Programme on Lie Groups
Mini Course on Lie Algebras

K.K. Choi and W.K. To

March - June, 1999

Contents

Chapter 1

Lie Algebras: Definitions and Examples

1.1 Lie Algebras.

Suppose k is a field and every vector space is of finite dimension.

Definition 1.1. *A k-module (k-vector space) A is said to be* **k-algebra** *if there exists a k-bilinear map $A \times A \longrightarrow A$, i.e., if there exists a k-homomorphism $A \otimes_k A \longrightarrow A$, i.e.,*

$$(\alpha X + \beta Y, Z) = \alpha(X, Z) + \beta(Y, Z)$$

and

$$(Z, \alpha X + \beta Y) = \alpha(Z, X) + \beta(Z, Y)$$

for any α, β in k and X, Y, Z in A.

Definition 1.2. *A* **Lie algebra** $\mathfrak{g}$ *over k is a k-algebra whose k-bilinear map*

$$\begin{array}{ccc} A \times A & \longrightarrow & A \\ (X, Y) & \longmapsto & [X, Y] \end{array}$$

satisfies :

(i) $[X, X] = 0, \quad \forall X \in \mathfrak{g}$ *and*

(ii) **Jacobi identity**, *i.e.,*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0, \quad \forall X, Y, Z \in \mathfrak{g}.$$

Here $[\ ,\ ]$ *is called the* **Lie bracket** *of $\mathfrak{g}$ and the dimension of the vector space $\mathfrak{g}$ over k is the* **dimension** *of $\mathfrak{g}$.*

Remarks.

1. $[0, X] = [X, 0] = 0, \quad \forall X \in \mathfrak{g}$.
2. $[X, Y] = -[Y, X], \quad \forall X, Y \in \mathfrak{g}$ which is equivalent to the condition (i) in Definition 1.2 when the characteristic of k is not 2.
3. In general, $\mathfrak{g}$ may not be associative (see Example 2 below).

Examples.

(1) The trivial example is the one defining the k-bilinear map to be the zero map, i.e., $[X, Y] = 0$ for all X, Y in $\mathfrak{g}$. This Lie algebra is called **abelian**.

(2) Let $\mathfrak{g}$ be an associative algebra over k with multiplication $*$. Define

$$[X, Y] := X * Y - Y * X, \quad \forall X, Y \in \mathfrak{g}.$$

Then $\mathfrak{g}$ is a Lie algebra : Clearly, $[X, Y]$ is a k-bilinear map and condition (i) in Definition 1.2 is trivial. It remains to show that Jacobi identity is also satisfied. For,

$$\begin{aligned} & [X, [Y, Z]] \\ = \; & X * [Y, Z] - [Y, Z] * X \\ = \; & X * (Y * Z - Z * Y) - (Y * Z - Z * Y) * X \\ = \; & X * Y * Z - X * Z * Y - Y * Z * X + Z * Y * X. \end{aligned}$$

Hence

$$\begin{aligned} & [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] \\ = \; & (X * Y * Z - X * Z * Y - Y * Z * X + Z * Y * X) + (Y * Z * X - Y * X * Z \\ & -Z * X * Y + X * Z * Y) + (Z * X * Y - Z * Y * X - X * Y * Z + Y * X * Z) \\ = \; & 0. \end{aligned}$$

For instance, the set $M_n(k)$ of all $n \times n$ matrices over k and the set $End(V)$ of all endomorphisms of k-vector space V are Lie algebras.

(3) On $\mathbb{C}^3$, we define the Lie bracket by the cross product :

$$[\boldsymbol{X}, \boldsymbol{Y}] = \boldsymbol{X} \times \boldsymbol{Y} \quad \forall \boldsymbol{X}, \boldsymbol{Y} \in \mathbb{C}^3$$

i.e., if $\boldsymbol{X} = [x_1, x_2, x_3]^T, \boldsymbol{Y} = [y_1, y_2, y_3]^T$ are vectors in $\mathbb{C}^3$, then

$$\begin{aligned} [\boldsymbol{X}, \boldsymbol{Y}] \; & := \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \\ & = \begin{bmatrix} \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} & -\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} & \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \end{bmatrix}^T \\ & = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ -(x_1 y_3 - x_3 y_1) \\ x_1 y_2 - x_2 y_1 \end{bmatrix}. \end{aligned}$$

The conditions for k-bilinearity and (i) are trivial. For Jacobi identity, since

$$\boldsymbol{X} \times (\boldsymbol{Y} \times \boldsymbol{Z}) = (\boldsymbol{X} \cdot \boldsymbol{Z})\boldsymbol{Y} - (\boldsymbol{X} \cdot \boldsymbol{Y})\boldsymbol{Z}$$

so

$$\begin{aligned}
& [\boldsymbol{X},[\boldsymbol{Y},\boldsymbol{Z}]] + [\boldsymbol{Y},[\boldsymbol{Z},\boldsymbol{X}]] + [\boldsymbol{Z},[\boldsymbol{X},\boldsymbol{Y}]] \\
= \ & \boldsymbol{X} \times (\boldsymbol{Y} \times \boldsymbol{Z}) + \boldsymbol{Y} \times (\boldsymbol{Z} \times \boldsymbol{X}) + \boldsymbol{Z} \times (\boldsymbol{X} \times \boldsymbol{Y}) \\
= \ & (\boldsymbol{X} \cdot \boldsymbol{Z})\boldsymbol{Y} - (\boldsymbol{X} \cdot \boldsymbol{Y})\boldsymbol{Z} + (\boldsymbol{Y} \cdot \boldsymbol{X})\boldsymbol{Z} - (\boldsymbol{Y} \cdot \boldsymbol{Z})\boldsymbol{X} + (\boldsymbol{Z} \cdot \boldsymbol{Y})\boldsymbol{X} - (\boldsymbol{Z} \cdot \boldsymbol{X})\boldsymbol{Y} \\
= \ & 0.
\end{aligned}$$

Thus, $\mathbb{C}^3$ with the cross product is a Lie algebra.

(4) Let $\mathcal{O}(3)$ be the set of all skew-symmetric 3×3 matrices with $\mathbb{C}$-entries, i.e.,

$$\begin{aligned}
\mathcal{O}(3) &= \left\{X \in M_3(\mathbb{C}) : X^T = -X\right\} \\
&= \left\{ \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix} : a, b, c \in \mathbb{C} \right\}.
\end{aligned}$$

Define

$$[X,Y] := XY - YX \qquad \forall X, Y \in \mathcal{O}(3).$$

To show that $\mathcal{O}(3)$ is a Lie algebra, since $\mathcal{O}(3)$ is a subset of $M_3(\mathbb{C})$, it remains to prove that $\mathcal{O}(3)$ is closed under $[\ ,\ ]$. For any X and Y in $\mathcal{O}(3)$, write

$$X = \begin{bmatrix} 0 & -x_1 & -x_2 \\ x_1 & 0 & -x_3 \\ x_2 & x_3 & 0 \end{bmatrix}, \qquad Y = \begin{bmatrix} 0 & -y_1 & -y_2 \\ y_1 & 0 & -y_3 \\ y_2 & y_3 & 0 \end{bmatrix}.$$

Then

$$[X,Y] = \begin{bmatrix} 0 & -\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} & \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} \\ \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} & 0 & -\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \\ -\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} & \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} & 0 \end{bmatrix} \in \mathcal{O}(3).$$

Notice that if we identify $\mathbb{C}^3$ with $\mathcal{O}(3)$ by

$$\begin{aligned}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{C}^3 \quad &\longleftrightarrow \quad \begin{bmatrix} 0 & -x_1 & -x_2 \\ x_1 & 0 & -x_3 \\ x_2 & x_3 & 0 \end{bmatrix} \in \mathcal{O}(3) \\
[\boldsymbol{X},\boldsymbol{Y}] = \boldsymbol{X} \times \boldsymbol{Y} \quad &\longleftrightarrow \quad [X,Y] = XY - YX,
\end{aligned}$$

then these two Lie algebras have the same structure.

(5) **Lie algebras of vector fields on $\mathbb{R}^n$.** Let $U \subset \mathbb{R}^n$ be an open subset. Consider the set $D(U)$ of all the smooth vector fields on U given by

$$D(U) := \left\{ X = \sum_{i=1}^{n} a^i(x) \frac{\partial}{\partial x_i} \ : \ a^i \in C^\infty(U) \right\}.$$

Here $\mathbf{x} = (x_1, \cdots, x_n)$ denotes the Euclidean coordinate functions on $\mathbb{R}^n$. For $X = \sum_{i=1}^n a^i(x) \frac{\partial}{\partial x_i} \in D(U)$ and $f \in C^\infty(U)$, define

$$X(f) := \sum_{i=1}^{n} a^i(x) \frac{\partial f}{\partial x_i} \in C^\infty(U).$$

Lemma 1.3. *$X(f) = 0$ for all $f \in C^\infty(U)$ if and only if $X \equiv 0$.*

Proof. The '$\Leftarrow$' part is trivial. To prove the '$\Rightarrow$' part, for $1 \le i \le n$, let $f = x_i$. Then by assumption, $a^i(x) = X(x_i) = 0$ for all i, and thus $X \equiv 0$. $\square$

For $X = \sum_{i=1}^n a^i(x) \frac{\partial}{\partial x_i}$, $Y = \sum_{i=1}^n b^i(x) \frac{\partial}{\partial x_i} \in D(U)$, we define their **Lie bracket** by

$$[X, Y] := \sum_{i,j=1}^{n} \left(a^i \frac{\partial b^j}{\partial x_i} - b^i \frac{\partial a^j}{\partial x_i} \right) \frac{\partial}{\partial x_j} \in D(U).$$

Lemma 1.4. *$[X, Y](f) = X(Y(f)) - Y(X(f))$ for all $f \in C^\infty(U)$.*

Proof.

$$\begin{aligned} X(Y(f)) &= \sum_{i=1}^{n} a^i \frac{\partial}{\partial x_i} \left(\sum_{j=1}^{n} b^j \frac{\partial f}{\partial x_j} \right) \\ &= \sum_{i,j=1}^{n} \left(a^i \frac{\partial b^j}{\partial x_i} \frac{\partial f}{\partial x_j} + a^i b^j \frac{\partial^2 f}{\partial x_i \partial x_j} \right) \end{aligned}$$

and similarly

$$Y(X(f)) = \sum_{i,j=1}^{n} \left(b^i \frac{\partial a^j}{\partial x_i} \frac{\partial f}{\partial x_j} + b^i a^j \frac{\partial^2 f}{\partial x_i \partial x_j} \right).$$

Therefore,

$$X(Y(f)) - Y(X(f)) = \sum_{i,j=1}^{n} \left(a^i \frac{\partial b^j}{\partial x_i} - b^i \frac{\partial a^j}{\partial x_i} \right) \frac{\partial f}{\partial x_j} = [X, Y](f).$$

$\square$

Proposition 1.5. $(D(U), [\ ,\ ])$ *forms a Lie algebra.*

Proof. It is obvious that $[\ ,\ ]$ is bilinear and $[X, X] = 0$ for all $X \in D(U)$. To prove the Jacobi identity, by Lemma 1.3, it suffices to check that

$$([X,[Y,Z]] + [Y,[X,Z]] + [Z,[X,Y]])f = 0 \quad \text{for all } f.$$

Write $XYZf := X(Y(Z(f)))$, etc. We have

$$\begin{aligned} [X,[Y,Z]](f) &= X[Y,Z]f - [Y,Z]Xf \\ &= X((YZ-ZY)f) - (YZ-ZY)(Xf) \\ &= (XYZ - XZY - YZX + ZYX)f, \end{aligned}$$

and similarly

$$\begin{aligned} [Y,[Z,X]](f) &= (YZX - YXZ - ZXY + XZY)(f), \\ [Z,[X,Y]](f) &= (ZXY - ZYX - XYZ + YXZ)(f). \end{aligned}$$

Upon summing up these three identities, the Jacobi identity follows. $\square$

(6) **Lie Algebra of vector fields on manifolds.** Let M be an n-dimensional manifold. A smooth vector field X on M is simply a smooth section of the tangent bundle TM, i.e., a function $X : M \to TM$ such that $\pi \circ X = id_M$, where $\pi : TM \to M$ is the projection map. Denote the set of smooth vector fields on M by $D(M)$. A smooth vector field X is given on each coordinate open subset $U \subset M$ by an expression of the form $\sum_{i=1}^n a^i(x)\frac{\partial}{\partial x_i}$, where $\mathbf{x} = (x_1, \cdots, x_n)$ denotes the coordinate functions on U. Moreover, suppose X is given by $\sum_{i=1}^n a^i(x)\frac{\partial}{\partial x_i}$ on U and $\sum_{i=1}^n b^i(y)\frac{\partial}{\partial y_i}$ on V. Then the following compatibility condition is satisfied: on $U \cap V$, one has

$$a^i(x) = \sum_{j=1}^n b^j(y)\frac{\partial x_i}{\partial y_j} \quad i = 1, \cdots, n.$$

Let $X \in D(M)$. Then X acts on $C^\infty(M)$ as follows. For any $f \in C^\infty(M)$, $X(f) \in C^\infty(M)$ is given on a coordinate open set U by $X(f) = \sum_{i=1}^n a^i(x)\frac{\partial f}{\partial x_i}$, where $X = \sum_{i=1}^n a^i(x)\frac{\partial}{\partial x_i}$ on U. By the compatibility condition and the chain rule, one has

$$\sum_{i=1}^n a^i(x)\frac{\partial f}{\partial x_i} = \sum_{i=1}^n\sum_{j=1}^n b^j(y)\frac{\partial x_i}{\partial y_j}\frac{\partial f}{\partial x_i} = \sum_{j=1}^n b^j(y)\frac{\partial f}{\partial y_j}$$

on $U \cap V$, which implies $X(f)$ is a well-defined smooth function on M. As in Lemma 1.3, we have

Lemma 1.6. $X(f) = 0$ *for all* $f \in C^\infty(M)$ *if and only if* $X \equiv 0$.

For $X, Y \in D(M)$, we define their **Lie bracket** $[X, Y] \in D(M)$ given by

$$[X,Y] = \sum_{i,j=1}^{n} \left(a^i \frac{\partial b^j}{\partial x_i} - b^i \frac{\partial a^j}{\partial x_i} \right) \frac{\partial}{\partial x_j}$$

on any coordinate open subset U (with $X = \sum_{i=1}^n a^i(x) \frac{\partial}{\partial x_i}$ and $Y = \sum_{i=1}^n b^i(x) \frac{\partial}{\partial x_i}$) on U. To see that the above definition is well-defined, we consider two coordinate open subsets U and V with coordinate functions $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \cdots, y_n)$ respectively. Regard the intersection $U \cap V$ itself as a manifold with two coordinate system provided by $\mathbf{x}$ and $\mathbf{y}$. Then the expressions for $[X, Y]$ in these two coordiante systems $\mathbf{x}$ and $\mathbf{y}$ define two vector fields on $U \cap V$. On the other hand, by Lemma 1.6, the action of these two vector fields on $C^\infty(U \cap V)$ are the same, and thus by Lemma 1.6 (applied to $U \cap V$), these two vector fields are the same, which means that $[X, Y]$ are well-defined.

As in Example 5, one easily sees that for $X, Y \in D(M)$ and $f \in C^\infty(M)$, we have $[X,Y](f) = X(Y(f)) - Y(X(f))$ and $(D(M), [\ ,\ ])$ also forms a Lie algebra.

Definition 1.7. *Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras.*

(i) A **homomorphism** *from $\mathfrak{g}$ into $\mathfrak{h}$ is a mapping*

$$\phi : \mathfrak{g} \longrightarrow \mathfrak{h}$$

which is k-linear and preserves $[\ ,\ ]$, i.e.,

(a) $\phi(\alpha X + \beta Y) = \alpha\phi(X) + \beta\phi(Y), \quad \forall \alpha, \beta \in k, \forall X, Y \in \mathfrak{g},$

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

(ii) The image of ϕ is called a **homomorphic image** *of $\mathfrak{g}$.*

(iii) An **isomorphism** *from $\mathfrak{g}$ onto $\mathfrak{h}$ is a bijection from $\mathfrak{g}$ onto $\mathfrak{h}$ which is also a homomorphism.*

(iv) $\mathfrak{g}$ is **isomorphic** *to $\mathfrak{h}$ if there is an isomorphism from $\mathfrak{g}$ onto $\mathfrak{h}$.*

For example, let $\phi : \mathbb{C}^3 \longrightarrow \mathcal{O}(3)$ be

$$\phi\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} 0 & -x_1 & -x_2 \\ x_1 & 0 & -x_3 \\ x_2 & x_3 & 0 \end{bmatrix}.$$

Then we have already shown that ϕ is an isomorphism. So $\mathbb{C}^3 \cong \mathcal{O}(3)$.

Since $[\ ,\ ]$ is a bilinear map and $\mathfrak{g}$ is a r-dimensional k-vector space, so the value of $[\ ,\ ]$ is determined entirely by its value on the elements in a basis of $\mathfrak{g}$.

Let $\{X_1, X_2, \cdots, X_r\}$ be a basis of $\mathfrak{g}$ over k and write

$$[X_i, X_j] := \sum_{l=1}^{r} C_{ij}^l X_l, \qquad C_{ij}^l \in k, 1 \le i, j \le r. \tag{1.1}$$

Then $[\ ,\ ]$ is uniquely determined by these r^3 constants C_{ij}^l. These constants are called the **structure constants** of $\mathfrak{g}$. Now condition (i) in Definition 1.2 is equivalent to

$$\begin{aligned} & [X, Y] = -[Y, X], \qquad \forall X, Y \in \mathfrak{g} \\ \Leftrightarrow\ & \sum_{l=1}^{r} C_{ij}^l X_l = -\sum_{l=1}^{r} C_{ji}^l X_l, \qquad 1 \le i, j \le r \\ \Leftrightarrow\ & C_{ij}^l = -C_{ji}^l \qquad 1 \le i, j, l \le r. \end{aligned} \tag{1.2}$$

If we denote the $r \times r$ matrix C_l by

$$C_l := [C_{ij}^l]_{1 \le i,j \le r}$$

then condition (i) is equivalent to

$$C_l^T = -C_l, \qquad 1 \le l \le r$$

i.e., C_l is skew-symmetric. For Jacobi identity,

$$\begin{aligned} [X_i, [X_j, X_n]] &= [X_i, \sum_{l=1}^{r} C_{jn}^l X_l] \\ &= \sum_{l=1}^{r} C_{jn}^l [X_i, X_l] \\ &= \sum_{l=1}^{r} C_{jn}^l \sum_{s=1}^{r} C_{il}^s X_s \\ &= \sum_{s=1}^{r} \left\{ \sum_{l=1}^{r} C_{jn}^l C_{il}^s \right\} X_s \end{aligned}$$

for any $1 \le i, j, n \le r$. So Jacobi identity is equivalent to

$$\begin{aligned} & \sum_{s=1}^{r} \left(\sum_{l=1}^{r} (C_{jn}^l C_{il}^s + C_{ni}^l C_{jl}^s + C_{ij}^l C_{nl}^s) \right) X_s = 0 \\ \Leftrightarrow\ & \sum_{l=1}^{r} (C_{jn}^l C_{il}^s + C_{ni}^l C_{jl}^s + C_{ij}^l C_{nl}^s) = 0, \ \ 1 \le s \le r. \end{aligned} \tag{1.3}$$

Conversely, if $\mathfrak{g}$ is any r-dimensional k-module and r^3 constants $\{C_{ij}^l\}$ satisfies (1.2) and (1.3), then define the Lie bracket on $\mathfrak{g}$ by (1.1) and the conditions in Definition 1.2 are satisfied.

Notice that the structure constants depend on the choice of the basis of $\mathfrak{g}$.

Proposition 1.8. *Two Lie algebras* $\mathfrak{g}$ *and* $\mathfrak{h}$ *are isomorphic if and only if they have the same structure constants under suitable bases.*

Proof. Exercise. □

Definition 1.9. *If* A *is an algebra over* k*, then a* **derivation** *on* A *is a* k*-linear map* $D : A \longrightarrow A$ *satisfying*

$$D(XY) = (DX)Y + X(DY).$$

This is kind of product rule. For instance, if A is the set of all differentiable functions on $\mathbb{R}$, then $\frac{d}{dx}$ is a derivation on A.

Let the set $Der(A)$ be the set of all derivations on an algebra A. Then $Der(A)$ is a Lie algebra with the Lie bracket

$$[D, D'] = DD' - D'D.$$

To show this, we first we need to prove that $[D, D']$ is actually a derivation on A. For any $X, Y \in A$,

$$\begin{aligned}
& [D, D'](XY) \\
= \ & DD'(XY) - D'D(XY) \\
= \ & D((D'X)Y + X(D'Y)) - D'((DX)Y + X(DY)) \\
= \ & D((D'X)Y) + D(X(D'Y)) - D'((DX)Y) - D'(X(DY)) \\
= \ & (D(D'X))Y + (D'X)(DY) + (DX)(D'Y) + X(D(D'Y)) \\
& -(D'(DX))Y - (DX)(D'Y) - (D'X)(DY) - X(D'(DY)) \\
= \ & (DD'X - D'DX)Y + X(DD'Y - D'DY) \\
= \ & ((DD' - D'D)X)Y + X((DD' - D'D)Y) \\
= \ & ([D, D']X)Y + X([D, D']Y).
\end{aligned}$$

So $[D, D']$ is indeed a derivation on A. Note that $DerA$ is a subset of $End(A)$. Thus $DerA$ is a Lie algebra.

Definition 1.10. *Let* $\mathfrak{g}$ *be a Lie algebra. The* **adjoint representation** *of* $\mathfrak{g}$ *on* $\mathfrak{g}$ *is the representation*

$$ad : \mathfrak{g} \longrightarrow End(\mathfrak{g}),$$

defined by

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

Lemma 1.11. *We have*

(i) $ad(X)$ *is a derivation of* $\mathfrak{g}$ *for all* X *in* $\mathfrak{g}$*.*

(ii) The map $ad : \mathfrak{g} \longrightarrow End(\mathfrak{g})$ *is a Lie homomorphism of* $\mathfrak{g}$ *into* $Der(\mathfrak{g})$*.*

Proof. (i) By Jacobi identity, we have

$$\begin{aligned} ad(X)([Y,Z]) &= [X,[Y,Z]] \\ &= -[Z,[X,Y]] - [Y,[Z,X]] \\ &= [[X,Y],Z] + [Y,[X,Z]] \\ &= [(adX)Y,Z] + [Y,(adX)Z]. \end{aligned}$$

So $ad(X)$ is a derivation.

(ii) It remains to show that ad preserves Lie bracket, i.e., $ad[X,Y] = [adX, adY]$. For any $Z \in \mathfrak{g}$,

$$\begin{aligned} (ad[X,Y])Z &= [[X,Y],Z] \\ &= -[Z,[X,Y]] \\ &= [X,[Y,Z]] - [Y,[X,Z]] \\ &= [X,(adY)Z] - [Y,(adX)Z] \\ &= (adX)((adY)Z) - (adY)((adX)Z) \\ &= ((adX)(adY) - (adY)(adX))Z \\ &= [adX, adY]Z. \end{aligned}$$

□

Definition 1.12. *The* **centralizers** *of $\mathfrak{g}$ are all $X \in \mathfrak{g}$ such that $[X,Y] = ad(X)Y = 0$ for all $Y \in \mathfrak{g}$. The set of all centralizers of $\mathfrak{g}$ is called the* **center** *of $\mathfrak{g}$ and is denoted by $\mathfrak{C}$.*

Since for any X,Y in $\mathfrak{C}$, we have $[X,Y] = 0$. So $[[X,Y],Z] = 0$ for all $Z \in \mathfrak{g}$ and hence $[X,Y]$ belongs to $\mathfrak{C}$. It follows that $\mathfrak{C} \subseteq \mathfrak{g}$ is an abelian Lie algebra.

1.2 Subalgebras, Ideals and Quotient Algebras.

Definition 1.13. *If $\mathfrak{m}$ and $\mathfrak{n}$ are subsets of $\mathfrak{g}$, then $\mathfrak{m}+\mathfrak{n}$ and $[\mathfrak{m},\mathfrak{n}]$ are the modules generated by elements of the form $m+n$ and $[m,n]$ respectively, where $m \in \mathfrak{m}$ and $n \in \mathfrak{n}$.*

Lemma 1.14. *If $\mathfrak{m}, \mathfrak{n}$ and $\mathfrak{p}$ are subsets of $\mathfrak{g}$, then*

(i) $[\mathfrak{m}+\mathfrak{n}, \mathfrak{p}] \subseteq [\mathfrak{m}+\mathfrak{p}] + [\mathfrak{n}+\mathfrak{p}]$.

(ii) $[\mathfrak{m},\mathfrak{n}] = [\mathfrak{n},\mathfrak{m}]$.

(iii) $[\mathfrak{m},[\mathfrak{n},\mathfrak{p}]] \subseteq [\mathfrak{n},[\mathfrak{p},\mathfrak{m}]] + [\mathfrak{p},[\mathfrak{m},\mathfrak{n}]]$.

Proof. (i) is trivial. (ii) and (iii) follow directly from the condition (i) in Definition 1.2 and Jacobi identity respectively. □

Definition 1.15. *Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a subspace. Then* *(i)* $\mathfrak{h}$ *is a* **subalgebra** *if* $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$.

(ii) $\mathfrak{h}$ *is an* **ideal** *if* $[\mathfrak{h}, \mathfrak{g}] \subseteq \mathfrak{h}$.

A subalgebra $\mathfrak{h}$ is a subset of a Lie algebra $\mathfrak{g}$ and itself is also a Lie algebra.

Lemma 1.16.

(i) Every ideal is a subalgebra.

(ii) If $\mathfrak{h}$ *and* $\mathfrak{p}$ *are ideals, so are* $\mathfrak{h} + \mathfrak{p}$ *and* $\mathfrak{h} \cap \mathfrak{p}$.

Proof. Exercise. □

Example.

(7) Recall that $M_n(k)$ is a Lie algebra with Lie bracket $[X, Y] = XY - YX$. Subalgebras of $M_n(k)$ are called **matrix or (linear) Lie algebras.** Some important examples are the following.

(a) the set $\{\lambda I \in M_n(k) : \lambda \in k\}$ is an ideal of dimension 1.

(b) the set of all diagonal matrices in $M_n(k)$ is an **abelian** subalgebra of dimension n.

(c) The set $sl(n, k)$ of all matrices in $M_n(k)$ with zero trace is an ideal of dimension $n^2 - 1$.

(d) The set $o(n, k)$ of all skew-symmetric matrices in $M_n(k)$ is a subalgebra of dimension $\frac{n(n-1)}{2}$.

Lemma 1.17. *Let* M *be an element in* $M_n(k)$. *The set of all matrices* X *in* $M_n(k)$ *satisfying*

$$X^T M + MX = 0$$

forms a Lie subalgebra.

Proof. It suffices to show that this set is closed under $[\ .\ ]$. In fact, if $X^T M + MX = 0$ and $Y^T M + MY = 0$, then

$$\begin{aligned}
[X, Y]^T M + M[X, Y] &= (XY - YX)^T M + M(XY - YX) \\
&= Y^T X^T M - X^T Y^T M + MXY - MYX \\
&= -Y^T MX + X^T MY - X^T MY + Y^T MX \\
&= 0.
\end{aligned}$$

This proves the lemma. □

Example.

(8) If $n = 2m$, then the set $sp(n,k) := \{X \in M_n(k) : X^T F + FX = 0\}$ is a subalgebra of dimension $\frac{n(n+1)}{2}$ where

$$F := \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}.$$

In fact,

$$sp(n,k) = \left\{ X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & -X_{11}^T \end{bmatrix} \in M_n(k) : X_{12}^T = X_{12}, X_{21}^T = X_{21}. \right\}$$

Similarly, we let

$$\begin{aligned} B_n &:= \{X \in M_{2n}(k) : X^T K + KX = 0\} \\ D_n &:= \{X \in M_{2n+1}(k) : X^T S + SX = 0\} \end{aligned}$$

where

$$K := \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \quad \text{and} \quad S := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & I_n \\ 0 & I_n & 0 \end{bmatrix}.$$

Usually, we also let $A_n = sl(n+1,k)$ and $C_n = sp(2n,k)$. The Lie algebras A_n, B_n, C_n and D_n are called the **classical Lie algebras**.

Exercise. Show that B_n and D_n are subalgebras of dimensions $2n^2 - n$ and $2n^2 + n$. Moreover,

$$\begin{aligned} B_n &= \left\{ X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & -X_{11}^T \end{bmatrix} : X_{12}^T = -X_{12}, X_{21}^T = -X_{21} \right\} \\ D_n &= \left\{ X = \begin{bmatrix} 0 & \boldsymbol{u} & \boldsymbol{v} \\ -\boldsymbol{v}^T & X_{11} & X_{12} \\ -u^T & X_{21} & -X_{11}^T \end{bmatrix} : X_{12}^T = -X_{12}, X_{21}^T = -X_{21} \right\}. \end{aligned}$$

Definition 1.18. *Let $\mathfrak{h} \subseteq \mathfrak{g}$ be an ideal. Then $\mathfrak{g}/\mathfrak{h}$ is well-defined. Denote the image of $X \in \mathfrak{g}$ under the natural projection*

$$f : \mathfrak{g} \longrightarrow \mathfrak{g}/\mathfrak{h}$$

by $\overline{X}$, and define, on $\mathfrak{g}/\mathfrak{h}$,

$$[\overline{X}, \overline{Y}] := \overline{[X,Y]},$$

then $\mathfrak{g}/\mathfrak{h}$ is a Lie algebra called the **quotient algebra** *of $\mathfrak{g}$ with respect to $\mathfrak{h}$.*

Proposition 1.19. *Let $\mathfrak{h} \subseteq \mathfrak{g}$ be an ideal. Then the natural projection*

$$f : \mathfrak{g} \longrightarrow \mathfrak{g}/\mathfrak{h}$$

is a Lie homomorphism called the **canonical homomorphism**. *Conversely, if $f : \mathfrak{g} \longrightarrow \mathfrak{g}_1$ is a Lie homomorphism with kernel $\mathfrak{h} \subset \mathfrak{g}$, then $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ and the map*

$$\overline{f} : \mathfrak{g}/\mathfrak{h} \longrightarrow f(\mathfrak{g}) \subseteq \mathfrak{g}_1$$

defined by $\overline{f}(\overline{X}) = f(X)$ is an isomorphism called the **canonical isomorphism induced by** f.

Proof. Exercise. □

Remark. Let $\mathfrak{n} \subseteq \mathfrak{m} \subseteq \mathfrak{g}$ be ideals of $\mathfrak{g}$. Then $[\mathfrak{m}, \mathfrak{g}] \subseteq \mathfrak{n}$ if and only if $\mathfrak{m}/\mathfrak{n}$ is contained in the center of $\mathfrak{g}/\mathfrak{n}$.

1.3 Simple Lie Algebras.

Definition 1.20. *If $\mathfrak{g}$ has only $\{0\}$ and $\mathfrak{g}$ as its ideals, then $\mathfrak{g}$ is called a* **simple Lie algebra**.

Clearly, all Lie algebras of dimension 1 are simple. Since the Lie bracket is always zero in abelian Lie algebra, so all abelian Lie algebras of dimension greater than 1 are **not** simple. Thus, except for 1-dimensional Lie algebras, simple Lie algebras are **not** abelian.
Remark. The Lie algebras A_n, B_n, C_n, D_n and $o(n,k)$ are simple Lie algebras.

1.4 Direct Sum of Lie algebras.

Definition 1.21. *If $\mathfrak{g}_1, \mathfrak{g}_2, \cdots, \mathfrak{g}_n$ are ideals of $\mathfrak{g}$ such that any X in $\mathfrak{g}$ can be* **uniquely** *written as*

$$X = X_1 + X_2 + \cdots + X_n, \qquad X_i \in \mathfrak{g}_i,$$

then $\mathfrak{g}$ is said to be the **direct sum** *of $\mathfrak{g}_1, \mathfrak{g}_2, \cdots, \mathfrak{g}_n$, and is written as*

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_n.$$

The last expression is called a **direct sum decomposition** *of $\mathfrak{g}$.*

Lemma 1.22. *If $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_n$, then*

(i) $\mathfrak{g}_i \cap \mathfrak{g}_j = \{0\}$ *if* $i \neq j$;

(ii) $[\mathfrak{g}_i, \mathfrak{g}_j] = \{0\}$ *if* $i \neq j$;

(iii) ideals of $\mathfrak{g}_i$ are also ideals of $\mathfrak{g}$.

