5 (Exercise). For the systems $A_n, D_n, E_6, E_7, E_8$, Write down the Weyl vector. Compute the lowest root and draw the affine diagram.

4.10.6. Theorem. The irreducible simply laced root systems are the ones of type $A_n, D_n, E_6, E_7, E_8$.

Sketch of proof. We need to argue that the connected simply laced Dynkin diagrams are the ones mentioned. These are called the spherical simply laced diagrams. Extend the spherical diagrams to affine diagrams by adding the lowest root. Let $\Delta = \{s_1, \ldots, s_r\}$ be one of these diagrams. Let $s_0$ be the lowest root. Clearly $-s_0$ is a positive root. Writing $-s_0$ as a linear combination of the simple roots, we obtain a relation of the form

$$v = n_0 s_0 + n_1 s_1 + \cdots + n_r s_r = 0$$

where $n_j \in \mathbb{Z}_{\geq 0}$ and $n_0 = 1$. We leave it as an exercise to show that $\langle s_0, s_j \rangle$ is either 0 or $-1$ for each simple root $s_j$. Thus we obtain a graph on the vertex set $\Delta \cup \{s_0\}$ by joining $s_0$ to the simple roots by the same rules used for drawing the Dynkin diagrams. The graph thus obtained is called the affine diagram. Each affine diagram has a “balanced numbering” $\{n_0, n_1, \ldots, n_r\}$ on its vertices. Here “balanced” means the condition that the number $n_j$ at any vertex times 2 is equal to the sum of the numbers at all the neighboring vertices to $s_j$. Note that this balancing conditions are just saying that $\langle v, s_j \rangle = 0$ (since $v = 0$). In fact the simply laced affine diagrams are exactly the connected graphs that admit a balanced numbering. Show that the positivity of the inner product implies that a spherical diagram cannot contain an affine diagram. Then a short combinatorial argument shows that if a connected diagram does not contain an affine diagram, then it must be one of the spherical diagrams. \hfill $\square$
4.10.7. Recovering a Lie algebra from the simply laced root system. Starting from a Dynkin diagram, one can re-construct the Lie algebra. This is Serre’s theorem (see Humphreys or Fulton and Harris). We shall describe a slightly different construction that works only for the simply laced cases but has the advantage that it “works over \( \mathbb{Z} \)”. This construction goes directly from the root system back to the Lie algebra.

Let \( \Lambda \) be a root lattice (i.e. an even lattice generated by the norm 2 vectors). Define \( q : \Lambda \to \mathbb{F}_2 \) by \( q(x) = (x, x)/2 \bmod 2 \). The associated bilinear form is \( b(x, y) = q(x + y) - q(x) - q(y) = (x, y) \bmod 2 \). This bilinear form determines a \( \mathbb{Z}/2 \) extension of \( \Lambda \), denoted \( \hat{\Lambda} \). In other words, there exists a short exact sequence

\[
1 \to \{\pm 1\} \to \hat{\Lambda} \to \Lambda \to 0
\]

where \( \hat{\Lambda} \) is a multiplicative abelian group which has two elements, denoted \( \pm e^v \), mapping to each \( v \in \Lambda \), with \( \pm 1 \) central, and the multiplication has the form

\[
e^{-v} e^v = (-1)^{c(u,v)} e^{u+v}
\]

where \( c(u, v) \in \mathbb{F}_2 \) satisfies \( c(u, v) - c(v, u) = b(u, v) \) and \( c(u, -u) = 0 \) for all \( u \in \Lambda \). In other words, the multiplication satisfies the condition

\[
e^{-v} e^v = (-1)^{c(u,v)} e^v e^u.
\]

Further \( e^0 = 1 \) is the unit of \( \hat{\Lambda} \). Note that \( e^u e^{-u} = e^0 = 1 \) since \( c(u, -u) = 0 \). So \( (e^u)^{-1} = e^{-u} \). For the moment, assume such an extension \( \hat{\Lambda} \) exists. Let \( \hat{\Phi} \) be the preimage of \( \Phi \) in \( \hat{\Lambda} \). Define

\[
\mathfrak{g}_\mathbb{Z} = \Lambda \oplus (\oplus_{x \in \hat{\Phi}} \mathbb{Z}[x]) / \langle [x] = -[x] \text{ for all } x \in \Phi \rangle
\]

\[
\simeq \Lambda \oplus (\oplus_{x \in \hat{\Phi}} \mathbb{Z}[x]).
\]