Proof. (i) is the usual property of direct sum. For (ii), since $[\mathfrak{g}_i, \mathfrak{g}_j] \subseteq [\mathfrak{g}_i, \mathfrak{g}] \subseteq \mathfrak{g}_i$, so $[\mathfrak{g}_i, \mathfrak{g}_j] \subseteq \mathfrak{g}_i \cap \mathfrak{g}_j$. Thus (ii) follows from (i). Lastly, if $\mathfrak{h} \subseteq \mathfrak{g}_i$ is an ideal of $\mathfrak{g}_i$, then $[\mathfrak{h}, \mathfrak{g}_j] = \{0\}$ for $i \neq j$ and hence

$$[\mathfrak{h}, \mathfrak{g}] \subseteq \sum_{j=1}^{n} [\mathfrak{h}, \mathfrak{g}_j] \subseteq [\mathfrak{h}, \mathfrak{g}_i] \subseteq \mathfrak{h}.$$

This proves (iii). □

Lemma 1.23. *If* $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_r \oplus \mathfrak{h}$ *and* $\mathfrak{h} = \mathfrak{g}_{r+1} \oplus \mathfrak{g}_{r+2} \oplus \cdots \oplus \mathfrak{g}_m$ *then* $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_m$.

Proof. Exercise. □

Remark. If X is in $M_n(k)$ with trace λ, then $X - \frac{\lambda}{n} I_n$ has zero trace and hence belongs to $sl(n, k)$. Thus, we have

$$M_n(k) = sl(n, k) \oplus \{\lambda I_n : \lambda \in k\}.$$

Chapter 2

Nilpotent and Solvable Lie Algebras

2.1 Nilpotent Lie Algebras.

Let $\mathfrak{g}$ be a Lie algebra. Define

$$\begin{aligned} \mathfrak{C}^{(0)}\mathfrak{g} &:= \mathfrak{g} \\ \mathfrak{C}^{(n+1)}\mathfrak{g} &:= [\mathfrak{g}, \mathfrak{C}^{(n)}\mathfrak{g}] \quad \forall n \geq 0. \end{aligned}$$

It is easy to show by induction on n that $\mathfrak{C}^{(n)}\mathfrak{g}$ are ideals of $\mathfrak{g}$ and hence we get a series of ideals of $\mathfrak{g}$ satisfying

$$\mathfrak{g} = \mathfrak{C}^{(0)}\mathfrak{g} \supseteq \mathfrak{C}^{(1)}\mathfrak{g} \supseteq \cdots \supseteq \mathfrak{C}^{(n)}\mathfrak{g} \supseteq \cdots.$$

Definition 2.1. *This series is called the* **descending central series** *of* $\mathfrak{g}$*. A Lie algebra* $\mathfrak{g}$ *is called* **nilpotent** *if its descending central series terminates, i.e, there is* $n > 0$ *such that* $\mathfrak{C}^{(n)}\mathfrak{g} = \{0\}$*. The smallest such* n *is called the* **nilpotent length** *of* $\mathfrak{g}$*.*

Proposition 2.2. *The following are equivalent:*

(i) $\mathfrak{g}$ *is nilpotent Lie algebra.*

(ii) There is $n > 0$ *such that*

$$[X_n, [\cdots, [X_1, X_0]] \cdots] = 0$$

for any $X_0, X_1, \cdots, X_n$ *in* $\mathfrak{g}$*.*

(iii) $\mathfrak{g}$ *is a* **succession of central extensions of ideals** $\mathfrak{g}_i$ *of* $\mathfrak{g}$*, i.e., there exists a chain of ideals* $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_n = \{0\}$ *such that* $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ *is in the center of* $\mathfrak{g}/\mathfrak{g}_{i+1}$ $\forall i$*, i.e.,*$[\mathfrak{g}, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i+1}$ $\forall i$*.*

Proof.

(i) $\Leftrightarrow$ (ii) The idea $\mathfrak{C}^{(n)}\mathfrak{g}$ is generated by elements $[X_n, [\cdots, [X_1, X_0]]\cdots]$.

(i) $\Rightarrow$ (iii) Let $\mathfrak{g}_i := \mathfrak{C}^{(i)}\mathfrak{g}\ \forall i \geq 0$. Then

$$\mathfrak{g} = \mathfrak{g}_0 \supseteq \mathfrak{g}_1 \supseteq \cdots \supseteq \mathfrak{g}_n = \{0\}.$$

Moreover, $[\mathfrak{g}, \mathfrak{g}_i] = [\mathfrak{g}, \mathfrak{C}^{(i)}\mathfrak{g}] = \mathfrak{C}^{(i+1)}\mathfrak{g} = \mathfrak{g}_{i+1}\ \forall i$.

(iii) $\Rightarrow$ (i) Clearly, $\mathfrak{C}^{(1)}\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{g}_0] \subseteq \mathfrak{g}_1$. For all $i \geq 0$, $\mathfrak{C}^{(i+1)}\mathfrak{g} = [\mathfrak{g}, \mathfrak{C}^{(i)}\mathfrak{g}] \subseteq [\mathfrak{g}, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i+1}$. It follows that $\mathfrak{C}^{(n)}\mathfrak{g} \subseteq \mathfrak{g}_n = \{0\}$.

$\square$

Remark. The descending central series $\{\mathfrak{C}^{(i)}\mathfrak{g}\}$ is the most rapidly descending chain with the properties described in (iii), i.e., if $\{\mathfrak{g}_i\}$ is as in (iii), then $\mathfrak{C}^{(i)}\mathfrak{g} \subseteq \mathfrak{g}_i\ \forall i$.

Example. Let

$$\begin{aligned} \mathfrak{t}(n,k) &:= \text{the set of all upper triangular matrices in } k \\ \mathfrak{n}(n,k) &:= \{X \in \mathfrak{t}(n,k) :\ X \text{ has equal diagonal elements}\}. \end{aligned}$$

Clearly, $\mathfrak{n}(n,k) \subset \mathfrak{t}(n,k)$ are Lie subalgebras of $M_n(k)$. For $1 \leq i \leq n$, define $\mathfrak{n}_{i-1}$ as the set of elements of $\mathfrak{n}(n,k)$ of the form

$$\begin{pmatrix} 0 & \cdots & 0 & x_{1i} & \cdots & x_{1n} \\ & \ddots & & \ddots & \ddots & \vdots \\ & & \ddots & & \ddots & x_{n-i+1,n} \\ & & & \ddots & & 0 \\ & & & & \ddots & \vdots \\ 0 & & & & & 0 \end{pmatrix}$$

and $\mathfrak{n}_n = \{0\}$. Then $\{\mathfrak{n}_i\}$ satisfies (iii) of Proposition 2.1 and so $\mathfrak{n}(n,k)$ is nilpotent.

Proposition 2.3.

(i) Subalgebras of nilpotent Lie algebras are nilpotent.

(ii) Homomorphic images of nilpotent Lie algebras are nilpotent.

(iii) Quotient algebras of nilpotent Lie algebras with respect to any ideals are nilpotent.

(iv) Direct sums of nilpotent algebras are nilpotent.

Proof. Exercise. $\square$

Remark. The conditions that $\mathfrak{h} \subset \mathfrak{g}$ is an ideal such that $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are nilpotent may not imply that g is also nilpotent. For example, let $\mathfrak{g} = \mathfrak{t}(n, k)$ and $\mathfrak{h}$ be the set of all diagonal matrices in $\mathfrak{g}$.

Example. Let V be a vector space and $F = (V_i)$ a **flag** in V, i.e, a sequence

$$\{0\} = V_0 \subset V_1 \subset \cdots \subset V_n = V$$

of subspaces of V such that $\dim V_i = i$ for all $i = 0, 1, \cdots, n$. Let

$$\mathfrak{n}(F) := \{u \in \mathrm{End}(V) : uV_i \subseteq V_{i-1} \ \forall i \geq 1\}.$$

$\mathfrak{n}(F)$ is an associative subalgebra of $\mathrm{End}(V)$ and is a Lie algebra under the bracket $[X, Y] = XY - YX$. In terms of a basis $\{V_i\}$ adapted to F, $\mathfrak{n}(F)$ consists of those endomorphisms which are strictly upper triangular. Thus $\mathfrak{n}(F) \subset \mathfrak{n}(n, k)$ and so it is nilpotent.

2.2 Main Theorems on Nilpotent Lie Algebras.

Let A be a matrix in $M_n(k)$. Then A is called nilpotent if and only if there is $n \geq 1$ such that $A^n = 0$. For example, every strictly upper or lower triangular matrices are nilpotent. The main theorem on nilpotent Lie algebra is

Theorem 2.4. *$\mathfrak{g}$ is nilpotent if and only if $ad(X)$ is nilpotent for all X in $\mathfrak{g}$.*

Theorem 2.5 (Engel's Theorem). *Let $\rho : \mathfrak{g} \longrightarrow End(V)$ be a linear representation of $\mathfrak{g}$ on a vector space V such that $\rho(X)$ is nilpotent $\forall X \in \mathfrak{g}$. Then there exists a flag $F = (V_i)$ in V such that $\rho(\mathfrak{g}) \subset \mathfrak{n}(F)$.*

Remarks.

(i) The converse of Engel's Theorem is trivial since every strictly upper matrix is nilpotent.

(ii) The meaning of Engel's Theorem is that if $\forall X \in \mathfrak{g}, \exists$ a flag $F(X) = (V_i(X))$ such that $\rho(X)V_i(X) \subseteq V_{i-1}(X)$ $\forall i$, then $\exists$ a flag $F = (V_i)$ which works for every $X \in \mathfrak{g}$. So in some sense it asserts that pointwise nilpotency implies uniform nilpotency.

(iii) Engel's Theorem $\Rightarrow$ Theorem 2.4:

($\Rightarrow$) If $\mathfrak{g}$ is nilpotent, from Proposition 2.2 (ii), there is $m > 0$ such that

$$[X_m, [X_{m-1}, \cdots, [X_2, X_1]] \cdots] = 0 \ \ \forall X_1, \cdots, X_m \in \mathfrak{g}.$$

In particular, for this m and any X in $\mathfrak{g}$, we have

$$(ad(X))^{m-1} Y = 0 \ \ \forall Y \in \mathfrak{g}.$$

So $ad(X)$ is nilpotent. In fact, we have proved that there is a unique m such that $(ad(X))^m = 0$ for all X in $\mathfrak{g}$.

($\Leftarrow$) If $ad(X)$ is nilpotent for all X in $\mathfrak{g}$, then applying Engel's Theorem with $\rho(X) = ad(X)$, there exists a flag $\{0\} = \mathfrak{g}_0 \subseteq \mathfrak{g}_1 \subseteq \cdots \subseteq \mathfrak{g}_n = \mathfrak{g}$ of subspaces of $\mathfrak{g}$ such that $(ad(X))\mathfrak{g}_i \subseteq \mathfrak{g}_{i-1}$ for all i and $X \in \mathfrak{g}$, i.e, $[\mathfrak{g}, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i-1}$ for all i. Hence by Proposition 2.2 (iii), $\mathfrak{g}$ is nilpotent.

Theorem 2.6. *Under the same hypothesis as that of Engel's Theorem, if* $V \neq \{0\}$, *then there exists* $0 \neq v \in V$ *such that* $\rho(X)v = 0\ \forall X \in \mathfrak{g}$

Remarks.

(i) Roughly speaking, Theorem 2.6 says that the existence of "pointwise eigenvector" implies the existence of "uniform eigenvector".

(ii) Engel's Theorem is equivalent to Theorem 2.6:

($\Rightarrow$) Since there is a flag $F = (V_i)$ such that $\rho(\mathfrak{g}) \subseteq \mathfrak{n}(F)$, so if $v \neq 0 \in V_1$, then $\rho(X)v = 0$ for all X in $\mathfrak{g}$.

($\Leftarrow$) Use induction on the dimension of V.

We now come to the proof of Theorem 2.6.

Proof of Theorem 2.6: For the sake of simplicity, we assume that $\mathfrak{g} \subseteq \mathrm{End}(V)$. We prove the theorem by induction on $\dim \mathfrak{g}$. It is trivial for the case of $\dim \mathfrak{g} = 0$. Suppose that the theorem holds for any Lie algebra with dimension $< r (r \geq 1)$.

Suppose that $\dim \mathfrak{g} = r$ and there is an ideal $\mathfrak{h} \subset \mathfrak{g}$ of $\dim \mathfrak{h} = r - 1$. By the induction hypothesis, there is $v_1 \neq 0 \in V$ such that $Xv_1 = 0$ for all $X \in \mathfrak{h}$. Let

$$U := \{u \in V : Xu = 0\ \forall X \in \mathfrak{h}\}.$$

Then $U \neq \{0\}$ since $v_1 \in U$. For any $Y \in \mathfrak{g}\backslash\mathfrak{h}$, we claim that U is invariant under Y. In fact, for any $u \in U$ and any $X \in \mathfrak{h}$, we have

$$\begin{aligned} XYu &= YXu + [X,Y]u \\ &= Y \cdot 0 + 0 \\ &= 0 \end{aligned}$$

because $[X, Y] \in \mathfrak{h}$. Hence, Yu belongs to U and prove our claim. Thus we can consider the restriction of Y on U and let

$$Y_1 := Y|_U : U \longrightarrow U$$

be the linear transformation for this restriction.

Since Y is nilpotent, so is Y_1. It follows that there is $v \neq 0$ in U such that $Yv = Y_1 v = 0$. For this vector v, since v belongs to U, so we have

$$Yv = 0 \text{ and } Xv = 0\ \forall X \in \mathfrak{h}.$$

However, $\mathfrak{h}$ is of dimension $r - 1$. Hence $\mathfrak{h}$ and Y span the whole space $\mathfrak{g}$. Therefore, for any x in $\mathfrak{g}$, we write $X = X_1 + \lambda Y$ for some X_1 in $\mathfrak{h}$ and λ in k. Thus,

$$Xv = X_1 v + \lambda Y v = 0.$$

This proves our theorem.

It remains to show that there is an ideal $\mathfrak{h} \subseteq \mathfrak{g}$ of dimension $r-1$.

Now let $\mathfrak{h} \subseteq \mathfrak{g}$ be a subalgebra of $\mathfrak{g}$ with $\dim \mathfrak{h} < \dim \mathfrak{g} = r$. Now define

$$\begin{array}{rcl} \mathfrak{h} \times \mathfrak{g}/\mathfrak{h} & \longrightarrow & \mathfrak{g}/\mathfrak{h} \\ (X, \overline{Y}) & \longmapsto & ad(\overline{X})\overline{Y} = \overline{[X,Y]}. \end{array}$$

So $\mathfrak{g}/\mathfrak{h}$ is a $\mathfrak{h}$-vector space and $ad(\overline{X})$ is nilpotent for $X \in \mathfrak{h}$. Since $\dim \mathfrak{h} < \dim \mathfrak{g}$, so induction assumption applies and get that there is $\overline{Y} \neq \overline{0}$ in $\mathfrak{g}/\mathfrak{h}$ such that $ad(\overline{X})\overline{Y} = \overline{0}$ for any X in $\mathfrak{h}$. It follows that for all X in $\mathfrak{h}$, we have

$$\begin{array}{rcl} (adY)X & = & [Y,X] \\ & = & -[X,Y] \\ & = & -(adX)Y \\ & \in & \mathfrak{h} \end{array}$$

because $(ad\overline{X})\overline{Y} = \overline{0}$. Since $Y \notin \mathfrak{h}$, so $\{Y, \mathfrak{h}\}$ generates a subalgebra $\widetilde{Y}$ of $\mathfrak{g}$ of dimension equals to $\dim \mathfrak{h} + 1$ which contains $\mathfrak{h}$ as an ideal because $[Y, X] \in \mathfrak{h}$ for any X in $\mathfrak{h}$. So $[\widetilde{Y}, \mathfrak{h}] \subseteq \mathfrak{h}$. Finally, we start from $\dim \mathfrak{h} = 0$, we obtain an ideal of $\mathfrak{g}$ of dimension $r-1$. □

Corollary 2.7. *Let $\rho : \mathfrak{g} \longrightarrow$ End (V) be a linear representation of $\mathfrak{g}$ on a vector space V such that $\rho(X)$ is nilpotent for all X in $\mathfrak{g}$, then $\rho(\mathfrak{g})$ is isomorphic to a subalgebra of $\mathfrak{n}(n,k)$, in particular, $\rho(\mathfrak{g})$ is nilpotent, i.e., every algebra of nilpotent matrices is isomorphic to a subalgebra of $\mathfrak{n}(n,k)$.*

2.3 Solvable Lie Algebras.

Let $\mathfrak{g}$ be a Lie algebra. Define

$$\begin{array}{rcl} \mathfrak{D}^{(0)}\mathfrak{g} & := & \mathfrak{g}, \\ \mathfrak{D}^{(n+1)}\mathfrak{g} & := & [\mathfrak{D}^{(n)}\mathfrak{g}, \mathfrak{D}^{(n)}\mathfrak{g}], \quad \forall n \geq 0. \end{array}$$

It can be shown by induction on n that $\mathfrak{D}^{(n)}\mathfrak{g}$ are ideals of $\mathfrak{g}$ and hence we get a series of ideals satisfying

$$\mathfrak{g} = \mathfrak{D}^{(0)}\mathfrak{g} \supseteq \mathfrak{D}^{(1)}\mathfrak{g} \supseteq \cdots \supseteq \mathfrak{D}^{(n)}\mathfrak{g} \supseteq \cdots.$$

Definition 2.8. *The series is called the* **derived series** *of $\mathfrak{g}$. A Lie algebra $\mathfrak{g}$ is called* **solvable** *if its derived series terminates, i.e., there is $n > 0$ such that $\mathfrak{D}^{(n)}\mathfrak{g} = \{0\}$. The smallest such n is called the* **solvable length** *of $\mathfrak{g}$.*

Remark. We sometimes simply denote $\mathfrak{D}\mathfrak{g} := \mathfrak{D}^{(1)}\mathfrak{g}$. Since $\mathfrak{D}^{(i+1)}\mathfrak{g} = \mathfrak{D}^{(i)}(\mathfrak{D}\mathfrak{g})$, it follows that $\mathfrak{g}$ is solvable if and only if $\mathfrak{D}\mathfrak{g}$ is solvable.

Proposition 2.9. *The following are equivalent:*

(i) *$\mathfrak{g}$ is solvable.*

(ii) *There is $n > 0$ such that for every family of 2^n elements in $\mathfrak{g}$,*

$$[[[\cdots],[\cdots]],[[\cdots],[\cdots]]] = 0.$$

(iii) *$\mathfrak{g}$ is a* **successive abelian extension of ideals** *$\mathfrak{g}_i$ of $\mathfrak{g}$, i.e., there exists a chain of ideals $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \mathfrak{g}_n = \{0\}$ such that $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is abelian for all i, i.e. $[\mathfrak{g}_i, \mathfrak{g}_i] \subset \mathfrak{g}_{i+1}$, $\forall i$.*

Proof. To prove (i) from (iii), let $\mathfrak{g}_i$ be as given. Then one can prove by induction on n that $\mathfrak{D}^{(i)}\mathfrak{g} \subset \mathfrak{g}_i$ for all i (induction step: $\mathfrak{D}^{(k+1)}\mathfrak{g} = [\mathfrak{D}^{(k)}\mathfrak{g}, \mathfrak{D}^{(k)}\mathfrak{g}] \subset [\mathfrak{g}_k, \mathfrak{g}_k] \subset \mathfrak{g}_{k+1}$). Then $\mathfrak{g}_n = \{0\}$ implies $\mathfrak{D}^{(n)}\mathfrak{g} = \{0\}$. □

Proposition 2.10.

(i) *Subalgebras of solvable Lie algebras are solvable.*

(ii) *Homomorphic images of solvable Lie algebras are solvable.*

(iii) *Quotient algebras of solvable Lie algebras (by ideals) are solvable.*

(iv) *Direct sum of solvable Lie algebras are solvable.*

(v) *If $\mathfrak{h} \subset \mathfrak{g}$ is an ideal such that $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are solvable, then $\mathfrak{g}$ is also solvable.*

Proof. Exercise. □

Remark. Nilpotent algebras are solvable (since $\mathfrak{D}^{(i)}\mathfrak{g} \subset \mathfrak{C}^{(i)}\mathfrak{g}$).

Examples.

(1) Let $\mathfrak{t}(n, k)$ (upper triangular $n \times n$ matrices with entries in k) and $\mathfrak{n}_i$ be as in the example in §2.1. Then it is easy to check that $\mathfrak{D}^{(i)}\mathfrak{t}(n, k) \subset \mathfrak{n}_i$ for all i. Thus, $\mathfrak{t}(n, k)$ is solvable Lie algebra (with solvable length $\leq n$).

(2) Let $F = (V_i)$ be a flag of a vector space V. Let

$$S(F) := \{X \in End(V) : XV_i \subset V_i, \quad \forall i\}.$$

Let $\{v_1, \cdots, v_n\}$ be an adapted basis of F. Then with respect to this basis, every element in $S(F)$ is an upper triangular matrix, i.e.,

$$S(F) \subset \mathfrak{t}(n, k).$$

Thus $S(F)$ is solvable.

2.4 Main Theorems on Solvable Lie Algebras.

In this section, we assume that k is algebraically closed.

Theorem 2.11. *Let $\mathfrak{g}$ be a solvable Lie algebra, and $\rho : \mathfrak{g} \to End(V)$ be a linear representation. Then there exists a flag $F = (V_i)$ of V such that $\rho(\mathfrak{g}) \subset S(F)$. (i.e., with respect to an adapted basis of F, $\rho(\mathfrak{g}) \subset \mathfrak{t}(n,k)$.)*

Theorem 2.12 (Lie). *Under the hypothesis of Theorem 2.11, if $V \neq \{0\}$, then $\exists 0 \neq v \in V$ such that v is a common eigenvector of $\rho(X)$, $X \in \mathfrak{g}$. (i.e. $\exists \lambda = \lambda_v : \mathfrak{g} \to k$ such that $\rho(X)v = \lambda(X)v \ \ \forall X \in \mathfrak{g}$. Note that λ is linear.)*

Proof of Theorem 2.11 assuming Theorem 2.12. By Theorem 2.12, there exists a common vector v_1 of $\rho(X)$, $X \in \mathfrak{g}$, and let $V_1 = Span(v_1)$. Let $\bar{\rho} : \mathfrak{g} \to End(V/V_1)$ be the homomorphism induced by ρ (i.e. $\bar{\rho}(X)(\bar{v}) = \overline{\rho(X)(v)}$ for all $X \in \mathfrak{g}$ and $\bar{v} \in V/V_1$, where $\bar{v}$ denotes the class of $v \in V$ in V/V_1). Then by Theorem 2.12 again, there exists a common eigenvector $\bar{v}_2$ of $\bar{\rho}(X)$, $X \in \mathfrak{g}$. Then $v_2 \in V - V_1$. Let $V_2 = Span(v_1, v_2)$. Note that $\bar{\rho}(X)\bar{v}_2 = \lambda(X)\bar{v}_2$ implies $\rho(X)v_2 = \lambda(X)v_2 + \eta(X)v_1$. Repeating the above argument, we get a flag $F = (V_i)$ with $V_i = Span(v_1, \cdots, v_i)$ with $\rho(\mathfrak{g}) \in S(F)$.

Proof of Theorem 2.12. We prove Theorem 2.12 by induction on $\dim \mathfrak{g}$.
Theorem 2.12 is trivially true when $\dim \mathfrak{g} = 0$.

When $\dim \mathfrak{g} > 0$, since $\mathfrak{g}$ is solvable, it follows that $\mathfrak{D}\mathfrak{g} \neq \mathfrak{g}$. Take a codimension one subspace $\mathfrak{h}$ such that $\mathfrak{D}\mathfrak{g} \subset \mathfrak{h}$. Then $\mathfrak{h}$ is an ideal (since $[\mathfrak{g}, \mathfrak{h}] \subset [\mathfrak{g}, \mathfrak{g}] = \mathfrak{D}\mathfrak{g} \subset \mathfrak{h}$), and thus $\mathfrak{h}$ is solvable.
By induction assumption (applied to $\mathfrak{h}$), $\exists 0 \neq v \in V$ and $\lambda_v : \mathfrak{h} \to k$ such that

$$\rho(X)v = \lambda_v(X)v \ \ \forall X \in \mathfrak{h},$$

i.e., v is a common eigenvector of $\rho(\mathfrak{h})$.
Consider the set (of common eigenvectors of $\rho(\mathfrak{h})$) given by

$$W = \{w \in V \mid \exists \lambda_w \text{ such that } \rho(X)w = \lambda_w(X)w \ \ \forall X \in \mathfrak{h}\}.$$

Then $W \neq \{0\}$.
Assume that W is an invariant subspace of $\rho(\mathfrak{g})$ (i.e. $\rho(X)(W) \subset W \ \ \forall X \in \mathfrak{g}$.). Take $Y \in \mathfrak{g} - \mathfrak{h}$. Then since k is algebraically closed, it follows that $\rho(Y)|_W \in End(W)$ has a non-zero eigenvector w for $\rho(Y)$. Then w is a common eigenvector for $\rho(Span(\mathfrak{h}, Y)) = \rho(\mathfrak{g})$, which will finish the proof of Theorem 2.11. Thus it remains to check that W is invariant under $\rho(\mathfrak{g})$.
Check: For simplicity, we assume $\mathfrak{g} \subseteq End(V)$. Then

W is invariant under $\rho(\mathfrak{g})$

$\Leftrightarrow$ for each $w \in W$, $X \in \mathfrak{g}$, Xw is a common eigenvector of $\rho(\mathfrak{h})$

$\Leftrightarrow$ for each $w \in W$, $X \in \mathfrak{g}$, $\exists \lambda_{Xw} : \mathfrak{h} \to k$ such that $HXw = \lambda_{Xw}(H)Xw$ for all $H \in \mathfrak{h}$.

But

$$\begin{aligned} HXw &= XHw + [H,X]w \\ &= \lambda_w(H)Xw + \lambda_w([H,X])w. \end{aligned} \tag{2.1}$$

Thus it remains to check that $\lambda_w([H,X]) = 0$ (which will imply W is invariant under $\rho(\mathfrak{g})$ with $\lambda_{Xw} = \lambda_w$.)
Check: Let i be the largest integer such that $\{w, Xw, \cdots, X^{i-1}w\}$ is linearly independent and let $V_i := < w, Xw, \cdots, X^{i-1}w >$.
Then $X^i w \in V_i$ and thus V_i is invariant under X. From (2.1), we have

$$\begin{aligned} HX^2w &= XHXw + [H,X]Xw \\ &= XHXw + X[H,X]w + [[H,X],X]w \\ &= \lambda_w(H)X^2w + \lambda_w([H,X])Xw + \lambda_w([H,X])Xw + \lambda_w([[H,X],X])w. \end{aligned}$$

Similarly we can show that V_i is invariant under $\rho(\mathfrak{h})$, and for any $H \in \mathfrak{h}$, with respect to the basis $\{w, Xw, \cdots, X^{i-1}w\}$, the matrix of $\rho(H)|_{V_i}$ is an upper triangular matrix with diagonal entries all equal to $\lambda_w(H)$. Thus we have $tr_{V_i} H = \lambda_w(H) \cdot \dim V_i$ for any $H \in \mathfrak{h}$.
Replacing H by $[H,X] (\in \mathfrak{h})$, it follows that

$$\begin{aligned} \lambda_w([H,X]) &= tr_{V_i}[H,X]/\dim V_i \\ &= tr_{V_i}(HX - XH)/\dim V_i \\ &= 0. \end{aligned}$$

Thus we have finished the proof of Theorem 2.11.

Corollary 2.13. *For any Lie algebra $\mathfrak{g}$ over k, $\mathfrak{g}$ is solvable $\Leftrightarrow$ $\mathfrak{D}\mathfrak{g}$ is nilpotent. Hence $\mathfrak{D}\mathfrak{g}$ is solvable $\Leftrightarrow$ $\mathfrak{D}\mathfrak{g}$ is nilpotent.*

Proof. We only have to prove the first statement, which will imply the second statement.
Consider the adjoint representation $ad : \mathfrak{g} \to End(\mathfrak{g})$ given by $ad(X)(Y) = [X,Y]$.
Since $\mathfrak{g}$ is solvable, it follows that $ad(\mathfrak{g})$ is also solvable. By Theorem 2.11, $\exists$ a flag $F = (\mathfrak{g}_i)$ of $\mathfrak{g}$ ($\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_n = \{0\}$) such that $ad(\mathfrak{g}) \subset S(F)$ (i.e., $[\mathfrak{g}, \mathfrak{g}_i] \subset \mathfrak{g}_i$, and thus each $\mathfrak{g}_i$ is an ideal).
Now for $X = [Y,Z] \in \mathfrak{D}\mathfrak{g}$, we have

$$ad(X)(\mathfrak{g}_i/\mathfrak{g}_{i+1}) = ad([Y,Z])(\mathfrak{g}_i/\mathfrak{g}_{i+1}) = [ad(X), ad(Y)](\mathfrak{g}_i/\mathfrak{g}_{i+1}) = 0,$$

since $End(\mathfrak{g}_i/\mathfrak{g}_{i+1}) \cong k$ as $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is one-dimensional.
Thus, $ad(X)$ is nilpotent on $\mathfrak{g}$ (and hence on $\mathfrak{D}\mathfrak{g}$).
By Engel's theorem, $\mathfrak{D}\mathfrak{g}$ is nilpotent. ☐

Example. The only non-zero solvable ideal of $M_2(\mathbb{C})$ is $\{\lambda I \mid \lambda \in \mathbb{C}\}$.

Proof. Recall the basis of $M_2(\mathbb{C}) \cong sl_2(\mathbb{C}) \oplus \mathbb{C}I$ given by

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Recall also that

$$[X,Y] = H, \quad [H,X] = 2X, \quad [H,Y] = -2Y.$$

It is easy to check that $\{\lambda I\}$ is a solvable ideal. Let $\mathfrak{h} \neq \{\lambda I\}$ be a solvable ideal of $M_2(\mathbb{C})$. Then $\mathfrak{h} + \{\lambda I\} \neq \{\lambda I\}$ is also a solvable ideal of $M_2(\mathbb{C})$. Thus WLOG, we may assume that $\mathfrak{h} \supset \{\lambda I\}$.
Take $v \in \mathfrak{h}\backslash\{\lambda I\}$. Then $v = \alpha X + \beta Y + \gamma H + \delta I$, where not all α, β, γ are 0. Since $I \in \mathfrak{h}$, it follows that we may assume $\delta = 0$. Suppose $\alpha \neq 0$. Then $[H, v] = 2\alpha X - 2\beta Y \in \mathfrak{h}$. Then $[2\alpha X - 2\beta Y, Y] = 2\alpha H \in \mathfrak{h}$, which implies $H \in \mathfrak{h}$. Then $[H, X] = 2X \in \mathfrak{h}$, and $[H, Y] = -2Y \in \mathfrak{h}$. Thus, $\mathfrak{h} = M_2(\mathbb{C})$. But then $\mathfrak{D}^{(1)} M_2(\mathbb{C}) = sl_2(\mathbb{C})$. Thus $\mathfrak{D}^{(2)} M_2(\mathbb{C}) = \mathfrak{D}^{(1)} sl_2(\mathbb{C}) = sl_2(\mathbb{C})$, and $\mathfrak{D}^{(n)} M_2(\mathbb{C}) = sl_2(\mathbb{C})$ for all $n \geq 2$, and this contradicts the assumption that $\mathfrak{h}$ is solvable. The other cases when $\beta \neq 0$ and when $\gamma \neq 0$ are similar. ☐

2.5 Cartan's criterion for solvable Lie algebras.

For simplicity, we assume that $k = \mathbb{C}$.

Definition 2.14. *An element $s \in End(V)$ is said to be* **semi-simple** *if s is diagonalizable.*

Lemma 2.15. *Each $u \in End(V)$ admits a unique decomposition of the form*

$$u = s + n \quad \text{with} \quad sn = ns$$

for some semi-simple s and nilpotent $n \in End(V)$ (which will simply be called the canonical decomposition of u).

Proof. Let $\Pi_i(t - \lambda_i)^{m_i}$ be the characteristic polynomial of u, and consider the decomposition of V with respect to u given by $V = \oplus_i V_i$, where $V_i = ker(u - \lambda_i I)^{m_i}$. From the Jordan canonical form of u, it is easy to see that one can let s to be given by $\lambda_i I_{V_i}$ on each V_i, and $n = u - s$.
To prove the uniqueness, suppose $u = s + n$, $ns = sn$ is a canonical decomposition of u.
Then s commutes with n (and s) $\Rightarrow$ s commutes with u (and hence with each $(u - \lambda_i)^{m_i}$). Thus, $sV_i \subset V_i$ for each i. Then it is easy to see that s is also semi-simple on each V_i. Since $su = us$, u will keep each eigenspace of s in V_i invariant. Since $s - u$ is nilpotent, it follows that each eigenvalue of s on V_i is equal to λ_i. Hence $s = \lambda_i I_{V_i}$ on V_i. Thus s is uniquely determined, and n is thus also uniquely determined. ☐

Theorem 2.16 (Cartan's criterion). *Let $\mathfrak{g}$ be a Lie subalgebra of $End(V)$. Then the following are equivalent:*

(i) *$\mathfrak{g}$ is solvable;*

(ii) $tr(XY) = 0 \quad \forall X \in \mathfrak{g}, Y \in \mathfrak{D}\mathfrak{g}$.

Proof. $(i) \Rightarrow (ii)$. Choose a flag $F = (V_i)$ such that $\mathfrak{g} \subset S(F)$. Then for any $X \in \mathfrak{g}$, $Y = [Z, W] \in \mathfrak{D}\mathfrak{g}$, we have

$$tr_V(XY) = \sum_i tr_{V_{i+1}/V_i}(XY) = \sum_i tr_{V_{i+1}/V_i}(X[Z,W]) = 0.$$

$(ii) \Rightarrow (i)$. It suffices to show that $\mathfrak{D}\mathfrak{g}$ is solvable.
Thus it suffices to show that $\mathfrak{D}\mathfrak{g}$ is nilpotent
By Engel's theorem, it suffices to show that each $u \in \mathfrak{D}\mathfrak{g}$ is nilpotent.
Write $u = s + n$ the canonical decomposition of u, where s and n are the semi-simple and nilpotent parts of u respectively.
Consider the Jordan decomposition of V with respect to u given by $V = \oplus_i V_i$, $V_i = ker(u - \lambda_i I)^{m_i}$ as in Lemma 2.15 above.
Define $\bar{s} \in End(V)$ given by

$$\bar{s} = \oplus \overline{\lambda_i} I_{V_i}.$$

Observe that

$$\begin{aligned} 0 = tr(u\bar{s}) = &\sum_i \dim V_i \cdot |\lambda_i|^2 \\ \Rightarrow \quad & \lambda_i = 0 \\ \Rightarrow \quad & s = 0 \\ \Rightarrow \quad & u \text{ is nilpotent.} \end{aligned}$$

Thus it suffices to show that $tr(u\bar{s}) = 0$.
Since $u \in \mathfrak{D}\mathfrak{g}$, we may write $u = \sum_\alpha c_\alpha [x_\alpha, y_\alpha]$.
Since $tr([a,b]c) = tr(abc - bac) = tr(b[c,a])$, we have

$$\begin{aligned} tr(u\bar{s}) &= \sum_\alpha c_\alpha \cdot tr([x_\alpha, y_\alpha]\bar{s}) \\ &= \sum_\alpha tr(y_\alpha[\bar{s}, x_\alpha]). \end{aligned}$$

By (ii), it suffices to show that $[\bar{s}, x_\alpha] \in \mathfrak{D}\mathfrak{g}$ (Problem: $\bar{s}$ may not be in $\mathfrak{g}$.) By Lemma 2.17 below, we have

$$\begin{aligned} & [\bar{s}, x_\alpha] \\ = \quad & ad(\bar{s})(x_\alpha) \\ = \quad & \overline{ad(s)}(x_\alpha) \quad \text{(by (ii))} \\ = \quad & P(ad(u))(x_\alpha) \quad \exists P(t) \text{ with } P(0) = 0 \quad \text{(by (i), (iii))} \\ \in \quad & \mathfrak{D}\mathfrak{g}, \end{aligned}$$

which finishes the proof of Theorem 2.16 $\square$

Thus we still owe the reader the following

Lemma 2.17. *Write $u = s + n \in End(V)$ the canonical decomposition. Then*

(i) $ad(s)$ is the semi-simple part of $ad(u)$,

(ii) $ad(\bar{s}) = \overline{ad(s)}$,

(iii) For any $U \in End(W)$, write $U = S + N$, then $\overline{S} = P(U)$ for some polynomial $P(t)$ such that $P(0) = 0$.

Proof. Recall that for $u \in End(V)$, $ad(u) \in End(End(V))$ is given by $ad(u)x = [u, x]$.
Let $\tilde{V} := V \otimes V^*$.
Then $End(V)$ acts on $\tilde{V}$ via the diagonal action as follows: for any $u \in End(V)$, u acts on $\tilde{V}$ by

$$\tilde{u} = u \otimes 1 - 1 \otimes u^*,$$

where $u^* \in End(V^*)$ is the transpose given by

$$\langle u^* v^*, w\rangle = \langle v^*, uw\rangle, \quad v^* \in V^*, w \in V.$$

[In fact, in terms of a basis $\{x_1, \cdots, x_n\}$ of V and its dual basis $\{x_1^*, \cdots, x_n^*\}$ of V^*, one has $u_{ij} = \langle x_j^*, ux_i\rangle$, and

$$\begin{aligned} u_{ij}^* &= \langle x_j^{**}, u^* x_i^*\rangle \\ &= \langle u^* x_i^*, x_j\rangle \\ &= \langle x_i^*, ux_j\rangle \\ &= u_{ji}.] \end{aligned}$$

In particular, one has $(uv)^* = v^* u^*$.
Notice that $\tilde{V} \cong End(V)$ with the isomorphism given by $x \otimes y^* \leftrightarrow \langle y^*, \ \rangle x$.
Under this isomorphism, one has $End(\tilde{V}) \cong End(End(V))$ with $\tilde{u} \leftrightarrow ad(u)$.
[Check: Take $v = x \otimes y^* \in \tilde{V} \leftrightarrow \langle y^*, \ \rangle x \in End(V)$.

$$\begin{aligned} & \tilde{u}(x \otimes y^*) \\ = \;& (u \otimes 1 - 1 \otimes u^*)(x \otimes y^*) \\ = \;& u(x) \otimes y^* - x \otimes u^* y^* \\ \leftrightarrow \;& \langle y^*, \ \rangle u(x) - \langle u^* y^*, \ \rangle x \\ = \;& u(\langle y^*, \ \rangle x) - \langle y^*, u(\ )\rangle x \\ = \;& uv - vu \\ = \;& ad(u)(v). \end{aligned}$$

Now to prove (i), it suffices to check that $\tilde{u} = \tilde{s} + \tilde{n}$ is the canonical decomposition of $\tilde{u}$ with semi-simple part $\tilde{s}$ and nilpotent part $\tilde{n}$.
Let $\{x_i\}$ be a basis of V consisting of eigenvectors of s with eigenvalues $\{\lambda_i\}$.

Let $\{x_i^*\}$ be its dual basis of V^*.
Then it is easy to see that $\{x_i \otimes x_j^*\}$ is a basis of $\tilde{V}$ consisting of eigenvectors of $\tilde{s}$ with each $x_i \otimes x_j^*$ corresponding to the eigenvalue $\lambda_i - \lambda_j$. Thus $\tilde{s}$ is semi-simple.
By the binomial theorem, it is easy to check that $\tilde{n}^k = (n \otimes 1 - 1 \otimes n^*)^k$ is nilpotent.
Also, one has

$$\begin{aligned}
\tilde{s}\tilde{n} &= (s \otimes 1 - 1 \otimes s^*)(n \otimes 1 - 1 \otimes n^*) \\
&= sn \otimes 1 - s \otimes n^* - n \otimes s^* + 1 \otimes s^* n^* \\
&= ns \otimes 1 - s \otimes n^* - n \otimes s^* + 1 \otimes n^* s^* \\
&= \tilde{n}\tilde{s}.
\end{aligned}$$

Thus $\tilde{u} = \tilde{s} + \tilde{n}$ is the canonical decomposition of $\tilde{u}$, which gives (i).
Let $\{x_i \otimes x_j^*\}$ be as above. Then it is easy to check that the basis consists of eigenvectors of both $ad(\bar{s})$ and $\overline{ad(s)}$ with each $x_i \otimes x_j^*$ corresponding to $\bar{\lambda}_i - \bar{\lambda}_j$, which gives (ii).
Let $\Pi_i(T - \lambda_i)^{m_i}$ be the characteristic polynomial of U. Then since $\mathbb{C}[T]$ is a Euclidean domain, it is easy to see that the ideals (T) and $((T - \lambda_i)^{m_i})$ for various i are pairwise relatively prime. Thus by the Chinese remainder theorem, one can find a polynomial $P(T)$ such that

$$\begin{aligned}
P(T) &\equiv \overline{\lambda_i} \mod (T - \lambda_i)^{m_i} \forall i, \text{ and} \\
P(T) &\equiv 0 \mod T.
\end{aligned}$$

Then $P(0) = 0$, and on each $W_i = ker(U - \lambda_i I)^{m_i}$, $P(U) = \overline{\lambda_i} I + Q(U)(U - \lambda_i I)^{m_i} = \overline{\lambda_i} I$, where $Q(T)$ is a polynomial. Hence we have $P(U) = \overline{S}$, which gives (iii). $\square$

Chapter 3

Semi-simple Lie Algebras.

Throughout this chapter, k is an algebraically closed field of characteristic zero and all algebras and vector spaces are finite dimensional over k.

3.1 Semi-simple Lie Algebras.

Definition 3.1. *A Lie algebra is* **semi-simple** *if and only if it does not contain any non-zero solvable ideal.*

Proposition 3.2. $\mathfrak{g}$ *is semi-simple if and only if* $\mathfrak{g}$ *doesn't contain any non-zero abelian ideal.*

Proof. Since any non-zero abelian ideal is solvable, so the "only if " part is trivial. Conversely, if $\mathfrak{n}$ is a non-zero solvable ideal, then there exists $n > 0$ such that $\mathfrak{D}^{(n-1)}\mathfrak{n} \neq \{0\}$ and $\mathfrak{D}^{(n)}\mathfrak{n} = \{0\}$. Thus $\mathfrak{D}^{(n-1)}\mathfrak{n}$ is a non-zero abelian ideal. $\square$

Remarks.

1. Solvable Lie algebras are not semi-simple because $\mathfrak{g}$ itself is a non-zero solvable ideal of $\mathfrak{g}$.
2. Abelian Lie algebras are not semi-simple because of Proposition 3.2.
3. The 1-dimensional Lie algebras are the only Lie algebras which are simple but not semi-simple.
4. Hence from now on, excluding the case $n = 1$, we say that simple Lie algebras are semi-simple.
5. The center of a Lie algbera is solvable. Hence semi-simple Lie algebra has center zero.

Proposition 3.3. *Every Lie algebra has a largest solvable ideal* $\mathfrak{r}$ *called the* **radical** *of* $\mathfrak{g}$.

Proof. It follows from the fact that if $\mathfrak{n}_1$ and $\mathfrak{n}_2$ are solvable, then so is $\mathfrak{n}_1 + \mathfrak{n}_2$. In fact, it can be shown by induction that

$$\mathfrak{D}^{(i)}(\mathfrak{n}_1 + \mathfrak{n}_2) \subseteq \mathfrak{D}^{(i)}(\mathfrak{n}_1) + \mathfrak{n}_2.$$

It follows that if $\mathfrak{D}^{(n_1)}\mathfrak{n}_1 = \mathfrak{D}^{(n_2)}\mathfrak{n}_2 = \{0\}$, then

$$\mathfrak{D}^{(n_1+n_2)}(\mathfrak{n}_1 + \mathfrak{n}_2) = \{0\}.$$

□

Corollary 3.4. $\mathfrak{g}$ *is semi-simple if and only if the radical* $\mathfrak{r}$ *of* $\mathfrak{g}$ *is zero.*

Proposition 3.5. $\mathfrak{g}/\mathfrak{r}$ *is always semi-simple.*

Proof. Let $\overline{\mathfrak{n}}$ be a solvable ideal of $\mathfrak{g}/\mathfrak{r}$. If $\mathfrak{n}$ is the set that mapped to $\overline{\mathfrak{n}}$ under the canonical mapping, then $\mathfrak{n}$ is solvable because $\mathfrak{D}\mathfrak{n} \subseteq \mathfrak{r}$ and $\mathfrak{r}$ is solvable. Thus, $\mathfrak{n} \subseteq \mathfrak{r}$ and hence $\overline{\mathfrak{n}} = \{\overline{0}\}$. It follows that $\mathfrak{g}/\mathfrak{r}$ is semi-simple. □

Definition 3.6. *The* **Cartan-Killing form** *(or simply, the* **Killing form***) on* $\mathfrak{g}$ *is the form*

$$(\ ,\ ) : \mathfrak{g} \times \mathfrak{g} \longrightarrow k$$

defined by

$$(X, Y) = tr(ad(X)ad(Y)), \quad X, Y \in \mathfrak{g}.$$

Lemma 3.7. *We have*

(i) $(\ ,\ )$ *is symmetric and bilinear.*

(ii) $(\ ,\ )$ *satisfies*

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

(iii) The Killing form of a nilpotent Lie algebra is identical to zero.

(iv) Automorphisms of $\mathfrak{g}$ *preserve* $(\ ,\ )$.

Proof. Exercise. □

Proposition 3.8. *If* $\mathfrak{h} \subseteq \mathfrak{g}$ *is an ideal and* $(\ ,\ )_{\mathfrak{h}}$ *is the Killing form on* $\mathfrak{h}$*, then*

$$(X, Y) = (X, Y)_{\mathfrak{h}} \ \ \forall X, Y \in \mathfrak{h}.$$

Proof. Exercise. □

Theorem 3.9 (Cartan's criterion on semi-semplicity). *$\mathfrak{g}$ is semi-simple if and only if the Killing form is non-degenerate, i.e., if $(X, Y) = 0$ for all $Y \in \mathfrak{g}$, then $X = 0$.*

Proof. We first suppose $(\ ,\ )$ is non-degenerate and $\mathfrak{a}$ is an abelian ideal of $\mathfrak{g}$ and $A \in \mathfrak{a}$. Then for any $X, Y \in \mathfrak{g}$, we have

$$(ad(A)ad(X))^2 Y = [A, [X, [A, [X, Y]]]] = 0.$$

Hence, $(ad(A)ad(X))^2 = 0$ and so $(A, X) = 0$ for all $X \in \mathfrak{g}$. Since the Killing form is non-degenerate, so $A = 0$ and hence $\mathfrak{a} = \{0\}$. Therefore, $\mathfrak{g}$ is semi-simple.

Conversely, suppose $(\ ,\ )$ is degenerate and let

$$\mathfrak{n} := \{X \in \mathfrak{g} : (X, Y) = 0,\ \forall Y \in \mathfrak{g}\}.$$

Then $\mathfrak{n} \neq \{0\}$. From (ii) of Lemma 3.7, $\mathfrak{n}$ is an ideal of $\mathfrak{g}$. Thus from Proposition 3.8, we have

$$(X, Y)_{\mathfrak{n}} = 0,\ \ \forall X, Y \in \mathfrak{n}$$

and hence by the Cartan's criterion on solvablility, $\mathfrak{n}$ is solvable. Therefore, $\mathfrak{g}$ is not semi-simple. □

Corollary 3.10. *Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$ an ideal of $\mathfrak{g}$. Let*

$$\mathfrak{h}^{\perp} := \{X \in \mathfrak{g} : (X, Y) = 0,\ \forall Y \in \mathfrak{h}\}.$$

Then $\mathfrak{h}^{\perp}$ is an ideal of $\mathfrak{g}$ and $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$.

Proof. From (ii) of Lemma 3.7, $\mathfrak{h}^{\perp}$ is an ideal of $\mathfrak{g}$. From the non-degeneracy of the Killing form and the finite-dimensionality of $\mathfrak{g}$ it follows that

$$\dim \mathfrak{g} = \dim \mathfrak{h} + \dim \mathfrak{h}^{\perp}.$$

So it remains to show that $\mathfrak{h} \cap \mathfrak{h}^{\perp} = \{0\}$. From Proposition 3.8, we have

$$(X, Y)_{\mathfrak{h} \cap \mathfrak{h}^{\perp}} = 0,\ \ \forall X, Y \in \mathfrak{h} \cap \mathfrak{h}^{\perp}$$

and hence by the Cartan's criterion on solvablility, $\mathfrak{h} \cap \mathfrak{h}^{\perp}$ is solvable. Since $\mathfrak{g}$ is semi-simple, so $\mathfrak{h} \cap \mathfrak{h}^{\perp} = \{0\}$. □

Remark. By Corollary 3.10, every ideal in $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ and so both $\mathfrak{h}$ and $\mathfrak{h}^{\perp}$ are semi-simple. In particular, every ideal of $\mathfrak{g}$ is a direct summand of $\mathfrak{g}$ and so is also semi-simple.

Corollary 3.11 (Cartan). *Every semi-simple Lie algebra is the direct sum of all its simple ideals, and the simple ideals are mutually orthogonal with respect to the Killing form.*

Proof. By induction on the dimension of the Lie algebra. □

Corollary 3.12. *If $\mathfrak{g}$ is semi-simple, then $\mathfrak{g} = \mathfrak{D}\mathfrak{g}$.*

Proof. It is because $(\ ,\ )$ is non-degenerate and hence $\mathfrak{D}\mathfrak{g}^{\perp} \subseteq \mathfrak{C}(\mathfrak{g}) = \{0\}$. $\square$

Corollary 3.13. *If $\mathfrak{g} = \oplus \mathfrak{g}_i$ is a direct sum of simple ideals, then every ideal of $\mathfrak{g}$ is a direct sum of some of the $\mathfrak{g}_i$'s.*

Proof. If $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, then

$$\mathfrak{h} = \bigoplus_{\mathfrak{h} \cap \mathfrak{g}_i \neq \{0\}} \mathfrak{g}_i.$$

$\square$

Remark. $sl(n,k)(n \geq 2), o(n,k)(n \geq 3)$ and $sp(2n,k)(n \geq 1)$ are semi-simple.

Chapter 4

Cartan Subalgebras

4.1 Cartan Subalgebras

Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a subalgebra and let the **normalizer** of $\mathfrak{h}$ in $\mathfrak{g}$ to be

$$\mathfrak{n}(\mathfrak{h}) := \{X \in \mathfrak{g} : (adX)\mathfrak{h} \subseteq \mathfrak{h}\} = \{X \in \mathfrak{g} : [X, \mathfrak{h}] \subseteq \mathfrak{h}\}.$$

Clearly, since $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$, so $\mathfrak{h} \subseteq \mathfrak{n}(\mathfrak{h})$.

Definition 4.1. *A subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is called a* **Cartan Subalgebra (CSA)** *if*

(i) $\mathfrak{h}$ *is nilpotent and*

(ii) $\mathfrak{h}$ *is the normalizer of itself in* $\mathfrak{g}$, *i.e.,* $\mathfrak{h} = \mathfrak{n}(\mathfrak{h})$.

Lemma 4.2. *$A \in End(V)$ is nilpotent if and only if all the eigenvalues of A are zero.*

Proof. This is because for any $A \in End(V)$, there is a basis for V such that the corresponding matrix is an upper triangular with eigenvalues as diagonal elements. □

Let $\mathfrak{h}^*$ be the set of all linear maps from $\mathfrak{h}$ into k. For any $\alpha \in \mathfrak{h}^*$, define

$$\mathfrak{g}_{\mathfrak{h}}^{\alpha} := \{X \in \mathfrak{g} : \exists\, n > 0 \; s.t. \; (adH - \alpha(H)I)^n X = 0, \; \forall H \in \mathfrak{h}\}.$$

$\mathfrak{g}_{\mathfrak{h}}^{\alpha}$ are subspaces of $\mathfrak{g}$ but in general are **not** subalgebra. However, from Lemma 4.5 below, $\mathfrak{g}_{\mathfrak{h}}^{0}$ is a subalgebra of $\mathfrak{g}$. If $\mathfrak{h} \subseteq \mathfrak{g}$ is nilpotent, then there is $m > 0$ such that $(adH)^m = 0$ on $\mathfrak{h}$ for all $H \in \mathfrak{h}$. Thus if $X \in \mathfrak{n}(\mathfrak{h})$, then $(adH)X \in \mathfrak{h}$ for all $H \in \mathfrak{h}$. It follows that $(adH)^{m+1}X = 0$ for all $H \in \mathfrak{h}$ and hence $\mathfrak{n}(\mathfrak{h}) \subseteq \mathfrak{g}_{\mathfrak{h}}^{0}$.

Lemma 4.3. *Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a nilpotent subalgebra. Then $\mathfrak{h}$ is CSA if and only if $\mathfrak{h} = \mathfrak{g}_{\mathfrak{h}}^{0}$.*

Proof. Since $\mathfrak{h}$ is nilpotent, so we have $\mathfrak{h} \subseteq \mathfrak{n}(\mathfrak{h}) \subseteq \mathfrak{g}^0_{\mathfrak{h}}$. If $\mathfrak{h} = \mathfrak{g}^0_{\mathfrak{h}}$, then $\mathfrak{h} = \mathfrak{n}(\mathfrak{h})$ and hence $\mathfrak{h}$ is a CSA. Conversely, if $\mathfrak{h}$ is a CSA, then according to Engel's theorem (since $\mathfrak{g}^0_{\mathfrak{h}}$ is invariant under $\mathfrak{h}$), there is a basis of $\mathfrak{g}^0_{\mathfrak{h}}$ with respect to which all the matrices of the transformations induced by $adH(H \in \mathfrak{h})$ on $\mathfrak{g}^0_{\mathfrak{h}}$ have the form

$$\begin{bmatrix} 0 & * & * \\ & \ddots & * \\ 0 & & 0 \end{bmatrix}.$$

Hence if $X \in \mathfrak{g}^0_{\mathfrak{h}}$, then we have $(adH_1)(adH_2)\cdots(adH_r)X = 0$ for all $H_1, \cdots, H_r \in \mathfrak{h}$ where $r = \dim \mathfrak{g}^0_{\mathfrak{h}}$. Thus,

$$[H_1, (adH_2)\cdots(adH_r)X] = 0, \ \ \forall H_1, \cdots H_r \in \mathfrak{h}.$$

and hence $(adH_2)\cdots(adH_r)X \in \mathfrak{n}(\mathfrak{h}) = \mathfrak{h}$ for all $H_2, \cdots, H_r$ in $\mathfrak{h}$. By repeating the same argument, we have $(adH_r)X \in \mathfrak{h}$ for all $H_r \in \mathfrak{h}$ and hence $X \in \mathfrak{h}$. It follows that $\mathfrak{h} = \mathfrak{g}^0_{\mathfrak{h}}$. □

Lemma 4.4. *If $\mathfrak{h} \subseteq \mathfrak{g}$ is a CSA, then*

(i) $\mathfrak{h}$ is a maximal nilpotent.

(ii) If $\mathfrak{h} \subset \mathfrak{k} \subset \mathfrak{g}$ for some subalgebra $\mathfrak{k}$ of $\mathfrak{g}$, then $\mathfrak{h}$ is also a CSA of $\mathfrak{k}$.

Proof. Suppose $\mathfrak{h}'$ is a nilpotent and $\mathfrak{h} \subseteq \mathfrak{h}' \subseteq \mathfrak{g}$. Since $\mathfrak{h}'$ is nilpotent, so there is $m > 0$ such that $(adH)^m = 0$ on $\mathfrak{h}'$ for any $H \in \mathfrak{h}'$. In particular, for any $H \in \mathfrak{h}, H' \in \mathfrak{h}'$, $(adH)^m H' = 0$. Thus, $H' \in \mathfrak{g}^0_{\mathfrak{h}} = \mathfrak{h}$ for all $H' \in \mathfrak{h}$ and hence $\mathfrak{h}' = \mathfrak{h}$. This proves (i). Part (ii) is trivial because

$$\mathfrak{h} \subseteq \{X \in \mathfrak{k} : (adX)\mathfrak{h} \subseteq \mathfrak{h}\} \subseteq \mathfrak{n}(\mathfrak{h}).$$

□

4.2 Existence of Cartan Subalgebra

Let $\mathfrak{g}$ be a Lie algebra and $X \in \mathfrak{g}$. If $\lambda \in \mathbb{C}$, we let

$$\mathfrak{g}^\lambda_X := \{Y \in \mathfrak{g} : (ad(X) - \lambda I)^n Y = 0 \ \text{ for some } n > 0\}$$

to be the nilspace of $ad(X) - \lambda I$. In particular, $\mathfrak{g}^0_X$ is the nilspace of $ad(X)$ and its dimension is the multiplicity of 0 as an eigenvalues of $ad(X)$.

Lemma 4.5. *Let $X \in \mathfrak{g}$. Then*

(a) $\mathfrak{g} = \oplus_\lambda \mathfrak{g}^\lambda_X$.

(b) $[\mathfrak{g}^\lambda_X, \mathfrak{g}^\mu_X] \subseteq \mathfrak{g}^{\lambda+\mu}_X$ if $\lambda, \mu \in \mathbb{C}$.

(c) $\mathfrak{g}_X^0$ *is a Lie subalgebra of* $\mathfrak{g}$.

(d) $(\mathfrak{g}_X^\lambda, \mathfrak{g}_X^\mu) = 0$ *if* $\lambda + \mu \neq 0$.

Proof. Part (a) is the usual property of vector space. To prove (b), we first show the following formula by induction:

$$(adX - \lambda - \mu)^n [Y, Z] = \sum_{i=0}^{n} \binom{n}{i} [(adX - \lambda)^i Y, (adX - \mu)^{n-i} Z]. \tag{4.1}$$

Then if $Y \in \mathfrak{g}_X^\lambda, Z \in \mathfrak{g}_X^\mu$ and take sufficiently large n, all the terms on the right hand side of (4.1) vanish. Hence $[Y, Z] \in \mathfrak{g}_X^{\lambda+\mu}$.

Part (c) follows directly from (b) by taking $\lambda = \mu = 0$.

For (d), we take a basis from each $\mathfrak{g}_X^\lambda$ and the union of these bases is a basis of $\mathfrak{g}$. If $Y \in \mathfrak{g}_X^\lambda$ and $Z \in \mathfrak{g}_X^\mu$, then from (b), we have

$$(adY\, adZ)\mathfrak{g}_X^\alpha \subseteq \mathfrak{g}_X^{\alpha+\lambda+\mu}.$$

Thus with respect to this basis, all diagonal elements of $adY\, adZ$ are zero and hence $(X, Y) = 0$. □

If $X \in \mathfrak{g}$, we will let

$$f_X(T) = \det(T - ad(X))$$

to be the characteristic polynomial of $ad(X)$ and if $n = \dim \mathfrak{g}$, then we write

$$f_X(T) = T^n + a_1(X)T^{n-1} + \cdots + a_{n-n_X}(X)T^{n_X}, \quad a_{n-n_X}(X) \neq 0.$$

If X has a coordinates $x_1, \cdots, x_n$ with respect to a fixed basis of $\mathfrak{g}$, we can view $a_i(X)$ as a function of the n variables $x_1, x_2, \cdots, x_n$ in k; it is a homogeneous polynomial of degree $n - i$ in $x_1, \cdots, x_n$.

The polynomial $f_X(T)$ is called the **Killing polynomial** of X and n_X of the eigenvalues of adX are zero. Since $[X, X] = 0$, so $n_X \geq 1$. Let

$$r := \min_{X \in \mathfrak{g}} n_X$$

to be the **rank** of $\mathfrak{g}$. So the rank of a nilpotent Lie algebra is equal to its dimension. If $X \in \mathfrak{g}$ and $n_X = r$, then X is said to be a **regular element**; otherwise a **singular element**. Note that for any $X \in \mathfrak{g}$, $n_X \geq r$ and hence adX has at least r zero eigenvalues.

The main objective of this section is to prove that every Lie algebra has a Cartan subalgebra. In fact, if X_0 is a regular element, then $\mathfrak{g}_{X_0}^0$ is a CSA.

Lemma 4.6. *Let* X_0 *be a fixed element of* $\mathfrak{g}$. *If for any* $Y \in \mathfrak{g}$, *there exist infinitely many* μ *such that* $ad(X_0 + \mu Y)$ *are all nilpotent, then* $\mathfrak{g}$ *is a nilpotent Lie algebra.*

Proof. Since a linear transformation is nilpotent if and only if all its eigenvalues are zero, so if we write

$$f_X(T) = T^n + a_1(X)T^{n-1} + \cdots + a_n(X)$$

then adX is nilpotent if and only if

$$a_1(X) = \cdots = a_n(X) = 0.$$

Hence if there exist infinitely many μ such that $ad(X_0 + \mu Y)$ are all nipotent, then

$$a_1(X_0 + \mu Y) = \cdots = a_n(X_0 + \mu Y) = 0$$

for infinitely many μ. Thus

$$a_1(X_0 + \mu Y) \equiv \cdots \equiv a_n(X_0 + \mu Y) \equiv 0, \ \ \forall \mu.$$

In particular, we let $Y = X - X_0$ and $\mu = 1$. Then,

$$a_1(X) \equiv \cdots \equiv a_n(X) \equiv 0, \ \ \forall X \in \mathfrak{g}.$$

Therefore adX is nilpotent and by Engel's theorem, $\mathfrak{g}$ is nilpotent. □

Theorem 4.7. *If X_0 is a regular element, then $\mathfrak{g}^0_{X_0}$ is a CSA.*

Proof. We first show that $\mathfrak{g}^0_{X_0}$ is nilpotent. From (a) of Lemma 4.5, we have

$$\mathfrak{g} = \mathfrak{g}^0_{X_0} \oplus \bigoplus_{\lambda \in \Delta} \mathfrak{g}^\lambda_{X_0}$$

where Δ is the set of non-zero eigenvalues of adX_0. Let $\tilde{\mathfrak{g}} = \oplus_{\lambda\in\Delta}\mathfrak{g}^\lambda_{X_0}$. For any $X \in \mathfrak{g}^0_{X_0}, [X, \tilde{\mathfrak{g}}] \subseteq \tilde{\mathfrak{g}}$ by (b) of Lemma 4.5. So we consider the linear transformation of adX restricted on $\tilde{\mathfrak{g}}$ and denote the determinant by $D(X)$. Since all the eigenvalues of adX_0 in $\tilde{\mathfrak{g}}$ are non-zero, so $D(X_0) \neq 0$. For any $Y \in \mathfrak{g}^0_{X_0}$, $D(X_0 + \mu Y)$ is a polynomial in μ and since $D(X_0) \neq 0$, so there exist infinitely many μ such that $D(X_0 + \mu Y) \neq 0$. Now since the rank of $\mathfrak{g}$ is r and $n_{X_0} = r$, so the linear transformation $ad(X_0+\mu Y)$ has at least r zero eigenvalues. However, $ad(X_0+\mu Y)$ has no zero eigenvalues in $\tilde{\mathfrak{g}}$ because $D(X_0+\mu Y) \neq 0$. So all the zero eigenvalues of $ad(X_0 + \mu Y)$ are in $\mathfrak{g}^0_{X_0}$. Since $\dim \mathfrak{g}^0_{X_0} = r$, so all the eigenvalues of $ad_{\mathfrak{g}^0_{X_0}}(X_0 + \mu Y)$ are zero and hence $ad_{\mathfrak{g}^0_{X_0}}(X_0 + \mu Y)$ is nilpotent. From Lemma 4.6, $\mathfrak{g}^0_{X_0}$ is nilpotent because $X_0 \in \mathfrak{g}^0_{X_0}$ and $X_0 + \mu Y \in \mathfrak{g}^0_{X_0}$.

Now let $\mathfrak{h} = \mathfrak{g}^0_{X_0}$. Then since $\mathfrak{h}$ is subalgebra, so $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$ and hence $\mathfrak{h} \subseteq \mathfrak{n}(\mathfrak{h})$. Conversely, if $X \in \mathfrak{n}(\mathfrak{h})$, then $[X, \mathfrak{h}] \subseteq \mathfrak{h}$ or equivalently,

$$(adH)X \in \mathfrak{h}, \ \ \forall H \in \mathfrak{h}.$$

In particular, $X_0 \in \mathfrak{h}$ and adX_0 is nilpotent. So there is $m > 0$ such that $(adX_0)^m = 0$. It follows that $(adX_0)^{m+1}X = 0$ and hence $X \in \mathfrak{g}^0_{X_0} = \mathfrak{h}$. Therefore, $\mathfrak{n}(\mathfrak{h}) = \mathfrak{h}$ and $\mathfrak{g}^0_{X_0}$ is CSA. □

Corollary 4.8. *Every Lie algebra $\mathfrak{g}$ contains a CSA.*

Corollary 4.9. *Let $\mathfrak{h}$ be a nilpotent subalgebra of $\mathfrak{g}$ that contains a regular element of $\mathfrak{g}$. Then*

$$\mathfrak{g}^0_{\mathfrak{h}} := \{X \in \mathfrak{g} : \exists\, n > 0 \; s.t. \; (adH)^n X = 0, \; \forall H \in \mathfrak{h}\}$$

is a CSA of $\mathfrak{g}$.