We shall simply write \([x] = x\) in \( \mathfrak{g}_\mathbb{Z} \). This wont cause any confusion since we have identified \([x] \) (minus in the group \( \hat{\Lambda} \) with \(-[x] \) (minus in \( \mathbb{Z} \)). We define a bracket on this \( \mathbb{Z} \)-module \( [\ , \] : \mathfrak{g}_\mathbb{Z} \times \mathfrak{g}_\mathbb{Z} \to \mathfrak{g}_\mathbb{Z} \) by the following rules:

\[
\left[ \lambda, \lambda' \right] = 0 \text{ for } \lambda, \lambda' \in \Lambda
\]

\[
\left[ \lambda, e^\alpha \right] = (\alpha, \lambda) e^\alpha \text{ for } \alpha \in \Phi, \lambda \in \Lambda,
\]

\[
\left[ e^\alpha, e^\beta \right] = \begin{cases} 
e^\alpha e^\beta & \text{if } \alpha + \beta \in \Phi, \\ \alpha & \text{if } \beta = -\alpha, \\ 0 & \text{otherwise.} \end{cases}
\]

Note that for two non-proportional roots \( \alpha, \beta \), the vector \( (\alpha + \beta) \) is a root if and only if \( (\alpha, \beta) = -1 \) and in this case \( c(\alpha, \beta) - c(\beta, \alpha) = -1 \). So

\[
\left[ e^\alpha, e^\beta \right] = \left( 1 \right)^{c(\alpha, \beta)} e^{\alpha + \beta} = \left( -1 \right)^{c(\beta, \alpha)} e^{\beta + \alpha} = -[e^\beta, e^\alpha].
\]

Verifying the Jacobi identity is a routine but tedious exercise since quite a few cases need to be considered. This shows that \( \mathfrak{g} = \mathfrak{g}_\mathbb{Z} \otimes \mathbb{C} \) is a complex Lie algebra with root system \( \Phi \) and it is visible that the killing form is non-degenerate, which implies \( \mathfrak{g} \) is semisimple. This completes the construction of the simply laced complex simple Lie algebras. The non-simply laced ones are obtained from the simply laced ones by “folding”, i.e. as fixed points of automorphisms of Dynkin diagrams. For example, the \( D_4 \) diagram has an order three symmetry which acts on the Lie algebra, and set of fixed points is the Lie algebra of type \( G_2 \). This identifies the Lie group \( G_2 \) as a subgroup of \( \text{SO}_8(\mathbb{C}) \).
4.10.8 (Construction of the double cover of the root lattice). It remains to construct \( \hat{\Lambda} \) the fits into the exact sequence

\[ 1 \to \{\pm 1\} \to \hat{\Lambda} \to \Lambda \to 0 \]

Extensions of \( \Lambda \) by abelian group \( M \) are classified by the second cohomology group \( H^2(\Lambda, M) \) and one can show that the elements of second cohomology are in bijection with the alternating forms \( \Lambda \times \Lambda \to M \), with the map \( H^2(\Lambda, M) \to \text{Alt}^2(\Lambda, M) \) taking a cocycle \( \omega \) to the alternating form \( (a, b) \to \omega(a, b) - \omega(b, a) \). In our case \( M = \mathbb{Z}/2 \) and \( b(a, b) = (a, b) \mod 2 \) is an alternating form (over \( \mathbb{Z}/2 \) alternating and symmetric are the same thing). So this \( b \) defines an extension. To write down an extension explicitly, choose a \( \mathbb{Z} \)-basis \( \alpha_1, \cdots, \alpha_n \) of \( \Lambda \). Define

\[
c(\alpha_i, \alpha_j) = \begin{cases} 
(\alpha_i, \alpha_j) \mod 2 & \text{if } i < j \\
0 & \text{otherwise.}
\end{cases}
\]

and extend \( c \) to \( \Lambda \times \Lambda \) as a bilinear form. Verify that \( c(u, v) - c(v, u) = (u, v) \mod 2 \) for all \( u, v \in \Lambda \). One can now verify that \( \hat{\Lambda} = \{ \pm e^v : v \in \Lambda \} \) with multiplication

\[
e^u e^v = (-1)^{c(u, v)} e^{u+v}
\]

forms a group. For example, bilinearity of \( c \) implies that \( c \) satisfies \( c(y, z) + c(x, y + z) = c(x, y) + c(x + y, z) \) which is equivalent to the product on \( \hat{\Lambda} \) is associative.