Proof. Let X_0 be a regular element of $\mathfrak{g}$ contained in $\mathfrak{h}$. Then $\mathfrak{g}^0_{X_0}$ is a CSA of $\mathfrak{g}$. Let $\mathfrak{h}_1 = \mathfrak{g}^0_{X_0}$. Since $\mathfrak{h}$ is nilpotent, and $X_0 \in \mathfrak{h}$, there is $m > 0$ such that $(adX_0)^m = 0$ on $\mathfrak{h}$. It follows that $\mathfrak{h} \subseteq \mathfrak{g}^0_{X_0} = \mathfrak{h}_1$. On the other hand,

$$\mathfrak{h}_1 = \mathfrak{g}^0_{X_0} \supseteq \mathfrak{g}^0_{\mathfrak{h}} \supseteq \mathfrak{g}^0_{\mathfrak{h}_1} \supseteq \mathfrak{h}_1$$

because $\mathfrak{h}_1$ is nilpotent. Hence $\mathfrak{h}_1 = \mathfrak{g}^0_{\mathfrak{h}} = \mathfrak{g}^0_{X_0}$ is CSA. □

Corollary 4.10. *Let $\mathfrak{h}$ be maximal nilpotent subalgebra. If $\mathfrak{h}$ contains a regular element X_0 of $\mathfrak{g}$, then $\mathfrak{h}$ is CSA and $\mathfrak{h} = \mathfrak{g}^0_{X_0}$.*

Proof. According to Theorem 4.7, $\mathfrak{g}^0_{X_0}$ is CSA. Since $\mathfrak{h}$ is nilpotent and $X_0 \in \mathfrak{h}$, so $\mathfrak{h} \subseteq \mathfrak{g}^0_{X_0}$. From the maximality of $\mathfrak{h}$, it follows that $\mathfrak{h} = \mathfrak{g}^0_{X_0}$. □

4.3 Conjugacy of Cartan subalgebras

Let $\mathfrak{g}$ be a Lie algebra, and denote its automorphism group by $\mathrm{Aut}(\mathfrak{g}) \subset \boldsymbol{g}\mathrm{l}(\mathfrak{g})$. Denote the adjoint group of $\mathfrak{g}$ by $G \subset \mathrm{Aut}(\mathfrak{g})$ generated by automorphisms of the form $\exp(\mathrm{ad}(y))$, $y \in \mathfrak{g}$. The adjoint group is a Lie subgroup of $GL(\mathfrak{g})$ with Lie algebra $\mathrm{ad}(\mathfrak{g}) \subset \boldsymbol{g}\mathrm{l}(\mathfrak{g})$.
(Using the Jacobi identity, one sees that each ad(y) is a derivation. Then it follows easily that $\exp(\mathrm{ad(y)}) = \sum_{\mathrm{k}\geq 0} \frac{1}{\mathrm{k!}}(\mathrm{ad(y)})^{\mathrm{k}}$ satisfies $\exp(\mathrm{ad(y)})[\mathrm{z},\mathrm{w}] = [\exp(\mathrm{ad(y)})\mathrm{z}, \exp(\mathrm{ad(y)})\mathrm{w}]$.)

Theorem 4.11. *G acts transitively on the set of CSA's.*

Corollary 4.12. *(i) Every CSA is of the form $\mathfrak{g}^0_X$ for some regular element X.*
(ii) All CSA have the same dimension equal to the rank of $\mathfrak{g}$.

Proof of Corollary 4.12 assuming Theorem 4.11: For any $g \in G$ and $X \in \mathfrak{g}$, since g is a Lie homomorphism, we have

$$\mathrm{ad(gX)} = [\mathrm{gX}, \cdot] = \mathrm{g}[\mathrm{X}, \mathrm{g}^{-1}\cdot] = \mathrm{g} \circ \mathrm{ad(X)} \circ \mathrm{g}^{-1}.$$

Thus,

$$\begin{aligned} \det(T - \mathrm{ad(gX)}) &= \det(T - g \circ \mathrm{ad(X)} \circ \mathrm{g}^{-1}) \\ &= \det(g) \cdot \det(T - \mathrm{ad(X)}) \cdot \det(\mathrm{g}^{-1}) \\ &= \det(T - \mathrm{ad(X)}). \end{aligned}$$

Thus gX is a regular element if and only if X is a regular element.
Now take a regular element $X \in \mathfrak{g}$ (it exists!), and recall that $\mathfrak{g}_X^0$ is a CSA. By Theorem 4.11, for any CSA $\mathfrak{h}$, there exists $g \in G$ such that $g\mathfrak{g}_X^0 = \mathfrak{h}$.
Obviously, $X \in \mathfrak{g}_X^0$, and thus the regular element $gX \in \mathfrak{h}$.
Then by Corollary 4.9, $\mathfrak{h} = \mathfrak{g}_{gX}^0$, and (i) follows.
Finally using (i) and the linear algebra fact that $\dim \mathfrak{g}_X^0 = \mathrm{rank}(\mathfrak{g})$ for any regular element X, one easily sees that (ii) holds.

We will split the proof of Theorem 4.11 into several lemmas.

Lemma 4.13. *The set $\mathfrak{g}_r$ of regular elements of $\mathfrak{g}$ is a connected dense open subset of $\mathfrak{g}$.*

Proof. Let $S := \mathfrak{g} \setminus \mathfrak{g}_r$, and let $r := \mathrm{rank}(\mathfrak{g})$. The characteristic polynomial of each X is of the form

$$\det(T - \mathrm{ad(X)}) = \mathrm{a_n(X)T^n} + \cdots + \mathrm{a_r(X)T^r},$$

where each $a_k(X)$ is a polynomial in the coordinates of X. Then S is equal to the zero set of the non-zero polynomial $a_r(X)$. Then the lemma is a consequence of the fact that S is Zariski closed in $\mathfrak{g}$.
First, one sees that $\mathfrak{g}_r$ is open and dense. To see that $\mathfrak{g}_r$ is connected, we take two points $X, Y \in \mathfrak{g}_r$, and consider a complex line D through X, Y. Then $D \cap S$ consists of finitely many points, and hence X and Y are connected in $D \cap \mathfrak{g}_r$ (and thus in $\mathfrak{g}_r$). □

Next we fix a CSA $\mathfrak{h}$, and for $X \in \mathfrak{h}$, we let $\overline{\mathrm{ad}}(X) : \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$ be the endomorphism induced by ad(X). Define

$$V := \{X \in \mathfrak{h} : \overline{\mathrm{ad}}(X) \text{ is invertible}\}.$$

Lemma 4.14. $V \neq \emptyset$.

Proof. Since $\mathfrak{h}$ is nilpotent, $\mathfrak{h}$ is also solvable.
By applying Lie's theorem to the $\mathfrak{h}$-module $\mathfrak{g}/\mathfrak{h}$, one can find a flag F given by

$$0 = D_0 \subset D_1 \subset \cdots \subset D_m = \mathfrak{g}/\mathfrak{h}$$

such that $\overline{\mathrm{ad}} \in S(F)$. Now $\mathfrak{h}$ acts on each one-dimensional space D_i/D_{i-1} via a linear form α_i, and we may write

$$\overline{\mathrm{ad}}(X)Z \equiv \alpha_i(X)Z \quad \text{for } \mathrm{X} \in \mathfrak{h},\ \mathrm{Z} \in \mathrm{D_i/D_{i-1}}.$$

The eigenvalues of $\overline{\mathrm{ad}}(X)$ are $\alpha_1(X), \cdots, \alpha_m(X)$. Hence it suffices to show that all α_i are not identically zero.
Suppose $\alpha_1, \cdots, \alpha_{k-1} \neq 0$, and $\alpha_k \equiv 0$.
Pick X_o such that $\alpha_1(X_o) \neq 0, \cdots, \alpha_{k-1}(X_o) \neq 0$, and $\alpha_k(X_o) = 0$. Then the endomorphism of D_{k-1} (resp. D_k) induced by $\overline{\mathrm{ad}}(X_o)$ is invertible (resp. has 0 as an eigenvalue with multiplicity 1).

The nilspace D of $\overline{\mathrm{ad}}(X_o)$ in D_k is therefore one dimensional and complementary to D_{k-1}.
Next we shall show that D is annihilated by each $\overline{\mathrm{ad}}(X)$, $X \in \mathfrak{h}$. Take a representative Z' of Z in $\mathfrak{g}$. By the Jacobi identity, we have

$$[X_o, [X, Z']] = [[X_o, X], Z'] + [X, [X_o, Z']].$$

Since $a_k(X_o) = 0$, it follows that the class of $[X_o, Z']$ in D_k/D_{k-1} is 0. Upon descending the above identity to D_k/D_{k-1}, we have $(\overline{\mathrm{ad}}(X_o))(\overline{\mathrm{ad}}(X)Z) = \overline{\mathrm{ad}}(\mathrm{ad(X_o)X)Z}$.
Similarly, we can show that $(\overline{\mathrm{ad}}(X_o))^n(\overline{\mathrm{ad}}(X)Z) = \overline{\mathrm{ad}}(\mathrm{(ad(X_o))^nX)Z}$.
Since $\mathfrak{h}$ is nilpotent, $\mathrm{((ad(X_o))^nX} = 0$ for sufficiently large n. It follows that $\overline{\mathrm{ad}}(X)Z \in D$. On the other hand, since $a_k(X) = 0$, we also have $\overline{\mathrm{ad}}(X)Z \in D_{k-1}$. Hence we have $\overline{\mathrm{ad}}(X)Z \in D_{k-1} \cap D = \{0\}$. Hence Z is annihilated by each $\overline{\mathrm{ad}}(X)$, $X \in \mathfrak{h}$.
Now pick a nonzero representative Z' of Z in $\mathfrak{g}$. Then $Z' \notin \mathfrak{h}$. The condition that $\overline{\mathrm{ad}}(X)(Z) = 0$, $X \in \mathfrak{h}$, is equivalent to saying that $[X, Z,] \in \mathfrak{h}$ for all $X \in \mathfrak{h}$. In other words, we have $Z \in \mathfrak{n}(\mathfrak{h}) \setminus \mathfrak{h}$, which contradicts the assumption that $\mathfrak{h}$ is a CSA. ☐

Lemma 4.15. *Let $W = G \cdot V$. Then W is open in $\mathfrak{g}$.*

Proof. Let $X \in V$. We shall use the implicit function theorem to show that W contains an open neighborhood of X in $\mathfrak{g}$.
Consider the map $\theta : G \times V \to \mathfrak{g}$ given by $\theta(g, X) = g \cdot X$ for $g \in G$ and $X \in V$. From the definition of $\mathfrak{g}$, one sees that $T_1G = \mathrm{ad}(\mathfrak{g}) \in \mathfrak{gl}(\mathfrak{g})$. We look at the jacobian $d\theta_{(1,X)} : \mathrm{ad}(\mathfrak{g}) \times \mathfrak{h} \to \mathfrak{g}$. For $H \in \mathfrak{h}$, we have $d\theta_{(1,X)}(0, H) = \frac{d}{dt}\Big|_{t=0}(1 \cdot (X + tH)) = H$.
For $Y \in T_1G$ with $Y = \mathrm{ad(Y')} \in \mathrm{ad}(\mathfrak{g})$, $Y' \in \mathfrak{g}$, we have

$$d\theta_{(1,X)}(Y, 0) = \frac{d}{dt}\Big|_{t=0}(\exp(t\mathrm{ad(Y')}) \cdot \mathrm{X}) = \frac{\mathrm{d}}{\mathrm{dt}}\Big|_{\mathrm{t=0}}(1 + [\mathrm{tY', X}] + \cdots) = [\mathrm{Y', X}].$$

Hence we see that $d\theta_{(1,X)}(\mathrm{ad}(\mathfrak{g}) \times \mathfrak{h}) = \mathrm{Im(ad(X))} + \mathfrak{h}$. Since $\overline{\mathrm{ad}}(X) : \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$ is invertible, it follows that $Im(\mathrm{ad(X)}) + \mathfrak{h} = \mathfrak{g}$. By the implicit function theorem, the map θ is open at $(1, X)$ (and hence open at any (g, X)). Thus we have finished the proof of the lemma. ☐

Lemma 4.16. *For any CSA $\mathfrak{h}$, there exists a regular element $X \in \mathfrak{g}$ such that $\mathfrak{h} = \mathfrak{g}^0_X$.*

Proof. For a CSA $\mathfrak{h}$, we let V be as defined before.
Lemma 4.14 and Lemma 4.15 shows that W is open and non-empty.
Together with Lemma 4.13, it follows that $W \cap \mathfrak{g}_r \neq \emptyset$. Thus W contains a regular element.
We have seen that gX is regular iff X is regular. Thus V contains a regular element, say X. Since $\mathfrak{h}$ is maximally nilpotent and $X \in \mathfrak{h}$, it follows from Theorem 4.9 that $\mathfrak{h} = \mathfrak{g}^0_X$.

By Lemma 4.16, we know that all CSA are of the form $\mathfrak{g}_X^0$. We define the following equivalence relation on $\mathfrak{g}_r$ given by

$$X \sim Y \Leftrightarrow g\mathfrak{g}_X^0 = \mathfrak{g}_Y^0 \text{ for some } \mathrm{g} \in \mathrm{G}.$$

$\square$

Lemma 4.17. *The equivalence classes of $\sim$ are open in $\mathfrak{g}_r$.*

Proof. We have to show that for $X \in \mathfrak{g}_r$, every Y sufficiently close to X is equivalent to X.
We apply our previous results to $\mathfrak{h} = \mathfrak{g}_X^0$. Then the corresponding set V contains X (this follows easily from the definition of $\mathfrak{g}_X^0$ and the Jordan form of ad(X)). Then by Lemma 4.15, $W = G \cdot V$ is open. Since $\mathfrak{h} = \mathfrak{g}_X^0$, by looking at the Jordan block decomposition of ad(X), one sees that $X \in V \subset W$.
Thus each Y sufficiently close to X is of the form gX' for some $X' \in V$ and $g \in G$.
Since $(\mathrm{ad(gX)})^{\mathrm{m}} = \mathrm{g} \circ (\mathrm{ad(X)})^{\mathrm{m}} \circ \mathrm{g}^{-1}$, it follows that $\mathfrak{g}_Y^0 = g\mathfrak{g}_{X'}^0 = g\mathfrak{h} = g\mathfrak{g}_X^0$. Hence X and Y are equivalent. $\square$

Finally we have

Completion of proof of Theorem 4.11: By Lemma 4.17, the equivalence classes of $\mathfrak{g}_r$ are all open (and thus also closed). But by Lemma 4.13, $\mathfrak{g}_r$ is connected. Hence there is only one equivalence class, and thus G acts transitively on the set of CSA.

4.4 Cartan Subalgebra of Semi-simple Lie Algebras

In this section, let $\mathfrak{g}$ be a semi-simple Lie algebra.

Theorem 4.18. *Let $\mathfrak{h}$ be a CSA of a semi-simple Lie algebra $\mathfrak{g}$. Then we have*

(i) The restriction of the Killing form of $\mathfrak{g}$ to $\mathfrak{h}$ is non-degenerate.

(ii) $\mathfrak{h}$ is abelian.

(iii) The centralizer

$$\mathfrak{c}(\mathfrak{h}) := \{X \in \mathfrak{g} : [X, Y] = 0,\ \forall Y \in \mathfrak{h}\}$$

of $\mathfrak{h}$ in $\mathfrak{g}$ is $\mathfrak{h}$ itself.

(iv) adH is semi-simple for all $H \in \mathfrak{h}$.

Proof. Since $\mathfrak{g}$ is semi-simple, so by Cartan's criterion, (,) is non-degenerate on $\mathfrak{g}$.

In the last section, we have proved that there is a regular element X such that $\mathfrak{h} = \mathfrak{g}_X^0$ and

$$\mathfrak{g} = \mathfrak{g}_X^0 \oplus \bigoplus_{\lambda \neq 0} \mathfrak{g}_X^\lambda .$$

From (d) of Lemma 4.5, we have $(\mathfrak{g}_X^\lambda, \mathfrak{g}_X^\mu) = 0$ if $\lambda + \mu \neq 0$. We therefore have a decomposition of $\mathfrak{g}$ into mutually orthogonal subspaces

$$\mathfrak{g} = \mathfrak{g}_X^0 \oplus \bigoplus_{\lambda \neq 0} \left(\mathfrak{g}_X^\lambda \oplus \mathfrak{g}_X^{-\lambda} \right) .$$

Thus if there is Y in $\mathfrak{h}$ such that $(Y, Z) = 0$ for all Z in $\mathfrak{h}$, then $(Y, Z) = 0$ for all Z in $\mathfrak{g}$. This contradicts (,) is non-degenerate on $\mathfrak{g}$. This proves (i).

For (ii), since $\mathfrak{h}$ is nilpotent, so we have $(X, Y) = 0$ for all $X \in \mathfrak{h}$ and $Y \in [\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$. From (i), $[\mathfrak{h}, \mathfrak{h}] = \{0\}$.

For (iii), it is clear that $\mathfrak{h} \subseteq \mathfrak{c}(\mathfrak{h})$ by (ii). However, $\mathfrak{c}(\mathfrak{h})$ is always contained in $\mathfrak{n}(\mathfrak{h})$. So $\mathfrak{h} = \mathfrak{c}(\mathfrak{h})$.

To prove (iv), we need the following result. If X is an element in a semi-simple Lie algebra $\mathfrak{g}$, then X is said to be **nilpotent** if adX is nilpotent and X is said to be **semi-simple** if adX is semi-simple. Then we have

Theorem 4.19. *If $\mathfrak{g}$ is semi-simple, every element X of $\mathfrak{g}$ can be written uniquely in the form $X = s + n$ with n nilpotent, s semi-simple and $[s, n] = 0$. Moreover, if $[Y, X] = 0$, then $[Y, s] = [Y, n] = 0$.*

Now let $X \in \mathfrak{h}$ and let $X = s + n$ as in the above theorem. If $Y \in \mathfrak{h}$, then from (ii), $[X, Y] = 0$ and hence $[Y, s] = [Y, n] = 0$. Thus s and n belong to $\mathfrak{c}(\mathfrak{h}) = \mathfrak{h}$. Now since $ad(n)$ is nilpotent, so $ad(Y)ad(n)$ is also nilpotent and hence $(Y, n) = 0$. It follows that $n = 0$ because $(\, , \,)_{\mathfrak{h}}$ is non-degenerate. Thus $X = s$ and so X is semi-simple. This completes the proof of the theorem. ☐

Corollary 4.20. *$\mathfrak{h}$ is a maximal abelian subalgebra of $\mathfrak{g}$.*

Proof. If $\mathfrak{h} \subseteq \mathfrak{m}$ is an abelian subalgebra of $\mathfrak{g}$, then $\mathfrak{m} \subseteq \mathfrak{c}(\mathfrak{h})$. Hence from (iii) of Theorem 4.18, $\mathfrak{h} = \mathfrak{m}$. ☐

Corollary 4.21. *Every regular element of $\mathfrak{g}$ is semi-simple.*

Proof. This is because every regular element is contained in some CSA and it follows from (iv) of Theorem 4.18. ☐

Theorem 4.22. *A subalgebra $\mathfrak{h}$ of a semi-simple Lie algebra $\mathfrak{g}$ is a CSA if and only if $\mathfrak{h}$ is a maximal abelian subalgebra and every elements in $\mathfrak{h}$ are semi-simple.*

Proof. The "only if" part follows immediately from (iv) of Theorem 4.18 and Corollary 4.20. Conversely, suppose $\mathfrak{h}$ is a maximal abelian aubalgebra and every elements in $\mathfrak{h}$ are semi-simple. Since $\mathfrak{h}$ is nilpotent, we consider the decomposition

$$\mathfrak{g} = \mathfrak{g}_{\mathfrak{h}}^0 \oplus \bigoplus_{\lambda \neq 0} \mathfrak{g}_{\mathfrak{h}}^{\lambda}$$

of $\mathfrak{g}$ under $\mathfrak{h}$. For any $X \in \mathfrak{g}_{\mathfrak{h}}^0$, there is $m > 0$ such that $(adH)^m X = 0$ for all $H \in \mathfrak{h}$. However, since adH is semi-simple, so $(adH)X = [H, X] = 0$ for all $H \in \mathfrak{h}$. Thus $< \mathfrak{h}, X >$ is a abelian subalgebra containing $\mathfrak{h}$. By the maximality, $X \in \mathfrak{h}$ and hence $\mathfrak{h} = \mathfrak{g}_{\mathfrak{h}}^0$. So $\mathfrak{h}$ is a CSA of $\mathfrak{g}$. $\square$

4.5 Roots System and Root Space Decomposition

Let $k = \mathbb{C}$. Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ be a CSA. Since adH is semi-simple for all $H \in \mathfrak{h}$, so if $(adH - \alpha(H)I)^n X = 0$ for some $n > 0$ then $(adH)X = \alpha(H)X$. For any $\alpha \in \mathfrak{h}^*$, define

$$\mathfrak{g}_\alpha := \{X \in \mathfrak{g} : (adH)X = \alpha(H)X, \ \ \forall H \in \mathfrak{g}\}.$$

As before, we have

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

for any $\alpha, \beta \in \mathfrak{h}^*$. Also

$$\mathfrak{g}_0 = \{X \in \mathfrak{g} : [H, X] = 0, \ \ \forall H \in \mathfrak{h}\} = \mathfrak{C}(\mathfrak{h}) = \mathfrak{h}.$$

Definition 4.23. *$\alpha \in \mathfrak{h}^*$ is called a* **root** *of $\mathfrak{g}$ with respect to $\mathfrak{h}$, or simply, a root, if $\alpha \neq 0$ and $\mathfrak{g}_\alpha \neq \{0\}$. The collection of all roots is denoted by Δ and $\Sigma := \Delta \cup \{0\}$. If $\alpha \in \Delta$, then $\mathfrak{g}_\alpha$ is said to be the* **root space** *corresponding to α. Note that there are only finitely many roots. A root α is* **simple** *if $\dim \mathfrak{g}_\alpha = 1$.*

For any $\alpha \in \mathfrak{h}^*$, there exists a unique $H_\alpha \in \mathfrak{h}$ such that

$$(H, H_\alpha) = \alpha(H), \ \ \forall H \in \mathfrak{h}.$$

In fact, every $H \in \mathfrak{h}$ defines a linear function on $\mathfrak{h}$ by

$$\phi_H(X) := (H, X), \ \ \forall H \in \mathfrak{h}.$$

Since the restriction of (X, Y) to $\mathfrak{h}$ is non-degenerate, if $H_1 \neq H_2$ then $\phi_{H_1} \neq \phi_{H_2}$. So $\alpha \in \mathfrak{h}^* \longrightarrow H_\alpha \in \mathfrak{h}$ is a isomorphism.

As before, we have

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha.$$

Lemma 4.24. *If $X, Y \in \mathfrak{h}$, then*
$$(X, Y) = \sum_{\alpha \in \Delta} \nu_\alpha \alpha(X)\alpha(Y)$$
where $\nu_\alpha = \dim \mathfrak{g}_\alpha$.

Proof. If $X \in \mathfrak{g}$, then all the eigenvalues of adX on $\mathfrak{g}_\alpha$ are $\alpha(X)$ and hence
$$Tr(ad_{\mathfrak{g}_\alpha} X)^2 = \nu_\alpha \alpha(X)^2$$
and so
$$(X, X) = Tr(ad_{\mathfrak{g}} X)^2 = \sum_{\alpha \in \Delta} \nu_\alpha \alpha(X)^2.$$
Now
$$\begin{aligned}(X+Y, X+Y) &= \sum_{\alpha \in \Delta} \nu_\alpha \alpha(X+Y)^2 \\ &= \sum_{\alpha \in \Delta} \nu_\alpha (\alpha(X) + \alpha(Y))^2 \\ &= (X, X) + 2 \sum_{\alpha \in \Delta} \nu_\alpha \alpha(X)\alpha(Y) + (Y, Y).\end{aligned}$$
On the other hand, $(X+Y, X+Y) = (X, X) + 2(X, Y) + (Y, Y)$. This proves the lemma. $\square$

Lemma 4.25. *Let $\phi, \alpha, -\alpha \in \Sigma$. If p and q are non-negative integers such that $[\mathfrak{g}_{-\alpha}, \mathfrak{g}_{\phi - p\alpha}] = [\mathfrak{g}_\alpha, \mathfrak{g}_{\phi + q\alpha}] = 0$, then for any $Z \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$,*
$$\phi(Z) = \frac{-\sum_{i=-p}^{q} i\nu_{\phi + i\alpha}}{\sum_{i=-p}^{q} \nu_{\phi + i\alpha}} \alpha(Z). \tag{4.2}$$
In particular, $\phi(Z) = 0$ for any $Z \in [\mathfrak{h}, \mathfrak{h}]$.

Proof. Since both sides of (4.2) are linear functions on Z, so it suffices to consider $Z = [X_\alpha, X_{-\alpha}], X_\alpha \in \mathfrak{g}_\alpha, X_{-\alpha} \in \mathfrak{g}_{-\alpha}$. Since $[\mathfrak{g}_{-\alpha}, \mathfrak{g}_{\phi - p\alpha}] = [\mathfrak{g}_\alpha, \mathfrak{g}_{\phi + q\alpha}] = 0$, the subspace
$$\tilde{\mathfrak{g}} = \bigoplus_{i=-p}^{q} \mathfrak{g}_{\phi + i\alpha}$$
is invariant under adX_α and $adX_{-\alpha}$. So
$$Tr(ad_{\tilde{\mathfrak{g}}} Z) = Tr(ad_{\tilde{\mathfrak{g}}}[X_\alpha, X_{-\alpha}]) = Tr(adX_\alpha adX_{-\alpha} - adX_{-\alpha} adX_\alpha) = 0.$$
On the other hand, the eigenvalues of adZ in $\mathfrak{g}_{\phi + i\alpha}$ are all equal to $(\phi + i\alpha)(Z) = \phi(Z) + i\alpha(Z)$. Thus,
$$0 = Tr(ad_{\tilde{\mathfrak{g}}} Z) = \sum_{i=-p}^{q} \nu_{\phi + i\alpha}(\phi(Z) + i\alpha(Z)).$$
This completes the proof. $\square$

Theorem 4.26. *We have*

(i) $(\mathfrak{h}, \mathfrak{g}_\alpha) = 0$ *for all* $\alpha \in \Delta$.

(ii) $(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ *for all* $\alpha, \beta \in \Delta$ *with* $\alpha + \beta \neq 0$.

(iii) $(\ , \ )$ *is non-singular on* $\mathfrak{h} \times \mathfrak{h}$ *and* $\mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha}, \alpha \in \Delta$.

(iv) $\Delta = -\Delta$.

(v) $\mathfrak{g}$ *has* r *independent roots where* $r = \dim \mathfrak{h}$.

(vi) $[X, Y] = (X, Y) H_\alpha$ *for all* $X \in \mathfrak{g}_\alpha$ *and* $Y \in \mathfrak{g}_{-\alpha}$. *Hence in particular,* $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathbb{C} \cdot H_\alpha$.

(vii) $\alpha(H_\alpha) \neq 0$ *for any* $\alpha \in \Delta$.

(viii) *If* $\alpha \in \Delta$, *then* $\pm\alpha$ *are the only multiples of* α *which are in* Δ.

(ix) *Each* $\mathfrak{g}_\alpha$ *is 1-dimensional.*

(x) *For any* $\alpha, \beta \in \Delta$ *with* $\alpha + \beta \neq 0$,

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \begin{cases} 0, & \text{if } \alpha + \beta \notin \Delta, \\ \mathfrak{g}_{\alpha+\beta}, & \text{if } \alpha + \beta \in \Delta. \end{cases}$$

Proof. Part (i)-(iii) have been proved before. For (iv), if $\alpha \in \Delta$ and $-\alpha \notin \Delta$, then $(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ for all $\beta \in \Delta$ and $(\mathfrak{g}_\alpha, \mathfrak{h}) = 0$. It follows that $(\mathfrak{g}_\alpha, \mathfrak{g}) = 0$. This contradicts that (X, Y) is non-degenerate. To prove (v), suppose that $\mathfrak{g}$ has less than r independent roots, then there is a non-zero $H \in \mathfrak{h}$ such that $\alpha(H) = 0$ for all $\alpha \in \Delta$. Thus from Lemma 4.24, we have

$$(H, H') = \sum_{\alpha \in \Delta} \nu_\alpha \alpha(H) \alpha(H') = 0, \quad \forall H' \in \mathfrak{h}.$$

This again contradicts that (X, Y) is non-degenerate on $\mathfrak{h}$. To prove (vi), if H is any element in $\mathfrak{h}$ and $X \in \mathfrak{g}_\alpha, Y \in \mathfrak{g}_{-\alpha}$, we have

$$\begin{aligned} ([X, Y], H) &= (H, [X, Y]) \\ &= (Y, [H, X]) \\ &= (Y, \alpha(H) X) \\ &= \alpha(H)(Y, X) \\ &= (X, Y)(H_\alpha, H) \\ &= ((X, Y) H_\alpha, H). \end{aligned}$$

For any $X_\alpha \in \mathfrak{g}_\alpha$, there is $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $(X_\alpha, X_{-\alpha}) \neq 0$, otherwise, $(X_\alpha, \mathfrak{g}_{-\alpha}) = 0$ and hence $(X_\alpha, \mathfrak{g}) = 0$. So there is $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that

$(X_\alpha, X_{-\alpha}) = 1$. Hence $H_\alpha = [X_\alpha, X_{-\alpha}] \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{h}$. From Lemma 4.25, if $\alpha(H_\alpha) = 0$, then $\phi(H_\alpha) = 0$ for all $\phi \in \Delta$. Hence by Lemma 4.24, we have

$$(H, H_\alpha) = \sum_{\phi \in \Delta} \nu_\phi \phi(H)\phi(H_\alpha) = 0, \ \ \forall H \in \mathfrak{h}.$$

This contradicts that (X, Y) restricted to $\mathfrak{h}$ is non-degenerate. So $\alpha(H_\alpha) \neq 0$ for all $\alpha \in \Delta$. This proves (vii).

We prove (viii) and (ix) at the same time. Let p be the largest positive integer such that $-p\alpha$ is a root and also let $E_\alpha \in \mathfrak{g}_\alpha$, i.e. $\mathfrak{g}_{-(p+1)\alpha} = \{0\}$. Consider

$$\tilde{\mathfrak{g}} = < E_\alpha > \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_{-\alpha} \oplus \mathfrak{g}_{-2\alpha} \oplus \cdots \oplus \mathfrak{g}_{-p\alpha}.$$

We choose $E_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $(E_\alpha, E_{-\alpha}) = 1$ and hence $[E_\alpha, E_{-\alpha}] = H_\alpha$. Then $\tilde{\mathfrak{g}}$ is invariant under adE_α and $adE_{-\alpha}$, and so it is also invariant under adH_α. Now

$$Tr(ad_{\tilde{\mathfrak{g}}} H_\alpha) = Tr(adE_\alpha adE_{-\alpha} - adE_{-\alpha} adE_\alpha) = 0.$$

On the other hand, if $m_j = \dim \mathfrak{g}_{-j\alpha}$, then

$$\begin{aligned} Tr(ad_{\tilde{\mathfrak{g}}} H_\alpha) &= \alpha(H_\alpha) - m_1\alpha(H_\alpha) - 2m_2\alpha(H_\alpha) - \cdots - pm_p\alpha(H_\alpha) \\ &= \alpha(H_\alpha)(1 - m_1 - 2m_2 - \cdots pm_p). \end{aligned}$$

Since $\alpha(H_\alpha) \neq 0$, hence $1 - m_1 - 2m_2 - \cdots pm_p = 0$. This imples that $m_1 = 1$ and $m_2 = \cdots = m_p = 0$. This proves that $\dim \mathfrak{g}_{-\alpha} = 1$ and $-2\alpha, \cdots$ are not roots. Now replace $-\alpha$ by α in the argument above, it follows that α is also a simple root and $2\alpha, \cdots$ are not roots. This completes the proof of (viii) and (ix).

Before proving (x), we first prove the following lemma.

Lemma 4.27. *Let $\alpha \in \Delta$, $\phi \in \Sigma$ and p, q be non-negative integers such that $\phi + i\alpha \in \Sigma$ for any $-p \leq i \leq q$ and $\phi - (p+1)\alpha \notin \Sigma$, $\phi + (q+1)\alpha \notin \Sigma$. Then*

$$\frac{2(H_\phi, H_\alpha)}{(H_\alpha, H_\alpha)} = -(q - p) \tag{4.3}$$

and $\phi + i\alpha$ are not roots for any $i > q$ or $i < -p$. In particular,

$$\phi - \frac{2(H_\phi, H_\alpha)}{(H_\alpha, H_\alpha)}\alpha \in \Sigma.$$

Moreover, if $\phi \in \Delta$ and $E_\phi \in \mathfrak{g}_\phi, E_\alpha \in \mathfrak{g}_\alpha$ and $E_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $[E_\alpha, E_{-\alpha}] = H_\alpha$, then we have

$$(adE_\alpha)^i E_\phi \neq 0 \ \ (1 \leq i \leq q) \qquad \textit{and} \qquad (adE_\alpha)^{q+1} E_\phi = 0$$

and

$$(adE_{-\alpha})^i E_\phi \neq 0 \ \ (1 \leq i \leq p) \qquad \textit{and} \qquad (adE_{-\alpha})^{p+1} E_\phi = 0.$$

Proof. Without loss of generality, we may assume that $\phi+i\alpha \neq 0$ for $-p \leq i \leq q$. So $\mathfrak{g}_{\phi+i\alpha}(-p \leq i \leq q)$ are 1-dimensional. Let $\tilde{\mathfrak{g}} = \oplus_{i=-p}^{q} \mathfrak{g}_{\phi+i\alpha}$. As before, $\tilde{\mathfrak{g}}$ is invariant under $adE_{-\alpha}$ and adE_{α}. Thus we have

$$\begin{aligned} 0 &= Tr_{\tilde{\mathfrak{g}}}[adE_{\alpha}, adE_{-\alpha}] = Tr_{\tilde{\mathfrak{g}}} adH_{\alpha} = \sum_{i=-p}^{q} (\phi + i\alpha)(H_{\alpha}) \\ &= (p+q+1)\phi(H_{\alpha}) + \frac{(p+q+1)(q-p)}{2}\alpha(H_{\alpha}). \end{aligned}$$

Thus

$$\frac{2(H_{\phi}, H_{\alpha})}{(H_{\alpha}, H_{\alpha})} = \frac{2\phi(H_{\alpha})}{\alpha(H_{\alpha})} = -(q-p).$$

Now if q' is the largest non-negative integer such that $(adE_{\alpha})^i E_{\phi} \neq 0$, then the subspace $\tilde{\mathfrak{g}'} = \oplus_{i=-p}^{q'} \mathfrak{g}_{\phi+i\alpha}$ is invariant under $adE_{-\alpha}$ and adE_{α}. Thus we also have

$$\frac{2(H_{\phi}, H_{\alpha})}{(H_{\alpha}, H_{\alpha})} = -(q'-p).$$

It follows that $q = q'$. Similarly, it can be proved that p is the largest non-negative integer such that $(adE_{-\alpha})^i E_{\phi} \neq 0$.

Finally if q'' is the largest non-negative integer such that $\phi + q''\alpha \in \Delta$, then $q'' \geq q$. Suppose $q'' > q$. Let $\phi_0 = \phi + q''\alpha$ and p_0, q_0 be largest non-negative integers such that $\phi_0 + i\alpha \in \Delta$ for all $-p_0 \leq i \leq q_0$ and $\phi_0 - (p_0+1)\alpha \notin \Sigma, \phi_0 + (q_0+1)\alpha \notin \Sigma$. Then $q_0 = 0$ and $p_0 < q'' - q - 1$. By the same argument above, we have

$$\frac{2(H_{\phi_0}, H_{\alpha})}{(H_{\alpha}, H_{\alpha})} = -(q_0 - p_0) = p_0 < q'' - q - 1.$$

Now $(H_{\phi_0}, H_{\alpha}) = (H_{\phi}, H_{\alpha}) + q''(H_{\alpha}, H_{\alpha})$ and hence

$$\frac{2(H_{\phi_0}, H_{\alpha})}{(H_{\alpha}, H_{\alpha})} = -(q-p) + 2q''.$$

Thus $-(q-p) + 2q'' < q'' - q - 1$ and $q'' + p < -1$. Hence $q'' = q$. Similarly, it can be proved that $\phi + i\alpha$ are not roots for $i < -p$. This completes the proof of the lemma. □

Now to complete the proof of (x), we put $\phi = \beta$ in the above lemma to get $0 \neq [E_{\alpha}, E_{\beta}] \in [\mathfrak{g}_{\alpha}, \mathfrak{g}_{\beta}]$ because $q \geq 1$ if $\alpha + \beta \in \Delta$. Since $0 \neq [\mathfrak{g}_{\alpha}, \mathfrak{g}_{\beta}] \subseteq \mathfrak{g}_{\alpha+\beta}$ and $\dim \mathfrak{g}_{\alpha+\beta} = 1$, so they must be the same. □

Recall the isomorphism from $\mathfrak{h}^*$ onto $\mathfrak{h}$ by

$$\begin{aligned} \mathfrak{h}^* &\longrightarrow \mathfrak{h} \\ \alpha &\longrightarrow H_{\alpha} \end{aligned}$$

given by

$$\alpha(H) = (H, H_\alpha), \ \ \forall H \in \mathfrak{h}.$$

Throught this isomorphism, $(\ , \ )_{\mathfrak{h}}$ pulls back to a non-singular symmetric bilinear form $< \ , \ >$ on $\mathfrak{h}^*$ defined by

$$< \alpha, \beta > := (H_\alpha, H_\beta).$$

Note that if $\alpha \in \Delta$, then $< \alpha, \alpha > = (H_\alpha, H_\alpha) = \alpha(H_\alpha) \neq 0$.

Theorem 4.28. *We have*

(i) For any $\alpha, \beta \in \Delta$,

$$\begin{aligned} &< \alpha, \alpha > > 0, \\ &< \alpha, \alpha >, < \alpha, \beta > \in \mathbb{Q}, \\ &\frac{2 < \alpha, \beta >}{< \alpha, \alpha >} \in \{0, \pm 1, \pm 2, \pm 3\}. \end{aligned}$$

(ii) If $\alpha_1, \alpha_2, \cdots, \alpha_r$ are r independent roots, then any $\phi \in \Delta$ can be written as a linear combination of them with rational coefficients.

Proof. To prove (i), for any $\alpha, \beta \in \Delta$,

$$< \alpha, \alpha > = (H_\alpha, H_\alpha) = \sum_{\phi \in \Delta} \nu_\phi \phi(H_\alpha)^2 = \sum_{\phi \in \Delta} \nu_\phi (H_\alpha, H_\phi)^2 = \sum_{\phi \in \Delta} \nu_\phi < \alpha, \phi >^2 .$$

By Lemma 4.27, we have

$$< \alpha, \phi > = (H_\alpha, H_\phi) = -\frac{1}{2}(q_{\phi,\alpha} - p_{\phi,\alpha})(H_\alpha, H_\alpha) = -\frac{1}{2}(q_{\phi,\alpha} - p_{\phi,\alpha}) < \alpha, \alpha > \tag{4.4}$$

where $p_{\phi,\alpha}$ and $q_{\phi,\alpha}$ are non-negative integers. It now follows that

$$< \alpha, \alpha > = \frac{1}{4} \sum_{\phi \in \Delta} \nu_\phi (q_{\phi,\alpha} - p_{\phi,\alpha})^2 < \alpha, \alpha >^2,$$

and hence

$$< \alpha, \alpha > = \frac{4}{\sum_{\phi \in \Delta} \nu_\phi (q_{\phi,\alpha} - p_{\phi,\alpha})^2} \in \mathbb{Q}^+$$

and by (4.4), $< \alpha, \beta > \in \mathbb{Q}$. Hence $< \alpha, \alpha > > 0$. Note that $\frac{2 < \alpha, \beta >}{< \alpha, \alpha >} \cdot \frac{2 < \alpha, \beta >}{< \beta, \beta >}$ is a positive integer and

$$\left| \frac{2 < \alpha, \beta >}{< \alpha, \alpha >} \cdot \frac{2 < \alpha, \beta >}{< \beta, \beta >} \right| \leq 4$$

because of Schwarz inequality. Hence

$$\frac{2 < \alpha, \beta >}{< \alpha, \alpha >} \in \{0, \pm 1, \pm 2, \pm 3\}$$

because if $\alpha = \lambda\beta \in \Delta$, then $\lambda = \pm 1$.

To prove (ii), for any $\phi \in \Delta$, we can write

$$\phi = \sum_{i=1}^{r} a_i \alpha_i, \quad \alpha_i \in \mathbb{C}.$$

Consider the equations

$$< \phi, \alpha_k > = \sum_{i=1}^{r} a_i < \alpha_i, \alpha_k >, \quad 1 \leq k \leq r.$$

as a system of linear equations of $a_i's$. The determinant $| < \alpha_i, \alpha_k > |$ is non-zero and thus $a_i's$ are uniquely determined. Since the coefficients of the system are rational numbers, so $a_i's$ are rational numbers.

□

Definition 4.29. *A set $\Delta_+ \subset \Delta$ is called a* **positive system of roots** *if*

(i) $\Delta_+ \cap (-\Delta_+) = \phi$,

(ii) $\Delta_+ \cup (-\Delta_+) = \Delta$,

(iii) if $\alpha, \beta \in \Delta_+$ and $\alpha + \beta \in \Delta$, then $\alpha + \beta \in \Delta_+$.

Note that Δ_+ contains exactly half of the elements in Δ.

Definition 4.30. *A set $\Delta_s \subseteq \Delta$ is said to be a* **simple system of roots** *if*

(i) $\Delta_s = \{\alpha_1, \alpha_2, \cdots, \alpha_r\}$ has r elements, where $r = \dim \mathfrak{h}^*$ and

(ii) for any $\alpha \in \Delta$, there are integers $m_i, 1 \leq i \leq r$, either all $m_i \geq 0$ or all $m_i \leq 0$ such that $\alpha = \sum_{i=1}^{r} m_i \alpha_i$.

Note that Δ_s is a basis for $\mathfrak{h}^*$ and the $m_i's$ corresponding to a given $\alpha \in \Delta$ are uniquely determined. Furthermore, if $\Delta_s = \{\alpha_1, \cdots, \alpha_r\}$ is a simple system, then

$$\Delta_+ = \left\{ \sum_{i=1}^{r} m_i \alpha_i : m_i \in \mathbb{Z}, m_i \geq 0 \right\}$$

is easily seen to be a positive system, called the **positive system corresponding to Δ_s**.

Theorem 4.31. *Simple systems always exist. Hence positive systems always exist.*

Proof. Fix a ordered basis $\{e_1, e_2, \cdots e_r\}$ of Σ. For any elements α and β in Δ, we write

$$\alpha = a_1 e_1 + a_2 e_2 + \cdots + a_r e_r, \qquad \beta = b_1 e_1 + b_2 e_2 + \cdots + b_r e_r, \;\; a_i, b_i \in \mathbb{Q}.$$

$\alpha > \beta$ if and only if $\exists\, i$ such that

$$a_1 = b_1, a_2 = b_2, \cdots, a_i = b_i \qquad \text{and} \qquad a_{i+1} > b_{i+1}.$$

If $\alpha > 0$, then α is said to be **positive**; if $\alpha < 0$, then α is said to be **negative**. If a positive element in Δ is not the sum of two other positive elements in Δ, then we call it a **prime element**. Let $\Pi = \{\alpha_1, \cdots, \alpha_m\}$ be the collection of all prime elements of Δ. We will show that Π is independent, and for any $\alpha \in \Delta, \exists\, m_i$ $(1 \leq i \leq m)$, either all $m_i \geq 0$ or all $m_i \leq 0$ such that $\alpha = \sum_{i=1}^m m_i \alpha_i$. Hence $r = m$ and Π is a simple system of roots.

If $< \alpha_i, \alpha_j >> 0$, then $\frac{2<\alpha_i,\alpha_j>}{<\alpha_i,\alpha_i>}$ and $\frac{2<\alpha_i,\alpha_j>}{<\alpha_j,\alpha_j>}$ are positive integers. So by Lemma 4.27, $\alpha_i - \alpha_j$ and $\alpha_j - \alpha_i \in \Delta$. Since $\alpha_i \neq \alpha_j$, so either $\alpha_i - \alpha_j > 0$ or $\alpha_j - \alpha_i > 0$. Without loss of generality, we assume that $\alpha_i - \alpha_j > 0$. Then let $\phi = \alpha_i - \alpha_j > 0$ and so $\alpha_i = \phi + \alpha_j$. This means that α_i is a sum of two positive elements in Δ contradicting α_i is a prime element. Therefore $< \alpha_i, \alpha_j > \leq 0$.

We next prove that Π is independent. Suppose not, without loss of generality, assume

$$\alpha_m = \sum_{i=1}^{m-1} \lambda_i \alpha_i.$$

Write

$$\alpha_m = \sum_{\lambda_i > 0} \lambda_i \alpha_i + \sum_{\lambda_i < 0} \lambda_i \alpha_i := Y + Z.$$

Since $\alpha_m > 0$ and $Z \leq 0$ because $\alpha_i > 0$, so $Y \neq 0$. Hence $Y > 0$. Since $< \alpha_i, \alpha_j > \leq 0$, so

$$\begin{aligned} < Y, Z > &= \left\langle \sum_{\lambda_i > 0} \lambda_i \alpha_i, \sum_{\lambda_i < 0} \lambda_i \alpha_i \right\rangle \\ &= \sum_{\lambda_i > 0} \sum_{\lambda_i < 0} \lambda_i \lambda_j < \alpha_i, \alpha_j > \geq 0. \end{aligned}$$

Thus,

$$< \alpha_m, Y > = < Y, Y > + < Y, Z > > 0.$$

On the other hand,

$$< \alpha_m, Y > = \sum_{\lambda_i > 0} \lambda_i < \alpha_m, \alpha_i > \leq 0.$$

This shows that Π is independent.

Since Δ is finite, so we can order all the positive elements in Δ increasingly with respect to the ordering $>$. For any $\alpha \in \Delta$, α is either > 0 or < 0. Without loss of generality, we assume $\alpha > 0$. If α is a prime element, then $\alpha = \alpha_i$ for some i, so condition (ii) is satisfied. Note that the first positive element α' in the ordering must be a prime element, otherwise $\alpha' = \beta_1 + \beta_2$ with $\beta_1, \beta_2 > 0$ and $\alpha' > \beta_i$ which contradicts α' is the first positive element. We prove by induction. So if α is not a prime element, then $\alpha = \beta + \gamma$ with $\beta, \gamma > 0$. So $\alpha > \beta$ and $\alpha > \gamma$. By induction, β and γ satisfy condition (ii) and hence α satisfies condition (ii). This completes the proof. $\square$

Theorem 4.32. *If $\Delta_s = \{\alpha_1, \cdots, \alpha_r\}$ is a simple system and if $i \neq j$, then $\alpha_i - \alpha_j$ is not a root and*

$$< \alpha_i, \alpha_j > \leq 0.$$

Proof. Clearly $\alpha_i - \alpha_j$ is not a root by (ii). If $< \alpha_i, \alpha_j >> 0$, then $\frac{2<\alpha_i,\alpha_j>}{<\alpha_i,\alpha_i>}$ and $\frac{2<\alpha_i,\alpha_j>}{<\alpha_j,\alpha_j>}$ are positive integers. So $\alpha_i - \frac{2<\alpha_i,\alpha_j>}{<\alpha_j,\alpha_j>}\alpha_j \in \Delta$. This contradicts that $\alpha_i - m\alpha_j$ is not a root. So $< \alpha_i, \alpha_j > \leq 0$. $\square$

Theorem 4.33. *A simple system $\Delta_s = \{\alpha_1, \cdots, \alpha_r\}$ and $\frac{2<\alpha_i,\alpha_j>}{<\alpha_j,\alpha_j>}$ determine the root system Δ.*

Proof. We first define the positive root with respect to Δ_s. Then it suffices to determine the positive roots. Let the **level** of $\beta = \sum_i m_i\alpha_i$ be the positive integer $|\beta| := \sum_i |m_i|$. Thus the positive roots of level one is just $\alpha_i \in \Delta_s$. Now suppose we already know the positive roots of level $\leq n, n \geq 1$. We give a method for determining the positive roots β of the next level. We claim that $\beta = \alpha + \alpha_i$ for some $\alpha > 0$ of level n and $\alpha_i \in \Delta_s$. Hence the problem is to determine for a given $\alpha > 0$ of level n the $\alpha_i \in \Delta_s$ such that $\alpha + \alpha_i$ is a root. If $\alpha = \alpha_i$, then $\alpha + \alpha_i$ is not a root. Hence we may assume that $\alpha = \sum m_j\alpha_j$ and some $k_j > 0$ for $i \neq j$. $\alpha - \alpha_i, \alpha - 2\alpha_i, \cdots$ which are roots are positive of level less than n and hence one knows which of these are roots already. Thus we can determine the non-negative integer p such that $\alpha - l\alpha_i$ are roots for $1 \leq l \leq p$ but $\alpha - (p+1)\alpha_i$ is not a root. From Lemma 4.27, we have

$$q = p - \frac{2 < \alpha, \alpha_i >}{< \alpha_i, \alpha_i >} = p - \sum_j m_j \frac{2 < \alpha_j, \alpha_i >}{< \alpha_i, \alpha_i >}$$

hence q is determined and $\alpha + \alpha_i$ is a root if and only if $q > 0$. It remains to prove the claim. As in the proof of Theorem 4.31, if $< \beta, \alpha_i > \leq 0$ for all $1 \leq i \leq r$, then $\{\beta, \alpha_1, \cdots, \alpha_r\}$ is linearly independent. Hence there is i such that $< \beta, \alpha_i >> 0$. By Lemma 4.27, we have $2 < \beta, \alpha_i > / < \alpha_i, \alpha_i >= -(q-p) > 0$ and hence $p > 0$. It follows that $\beta - \alpha_i$ is a root. If $\beta - \alpha_i < 0$, then $\alpha_i - \beta > 0$ and $\alpha_i = \beta + (\alpha_i - \beta)$ is not a prime element but all α_i are clearly prime elements. So $\beta - \alpha_i > 0$ and $\beta = (\beta - \alpha_i) + \alpha_i$. This proves the claim. $\square$

Definition 4.34. *Let $r \in \mathbb{N}$. An $r \times r$ matrix $A = (a_{ij})$ is called a* **Cartan matrix** *if the following conditions are satisfied*

(i) $a_{ij} \in \mathbb{Z}$,

(ii) $a_{ij} \leq 0$ *if* $i \neq j$,

(iii) $a_{ii} = 2$ *for all* i,

(iv) $a_{ij} = 0$ *if and only if* $a_{ji} = 0$,

(v) $\det A \neq 0$,

Let V be a vector space over $\mathbb{C}$ with a basis $\{v_1, v_2, \cdots, v_r\}$ and for every i, let $s_i \in End(V)$ be such that

$$s_i(v_j) := v_j - a_{ij}v_i, \quad j = 1, \cdots, r,$$

then $\{s_1, s_2, \cdots, s_r\}$ generates a finite subgroup of $GL(r, V)$. This group is called the **Weyl group** *of A. r is said to be the* **rank** *of A. Note that* $\det(s_i) = -1$ *and s_i^2 is an identity endomorphism of V.*

Definition 4.35. *Two Cartan matrices $A = (a_{ij})$ and $B = (b_{ij})$ are said to be* **equivalent** *if they have the same rank and there is a permutation $i \longrightarrow i'$ such that $a_{ij} = b_{i'j'}$. A Cartan matrix $A = (a_{ij})$ of rank r is said to be* **reducible** *if there exists a partition of $\{1, \cdots, r\}$ into two non-empty sets Δ_{s_1} and Δ_{s_2} such that $a_{ij} = 0$ for all $i \in \Delta_{s_1}$ and $j \in \Delta_{s_2}$. If there is no such partition, A is said to be* **irreducible**. *Clearly, irreducibility is preserved by the equivalence relation defined above.*

Let $\Delta_s = \{\alpha_1, \cdots, \alpha_r\}$ be a simple system of roots of $(\mathfrak{g}, \mathfrak{h})$. Let

$$a_{ij} := \frac{2 < \alpha_i, \alpha_j >}{< \alpha_i, \alpha_i >}.$$

Then it is easy to check that $A = (a_{ij})$ is a Cartan matric of rank r. The Weyl group of A is said to be the **Weyl group** of $\mathfrak{g}$ with respect to $\mathfrak{h}$. If s_i be the generator of the Weyl group of $\mathfrak{g}$ with respect to $\mathfrak{h}$ defined above, then s_i preserves $< , >$. In fact,

$$\begin{aligned} & < s_i(\alpha_j), s_i(\alpha_l) > \\ = \; & < \alpha_j - a_{ij}\alpha_i, \alpha_l - a_{il}\alpha_i > \\ = \; & < \alpha_j, \alpha_l > -a_{ij} < \alpha_i, \alpha_l > -a_{il} < \alpha_j, \alpha_i > +a_{ij}a_{il} < \alpha_i, \alpha_i > \\ = \; & < \alpha_j, \alpha_l > . \end{aligned}$$

Clearly, if Δ_s is a simple system, then for any s_i in the Weyl group, $s_i(\Delta_s)$

is also a simple system. In fact, if $\alpha = \sum_{j=1}^r m_j\alpha_j$ with $m_j \geq 0$, then

$$\begin{aligned}\alpha &= s_i^2(\alpha)\\ &= \sum_{j=1}^r m_j s_i^2(\alpha_j)\\ &= \sum_{j=1}^r m_j s_i(\alpha_i - a_{ij}\alpha_i)\\ &= \sum_{j\neq i} m_j s_i(\alpha_j) + \left(m_i - \sum_{j=1}^r m_j a_{ij}\right)\alpha_i.\end{aligned}$$

Then $s_i(\Delta_s)$ is a simple system because $a_{ij} \leq 0$. It can also be shown that any two simple systems are conjugate under the Weyl group. Hence if $\mathfrak{g}$ is a semi-simple Lie algebra and $\mathfrak{h}$ is a CSA with root system Δ, then if $\Delta_{s_1} = \{\alpha_1, \cdots \alpha_r\}$ and $\Delta_{s_2} = \{\beta_1, \cdots \beta_r\}$ are two simple systems of Δ and $A = (a_{ij})$ and $B = (b_{ij})$ are the corresponding Cartan matrices, then there is s in the Weyl group such that $s(\Delta_{s_1}) = \Delta_{s_2}$. Hence there is a permutation ϕ such that $s(\alpha_j) = \beta_{\phi(j)}$. Then

$$a_{ij} = \frac{2 < \alpha_i, \alpha_j >}{< \alpha_i, \alpha_i >} = \frac{2 < s(\alpha_i), s(\alpha_j) >}{< s(\alpha_i), s(\alpha_i) >} = \frac{2 < \beta_{\phi(i)}, \beta_{\phi(j)} >}{< \beta_{\phi(i)}, \beta_{\phi(i)} >} = b_{\phi(i)\phi(j)}.$$

Hence, A and B are equivalent.

On the other hand, if $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two CSA of a semi-simple Lie algebra $\mathfrak{g}$ with root systems Δ_1 and Δ_2 and ϕ is an automorphism of $\mathfrak{g}$ such that $\phi(\mathfrak{h}_1) = \mathfrak{h}_2$, then for any $\alpha \in \Delta_1$, ϕ induces a root $\phi(\alpha) \in \Delta_2$ given by

$$\phi(\alpha)(X) := \alpha(\phi^{-1}(X)), \;\; \forall X \in \mathfrak{h}_2.$$

It follows that $< \alpha, \beta >=< \phi(\alpha), \phi(\beta) >$. Therefore, their Cartan matrices are also equivalent. Now by the conjugacy of CSA under the adjoint group, different choices of $\mathfrak{h}$ and Δ_s will only result in equivalent Catran matrices. Thus we can associate with $\mathfrak{g}$ with a unique equivalent class of Cartan matrices of rank r and we have

Theorem 4.36. *Two semi-simple Lie algebras over $\mathbb{C}$ are isomorphic if and only if the corresponding equivalence classes of Catran matrices are identical. Furthermore, a Lie algebra over $\mathbb{C}$ is simple if and only if the associated equivalence class of Cartan matrices consists entirely of irreducible elements.*

For the last assertion of the theorem, if Δ_s is a disjoint union of Δ_{s_1} and Δ_{s_2} and $< \alpha, \beta >= 0$ for any $\alpha \in \Delta_{s_1}$ and $\beta \in \Delta_{s_2}$, then Δ_{s_i} determine root systems Δ_i. Let

$$\mathfrak{h}_1 := \{H \in \mathfrak{h} : \alpha(H) = 0 \;\; \forall \alpha \in \Delta_2\}$$

and

$$\mathfrak{h}_2 := \{H \in \mathfrak{h} : \alpha(H) = 0 \ \ \forall \alpha \in \Delta_1\}.$$

Then $\mathfrak{h} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$. Also, if we let

$$\mathfrak{g}_i := \mathfrak{h}_i \oplus \bigoplus_{\alpha_i \in \Delta_i} \mathfrak{g}_{\mathfrak{h}}^{\alpha_i}$$

then $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$. Thus $\mathfrak{g}$ is not simple.

4.6 Classification of Simple Lie Algebras

Let $\mathfrak{g}$ be a semi-simple Lie algebra with a CSA $\mathfrak{h}$ and a simple system $\Delta_s = \{\alpha_1, \cdots, \alpha_\ell\}$ of roots. Also let $A = (a_{ij})$ be the associated Cartan matrix.

Definition 4.37. *The Dynkin diagram associated to such $\mathfrak{g}$, $\mathfrak{h}$ and Δ_s is the graph consisting of ℓ points $P_1, \cdots, P_\ell$, where each P_i is given a weight proportional to $\langle \alpha_i, \alpha_i \rangle$, and for $i \neq j$, P_i and P_j are connected by $A_{ij} = a_{ij}a_{ji}$ lines.*

Definition 4.38. *Two Dynkin diagrams $D = \{P_i\}$ and $D' = \{P_i'\}$ are said to be isomorphic if there is a bijection between $\{P_i\}$ and $\{P_i'\}$ such that corresponding points have the same weight (up to a scaling factor) and corresponding pair of points are connected by the same number of lines.*

Proposition 4.39. *(i) Two semi-simple Lie algebras are isomorphic if and only if their Dynkin diagrams (with respect to some CSA and simple system of roots) are isomorphic.*
(ii) A semi-simple Lie algebra is simple if and only if its Dynkin diagram is connected.

Proof. (i) The '$\Leftarrow$' part is obvious. To prove the '$\Rightarrow$' part, we first observe the useful formula that for $i \neq j$,

$$a_{ij} = -\sqrt{A_{ij}} \frac{\langle \alpha_j, \alpha_j \rangle^{\frac{1}{2}}}{\langle \alpha_i, \alpha_i \rangle^{\frac{1}{2}}}.$$

Thus the Cartan matrix is determined (up to equivalence) by the Dynkin diagram. Hence the Lie algebra is also determined (up to isomorphism) by the the Dynkin diagram.
(ii) follows from the simple observation that the Dynkin diagram is partitioned into two components $D = D_1 \cup D_2$ if and only if $a_{ij} = a_{ji} = 0$ for $P_i \in D_1$ and $P_j \in D_2$, i.e. the Dynkin diagram is not connected if and only if the Cartan matrix is reducible. ☐

We will reduce the problem of classifying all simple Lie algebras to the following two steps:

(A) Reduce the (connected) Dynkin diagrams to a list of possibilities (as far as possible).
(B) For each of those possibilities, we construct a corresponding simple Lie algebra whose Dynkin diagram is the given one.

(A) First we determine a list of possible Dynkin diagrams corresponding to simple Lie algebras.
We observe that such Dynkin diagram is necessarily connected.
Each point P_i in the Dynkin diagram corresponds to a simple root α_i, which can be considered as a vector in R^ℓ equipped with the Euclidean inner product given by the Killing form $\langle \cdot, \cdot \rangle$.
For $i \neq j$, the formula

$$A_{ij} = a_{ij}a_{ji} = 4\frac{\langle \alpha_i, \alpha_j \rangle^2}{\langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle}$$

implies that the angle θ_{ij} between α_i and α_j satisfies $4\cos^2\theta_{ij} = A_{ij}$. Since $a_{ij} \leq 0$, we have $\langle \alpha_i, \alpha_j \rangle \leq 0$, and thus $90^\circ \leq \theta_{ij} < 180^\circ$.
Since a_{ij} is an integer, A_{ij} is also an integer. Thus the possible values of A_{ij} are $0, 1, 2, 3$ with corresponding values of θ_{ij} given by 90°, 120°, 135°, 150° respectively.

Definition 4.40. *A linearly independent set of vectors $\alpha_1, \cdots, \alpha_\ell$ in R^ℓ is called an allowable configuration (a.c.) if the angle between any pair of vectors is 90°, 120°, 135° or 150°.*

Thus the simple roots of a simple Lie algebra gives rise to an a.c.

Definition 4.41. *A* **Coxeter graph** *is a graph such that each pair of points is connected by $0, 1, 2$ or 3 lines.*

By dropping the weights of a Dynkin diagram, we obtain a Coxeter graph.

To determine the Dynkin diagrams of simple Lie algebras, we will first determine the associated Coxeter graphs. As a matter of fact, we will first determine the connected Coxeter graphs of a.c. by making a number of observation as follows. In the sequel, Coxeter graphs will implicitly mean those of an a.c.)

(1) If S is a Coxeter graph, the graph obtained by suppressing a number of points and the lines incident with these points is the Coxeter graph of an a.c. obtained by dropping the vectors corresponding to these points.

(2) If ℓ is the number of points of a Coxeter graph, then the number n of pairs of connected points $< \ell$.

Proof. By normalizing the vectors, we may assume that each vector α_i of the a.c. is of unit length. Then let $\alpha = \sum_{1 \leq i \leq \ell} \alpha_i$. Then

$$0 < \langle \alpha, \alpha \rangle = \ell + 2\sum_{i<j} \langle \alpha_i, \alpha_j \rangle.$$

If $\langle \alpha_i, \alpha_j \rangle \neq 0$, then since $\theta_{ij} \geq 120°$, we have $2\langle \alpha_i, \alpha_j \rangle \leq -1$.
Thus we have $0 < \ell + n(-1)$ or $n < \ell$. □

(3) A Coxeter graph of an a.c. contains no cycles.

Proof. Suppose a Coxeter graph contains a cycle. By (1), we obtain a Coxeter graph of an a.c. by suppressing all the points not in the original Coxeter graph. But the resulting Coxeter graph violates (2). □

(4) The number of lines (counting multiplicities) emanating from a point of a Coxeter graph is at most 3.

Proof. Let α be a point, and let $\alpha_1, \cdots, \alpha_k$ be the points connected to α.
No two α_i are connected since there are no cycles. Thus $\langle \alpha_i, \alpha_j \rangle = 0$ for $i \neq j$.
Choose a unit vector $\alpha_0 \in \langle \alpha, \alpha_1, \cdots, \alpha_k \rangle$ such that $\langle \alpha_0, \alpha_i \rangle = 0$ for $i \geq 1$. Since α and the α_i are linearly independent, $\langle \alpha, \alpha_0 \rangle \neq 0$. Then we have

$$\langle \alpha, \alpha \rangle = \langle \alpha, \alpha_o \rangle^2 + \langle \alpha, \alpha_1 \rangle^2 + \cdots + \langle \alpha, \alpha_k \rangle^2 = 1.$$

Thus $4 \sum_{1 \leq i \leq k} 4\langle \alpha, \alpha_i \rangle^2 < 4$. But $4\langle \alpha, \alpha_i \rangle^2$ is the no. of lines connecting α and α_i, and (4) follows. □

(5) The only connected Coxeter graph of an a.c. containing a triple line is $G_2 : \circ \equiv \circ$.

Proof. (5) follows from (4). □

(6) Let π be an a.c. and let $\alpha_1, \cdots, \alpha_k$ be vectors of π such that the corresponding points of the corresponding Coxeter graph form a simple chain in the sense that each point is connected to the next one by a single line. Let π' be the collection of vectors of π which are not in the simple chain together with the vector $\alpha = \sum_{1 \leq i \leq k} \alpha_i$. Then π' is an a.c.
(In this case, the Coxeter graph of π' is obtained by shrinking the simple chain in the Coxeter graph of π to a point.)

Proof. We have $2\langle \alpha_i, \alpha_{i+1} \rangle = -1$. Hence

$$\langle \alpha, \alpha \rangle = k + 2 \sum_{i<j} \langle \alpha_i, \alpha_j \rangle = k - (k-1) = 1.$$

Thus α is a unit vector.
Also it is clear that the vectors of π' are linearly independent.
Since π has no cycles, it follows that for a point $\beta \in \pi$ with $\beta \neq \alpha_i$ for any i is connected to at most one of the α_i. Then

$$4\langle \alpha, \beta \rangle^2 = 4\langle \alpha, \alpha_i \rangle^2 = 0, 1, 2, 3.$$

This finishes the proof of (6). □

(7) No Coxeter graph of an a.c. contains a subgraph of one of the following three forms:

Proof. By (6), the existence of the above three forms would imply the existence of the following three forms:

But by (4), the above three forms are impossible. □

(8) The only possible connected Coxeter graphs have one of the following forms:

Proof. By (5), if a Coxeter graph S contains a triple line, it must be G_2.
If S contains a double line, then it contains no node (i.e. a point connected to three other points). Also it is clear that S contains only one double line. This reduces the possibilities to the second graph in (8) above.
Finally we consider the case when there are only single lines.
If there are no nodes, then it reduces to A_ℓ.
If there is a node, then it is clear that there is at most one node, and this reduces to the third graph above. $\square$

(9) The only connected Coxeter graphs of the second type in (8) are $B_\ell = C_\ell$ and F_4 as follows:

Proof. Let p and q be as in the second graph in (8).
Let $\alpha =: \sum_{1\leq i\leq p} i\alpha_i$, $\beta := \sum_{1\leq j\leq q} j\beta_j$.
Since $2\langle\alpha_i, \alpha_{i+1}\rangle = -1$ and $2\langle\beta_j, \beta_{j+1}\rangle = -1$, we have

$$\begin{aligned}
\langle\alpha, \alpha\rangle &= \sum_{1\leq i\leq p} i^2 - \sum_{1\leq i\leq p-1} i(i+1) \\
&= p^2 - \frac{p(p-1)}{2} \\
&= \frac{p(p+1)}{2}, \quad \text{and} \\
\langle\beta, \beta\rangle &= \frac{q(q+1)}{2}.
\end{aligned}$$

Also, we have $\langle\alpha, \beta\rangle^2 = \langle p\alpha_p, q\beta_q\rangle^2 = \frac{p^2q^2}{2}$.
By Schwarz inequality, we have

$$\begin{aligned}
\frac{p^2q^2}{2} &< \frac{p(p+1)}{2} \cdot \frac{q(q+1)}{2}, \quad \text{or} \\
(p+1)(q+1) &> 2pq, \quad \text{or} \\
(p-1)(q-1) &< 2.
\end{aligned}$$

The only possibilities are: $p = 1, q$ arbitrary or $q = 1$, p arbitrary or $p = q = 2$.

The first two cases differ only in notation, and we get $B_\ell = C_\ell$. The third case is F_4. $\square$

(10) The only connected Coxeter graph of the third type in (8) are D_ℓ, E_6, E_7 and E_8 as follows:

Proof. Let $\alpha = \sum_{1\le i\le p-1} i\alpha_i$, $\beta = \sum_{1\le j\le q-1} j\beta_j$, $\gamma = \sum_{1\le k\le r-1} k\gamma_k$. The vectors α, β, γ are mutually orthogonal and δ is not in the space spanned by these vectors. Hence if $\theta_1, \theta_2, \theta_3$ are the angles between δ and α, β, γ respectively. Then we must have

$$\cos^2\theta_1 + \cos^2\theta_2 + \cos^2\theta_3 < 1.$$

As in (9), we have $\langle \delta, \alpha \rangle = \langle \delta, (p-1)\alpha_{p-1} \rangle = -(p-1)/2$, and thus

$$\cos^2\theta_1 = \frac{\langle \delta, \alpha \rangle^2}{\langle \alpha, \alpha \rangle} = \frac{\frac{1}{4}(p-1)^2}{p(p-1)} = \frac{1}{2}(1 - \frac{1}{p}).$$

Similarly, we have

$$\cos^2\theta_2 = \frac{1}{2}(1 - \frac{1}{q}), \quad \cos^2\theta_3 = \frac{1}{2}(1 - \frac{1}{r}).$$

Thus we have

$$\frac{1}{2}\Big(1 - \frac{1}{p} + 1 - \frac{1}{q} + 1 - \frac{1}{r}\Big) < 1,$$

or

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1.$$

A simple check shows that the only possibilities are:

(i) $(p, q, r) = (2, 2, *)$, which corresponds to D_ℓ;
(ii) $(p, q, r) = (2, 3, 3)$, which corresponds to E_6;
(iii) $(p, q, r) = (2, 3, 4)$, which corresponds to E_7;
(iv) $(p, q, r) = (2, 3, 5)$, which corresponds to E_8. $\square$

Summarizing (1)-(10), we have

Theorem 4.42. *The only possible connected Coxeter graphs of a.c. are A_ℓ, $B_\ell = C_\ell$, D_ℓ, G_2, F_4, E_6, E_7, E_8.*

Now we reintroduce the weights to the points, and consider the possible candidates of Dynkin diagrams of simple Lie algebras as follows.

Consider a Dynkin diagram, suppose that two points α_i and α_j are joined by $A_{ij} = a_{ij}a_{ji}$ lines. Then since a_{ij} and a_{ji} are negative integers, and A_{ij} is either $1, 2$ or 3, it follows that $(a_{ij}, a_{ji}) = (-1, -A_{ij})$ or $(-A_{ij}, -1)$.

Then we have
$$\frac{\langle \alpha_i, \alpha_i \rangle}{\langle \alpha_j, \alpha_j \rangle} = \frac{a_{ji}}{a_{ij}} = A_{ij} \text{ or} \frac{1}{A_{ij}}.$$
In other words, if two points are joined by ℓ points, then the ratio of the square of their lengths is ℓ or $1/\ell$.

In particular, if two points are joined by one single line, then their lengths are equal. Using this information, we arrive easily at the following

Theorem 4.43. *The candidates for the Dynkin diagrams of simple Lie algebras are given by A_ℓ, B_ℓ, C_ℓ, D_ℓ, G_2, F_4, E_6, E_7 and E_8 as follows:*

4.7 The classical Lie Algebras

Recall from the previous section that the classification of simple Lie algebras consists of
(A) finding a list of candidates of Dynkin diagrams for simple Lie algebras, and
(B) constructing a simple Lie algebra for each candidate in the list whose Dynkin diagram is the given one.
In the previous section, we worked out step (A), and found a list of candidates for Dynkin diagrams corresponding to simple Lie algebras, which are

$$A_\ell,\ B_\ell,\ C_\ell,\ D_\ell,\ G_2,\ F_4,\ E_6,\ F_7,\ E_8.$$

In this section, we are going to undertake the step (B).
In view of time limitaions and for the sake of simplicity, we will only construct the simple Lie algebras corresponding to $A_\ell, B_\ell, C_\ell, D_\ell$, which are the *classical Lie algebras* $sl(\ell+1)$, $so(2\ell+1)$, $sp(2\ell)$, $so(2\ell)$ respectively.

Before we construct the classical Lie algebras, we quote the following lemma, which says roughly that if a Lie algebra admits a direct sum decomposition which looks like a root space decomposition, then it is the root space decomposition. Specifically, we have

Lemma 4.44. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$. Let $\mathfrak{h}$ be an abelian subalgebra of $\mathfrak{g}$, Δ a finite subset of $\mathfrak{h}^* \setminus \{0\}$. For each $\lambda \in \mathfrak{h}^*$, let*

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

Suppose that the following conditions are satisfied:
(i) Δ *spans* $\mathfrak{h}^*$,
(ii) $\Delta = -\Delta$, *and* $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \neq 0 \quad \forall \alpha \in \Delta$,
(iii) $\mathfrak{g} = \mathfrak{h} + \sum_{\alpha\in\Delta} \mathfrak{g}_\alpha$.
Then $\mathfrak{g}$ is semi-simple, $\mathfrak{h}$ is CSA, and (iii) is the root space decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$.

Proof. It is obvious that (iii) is a direct sum, $\mathfrak{h} = \mathfrak{g}_0$, and $[\mathfrak{g}_\lambda, \mathfrak{g}_{\lambda'}] \subset \mathfrak{g}_{\lambda+\lambda'}$. In particular, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{h}$ for $\alpha \in \Delta$. It is also clear that $\mathfrak{g}_\alpha \neq 0$ for $\alpha \in \Delta$.
Fix $\alpha \in \Delta$. By (ii), we can find $X'_\alpha \in \mathfrak{g}_\alpha$ and $X'_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $H'_\alpha = \left[X'_\alpha, X'_{-\alpha}\right] \neq 0$. Then as in (4.4), we can then prove that for any $\beta \in \Delta$, $\beta(H'_\alpha) = q_{\beta\alpha}\alpha(H'_\alpha)$ for some rational number $q_{\beta\alpha}$.
In particular, if $\alpha(H'_\alpha) = 0$, then $\beta(H'_\alpha) = 0$ for all $\beta \in \Delta$, and thus by (i), we must have $H'_\alpha = 0$, which is a contradiction. Hence we have $\alpha(H'_\alpha) = 0$.
Then as in (4.4), we can deduce that $\dim \mathfrak{g}_\alpha = 1$ for all $\alpha \in \Delta$.
Let $\mathfrak{r}$ be the radical of $\mathfrak{g}$. Since $\mathfrak{r}$ is an ideal, $\mathfrak{r}$ is invariant under $\mathrm{ad}(\mathfrak{h})$. Thus we have $\mathfrak{r} = \mathfrak{r} \cap \mathfrak{h} + \sum_{\alpha\in\Delta}(\mathfrak{r} \cap \mathfrak{g}_\alpha)$.
Then for any $\alpha \in \Delta$, we must have $\mathfrak{r} \cap \mathfrak{g}_\alpha = 0$. Otherwise, one can easily show that the semi-simple Lie algebra $< H_\alpha > +\mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha}$ lies in the solvable algebra $\mathfrak{r}$, which gives rise to a contradiction. Thus we have $\mathfrak{r} \subset \mathfrak{h}$.

If $\mathfrak{r} \neq 0$, then pick a nonzero $H \in \mathfrak{r}$. Then by (i), there exists $\alpha \in \Delta$ such that $\alpha(H) \neq 0$. Then one can show that X_α is a multiple of $[H, X_\alpha]$, and thus $X_\alpha \in \mathfrak{h}$, which is a contradiction.
Hence $\mathfrak{g}$ is semi-simple. It is clear from (i) to (iii) that $\mathfrak{h}$ is maximal abelian and each $H \in \mathfrak{h}$ is semi-simple. Thus $\mathfrak{h}$ is a CSA. Then it is easy to see that (iii) is the root space decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$. □

The Lie algbra A_ℓ ($\ell \geq 2$).

Consider the Lie algebra $\mathfrak{g} = sl(\ell+1)$.
Recall that $sl(\ell+1) = \{X \in M_{\ell+1}(\mathbb{C}) : TrX = 0\}$.
Let $\mathfrak{h}$ be the abelian subalgebra of all diagonal elements of $\mathfrak{g}$, i.e. $\mathfrak{h} = \{\mathrm{diag}(a_1, \cdots, a_{l+1}) : a_i \in \mathbb{C}\}$.
Write E_{ij} for the matrix whose (i,j)-th entry is 1 and other entries are 0.
One easily sees that

$$E_{ii} - E_{i+1,i+1} \quad (1 \leq i \leq \ell+1), \quad E_{ij} \quad (i \neq j, 1 \leq i, j \leq \ell+1)$$

form a basis of $\mathfrak{g}$.
Let $\lambda_1, \cdots, \lambda_{\ell+1} \in \mathfrak{h}^*$ be such that

$$\lambda_i(\mathrm{diag}(a_1, \cdots, a_{l+1})) = a_i.$$

Obviously, $\lambda_1 + \cdots + \lambda_{\ell+1} = 0$.
Since $[\mathrm{diag}(a_1, \cdots, a_\ell), E_{ij}] = (a_i - a_j)E_{ij}$, we have

$$[H, E_{ij}] = (\lambda_i - \lambda_j)(H)E_{ij}, \quad H \in \mathfrak{h}.$$

Let $\Delta := \{\lambda_i - \lambda_j : i \neq j, 1 \leq i, j \leq \ell+1\}$. Then

$$\mathfrak{g} = \mathfrak{h} + \sum_{\alpha \in \Delta} \mathfrak{g}_\alpha \quad (\mathfrak{g}_{\lambda_i - \lambda_j} = \mathbb{C} \cdot E_{ij} \quad (i \neq j)).$$

Also, $[E_{ii}, E_{jj}] = E_{ii} - E_{jj}$ for $i \neq j$. Thus by Lemma 4.44, $\mathfrak{g}$ is semi-simple, $\mathfrak{h}$ is a CSA, and the above decomposition of $\mathfrak{g}$ is the root space decomposition with respect to $\mathfrak{h}$.
Now we consider the Killing form (,) on $\mathfrak{g}$. Using the formula $(H, H') = \sum_{\alpha \in \Delta} \nu_\alpha \alpha(H)\alpha(H')$ for $H, H' \in \mathfrak{h}$, we see that for $H = \mathrm{diag}(a_1, \cdots, a_{l+1})$ and $H' = \mathrm{diag}(a'_1, \cdots, a'_{l+1})$ (so that $\sum \alpha_i = \sum \alpha'_i = 0$), we have

$$(H, H') = \sum_{i \neq j}(a_i - a_j)(a'_i - a'_j) = 2(\ell+1) \sum_{1 \leq i \leq \ell+1} a_i a'_i.$$

Recall that $\lambda(H) = (H, H_\lambda)$. Thus for $i \neq j$, we have

$$H_{\lambda_i - \lambda_j} = \frac{1}{2(\ell+1)}(E_{ii} - E_{jj}).$$

Consider the set $S = \{\alpha_1, \cdots, \alpha_\ell\}$, where $\alpha_i = \lambda_i - \lambda_{i+1}$.

Since $\lambda_i - \lambda_j = \alpha_i + \cdots + \alpha_{j-1}$, it follows that S forms a simple system of roots, and we may write

$$\Delta = \{\pm(\alpha_i + \alpha_{i+1} + \cdots + \alpha_j) : 1 \leq i < j \leq \ell\}.$$

We have

$$\langle \alpha_i, \alpha_i \rangle = \frac{1}{\ell + 1}.$$

Let $A = (a_{ij})$ be the Cartan matrix of $\mathfrak{g}$. The off-diagonal entries are easily seen to be

$$a_{ij} = 0 = \begin{cases} 0 & \text{if } |i-j| \geq 2 \\ -1 & \text{if } |i-j| = 1. \end{cases}$$

Then it is easy to see that the Dynkin diagram of $sl(\ell+1)$ is A_ℓ.

The Lie algbra D_ℓ ($\ell \geq 2$). The Dynkin diagram D_ℓ can be realized as the Dynkin diagram of the Lie algebra $so(2n)$.
Recall from Chapter 1 that the Lie algebra $\mathfrak{g} = so(2\ell)$ is given by

$$so(2\ell) = \{X \in M_{2\ell}(\mathbb{C}) : X^t K + KX = 0\}, \quad K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Recall that each $X \in so(2\ell)$ can be written as

$$X = \begin{bmatrix} A & B \\ C & -A^t \end{bmatrix}, \quad B^t = -B, \ C^t = -C.$$

We will simply denote X by (A, B, C).
A direct computation gives

$$[(A_1, B_1, C_1), (A_2, B_2, C_2)] = (A, B, C),$$

where

$$\begin{aligned} A &= [A_1, A_2] + B_1 C_2 - B_2 C_1, \\ B &= (A_1 B_2 - A_2 B_1) - (A_1 B_2 - A_2 B_1)^t, \\ C &= (C_1 A_2 - C_2 A_a) - (C_1 A_2 - C_2 A_a)^t. \end{aligned}$$

Write E_{ij} for the $\ell \times \ell$ matrix whose (i,j)-th entry is 1, and other entries are 0.
Also, for $1 \leq p < q \leq \ell$, we let $F_{pq} = E_{pq} - E_{qp}$.
Let $\mathfrak{h} \subset \mathfrak{g}$ be the abelian Lie subalgebra consisting of elements of the form $(A, 0, 0)$, where A is any diagonal matrix.
For $1 \leq i \leq \ell$, let $\lambda_i \in \mathfrak{h}^*$ be defined by

$$\lambda_i((\mathrm{diag}(a_1, \cdots, a_\ell), 0, 0)) = a_i.$$

Then for $1 \leq i, j \leq \ell$, $1 \leq p < q \leq \ell$, $H \in \mathfrak{h}$, one can check that

$$\begin{aligned} [H, (E_{ij}, 0, 0)] &= (\lambda_i - \lambda_j)(H)(E_{ij}, 0, 0), \\ [H, (0, F_{pq}, 0)] &= (\lambda_i + \lambda_j)(H)(0, F_{pq}, 0), \\ [H, (0, 0, F_{pq})] &= -(\lambda_i + \lambda_j)(H)(0, 0, F_{pq}). \end{aligned}$$

We also have

$$\mathfrak{g} = \mathfrak{h} + \sum_{i\neq j} \mathbb{C}\cdot(E_{ij},0,0) + \sum_{p<q} \mathbb{C}\cdot(0,F_{pq},0) + \sum_{p<q} \mathbb{C}\cdot(0,0,F_{pq}).$$

Moreover, we have, for $i \neq j$ and $p < q$,

$$\begin{aligned} [(E_{ij},0,0),(E_{ji},0,0)] &= (E_{ii}-E_{jj},0,0), \\ [(0,F_{pq},0),(0,0,F_{pq})] &= (-E_{pp}-E_{qq},0,0). \end{aligned}$$

Thus we can conclude as in the case of $sl(\ell+1)$ that $\mathfrak{g}$ is semi-simple of rank ℓ, $\mathfrak{h}$ is a CSA, and the above decomposition of $\mathfrak{g}$ is the root space decomposition with respect to $\mathfrak{h}$. Now,

$$\Delta = \{\pm(\lambda_i - \lambda_j) : 1 \leq i < j \leq \ell\} \cup \{\pm(\lambda_p + \lambda_q) : 1 \leq p < q \leq \ell\}.$$

Let $\alpha_i = \lambda_i - \lambda_{i+1}$, $1 \leq i \leq \ell - 1$, and let $\alpha_\ell = \lambda_{\ell-1} + \lambda_\ell$.
Then it is easy to see that $S = \{\alpha_1, \cdots, \alpha_\ell\}$ forms a simple system of roots. In fact, $\Delta = P \cup -P$, where

$$P = \{\lambda_i - \lambda_j : 1 \leq i < j \leq \ell\} \cup \{\lambda_p + \lambda_q : 1 \leq p < q \leq \ell\},$$

and the elements in P can be given in terms of the a_i as follows:

$$\begin{aligned} \lambda_i - \lambda_j &= \alpha_i + \cdots + \alpha_{j-1} \quad (1 \leq i < j \leq \ell), \\ \lambda_p + \lambda_q &= (\alpha_p + \cdots + \alpha_{\ell-2}) + (\alpha_q + \cdots + \alpha_\ell) \quad (1 \leq p < q \leq \ell). \end{aligned}$$

Let $H = (\mathrm{diag}(a_1, \cdots, a_\ell), 0, 0)$, $H' = (\mathrm{diag}(a'_1, \cdots, a'_\ell), 0, 0)$ be in $\mathfrak{h}$. One can compute that

$$(H, H') = 4(\ell - 1) \sum_{1 \leq i \leq \ell} a_i a'_i.$$

In particular, we have

$$\begin{aligned} H_{\alpha_i} &= \frac{1}{4(\ell-1)}(E_{ii} - E_{i+1,i+1}, 0, 0) \quad (1 \leq i \leq \ell - 1) \\ H_{\alpha_\ell} &= \frac{1}{4(\ell-1)}(E_{\ell-1,\ell-1} - E_{\ell\ell}, 0, 0). \end{aligned}$$

Thus we have

$$\langle \alpha_i, \alpha_i \rangle = \frac{1}{2(\ell-1)}.$$

Denote by $A = (a_{ij})$ the Cartan matrix of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and S. For $i \neq j$, we have

$$a_{ij} = \begin{cases} -1 & \text{if } |i-j| = 1, i, j < \ell, \text{or} \\ & \qquad (i,j) = (\ell-2,\ell), (\ell, \ell-2) \\ 0 & \text{otherwise.} \end{cases}$$

Thus the Dynkin diagram of $so(2\ell)$ is D_ℓ.

The Lie algbra C_ℓ $(\ell \geq 2)$. The Dynkin diagram C_ℓ can be realized as the Dynkin diagram of the Lie algebra $sp(2\ell)$.
Recall from Chapter 1 that the Lie algebra $\mathfrak{g} = sp(2\ell)$ is given by

$$sp(2\ell) = \{X \in M_{2\ell}(\mathbb{C}) : X^t F + FX = 0\}, \quad F = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Recall that each $X \in sp(2\ell)$ can be written as

$$X = \begin{bmatrix} A & B \\ C & -A^t \end{bmatrix}, \quad B^t = B, \ C^t = C.$$

We will simply denote X by (A, B, C) as before.
A direct computation gives

$$[(A_1, B_1, C_1), (A_2, B_2, C_2)] = (A, B, C),$$

where

$$\begin{aligned} A &= [A_1, A_2] + B_1 C_2 - B_2 C_1, \\ B &= (A_1 B_2 - A_2 B_1) + (A_1 B_2 - A_2 B_1)^t, \\ C &= (C_1 A_2 - C_2 A_a) + (C_1 A_2 - C_2 A_a)^t. \end{aligned}$$

Write E_{ij} for the $\ell \times \ell$ matrix whose (i,j)-th entry is 1, and othe entries are 0.
Also, for $1 \leq p < q \leq \ell$, we let $F_{pq} = E_{pq} + E_{qp}$.
Let $\mathfrak{h} \subset \mathfrak{g}$ be the abelian Lie subalgebra consisting of elements of the form $(A, 0, 0)$, where A is any diagonal matrix.
For $1 \leq i \leq \ell$, let $\lambda_i \in \mathfrak{h}^*$ be defined by

$$\lambda_i((\operatorname{diag}(a_1, \cdots, a_\ell), 0, 0)) = a_i.$$

Then for $1 \leq i, j \leq \ell$, $1 \leq p < q \leq \ell$, $H \in \mathfrak{h}$, one can check that

$$\begin{aligned} [H, (E_{ij}, 0, 0)] &= (\lambda_i - \lambda_j)(H)(E_{ij}, 0, 0), \\ [H, (0, F_{pq}, 0)] &= (\lambda_i + \lambda_j)(H)(0, F_{pq}, 0), \\ [H, (0, 0, F_{pq})] &= -(\lambda_i + \lambda_j)(H)(0, 0, F_{pq}). \end{aligned}$$

We also have

$$\mathfrak{g} = \mathfrak{h} + \sum_{1 \neq j} \mathbb{C} \cdot (E_{ij}, 0, 0) + \sum_{p<q} \mathbb{C} \cdot (0, F_{pq}, 0) + \sum_{p<q} \mathbb{C} \cdot (0, 0, F_{pq}).$$

Moreover, we have, for $i \neq j$ and $p < q$,

$$\begin{aligned} [(E_{ij}, 0, 0), (E_{ji}, 0, 0)] &= (E_{ii} - E_{jj}, 0, 0), \\ [(0, F_{pq}, 0), (0, 0, F_{pq})] &= (E_{pp} + E_{qq}, 0, 0). \end{aligned}$$

Thus we can conclude as in the case of $sp(\ell+1)$ that $\mathfrak{g}$ is semi-simple of rank ℓ, $\mathfrak{h}$ is a CSA, and the above decomposition of $\mathfrak{g}$ is the root space decomposition with respect to $\mathfrak{h}$. Now,

$$\Delta = \{\pm(\lambda_i - \lambda_j) : 1 \le i < j \le \ell\} \cup \{\pm(\lambda_p + \lambda_q) : 1 \le p \le q \le \ell\}.$$

Let $\alpha_i = \lambda_i - \lambda_{i+1}$, $1 \le i \le \ell - 1$, and let $\alpha_\ell = 2\lambda_\ell$.
Then it is easy to see that $S = \{\alpha_1, \cdots, \alpha_\ell\}$ forms a simple system of roots. In fact, $\Delta = P \cup -P$, where

$$P = \{\lambda_i - \lambda_j : 1 \le i < j \le \ell\} \cup \{\lambda_p + \lambda_q : 1 \le p \le q \le \ell\},$$

and the elements in P can be given in terms of the a_i as follows:

$$\begin{aligned} \lambda_i - \lambda_j &= \alpha_i + \cdots + \alpha_{j-1} \quad (1 \le i \le j \le \ell), \\ \lambda_p + \lambda_q &= (\alpha_p + \cdots + \alpha_{\ell-1}) + (\alpha_q + \cdots + \alpha_\ell) \quad (1 \le p \le q \le \ell). \end{aligned}$$

Let $H = (\mathrm{diag}(a_1, \cdots, a_\ell), 0, 0)$, $H' = (\mathrm{diag}(a_1', \cdots, a_\ell'), 0, 0)$ be in $\mathfrak{h}$. One can compute that

$$(H, H') = 4(\ell+1) \sum_{1 \le i \le \ell} a_i a_i'.$$

In particular, we have

$$\begin{aligned} H_{\alpha_i} &= \frac{1}{4(\ell+1)}(E_{ii} - E_{i+1,i+1}, 0, 0) \quad (1 \le i \le \ell - 1) \\ H_{\alpha_\ell} &= \frac{1}{4(\ell+1)}(E_{\ell-1,\ell-1} - E_{\ell\ell}, 0, 0). \end{aligned}$$

Thus we have

$$\langle \alpha_i, \alpha_i \rangle = \begin{cases} \frac{1}{2(\ell+1)}, & 1 \le i < \ell \\ \frac{1}{\ell+1}, & i = \ell. \end{cases}$$

Denote by $A = (a_{ij})$ the Cartan matrix of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and S. For $i \ne j$, we have

$$a_{ij} = \begin{cases} -1 & \text{if } |i-j| = 1, 1 \le i, j \le \ell - 1, \text{or} \\ & \qquad (i,j) = (\ell, \ell-1), \\ -2 & \text{if } (i,j) = (\ell-1, \ell), \\ 0 & \text{otherwise.} \end{cases}$$

Thus the Dynkin diagram of $sp(2\ell)$ is C_ℓ.

The Lie algbra B_ℓ ($\ell \ge 1$). The Dynkin diagram B_ℓ can be realized as the Dynkin diagram of the Lie algebra $so(2\ell+1)$.
Recall from Chapter 1 that the Lie algebra $\mathfrak{g} = so(2\ell+1)$ is given by

$$so(2\ell+1) = \{X \in M_{2\ell+1}(\mathbb{C}) : X^t K + KX = 0\}, \quad K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1_\ell \\ 0 & 1_\ell & 0 \end{bmatrix}.$$

Recall that each $X \in so(2\ell+1)$ can be written as

$$X = \begin{bmatrix} 0 & a & b \\ -b^t & A & B \\ -a^t & C & -A^t \end{bmatrix}, \quad B^t = -B,\ C^t = -C.$$

We will simply denote X by $(a, b : A, B, C)$.
Write E_{ij} for the $\ell \times \ell$ matrix whose (i,j)-th entry is 1, and other entries are 0.
Also, for $1 \le p < q \le \ell$, we let $F_{pq} = E_{pq} - E_{qp}$.
Let $\mathfrak{h} \subset \mathfrak{g}$ be the abelian Lie subalgebra consisting of elements of the form $(0, 0 : A, 0, 0)$, where A is any diagonal matrix.
For $1 \le i \le \ell$, let $\lambda_i \in \mathfrak{h}^*$ be defined by

$$\lambda_i((0, 0 : \mathrm{diag}(a_1, \cdots, a_\ell), 0, 0)) = a_i.$$

Denote by e_r the $1 \times \ell$ (row matrix) where the r-th coordinate is 1, and the other coordinates are 0. Then for $1 \le i, j \le \ell$, $1 \le p < q \le \ell$, $1 \le r \le \ell$, $H \in \mathfrak{h}$, one can check that

$$\begin{aligned} [H, (0, 0 : E_{ij}, 0, 0)] &= (\lambda_i - \lambda_j)(H)(0, 0 : E_{ij}, 0, 0), \\ [H, (0, 0 : 0, F_{pq}, 0)] &= (\lambda_p + \lambda_q)(H)(0, 0 : 0, F_{pq}, 0), \\ [H, (0, 0 : 0, 0, F_{pq})] &= -(\lambda_p + \lambda_q)(H)(0, 0 : 0, 0, F_{pq}) \\ [H, (0, e_r : 0, 0, 0)] &= \lambda_r(H)(0, e_r : 0, 0, 0), \\ [H, (e_r, 0 : 0, 0, 0)] &= -\lambda_r(H)(e_r, 0 : 0, 0, 0). \end{aligned}$$

Notice that the first three equations follows as in $so(2\ell)$, while the last two equations can be deduced from the identity

$$[(0, 0 : A, 0, 0), (a, b : 0, 0, 0)] = (-aA, bA^t : 0, 0, 0).$$

We also have

$$\begin{aligned} \mathfrak{g} &= \mathfrak{h} + \sum_{i \ne j} \mathbb{C} \cdot (0, 0 : E_{ij}, 0, 0) + \sum_{p<q} \mathbb{C} \cdot (0, 0 : 0, F_{pq}, 0) + \sum_{p<q} \mathbb{C} \cdot (0, 0 : 0, 0, F_{pq}) \\ &\quad + \sum_{1 \le r \le \ell} \mathbb{C} \cdot (0, e_r : 0, 0, 0) + \sum_{1 \le r \le \ell} \mathbb{C} \cdot (e_r, 0 : 0, 0, 0). \end{aligned}$$

Thus we can conclude as before that $\mathfrak{g}$ is semi-simple of rank ℓ, $\mathfrak{h}$ is a CSA, and the above decomposition of $\mathfrak{g}$ is the root space decomposition with respect to $\mathfrak{h}$.
Now,

$$\begin{aligned} \Delta &= \{\pm(\lambda_i - \lambda_j) : 1 \le i < j \le \ell\} \cup \{\pm(\lambda_p + \lambda_q) : 1 \le p < q \le \ell\} \\ &\quad \cup \{\pm(\lambda_r) : 1 \le r \le \ell\}. \end{aligned}$$

Let $\alpha_i = \lambda_i - \lambda_{i+1}$, $1 \le i \le \ell - 1$, and let $\alpha_\ell = \lambda_\ell$.

Then it is easy to see that $S = \{\alpha_1, \cdots, \alpha_\ell\}$ forms a simple system of roots. In fact, $\Delta = P \cup -P$, where

$$\begin{aligned} P &= \{\lambda_i - \lambda_j : 1 \le i < j \le \ell\} \cup \{\lambda_p + \lambda_q : 1 \le p < q \le \ell\} \\ &\quad \cup \{\lambda_r : 1 \le r \le \ell\}. \end{aligned}$$

and the elements in P can be given in terms of the a_i as follows:

$$\begin{aligned} \lambda_i - \lambda_j &= \alpha_i + \cdots + \alpha_{j-1} \quad (1 \le i < j \le \ell), \\ \lambda_p + \lambda_q &= (\alpha_p + \cdots + \alpha_\ell) + (\alpha_q + \cdots + \alpha_\ell) \quad (1 \le p < q \le \ell), \\ \lambda_r &= \alpha_r + \cdots + \alpha_\ell \quad (1 \le r \le \ell) \end{aligned}$$

Let $H = (0, 0 : \mathrm{diag}(a_1, \cdots, a_\ell), 0, 0)$, $H' = (0, 0 : \mathrm{diag}(a'_1, \cdots, a'_\ell), 0, 0)$ be in $\mathfrak{h}$. One can compute that

$$(H, H') = 2(2\ell - 1) \sum_{1 \le i \le \ell} a_i a'_i.$$

In particular, we have

$$\begin{aligned} H_{\alpha_i} &= \frac{1}{2(2\ell-1)}(0, 0 : E_{ii} - E_{i+1,i+1}, 0, 0) \quad (1 \le i \le \ell - 1) \\ H_{\alpha_\ell} &= \frac{1}{2(2\ell-1)}(0, 0 : E_{\ell\ell}, 0, 0). \end{aligned}$$

Thus we have

$$\langle \alpha_i, \alpha_i \rangle = \begin{cases} \frac{1}{2\ell-1}, & 1 \le i \le \ell - 1 \\ \frac{1}{2(2\ell-1)}, & i = \ell. \end{cases}$$

Denote by $A = (a_{ij})$ the Cartan matrix of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and S. For $i \ne j$, we have

$$a_{ij} = \begin{cases} -1 & \text{if } |i - j| = 1, i, j < \ell, \text{or} \\ & \qquad (i, j) = (\ell - 1, \ell) \\ -2 & \text{if } (i, j) = (\ell, \ell - 1) \\ 0 & \text{otherwise.} \end{cases}$$

Thus the Dynkin diagram of $so(2\ell + 1)$ is B_ℓ.

The Lie algbra G_2. We are going to construct a Lie subalgebra of B_3 whose Dynkin diagram is G_2. We will following the notation in B_ℓ with $\ell = 3$.
Recall the CSA $\mathfrak{h} = \{H_{\lambda_1\lambda_2\lambda_3} := (0, 0, \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3), 0, 0) : \lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}\}$ of B_3. Let

$$\mathfrak{h}' = \{H_{\lambda_1\lambda_2\lambda_3} : \lambda_1 + \lambda_2 + \lambda_3 = 0\}.$$

Then $\mathfrak{h}'$ is a subalgebra of $\mathfrak{h}$.
For each root $\beta \in \mathfrak{h}^*$ in B_3, we denote by E_β the corresponding root vector constructed in the previous section. For example, we have

$$E_{\lambda_i} = (0, e_i; 0, 0, 0), E_{\lambda_i - \lambda_j} = (0, 0 : E_{ij}, 0, 0, 0), \text{etc.}$$

Define

$$\begin{aligned} G_{\lambda_i} &= \sqrt{2}E_{\lambda_i} + E_{-\lambda_i-\lambda_k} \\ G_{-\lambda_i} &= \sqrt{2}E_{-\lambda_i} + E_{+\lambda_i+\lambda_k}, \end{aligned}$$

where (i,j,k) are even permutations of $1,2,3$, and let

$$G_{\lambda_i-\lambda_j} = E_{\lambda_i-\lambda_j}, \quad i \neq k.$$

For any $H_{\lambda_1\lambda_2\lambda_3} \in \mathfrak{h}'$, we have

$$[H_{\lambda_1\lambda_2\lambda_3}, G_\alpha] = \alpha G_\alpha,$$

where $\alpha = \pm\lambda_i \ \ (i=1,2,3)$ or $\alpha = \lambda_i - \lambda_j \ \ (i,j=1,2,3,\ i\neq k)$. Let

$$H_1 = H_{100}, \quad H_2 = H_{010}, \quad H_3 = H_{001}.$$

One can compute directly that

$$\begin{aligned} [G_{\lambda_i}, G_{-\lambda_i}] &= 3H_i - (H_1 + H_2 + H_3) \\ [G_{\lambda_i-\lambda_k}, G_{\lambda_k-\lambda_i}] &= H_i - H_k, \\ \left[G_{\lambda_i}, G_{\lambda_j}\right] &= -2G_{-k} \ \ (i,j,k) \text{ is even permutation}, \\ \left[G_{-\lambda_i}, G_{-\lambda_j}\right] &= 2G_{-k} \ \ (i,j,k) \text{ is even permutation}, \\ \left[G_{\lambda_i}, G_{-\lambda_j}\right] &= 3G_{\lambda_i-\lambda_j} \ \ (i\neq j), \\ \left[G_{\lambda_i-\lambda_j}, G_{\lambda_k}\right] &= \delta_{jk}G_{\lambda_i}, \\ \left[G_{\lambda_i-\lambda_j}, G_{-\lambda_k}\right] &= -\delta_{ik}G_{\lambda_j}, \\ \left[G_{\lambda_i-\lambda_j}, G_{\lambda_k-\lambda_\ell}\right] &= \delta_{jk}G_{\lambda_i-\lambda_\ell} - \delta_{il}G_{\lambda_k-\lambda_j} \ \ (i\neq \ell,\ j\neq k). \end{aligned}$$

Again one can conclude as before that $\mathfrak{h}'$ and those G_α generate a 14 dimensional semi-simple Lie algebra with a CSA $\mathfrak{h}'$, and the roots are given by $\pm\lambda_i$ $(i = 1,2,3)$ and $\lambda_i - \lambda_j$ $(i\neq j)$. Since $\lambda_1+\lambda_2+\lambda_3 = 0$, one can show that the roots can be written as

$$\pm\lambda_1, \ \ \pm\lambda_2, \ \ \pm(\lambda_1 \pm \lambda_2), \ \ \pm(2\lambda_1+\lambda_2), \ \ \pm(\lambda_1+2\lambda_2),$$

and one can see that $\{\lambda_1-\lambda_2, \lambda_2\}$ is a simple system of roots. For $H = (0,0:\mathrm{diag}(a_1,a_2,-a_1-a_2),0,0)$ and $H' = (0,0:\mathrm{diag}(a_1',a_2',-a_1'-a_2'),0,0)$ in $\mathfrak{h}'$, one can compute that

$$(H,H') = 16(a_1a_1' + a_2a_2') + 8(a_1a_2' + a_2a_1').$$

Then using $(H,H_\alpha) = \alpha(H)$, we can see that

$$\begin{aligned} H_{\lambda_1-\lambda_2} &= \frac{1}{8}(0,0:\mathrm{diag}(1,-1,0),0,0), \\ H_{\lambda_2} &= \frac{1}{24}(0,0:\mathrm{diag}(-1,2,-1),0,0). \end{aligned}$$

Thus, we have

$$(H_{\lambda_1-\lambda_2}, H_{\lambda_1-\lambda_2}) = \frac{1}{4}, \quad (H_{\lambda_2}, H_{\lambda_2}) = \frac{1}{12}.$$

Also, we have $(H_{\lambda_1-\lambda_2}, H_{\lambda_2}) = -\frac{1}{8}$.
Thus, $a_{12} = -1$, $a_{21} = -3$, and $A_{12} = 3$.
It follows that the Dynkin diagram of the above Lie algebra is G_2.
We remark that the Lie algebra G_2 can also be realized the algebra of derivations of the (non-associative) Cayley's octonion algebra (see e.g. N. Jacobson's book).

Finally, we mention that the dimensions of the simple Lie algebras are as follows:

$$\begin{aligned}
\dim A_\ell \ (= sl(\ell+1)) &= \ell(\ell+2), \quad (\ell \geq 1),\\
\dim B_\ell \ (= so(2\ell+1)) &= \ell(2\ell+1), \quad (\ell \geq 2),\\
\dim C_\ell \ (= sp(2\ell)) &= \ell(2\ell+1), \quad (\ell \geq 3),\\
\dim D_\ell \ (= so(2\ell)) &= \ell(2\ell-1), \quad (\ell \geq 4),\\
\dim G_2 &= 14,\\
\dim F_4 &= 52,\\
\dim E_6 &= 78,\\
\dim E_7 &= 133,\\
\dim E_8 &= 248.
\end{aligned}$$

Chapter 5

Poincaré-Birkhoff-Witt Theorem

Let $\mathfrak{g}$ be a Lie algebra over k. We define a **universal enveloping algebra** for $\mathfrak{g}$ to be an associative algebra $U(\mathfrak{g})$ with unity over k, together with a Lie algebra homomorphism $j : \mathfrak{g} \longrightarrow U(\mathfrak{g})$, satisfying the **universal mapping property**, namely, if A is any associative algebra with a unit over k and a Lie algebra homomorphism $\psi : \mathfrak{g} \longrightarrow A$, there is a unique associative algebra homomorphism $\Psi : U(\mathfrak{g}) \longrightarrow A$ so that

$$\psi(X) = \Psi(j(X)), \quad \forall X \in \mathfrak{g},$$

i.e., the following diagram is commutative

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "j"] \arrow[dr, "\psi"'] & U\mathfrak{g} \arrow[d, "\Psi"] \arrow[dl, phantom, "\circlearrowleft" near start] \\ & A \end{tikzcd}$$

The universal enveloping algebra is clearly unique if exists. To show the existence, we construct $U(\mathfrak{g})$ as follows. Consider the tensor algebra

$$T\mathfrak{g} := \oplus_{n \geq 0} T^n\mathfrak{g}, \quad \text{where } T^n\mathfrak{g} := \underbrace{\mathfrak{g} \otimes \mathfrak{g} \otimes \cdots \otimes \mathfrak{g}}_{n},$$

and the ideal $I :=< [X,Y] - X \otimes Y + Y \otimes X >$. Then the universal algebra is the associative algebra given by $U(\mathfrak{g}) := T\mathfrak{g}/I$. Denote the map

$$j : \mathfrak{g} \longrightarrow T^1\mathfrak{g} \longrightarrow T\mathfrak{g} \longrightarrow U(\mathfrak{g}).$$

It is easy to check that $U(\mathfrak{g})$ has the universal mapping property. It also follows that $j(\mathfrak{g})$ generates $U(\mathfrak{g})$. Thus, any Lie algebra $\mathfrak{g}$ can be embedded into an associative algebra so that the Lie bracket on $\mathfrak{g}$ comes from the associative multiplication in the larger algebra which can be constructed in the way that it has no additional relations beyond those implied by associativity and the Lie bracket relations from $\mathfrak{g}$.

We now let $k = \mathbb{C}$. Suppose $\mathfrak{g}$ is a Lie algebra over $\mathbb{C}$ and $U(\mathfrak{g})$ is the universal enveloping algebra of $\mathfrak{g}$ and $j : \mathfrak{g} \longrightarrow U(\mathfrak{g})$ is the canonical map. Then $U(\mathfrak{g})$ is generated by the elements $j(X)$ for $X \in \mathfrak{g}$. For $n \geq 0$, we let

$$U_n(\mathfrak{g}) := \langle j(X_1)j(X_2)\cdots j(X_m) : m \leq n, X_i \in \mathfrak{g}\rangle .$$

Then

$$\mathbb{C} = U_0(\mathfrak{g}) \subset U_1(\mathfrak{g}) \subset U_2(\mathfrak{g}) \subset \cdots$$

and $U(\mathfrak{g}) = \cup_{n=0}^{\infty} U_n(\mathfrak{g})$. We are going to prove

Theorem 5.1 (Poincaré-Birkhoff-Witt). *Suppose $\mathfrak{g}$ is a finite dimensional Lie algebra and $U(\mathfrak{g})$ is the universal enveloping algebra of $\mathfrak{g}$. Let $j : \mathfrak{g} \longrightarrow U(\mathfrak{g})$ be the canonical map. Then j is injective. So $X \in \mathfrak{g}$ may be identified with $j(X) \in U(\mathfrak{g})$. Furthermore, for any ordered basis $\{X_1, X_2, \cdots, X_d\}$ of $\mathfrak{g}$,*

$$\{j(X_{i_1})j(X_{i_2})\cdots j(X_{ik}) : 1 \leq i_j \leq d \text{ and } i_1 \leq i_2 \leq \cdots i_k, k \in \mathbb{N}\}$$

is a basis for $U(\mathfrak{g})$.

We prove the theorem by constructing a $\mathfrak{g}$-module (ρ, V) with an explicit basis and then show that this module is isomorphic to $U(\mathfrak{g})$, with $\mathfrak{g}$ acting by left multiplication.

Let $\{X_1, X_2, \cdots, X_d\}$ be an ordered basis for $\mathfrak{g}$. Let

$$S := \{M = \{i_1 \leq i_2 \leq \cdots \leq i_n\} : 1 \leq i_l \leq d, n \in \mathbb{N}\}$$

and

$$S_n := \{M = \{i_1 \leq i_2 \leq \cdots \leq i_k\} : 1 \leq i_l \leq d, k \leq n\} .$$

Thus

$$\{\phi\} = S_0 \subset S_1 \subset S_2 \subset \cdots$$

and $S = \cup_{n=0}^{\infty} S_n$. For any $1 \leq i \leq d$, $i \prec M$ if and only if $i \leq i_1$ where $M = \{i_1 \leq i_2 \leq \cdots \leq i_n\}$. If $i \prec M$ then

$$iM := \{i \leq i_1 \leq i_2 \leq \cdots \leq i_n\} \in S_{n+1}.$$

Let $j\sharp M$ be the non-decreasing rearrangement of $\{j, i_1, \cdot, i_n\} \in S_{n+1}$. Thus $j\sharp M = jM$ if $j \prec M$. Let V and V_n be the vector spaces over $\mathbb{C}$ generated by $\{Z_M : M \in S\}$ and $\{Z_M : M \in S_n\}$ respectively. Then

$$\{\lambda Z_\phi : \lambda \in \mathbb{C}\} = V_0 \subset V_1 \subset V_2 \subset \cdots$$

and $V = \cup_{n=0}^{\infty} V_n$.

Lemma 5.2. *There is a unique bilinear map*

$$\begin{array}{ccc} \mathfrak{g} \times V & \longrightarrow & V \\ (X, Z) & \longrightarrow & X \cdot Z \end{array}$$

satisfying for $1 \le i \le d$ and $M \in S$

(1) $X_i \cdot Z_\phi = Z_i$.

(2) $X_i \cdot Z_M = Z_{iM}$ *if* $i \prec M$.

(3) If $M = jN$ *with* $j \prec N$, *then for all* $i > j$,

$$X_i \cdot Z_M = X_i \cdot Z_{jN} = X_j \cdot (X_i \cdot Z_N) + [X_i, X_j] \cdot Z_N.$$

(4) $X_i \cdot Z_M - Z_{i\sharp M} \in V_{|M|}$.

In particular, $\mathfrak{g}(V_n) \subseteq V_{n+1}$ *for all* $n \ge 0$.

Proof. For any $1 \le i \le d$ and $M \in S$ with $M = \{i_1 \le i_2 \le \cdots \le i_n\}$, then either $i \prec M$ or $M = jN$ with $j \prec N$ and $j < i$ (in fact $j = i_1$). Then the uniqueness can be proved by induction on $|M|$ because of conditions (1)-(3). We now prove the existence. Using (1), we define the bilinear map from $\mathfrak{g} \times V_0 \longrightarrow V_1$ by $X_i \cdot Z_\phi := Z_i$. For this case, conditions (2) and (3) are vacuous. For (4), $X_i \cdot Z_\phi - Z_{i\sharp\phi} = Z_i - Z_i = 0 \in V_1$. So we can define a bilinear map from $\mathfrak{g} \times V_0 \longrightarrow V_1$ satisfying conditions (1)-(4). Let $n \ge 1$ and suppose we have a bilinear map $\mathfrak{g} \times V_{n-1} \longrightarrow V_n$ satisfying conditions (1)-(4). We are going to define $X_i \cdot Z_M$ for any $M \in S$ with $|M| = n$ and $1 \le i \le d$. From now on, condition (1) is vacuous.

Case I: ($i \prec M$). We define $X_i \cdot Z_M := Z_{iM}$ so that condition (2) is satisfied. (3) is vacuous. Since $i\sharp M = iM$, so

$$X_i \cdot Z_M - Z_{i\sharp M} = Z_{iM} - Z_{iM} = 0 \in V_n$$

and hence (4) holds.

Case II: ($i \not\prec M$). Thus $M = jN$ with $j \prec N$ and $j < i$ and $|N| = n - 1$. By the induction assumption for condition (4),

$$X_i \cdot Z_N = Z_{i\sharp N} + W$$

where $W \in V_{n-1}$. In order to define $X_i \cdot Z_M$ satsifying (3), we must have

$$X_i \cdot Z_M = X_i \cdot Z_{jN} = X_j \cdot (X_i \cdot Z_N) + [X_i, X_j] \cdot Z_N.$$

Since $[X_i, X_j] \cdot Z_N$ has already been defined, it remains to define $X_j \cdot (X_i \cdot Z_N)$. Note that $j \prec i\sharp N$. Thus if $X_j \cdot (X_i \cdot Z_N)$ has been defined, we must have

$$\begin{aligned} X_j \cdot (X_i \cdot Z_N) &= X_j \cdot (Z_{i\sharp N} + W) \\ &= X_j \cdot Z_{i\sharp N} + X_j \cdot W \\ &= Z_{j(i\sharp N)} + X_j \cdot W \end{aligned}$$

by case I. Since $W \in V_{n-1}$, so $X_j \cdot W$ has already been defined. Thus we define $X_j \cdot (X_i \cdot Z_N) := Z_{j(i\sharp N)} + X_j \cdot W$ and hence

$$X_i \cdot Z_M := Z_{j(i\sharp N)} + X_j \cdot W + [X_i, X_j] \cdot Z_N. \tag{5.1}$$

It remains to check conditions (2)-(4) are satisfied. In this case, (2) is vacuous and (3) is automatically satisfied by our definition. To check (4), we have

$$\begin{aligned} X_i \cdot Z_M - Z_{i\sharp M} &= X_i \cdot Z_{jN} - Z_{i\sharp jN} \\ &= X_i \cdot Z_{jN} - Z_{j(i\sharp N)} \\ &= -[X_i, X_j] \cdot Z_N - X_j \cdot W \\ &\in \mathfrak{g} \cdot V_{n-1} \subseteq V_n = V_{|M|} \end{aligned}$$

by (5.1). This completes the proof of the lemma.

□

Lemma 5.3. *For $v \in V$ and $X, Y \in \mathfrak{g}$, we have*

$$[X, Y] \cdot v = X \cdot (Y \cdot v) - Y \cdot (X \cdot v). \tag{5.2}$$

Hence (ρ, V) is a representation of $\mathfrak{g}$ where $\rho(X)v = X \cdot v$.

Proof. For simplicity, we write $X \cdot v$ by Xv. For any $v \in V$, v belongs to V_n for some $n \geq 0$. If $i \leq j$, then

$$\begin{aligned} X_i(X_j Z_\phi) - X_j(X_i Z_\phi) &= X_i Z_j - X_j Z_i \\ &= Z_{ij} - \{X_i(X_j Z_\phi) + [X_j, X_i] Z_\phi\} \\ &= Z_{ij} - (X_i Z_j + [X_j, X_i] Z_\phi) \\ &= [X_i, X_j] Z_\phi. \end{aligned} \tag{5.3}$$

If $i > j$, then since $[X_i, X_j] = -[X_j, X_i]$, so (5.3) still holds. This proves (5.2) when $n = 0$. We prove the theorem by induction on n.

Let $\mathfrak{g}_j :=< X_1, \cdots, X_j >$ for $1 \leq j \leq d$. We assume (5.2) is true for any $v \in V_{n-1}, n \geq 1$ and $X, Y \in \mathfrak{g}$. We want to prove that (5.2) is true for any $v \in V_n$ and $X, Y \in \mathfrak{g}$. To do this, we employ another induction on $j \geq 1$.

For $j = 1$, $\mathfrak{g}_1 =< X_1 >$. For any $|M| \leq n$, we have $1 \prec M$. It follows that for any $i > 1$,

$$X_i(X_1 Z_M) = X_i Z_{1M} = X_1(X_i Z_M) + [X_i, X_1] Z_M$$

by Lemma 5.2. This shows that (5.2) is true for any $v \in V_n$ and any $X \in \mathfrak{g}$, $Y \in \mathfrak{g}_1$. We then suppose (5.2) is true for any $v \in V_n$ and $X \in \mathfrak{g}$, $Y \in \mathfrak{g}_{j-1}$ with $j \geq 2$. Let $v = Z_M \in V_n$ and $X = X_i, Y = X_j$ with $i > j$. We also let $M = lN$ with $l \prec N$. If $j \leq l$, then $j \prec M$ and hence

$$X_i(X_j Z_M) = X_i(Z_{jM}) = X_j(X_i Z_M) + [X_i, X_j] Z_M.$$

So (5.2) holds for this case. If $j > l$, since $l \prec N$, so by condition (2) of Lemma 5.2, we have $X_l Z_N = Z_{lN} = Z_M$. Applying the induction assumption with $v = Z_N, X = [X_i, X_j]$ and $Y = X_l$, we have

$$\begin{aligned}
[X_i, X_j](X_l Z_N) &= X_l([X_i, X_j] Z_N) + [[X_i, X_j], X_l] Z_N \\
&= X_l(X_i(X_j Z_N) - X_j(X_i Z_N)) + [[X_i, X_l], X_j] Z_N + [X_i, [X_j, X_l]] Z_N
\end{aligned}$$

by Jacobi's identity. By induction assumption, we have

$$\begin{aligned}
[[X_i, X_l], X_j] Z_N &= [X_i, X_l](X_j Z_N) - X_j([X_i, X_l] Z_N) \\
&= X_i(X_l(X_j Z_N)) - X_l(X_i(X_j Z_N)) - X_j([X_i, X_l] Z_N).
\end{aligned}$$

Similarly,

$$[X_i, [X_j, X_l]] Z_N = X_i([X_j, X_l] Z_N) - X_j(X_l(X_i, Z_N)) + X_l(X_j(X_i Z_N)).$$

All together, we have

$$\begin{aligned}
[X_i, X_j](Z_M) &= X_i X_l X_j Z_N - X_j([X_i, X_l] Z_N) - X_j X_l X_i Z_N + X_i([X_j, X_l] Z_N) \\
&= X_i X_j X_l Z_N - X_j X_i X_l Z_N \\
&= X_i X_j Z_M - X_j X_i Z_M.
\end{aligned}$$

This proves (5.2) holds for $v = Z_M \in V_n$ and $X = X_i, Y = X_j$ with $i > j$. If $i < j$, then we interchange X_i and X_j. So (5.2) is still true for this case. This completes the proof of the lemma. $\square$

Proof of Poincaré-Birkhoff-Witt Theorem:

Let (ρ, V) be the representation of $\mathfrak{g}$ in Lemmas 5.2 and 5.3. By the universal mapping property, there is homomorphism R from $U(\mathfrak{g})$ into $End(V)$ such that $\rho = R \circ j$, i.e,

$$\rho(X) v = R(j(X)) v$$

for all $X \in \mathfrak{g}$ and $v \in V$.

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "j"] \arrow[dr, "\rho"'] & U\mathfrak{g} \arrow[d, "R"] \arrow[dl, phantom, "\circlearrowleft" near end] \\
 & End(V)
\end{tikzcd}$$

In particular, we have

$$R(j(X_i))Z_\phi = \rho(X_i)Z_\phi = X_i Z_\phi = Z_i, \tag{5.4}$$

for $1 \leq i \leq d$. Since $\{Z_i : 1 \leq i \leq d\}$ is linearly independent over $\mathbb{C}$, so $\{j(X_i) : 1 \leq i \leq d\}$ is also linearly independent. Therefore, j is injective.

Since $\{Z_M : M = \{i_1 \leq i_2 \leq \cdots \leq i_k\} \in S_n\}$ is a basis for V_n, so by (5.4) we have

$$R(j(X_{i_1})j(X_{i_2})\cdots j(X_{i_k}))Z_\phi = R(j(X_{i_1}))R(j(X_{i_2}))\cdots R(j(X_{i_k}))Z_\phi = Z_M$$

and hence $\{j(X_{i_1})j(X_{i_2})\cdots j(X_{i_k}) : M \in S_n\} \subset U_n(\mathfrak{g})$ is linearly independent. We claim that it spans $U_n(\mathfrak{g})$. Indeed, any product $j(X_i)j(X_l)$ with $i > l$ can be written as $j(X_l)j(X_i) + [j(X_l), j(X_i)]$, so by iterating this process we can write any monomial involving n or fewer elements from the basis $\{j(X_j)\}$ as a linear combination of the ordered monomials $\{j(X_M) : M \in S_n\}$. This completes the proof of Poincaré-Birkhoff-Witt Theorem.

Chapter 6

Semi-simple Representations

Definition 6.1. *Let $\mathfrak{g}$ be a Lie algebra over k. A* **$\mathfrak{g}$-module** *is a vector space V over k together with a Lie algebra homomorphism $\rho : \mathfrak{g} \to End(V)$, i.e.*

$$\rho([x,y])(v) = ([\rho(x), \rho(y)])(v) \quad \textit{for all } x, y \in \mathfrak{g}, v \in V.$$

We also say ρ is a **linear representation***, and V is the* **space of representation**.

When no confusion arises, we simply write $xv = \rho(x)(v)$.

Definition 6.2. *Let V_1, V_2 be $\mathfrak{g}$-modules. $f \in Hom_k(V_1, V_2)$ is said to be a* **$\mathfrak{g}$-homomorphism** *if $x(f(v)) = f(xv)$ for all $x \in \mathfrak{g}, v \in V_1$.*

Remarks:

1. *(Trivial representation)* Any vector space V admits the trivial representation, i.e. $xv = 0$ for all $x \in \mathfrak{g}$ and $v \in V$.

2. $\mathfrak{g}$ itself is a $\mathfrak{g}$-module via the adjoint representation.

3. *(Tensor product representation)* Let V_1, V_2 be two $\mathfrak{g}$-modules. Then $V_1 \otimes V_2$ can be made into a $\mathfrak{g}$-module via the diagonal action as follows:

$$x(v_1 \otimes v_2) := xv_1 \otimes v_2 + v_1 \otimes xv_2.$$

Check:

$$\begin{aligned}
& x(y(v_1 \otimes v_2)) - y(x(v_1 \otimes v_2)) \\
= \;& x(yv_1 \otimes v_2 + v_1 \otimes yv_2) - y(xv_1 \otimes v_2 + v_1 \otimes xv_2) \\
= \;& (xyv_1 \otimes v_2 + yv_1 \otimes xv_2) + (xv_1 \otimes yv_2 + v_1 \otimes xyv_2) \\
& -(yxv_1 \otimes v_2 + xv_1 \otimes yv_2) - (yv_1 \otimes xv_2 + v_1 \otimes xyxv_2) \\
= \;& (xy - yx)v_1 \otimes v_2 + v_1 \otimes (xy - yx)v_2 \\
= \;& ([x,y]v_1) \otimes v_2 + v_1 \otimes ([x,y]v_2) \\
= \;& [x,y](v_1 \otimes v_2).
\end{aligned}$$

4. $Hom_k(V_1, V_2)$ can be made into a $\mathfrak{g}$-module by letting

$$(xf)(v) := x(f(v)) - f(xv), \quad x \in \mathfrak{g}, f \in Hom(V_1, V_2), v \in V_1.$$

Check:

$$\begin{aligned}
& x(y(f))(v) - y(x(f))(v) \\
= \;& x((yf)(v)) - (yf)(xv)) - y((xf)(v)) - (xf)(yv)) \\
= \;& x(y(fv) - f(yv)) - (y(f(xv)) - f(yxv)) \\
& -y(x(fv) - f(xv)) + (x(f(yv)) - f(xyv)) \\
= \;& x(y(fv)) - y(x(fv)) - f((xy - yx)v) \\
= \;& [x,y](fv) - f([x,y]v) \\
= \;& ([x,y]f)v.
\end{aligned}$$

Thus $f \in Hom_k(V_1, V_2)$ is a $\mathfrak{g}$-homomorphism iff $xf = 0 \quad \forall x \in \mathfrak{g}$.

Definition 6.3.

(i) A $\mathfrak{g}$-module V is said to be **simple** *(or* **irreducible***) if $V \neq \{0\}$ and V has no submodules except $\{0\}$ and V.*

(ii) A $\mathfrak{g}$-module V is said to be **semi-simple** *(or* **completely reducible***) if any of the following equivalent conditions holds:*

(a) V is a direct sum of simple submodules,

(b) every submodule is a direct summand, i.e., has a complementary submodule, or

(c) any short exact sequence of $\mathfrak{g}$-modules

$$0 \longrightarrow W \xrightarrow{\alpha} V \xrightarrow{\beta} V/W \longrightarrow 0$$

splits, (i.e. there exists a $\mathfrak{g}$-homomorphism $\epsilon : V \longrightarrow W$ such that $\epsilon \circ \alpha = id_W$, and thus $V = \alpha(W) \oplus ker(\epsilon)$ with $x = \alpha\epsilon x + (x - \alpha\epsilon x)$; or equivalently, there exists a $\mathfrak{g}$-homomorphism $\epsilon' : V/W \longrightarrow V$ such that $\beta \circ \epsilon' = id_{V/W}$, in which case one has $V = \epsilon'(V/W) \oplus ker(\beta)$ with $x = \epsilon'\beta x + (x - \epsilon'\beta x)$).

Remarks.

1. The above conditions (a), (b) and (c) are equivalent. It is obvious that $(b) \Leftrightarrow (c)$. To see that $(a) \Rightarrow (b)$, we let $V = \oplus_{i \in I} V_i$ be a direct sum of simple submodules. Given any submodule W, the submodule $W + V_1$ is given by either $W \oplus V_1$ or W (depending on whether $W \cap V_1 = \{0\}$ or V_1). Repeating the argument and consider $(W + V_1) + V_2$ and so on, we see that $V = W \oplus_{i \in J} V_i$ for some $J \subset I$. To see that $(b) \Rightarrow (a)$, first we observe that any finite dimensional $\mathfrak{g}$-module V has a simple submodule (if V has no proper submodule, then V is simple; otherwise, take a proper submodule, and repeat the argument until we get a simple submodule). Now given V, we take a simple submodule V_1. If $V_1 = V$, then we are done. Otherwise, by (b), $V = V_1 \oplus V'$ for some complementary submodule V'. Take a simple submodule V_2 of V', and consider the submodule $V_1 \oplus V_2$, and repeat the argument.

2. A simple $\mathfrak{g}$-module is semi-simple.

3. $\mathfrak{g}$ may be a semi-simple $\mathfrak{g}$-module but not a semi-simple Lie algebra (e.g. $\mathfrak{g} = k$).

Theorem 6.4 (H. Weyl). *If $\mathfrak{g}$ is semi-simple, then all $\mathfrak{g}$-modules are semi-simple.*

Lemma 6.5. *Let $\mathfrak{g}$ be a semi-simple Lie algebra, and let $\rho : \mathfrak{g} \to End(V)$ be injective, i.e., (ρ, V) is faithful. Then*

(a) The bilinear form $B_\rho(X,Y) := tr_V(\rho(X)\rho(Y))$ is non-degenerate on $\mathfrak{g}$.

(b) Take two basis $\{e_i\}$ and $\{f_i\}$ of $\mathfrak{g}$ such that $B(e_i, f_j) = \delta_{ij}$. Let $b = \sum e_i f_i \in U\mathfrak{g}$. Then the element $\tilde{\rho}(b) \in End(V)$ commutes with $\rho(x)$ for all $x \in \mathfrak{g}$ where $\tilde{\rho} : U\mathfrak{g} \longrightarrow End(V)$ such that $\tilde{\rho} \circ j = \rho$. Here the element b is called the **Casimir element** *corresponding to B_ρ.*

(c) If V is irreducible, then $\tilde{\rho}(b)(V) = V$.

Proof. For simplicity, write $B_\rho(x,y) = B(x,y)$.

(a) Consider the set

$$\mathfrak{h} = \{X \in \mathfrak{g} : B_\rho(X,Y) = 0 \quad \forall Y \in \mathfrak{g}\}.$$

Using $tr(A[B,C]) = tr(B[C,A])$, it is easy to see that $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.

By Cartan's criterion, $\mathfrak{h}$ is solvable.

Since $\mathfrak{g}$ is semi-simple, $\mathfrak{h} = \{0\}$. Thus B_ρ is non-degenerate.

(b) First we consider the mapping induced by B given by

$$\begin{array}{ccc} \mathfrak{g} \otimes \mathfrak{g} & \xrightarrow{\Phi} & Hom(\mathfrak{g}, \mathfrak{g}) \\ x \otimes y & \longrightarrow & B(y, \cdot)x \end{array}$$

where $B(y,\cdot)x \in Hom(\mathfrak{g},\mathfrak{g})$. Since B is non-degenerate, so Φ is isomorphism. Moreover, $\Phi \in Hom(\mathfrak{g}\otimes\mathfrak{g}, Hom(\mathfrak{g},\mathfrak{g}))$ and recall that both $Hom(\mathfrak{g},\mathfrak{g})$ and $\mathfrak{g}\otimes\mathfrak{g}$ are $\mathfrak{g}$-modules. One can check that Φ is $\mathfrak{g}$-homomorphism, i.e., to show that for any $z \in \mathfrak{g}$, $z\Phi = 0 \in Hom(\mathfrak{g}\otimes\mathfrak{g}, Hom(\mathfrak{g},\mathfrak{g}))$.

Check: Using $tr(A[B,C]) = tr(B[C,A])$, one has for any $x\otimes y \in \mathfrak{g}\otimes\mathfrak{g}$,

$$\begin{aligned}
(z\Phi)(x\otimes y) &= z(\Phi(x\otimes y)) - \Phi(z(x\otimes y)) \\
&= z(B(y,\cdot)x) - \Phi([z,x]\otimes y + x\otimes [z,y]) \\
&= [z, B(y,\cdot)x] - B(y,[z,\cdot])x - B(y,\cdot)[z,x] - B([z,y],\cdot)x \\
&= 0.
\end{aligned}$$

Under this correspondence, it is easy to check that

$$\Phi\left(\sum_i e_i\otimes f_i\right) = \sum_i B(f_i,\cdot)e_i$$

is the identity map $id_{\mathfrak{g}}$ from $\mathfrak{g}$ to $\mathfrak{g}$ because $\sum_i B(f_i,e_j)e_i = \delta_{i,j}e_i = e_j$.

It is also easy to see that $x(id_{\mathfrak{g}}) = 0 \in Hom(\mathfrak{g},\mathfrak{g})$ for all $x \in \mathfrak{g}$.

Thus we have, in $U\mathfrak{g}$,

$$\begin{aligned}
x(\sum_i e_i\otimes f_i) &= 0 \quad \forall x \in \mathfrak{g} \\
\Rightarrow \sum_i([x,e_i]\otimes f_i + e_i\otimes[x,f_i]) &= 0 \quad \forall x \in \mathfrak{g} \\
\Rightarrow \sum_i((x\otimes e_i - e_i\otimes x)\otimes f_i + e_i\otimes(x\otimes f_i - f_i\otimes x)) &= 0 \quad \forall x \in \mathfrak{g} \\
\Rightarrow x\otimes b - b\otimes x &= 0 \quad \forall x \in \mathfrak{g}.
\end{aligned}$$

Since $\mathfrak{g}$ generates $U\mathfrak{g}$, it follows that b is in the center of $U\mathfrak{g}$.

Now $\tilde\rho(b) = \sum_i \rho(e_i)\rho(f_i)$. Thus,

$$tr_V\tilde\rho(b) = \sum_i tr_V(\rho(e_i)\rho(f_i)) = \sum_i B_\rho(e_i,f_i) = \dim V \neq 0.$$

Thus $\tilde\rho(b) \neq 0$. If V is irreducible, by Schur's lemma, namely, if (ρ, V) is an irreducible representation of a Lie algebra $\mathfrak{g}$ then the only endomorphisms of V commuting with all $\rho(\mathfrak{g})$ are the scalars, one has $\tilde\rho(b)V = V$.

□

Proof of Theorem 6.4.

Let $\rho : \mathfrak{g} \longrightarrow End(V)$ be a $\mathfrak{g}$-module.

We first prove the theorem for the special case that V has a $\mathfrak{g}$-submodule W of codimension one. So $\rho(x)W \subseteq W$ for all $x \in \mathfrak{g}$ and

$$\begin{aligned}
\rho : \mathfrak{g} &\longrightarrow End(W) \\
x &\longrightarrow \rho(x)|_W.
\end{aligned}$$

Since $\mathfrak{g}$ is semi-simple, so $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. Now if $\dim(V/W) = 1$, then we write

$$V/W := \{\lambda v + W : \lambda \in k\}$$

for some $v \in V$. Thus $\rho(x)v = \lambda_x v$ for some $\lambda_x \in k$. It follows that $\rho([x,y])v = (\rho(x)\rho(y) - \rho(y)\rho(x))v = \lambda_x\lambda_y v - \lambda_y\lambda_x v = 0$. So $\mathfrak{g}$ acts trivially on $V/W \cong k$. Hence $\rho(x)V \subseteq W$ for all $x \in \mathfrak{g}$.

We first suppose that $\rho : \mathfrak{g} \longrightarrow End(W)$ is injective. We have an exact sequence

$$0 \longrightarrow W \longrightarrow V \longrightarrow k \longrightarrow 0.$$

If W is not irreducible, let W' be a proper non-zero submodule of W. This gives an exact sequence

$$0 \longrightarrow W/W' \longrightarrow V/W' \longrightarrow k \longrightarrow 0.$$

Then by induction on $\dim W$, this sequence splits. So there is $W' \subseteq \widetilde{W} \subseteq V$ such that $\dim(\widetilde{W}/W') = 1$ and

$$W/W' \oplus \widetilde{W}/W' = V/W'.$$

This yields another exact sequence

$$0 \longrightarrow W' \longrightarrow \widetilde{W} \longrightarrow k \longrightarrow 0$$

because $\dim(\widetilde{W}/W') = 1$. By induction assumption again, there is $X \subseteq \widetilde{W}$ such that $\dim(X) = 1$ and

$$\widetilde{W} = W' \oplus X.$$

Then $W \cap X = \{0\}$ and hence $V = W \oplus X$ by considering their dimensions. So V is completely reducible.

If W is irreducible, then since $\rho(x)V \subseteq W$ for all $x \in \mathfrak{g}$, if $\tilde{\rho}(b)$ is the Casimir element, then since $\tilde{\rho}(b)$ commutes with $\rho(x)$ for all $x \in \mathfrak{g}$, so

$$\rho(x)(\tilde{\rho}(b)v) = \tilde{\rho}(b)(\rho(x)v), \quad \forall v \in V$$

and hence $\tilde{\rho}(b)$ is a $\mathfrak{g}$-endomorphism of V. Thus $ker(\tilde{\rho}(b))$ is an $\mathfrak{g}$-submodule of V. Since $\tilde{\rho}(b) = \sum_i \rho(e_i)\rho(f_i)$, so $\tilde{\rho}(b)$ also acts trivially on V/W and

$$\tilde{\rho}(b)W \subseteq \tilde{\rho}(b)V \subseteq W.$$

It follows that $tr_{V/W}(\tilde{\rho}(b)) = 0$. On the other hand, since W is irreducible, by Schur's lemma,

$$\tilde{\rho}(b)|_W = a \cdot 1_W$$

on W but $a \neq 0$, otherwise $tr_V(\tilde{\rho}(b)) = 0 \neq \dim V$. Therefore, $ker(\tilde{\rho}(b))$ is a one-dimensional $\mathfrak{g}$-submodule of V and $ker(\tilde{\rho}(b)) \cap W = \{0\}$. Hence $V = W \oplus ker(\tilde{\rho}(b))$ because $\dim(V/W) = 1$.

In case of $\rho : \mathfrak{g} \longrightarrow End(W)$ is not injective, we let $\mathfrak{h} = ker(\rho : \mathfrak{g} \longrightarrow End(W))$. Then $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. Since $\mathfrak{g}$ acts trivially on V/W, so for any $x \in \mathfrak{h}$, $\rho(x)V \subseteq W$ and $\rho(x)W = \{0\}$. It follows that $[\mathfrak{h}, \mathfrak{h}]$ also acts trivially on V. Since $\mathfrak{h}$ is also semi-simple and hence $[\mathfrak{h}, \mathfrak{h}] = \mathfrak{h}$, so $\mathfrak{h}$ also acts trivially on V. Thus $\mathfrak{g}/\mathfrak{h}$ acts on V and by construction, $\mathfrak{g}/\mathfrak{h}$ acts faithfully on W. Since $\mathfrak{g}/\mathfrak{h}$ is also semi-simple, so the exact sequence

$$0 \longrightarrow W \longrightarrow V \longrightarrow V/W \longrightarrow 0$$

splits. Hence V is completely reducible.

Finally, we reduce the general case to the special case. Let $W \subseteq V$ be a $\mathfrak{g}$-submodule of V and

$$0 \longrightarrow W \longrightarrow V \longrightarrow V/W \longrightarrow 0$$

be an exact sequence. Let

$$\mathcal{V} := \{\phi \in Hom(V, W) : \phi|_W = a \cdot 1_W\}$$

and

$$\mathcal{W} := \{\phi \in Hom(V, W) : \phi|_W = 0\}.$$

Then $\mathcal{W} \subseteq \mathcal{V}$ are subspaces of $Hom(V, W)$. In fact, they are $\mathfrak{g}$-submodules: we need to verify that $x\phi \in \mathcal{V}$ for all $x \in \mathfrak{g}, \phi \in \mathcal{V}$. Indeed,

$$\begin{aligned} (x\phi)(w) &= x(\phi(w)) - \phi(xw) \\ &= x(aw) - a(xw) = 0 \end{aligned}$$

for all $w \in W$. So $(x\phi)|_W = 0$ and hence $x\phi \in \mathcal{W}$. Clearly, $\dim(\mathcal{V}/\mathcal{W}) = 1$. So we have an exact sequence

$$0 \longrightarrow \mathcal{W} \longrightarrow \mathcal{V} \longrightarrow \mathcal{V}/\mathcal{W} \longrightarrow 0.$$

From the special case, there is $l \subseteq \mathcal{V}$ such that $\dim l = 1$ and

$$\mathcal{V} = l \oplus \mathcal{W}.$$

Let ϕ be in l such that $\phi|_W = 1_W$. Therefore,

$$0 \longrightarrow W \longrightarrow V \longrightarrow V/W \longrightarrow 0$$

splits and completes the whole proof. $\square$

Corollary 6.6. *Let $\mathfrak{h}$ be a semi-simple ideal of a Lie algebra $\mathfrak{g}$, then $\exists$ a unique ideal $\mathfrak{t}$ such that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{t}$.*
(Here 'semi-simple' means $\mathfrak{h}$ itself is a semi-simple algebra.)

Proof. Under the adjoint representation of $\mathfrak{h}$ on $\mathfrak{g}$, we can consider $\mathfrak{g}$ as a $\mathfrak{h}$-module, and $\mathfrak{h}$ is a $\mathfrak{h}$-submodule of $\mathfrak{g}$.
Since $\mathfrak{h}$ is semi-simple, it follows from Weyl's theorem that there exists a $\mathfrak{h}$-submodule $\mathfrak{t}$ of $\mathfrak{g}$ complementary to $\mathfrak{h}$, so that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{t}$ as $\mathfrak{h}$-modules, which then imlies $[\mathfrak{h}, \mathfrak{t}] \subset \mathfrak{h} \cap \mathfrak{t} = \{0\}$.
On the other hand, for any $g = h + t \in \mathfrak{g}$ with $h \in \mathfrak{h}$ and $t \in \mathfrak{t}$, it follows that $[h, \mathfrak{h}] = \{0\}$ and thus $h = 0$ (since the center of $\mathfrak{h} = \{0\}$), which implies $g \in \mathfrak{t}$.
Thus, $\mathfrak{t} = \{g \in \mathfrak{g} : [g, \mathfrak{h}] = \{0\}\}$.
It follows that $\mathfrak{t}$ is an ideal and is unique. $\square$

Corollary 6.7. *If $\mathfrak{g}$ is semi-simple, then $ad(\mathfrak{g}) = Der(\mathfrak{g})$, i.e. all derivations of a semi-simple Lie algebra are "inner".*

Proof. Using the formula

$$[D, adx] = D([x, \cdot]) - [x, D\cdot] = [Dx, \cdot] + [x, D\cdot] - [x, D\cdot] = ad(Dx),$$

it is easy to see that $ad(\mathfrak{g})$ is an ideal of $Der(\mathfrak{g})$.
Since $center(\mathfrak{g}) = \{0\}$, it follows that $\mathfrak{g} \cong ad(\mathfrak{g})$, and thus $ad(\mathfrak{g})$ is a semi-simple ideal.
By Corollary 6.6, there exists a complementary ideal $\mathfrak{t}$ such that $Der(g) = ad(\mathfrak{g}) \oplus \mathfrak{t}$ (with $[\mathfrak{t}, ad(\mathfrak{g})] = \{0\}$). Thus for any $D \in \mathfrak{t}$, we have $[D, ad(x)] = ad(Dx) = \{0\} \ \ \forall x \in \mathfrak{g}$. Since center$(\mathfrak{g}) = \{0\}$ it follows that $Dx = 0 \ \ \forall x \in \mathfrak{g}$.
Thus $D = 0$, and $\mathfrak{t} = \{0\}$. $\square$

Theorem 6.8 (Levi). *Let $\phi : \mathfrak{g} \to \mathfrak{s}$ be a surjective Lie algebra homomorphism of $\mathfrak{g}$ onto a semi-simple Lie algebra $\mathfrak{s}$. Then $\exists$ an injective homomorphism $\epsilon : \mathfrak{s} \to \mathfrak{g}$ such that*

$$\phi \circ \epsilon = id_{\mathfrak{s}}.$$

Proof. Let $\mathfrak{a} = ker(\phi)$ so that $\mathfrak{s} \cong \mathfrak{g}/\mathfrak{a}$. First we prove the theorem under the assumption that $\mathfrak{a}$ is a simple ideal of $\mathfrak{g}$.
Let $\mathfrak{r}$ be the radical of $\mathfrak{g}$. Then since ϕ is surjective, $\phi(\mathfrak{r})$ is a solvable ideal of $\mathfrak{s}$.
Since s is semi-simple, it follows that $\phi(\mathfrak{r}) = \{0\}$, i.e., $\mathfrak{r} \subset \mathfrak{a}$.
If $\mathfrak{r} = \{0\}$, then $\mathfrak{g}$ is semi-simple, then by an earlier theorem, we have $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^{\perp}$, which easily implies the theorem.
If $\mathfrak{r} \neq \{0\}$, then $\mathfrak{r} = \mathfrak{a}$ (since by assumption, $\mathfrak{a}$ is a simple ideal).
On the other hand, then $\mathfrak{a}$ is solvable $\Rightarrow [\mathfrak{a}, \mathfrak{a}] \neq \mathfrak{a} \Rightarrow [\mathfrak{a}, \mathfrak{a}] = \{0\}$. Thus $\mathfrak{a}$ is an abelian simple ideal.

(i) Now suppose $\mathfrak{g}$ act trivially on $\mathfrak{a}$ (i.e. $\mathfrak{a} \subset center(\mathfrak{g})$). Then $\mathfrak{g}$-module $\mathfrak{g}$ becomes a $\mathfrak{g}/\mathfrak{a}$-module. Since $\mathfrak{g}/\mathfrak{a} \cong \mathfrak{s}$ is semi-simple, it follows that $\mathfrak{g}$ is completely reducible as an $\mathfrak{s}$-module, and one can find an $\mathfrak{s}$-submodule (and thus also an ideal) complementary to $\mathfrak{a}$.

(ii) Suppose that $\mathfrak{g}$ does not act trivially on the simple abelian ideal $\mathfrak{a}$.

Consider the $\mathfrak{g}$-module $W := End_k(\mathfrak{g})$.

Claim: Suppose $\exists w \in W$ such that

(i) the map $a \to aw$ is a bijection $\mathfrak{a} \leftrightarrow \mathfrak{a}w$, and

(ii) $\mathfrak{g}w = \mathfrak{a}w$.

Then the stabilizer $i_w = \{x \in \mathfrak{g} : xw = 0\}$ is a Lie subalgebra of $\mathfrak{g}$, and $\mathfrak{g} = \mathfrak{a} \oplus i_w$ (as vector spaces), and thus ϵ exists.

Check: For $x, y \in i_w$, $[x,y]w = x(y(w)) - y(x(w)) = 0$. Thus i_w is a Lie subalgebra of $\mathfrak{g}$. The rest statements are also obvious.

Thus it remains to find such an w. Define

$$\begin{aligned} P &= \{ad_{\mathfrak{g}}a : a \in \mathfrak{a}\}, \\ Q &= \{\phi \in W : \phi(\mathfrak{g}) \subset \mathfrak{a}, \phi(\mathfrak{a}) = \{0\}\}, \\ R &= \{\phi \in W : \phi(\mathfrak{g}) \subset \mathfrak{a}, \phi(\mathfrak{a}) = \lambda_\phi \cdot id_{\mathfrak{a}}, \lambda_\phi \in k\}. \end{aligned}$$

Claim: $P \subset Q \subset R \subset W$ and P, Q, R are $\mathfrak{g}$-submodules of W.

Check: Since $\mathfrak{a}$ is an abelian ideal of $\mathfrak{g}$, it follows that $P \subset Q$. The other inclusions are obvious.

For $x \in \mathfrak{g}$ and $ad_{\mathfrak{g}}a \in P$, we have

$$x(ad_{\mathfrak{g}}a) = x([a,\cdot]) = [x,[a,\cdot]] - [a,[x,\cdot]] = [[x,a],\cdot] = ad_{\mathfrak{g}}[x,a].$$

Thus $\mathfrak{g}$ leaves P invariant, and P is a submodule of W.

For $x, y \in \mathfrak{g}$ and $\phi \in Q$, we have

$$(x\phi)(y) = [x, \phi(y)] - \phi([x,y]).$$

Since $\mathfrak{a}$ is an ideal, it follows that $(x\phi)(y) \in \mathfrak{a}$.

It is also easy to see that $(x\phi)(y) = 0$ if $y \in \mathfrak{a}$.

Thus Q is also a $\mathfrak{g}$-submodule of W. For $x \in \mathfrak{g}$, $\phi \in R$ and $a \in \mathfrak{a}$, we have

$$(x\phi)(a) = [x, \phi(a)] - \phi([x,a]) = [x, \lambda_\phi \cdot a] - \lambda_\phi \cdot [x,a] = 0.$$

Thus $\lambda_{x\phi} = 0$, and one can easily check that R is also a $\mathfrak{g}$-submodule of W.

Also one has a short exact sequence of $\mathfrak{g}$-modules

$$0 \to Q \xrightarrow{i} R \xrightarrow{\rho} k \to 0,$$

where k is the trivial $\mathfrak{g}$-module, $\rho(\phi) = \lambda_\phi$. The above calculation that $\lambda_{x\phi} = 0$ implies that ρ is a $\mathfrak{g}$-homomorphism.

The above exact sequence descends to a short exact sequence of $\mathfrak{g}$-modules

$$0 \to Q/P \to R/P \xrightarrow{\bar{\rho}} k \to 0.$$

For any $a \in \mathfrak{a}$ and $\phi \in R$, we have

$$(a\phi)(\cdot) = [a, \phi(\cdot)] - \phi([a,\cdot]) = 0 - \lambda_\phi \cdot ad_{\mathfrak{g}}a = ad_{\mathfrak{g}}(-\lambda_\phi a).$$

Thus we have $\mathfrak{a}R \subset P$, and thus the above exact sequence can be regarded as a short exact sequence of $\mathfrak{g}/\mathfrak{a}$- or $\mathfrak{s}$- modules. Then by Weyl's theorem, one find a line R'/P in R/P complementary to Q/P and invariant under $\mathfrak{s}$. Necessarily the action of $\mathfrak{s}$ on R'/P is trivial, since $\mathfrak{D}s = s$. Consider the short exact sequence of $\mathfrak{s}$-modules

$$0 \to P \to R' \to R'/P \to 0.$$

By Weyl's theorem, one can find a line ℓ in R' stable under $\mathfrak{s}$ and complementary to P. Then one can find $w \in \ell$ such that $w|_{\mathfrak{a}} = id_{\mathfrak{a}}$.

Check:

(i) For any $a \in \mathfrak{a}$, $aw = [a, w(\cdot)] - w([a, \cdot]) = -ad_{\mathfrak{g}}a$ (since $w \in R$ and $\mathfrak{a}$ is an abelian ideal $\Rightarrow [a, w(\cdot)] = \{0\}$.
Thus $aw = 0 \Rightarrow a \in center(\mathfrak{g})$.
But $\mathfrak{g}$ acts non-trivially on $\mathfrak{a} \Rightarrow center(\mathfrak{g}) \cap \mathfrak{a}$ is a proper ideal of $\mathfrak{g}$. Since $\mathfrak{a}$ is a simple ideal, thus $a = 0$.
Thus the map $a \to aw$ is injective.

(ii) $\mathfrak{s}$ (hence $\mathfrak{g}$) acts trivially on $R'/P \Rightarrow xw \in P$ for all $x \in \mathfrak{g} \Rightarrow \mathfrak{g}w = \mathfrak{a}w$. This finishes the proof of the theorem under the assumption that $\mathfrak{a}$ is a simple ideal of $\mathfrak{g}$.

In the general case, we prove the theorem by induction on $\dim \mathfrak{a}$.

When $\dim \mathfrak{a} = 0$, the theorem is trivially true because ϕ is bijective.

If $\mathfrak{a}$ is simple, then the theorem reduces to (a). If not, take an ideal $\mathfrak{a}'$ of $\mathfrak{g}$ such that $\mathfrak{a}' \subsetneqq \mathfrak{a}$. Then ϕ induces a surjective homomorphism $\bar{\phi} : \mathfrak{g}/\mathfrak{a}' \to \mathfrak{s}$. By induction assumption, there exists a Lie subalgebra $\mathfrak{g}'/\mathfrak{a}'$ isomorphic to $\mathfrak{s}$ and complementary to $\mathfrak{a}/\mathfrak{a}'$. Then consider

$$0 \to \mathfrak{a}' \to \mathfrak{g}' \to \mathfrak{g}'/\mathfrak{a}' \cong \mathfrak{s} \to 0.$$

Again by induction assumption, there exists a Lie subalgebra $\mathfrak{s}'$ of $\mathfrak{g}'$ (and hence of $\mathfrak{g}$) isomorphic to s and complementary to $\mathfrak{a}'$. Notice that $\mathfrak{s}'$ is complementary to $\mathfrak{a}$ since

$$(\mathfrak{s}' + \mathfrak{a})/\mathfrak{a} = (\mathfrak{s}' + \mathfrak{a}' + \mathfrak{a})/\mathfrak{a} = (\mathfrak{g}' + \mathfrak{a})/\mathfrak{a} = \mathfrak{g}/\mathfrak{a} \cong \mathfrak{s}.$$

□

Definition 6.9. *Let $\mathfrak{g}$ be a Lie algebra, $\mathfrak{h} \subset \mathfrak{g}$ an ideal, and $\mathfrak{t} \subset \mathfrak{g}$ a Lie subalgebra. Then $\mathfrak{g}$ is called the* **semi-direct product** *of $\mathfrak{t}$ by $\mathfrak{h}$ if the natural map $\mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$ induces an isomorphism $\mathfrak{t} \cong \mathfrak{g}/\mathfrak{h}$.*

In this case, $\mathfrak{g}$ is completely determined by $\mathfrak{h}$, $\mathfrak{t}$ and the Lie algebra homomorphism $\theta : \mathfrak{t} \to Der(\mathfrak{a})$ given by the adjoint map, and any such triple gives rise to a Lie algebra as the semi-direct product of $\mathfrak{t}$ by $\mathfrak{h}$ with $\theta(t) = ad_{\mathfrak{h}}t$. (Ex.)

Corollary 6.10. *Every Lie algebra $\mathfrak{g}$ is the semi-direct product of a semi-simple Lie subalgebra by the radical $\mathfrak{r}$.*

Proof. Apply Levi's theorem to $\mathfrak{g} \to \mathfrak{g}/\mathfrak{r}$. □

Corollary 6.11. *Any non-zero Lie algebra which is neither simple nor one-dimensional, is a semi-direct product of two Lie algebras of smaller dimension.*

Proof. (i) If $\mathfrak{g}$ is semi-simple (but not simple), then the $\mathfrak{g}$ is the direct sum of simple ideals, and thus the statement is true.

(ii) If $\mathfrak{g}$ is not semi-simple, and $\mathfrak{D}\mathfrak{g} = \mathfrak{g}$, then $\mathfrak{g}$ is not solvable, thus its radical $\mathfrak{r}$ is a proper ideal of $\mathfrak{g}$. Then we can apply Corollary 6.10.

(iii) If $\mathfrak{g}$ is not semi-simple, and $\mathfrak{D}\mathfrak{g} \neq \mathfrak{g}$, then take a vector subspace $\mathfrak{t}$ of $\mathfrak{g}$ of codimension one and such that $\mathfrak{t} \supset \mathfrak{D}\mathfrak{g}$. Since $\mathfrak{t} \supset \mathfrak{D}\mathfrak{g}$, $\mathfrak{t}$ is an ideal of $\mathfrak{g}$. Take any non-zero $x \in \mathfrak{g} \backslash \mathfrak{t}$. Then kx is necessarily a Lie subalgebra, and we have $\mathfrak{g} = \mathfrak{t} \oplus kx$.

□

Theorem 6.12. (A criterion for complete reducibility of a representation of a Lie algebra). *Let $\mathfrak{g} \subset End(V)$ be a Lie subalgebra, $V \neq \{0\}$. Then V is semi-simple if and only if*

(a) $\mathfrak{g} = \mathfrak{c} \times \mathfrak{s}$, where $\mathfrak{c}$ is abelian and $\mathfrak{s}$ is semi-simple, and

(b) the elements in $\mathfrak{c}$ can be diagonalized by a basis of V.

Remark. Since $\mathfrak{c}$ is abelian, the two conditions (a) and (b) mean that one single basis of V will diagonalize all elements of $\mathfrak{c}$.

Proof. '$\Rightarrow$' Let V be completely reducible. Let $\mathfrak{r}$ be the radical of $\mathfrak{g}$.
If $\mathfrak{r} = \{0\}$, then nothing needs to be proved.
If $\mathfrak{r} \neq \{0\}$, then by Lie's theorem, $\exists$ a line stable by $\mathfrak{r}$, which implies that there exists a linear form $\lambda : \mathfrak{r} \to k$ such that

$$V_\lambda = \{v \in V : xv = \lambda(x)v \quad \forall v \in \mathfrak{r}\}$$

is non-zero. Moreover, from the proof of Lie's theorem, V_λ is invariant under $\mathfrak{g}$, and is thus a $\mathfrak{g}$-submodule.
Then by complete reducibility of V, we can write

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}.$$

Under such decomposition, one sees that $\mathfrak{r}$ acts diagonally, and commutes with the action of $\mathfrak{g}$, i.e., $\mathfrak{r} \subset center(\mathfrak{g})$, and hence $\mathfrak{r} = center(\mathfrak{g})$. Take $\mathfrak{c}$ to be $\mathfrak{r}$. Then $\mathfrak{g}/\mathfrak{c}$ is semi-simple.
Apply Levi's theorem to $\mathfrak{g} \to \mathfrak{g}/\mathfrak{c}$, then we get a Lie subalgebra $\mathfrak{s}$ of $\mathfrak{g}$ complementary to $\mathfrak{c}$.

Since $[\mathfrak{g},\mathfrak{s}] = [\mathfrak{c}\oplus\mathfrak{s},\mathfrak{s}] = [\mathfrak{s},\mathfrak{s}] = \mathfrak{s}$, it follows that $\mathfrak{s}$ is also an ideal of $\mathfrak{g}$, and we proved (a) and (b).
'$\Leftarrow$' Take a basis of V as given in (b). Then one sees that V has a decomposition

$$V = V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_m}$$

as a direct sum of vector spaces, where the $\lambda_i' s$ are linear forms on $\mathfrak{c}$.
For each $v\in V_{\lambda_i}$, $h\in\mathfrak{c}$ and $x\in\mathfrak{g}$, $h(xv) = x(hv) = x(\lambda_i(h)v) = \lambda_i(h)\cdot xv$.
It follows that each V_{λ_i} is stable under $\mathfrak{g}$, and is thus a $\mathfrak{g}$-submodule of V.
Regard each V_{λ_i} as a $\mathfrak{s}$-module. Then since $\mathfrak{s}$ is semi-simple, V_{λ_i} is completely reducible as a $\mathfrak{s}$-module. Then one easily sees that V_{λ_i} is completely reducible as a $\mathfrak{g}$-module. This finishes the proof of the theorem. $\square$

Corollary 6.13. *The tensor products of semi-simple $\mathfrak{g}$-modules is semi-simple.*

Proof. We only need to consider the case of tensor product of two modules, since the general case can then be deduced by tensoring two modules repeatedly.

Let $\rho_i : \mathfrak{g}\to End(V_i)$, $i = 1,2$, be two semi-simple representations. For $i = 1,2$, we can consider V_i as a $\rho_i(\mathfrak{g})(=\mathfrak{g}/ker(\rho_i))$-module. Then by Theorem 6.12, $\rho_i(\mathfrak{g})\subset End(V_i)$ is of the form $\mathfrak{c}_i\times\mathfrak{s}_i$, and $\mathfrak{c}_i$ acts diagonally on V_i.
Thus both the radical and center of $\rho_i(\mathfrak{g})$ is equal to $\mathfrak{c}_i$.
Denote by $\mathfrak{r}$ the radical of $\mathfrak{g}$. Then since ρ_i is surjective on to $\rho(\mathfrak{g})$,it is easy to see that $\rho_i(\mathfrak{r})$ is the radical of $\rho_i(\mathfrak{g})$.
(*check:* Since $\rho_i : \mathfrak{g}\to\rho_i(\mathfrak{g})$ is surjective, it is easy to see that $\rho_i(\mathfrak{r})$ is a solvable ideal of $\rho_i(\mathfrak{g})$, and thus $\rho_i(\mathfrak{r})\subset\mathfrak{r}'$, where $\mathfrak{r}'$ denotes the radical of $\rho_i(\mathfrak{g})$.
On the other hand, $\mathfrak{g}/\mathfrak{r}$ is semi-simple, and ρ_i induces naturally a surjective homomorphism $\bar{\rho}_i : \mathfrak{g}/\mathfrak{r}\to\rho_i(\mathfrak{g})/\rho_i(\mathfrak{r})$. Thus $\rho_i(\mathfrak{g})/\rho_i(\mathfrak{r})$ is also semi-simple.
If $\mathfrak{r}'\neq\rho_i(\mathfrak{r})$, then $\mathfrak{r}'/\rho_i(\mathfrak{r})$ is a non-zero solvable ideal of $\rho_i(\mathfrak{g})/\rho_i(\mathfrak{r})$, which is a contradiction.)
Thus we have $\rho_i(\mathfrak{r}) = \mathfrak{c}_i$, and hence $\mathfrak{r}$ acts diagonally on V_i.
Now we consider the $\mathfrak{g}$-module $V_1\otimes V_2$ with corresponding $\mathfrak{g}$-homomorphism $\rho = \rho_1\otimes\rho_2$.
As we have seen, $\rho(\mathfrak{r})$ is the radical of $\rho(\mathfrak{g})$. Moreover, for $h\in\mathfrak{r}$, $x\in\mathfrak{g}$, we have

$$\begin{aligned} & [\rho(h),\rho(x)](v_1\otimes v_2) \\ = \;& \rho([h,x])(v_1\otimes v_2) \\ = \;& \rho_1([h,x])v_1\otimes v_2 + v_1\otimes\rho_1([h,x])v_2 \\ = \;& ([\rho_1(h),\rho_1(x)])v_1\otimes v_2 + v_1\otimes[\rho_2(h),\rho_2(x)])v_2 \\ = \;& 0. \end{aligned}$$

Here the last line follows from the fact that $\mathfrak{c}_i = \rho_i(\mathfrak{r})$ is the center of $\rho_i(\mathfrak{g})$.
Thus the radical $\rho(\mathfrak{r})$ is also the center of $\rho(\mathfrak{g})$. In particular, we have, $\rho(\mathfrak{g}) = \rho(\mathfrak{r})\times\mathfrak{s}$, where $\mathfrak{s}$ is semi-simple and $\rho(\mathfrak{r})$ is abelian.
Moreover, for $x\in\mathfrak{g}$, take a basis $\{e_\alpha\}$ of V_1 consisting of eigenvectors of $\rho_1(x)$ with corresponding eigenvalues $\{\lambda_\alpha\}$, and similarly a basis $\{f_\beta\}$ of V_2 consisting

of eigenvectors of $\rho_2(x)$ with corresponding eigenvalues $\{\lambda'_\beta\}$. Then

$$\begin{aligned} & \rho(e_\alpha \otimes f_\beta) \\ = \; & \rho_1(e_\alpha) \otimes f_\beta + e_\alpha \otimes \rho_2(f_\beta) \\ = \; & \lambda_\alpha \cdot e_\alpha \otimes f_\beta + e_\alpha \otimes \lambda'_\beta \cdot f_\beta. \end{aligned}$$

Thus $\rho(\mathfrak{t})$ also acts diagonally on $V_1 \otimes V_2$.
By Theorem 6.12 again, it follows that $V_1 \otimes V_2$ is also a semi-simple $\mathfrak{g}$-module. $\square$

Bibliography

[1] R. Goodman and N. R. Wallach, *Representations and Invariants of the Classical Groups*, Cambridge University Press, Cambridge 1998.

[2] N. Jacobson, *Lie Algebras*, Dover Publications, Inc., New York 1979.

[3] J. P. Serre, *Complex Semisimple Lie Algebras*, Springer-Verlag- Heidelberg-New York 1987.

[4] J. P. Serre, *Lie Algebras and Lie Groups*, W.A. Benjamin, Inc., New York-Amsterdam 1965.

[5] Z. X. Wan, *Lie Algebras*, Pergamon Press 1975.