NOTES FOR LIE GROUPS AND LIE ALGEBRA

TRANSCRIBED BY NING TANG   INSTRUCTOR: PROFESSOR FAN GAO

Contents

Date: June 14, 2022.

Syllabus

In this course, we denote $k = \mathbb{R}$ or $\mathbb{C}$ (rarely but possibly $\mathbb{Q}_p$). We will cover

- Lie groups G over k;
- Linear Lie groups over k (those inside $GL_n(\mathbb{R})$);
- Lie algebra $\mathfrak{g}$ over k.

The "Lie algebra Everest" for us to climb:

- Classification (over $\mathbb{C}$) related to Dynkin diagram;
- Finite dimensional representation of $\mathfrak{g}$, intuitively, a homomorphism $\rho : \mathfrak{g} \to M_{n\times n}(\mathbb{C})$.

We will briefly introduce the basic knowledge of Lie groups and devote most of our energies to develop the main tools of Lie algebras.

1. Lie groups

1.1. Lie groups: Definitions.

Definition 1.1. *We say M is a smooth manifold over k if transition maps between charts are given by C^∞ functions. Here the term C^∞ denotes*

$$\begin{cases} \text{infinitely differentiable, if } k = \mathbb{R}, \\ \text{analytic(holomorphic), if } k = \mathbb{C}. \end{cases}$$

Example 1.2. *Here are some examples for smooth manifolds.*

(1) $M = \mathbb{R}$, $k = \mathbb{R}$.
(2) $M = \mathbb{R}/\mathbb{Z}$, $k = \mathbb{R}$: *We should note that* $\mathbb{R}/\mathbb{Z} = \{z \in \mathbb{C} : |z| = 1\}$ *is not a smooth manifold over* $\mathbb{C}$.
(3) $M = \mathbb{R}^2/L$, $L \simeq \mathbb{Z}\times\mathbb{Z} = \mathbb{Z}v_1 \otimes \mathbb{Z}v_2$, $k = \mathbb{R}$.
(4) $\mathbb{R}^n, \mathbb{C}^n$, $M_n(k) = \{$*all* $n \times n$ *matrices over* $k\}$.
(5) Graphs of $f(x) = x^2$.

Definition 1.3. *A Lie group G over $k = \mathbb{R}$ or $\mathbb{C}$ is a smooth manifold with a group structure $(G, \cdot, e)$ such that the two maps*

$$G \times G \to G, (g, h) \mapsto g \cdot h \quad \text{and} \quad G \to G, g \mapsto g^{-1}$$

are smooth.

Definition 1.4. *Suppose G, H are two Lie groups, we denote*

$$Hom(G, H) = \{\varphi : G \to H : \varphi \text{ is a group homomorphism that is smooth}\}.$$

Example 1.5 (Key example)**.** *We define the* ***general linear group***

$$GL_r(k) = \{\text{all } r \times r \text{ invertible matrices over } k\}.$$

Now we check it is a Lie group. Obviously, $G = GL_r(k)$ is a smooth manifold, so it suffices to check

$$G \times G \to G, (g, h) \mapsto gh, \quad G \to G, g \mapsto g^{-1}$$

are smooth.

Let $g = (g_{ij})$, $h = (h_{ij})$, then since addition and multiplication are infinitely differentiable, we know $(g, h) \mapsto gh$ is smooth. On the other hand, $(g^{-1})_{ij} = \frac{(j,i)-\text{cofactor of } g}{\det g}$, which is also infinitely differentiable.

In fact, the first four examples in Example 1.2 are all key examples of Lie groups.

Definition 1.6. *A Lie subgroup of G over k is a closed subgroup which itself is a Lie group.*

Example 1.7. *(1) $SL_r(\mathbb{R}) = \{g \in GL_r(\mathbb{R}) : \det(g) = 1\}$ is a Lie subgroup of $GL_r(\mathbb{R})$. $SL_r(\mathbb{R})$ is obviously closed subgroup of GL_r since* $\det$ *is a continuous map.*

(2) $\mathbb{R}^\times \overset{\varphi}{\hookrightarrow} GL_r(\mathbb{R}) \to GL_r(\mathbb{R})/\mathbb{R}^\times$, where $GL_r(\mathbb{R})/\mathbb{R}^\times := GL_r(\mathbb{R})/\varphi(\mathbb{R}^\times)$ and we denote it by $PGL_r(\mathbb{R})$.

1.2. **Constructions: Stablizer subgroup and group actions.** We consider the action of $GL_r(\mathbb{R})$ on $\mathbb{R}^r$, denoted by

$$GL_r(\mathbb{R}) \circlearrowleft \mathbb{R}^r.$$

We are interested in $G_v := \{g \in G : gv = v\}$, which is a subgroup of G for $v \in \mathbb{R}^r$. For $G = GL_r(\mathbb{R})$, one can check $G \to \mathbb{R}^r, g \mapsto gv - v$ is smooth, then G_v is closed. So G_v is a Lie subgroup of G. We say G_v is a stablizer subgroup of G.

It is also natural to have the action $GL_r(\mathbb{R}) \circlearrowleft (\mathbb{R}^*)^r$, where $V^* = Hom_{\mathbb{R}}(V, k)$ is the dual of V. The action is defined as follows. For $g \in GL_r(\mathbb{R}), f \in (\mathbb{R}^*)^r$, $g\dot{f} \in (\mathbb{R}^*)^r$ is defined by $(g \cdot f)(v) := f(g^{-1}v)$ for all $v \in \mathbb{R}^r$.

Though $\mathbb{R}^* \simeq \mathbb{R}$ as two vector spaces, we still consider these two different actions since there maybe more structures in one of them. Taking $\mathbb{R}^4 \simeq M_2(\mathbb{R})$ as an example. The matrix multiplication on $M_2(\mathbb{R})$ does not inherited in $\mathbb{R}^4$.

Example 1.8. *(1) For $v = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $G_v = GL_2(k)$.*

(2) For $v = (0 \ \vdots \ 0 \ \ 1)^t \in \mathbb{R}^{r+1}$, $G = GL_{r+1}(\mathbb{R})$, we have

$$G_v = \left\{ \begin{pmatrix} a & 0 \\ c & 1 \end{pmatrix} : a \in GL_r(\mathbb{R}), c \in \mathbb{R}^r \right\}.$$

Now we consider some extensions. Note that in the following examples, we will not focus on checking the smoothness of the actions rigorously, but only introduce these examples in an intuitive way.

1.2.1. *First Extension.*

Example 1.9. *We consider the action $GL_r(\mathbb{R}) \circlearrowleft \otimes_{i=1}^k \mathbb{R}^r$, which is defined naturally by*

$$g(v_1 \otimes \cdots \otimes v_k) = (g \cdot v_1) \otimes \cdots \otimes (g \cdot v_k).$$

Here the tensor product of two free product of free abelian group V, W is defined by

$$V \otimes W = \mathbb{R}[V \times W] \Big/ \left\langle \begin{matrix} (v_1 + v_2, w) - (v_1, w) - (v_2, w), \\ (av, w) - a(v, w), \\ (v, aw) - a(v, w), \\ (v, w_1 + w_2) - (v, w_1) - (v, w_2) \end{matrix} \right\rangle .$$

Note that $\dim(V \times W) = \dim V + \dim W, \dim(V \otimes W) = \dim V \cdot \dim W$. *Here* $V \times W$ *sometimes also writes as* $V \oplus W$.

Example 1.10. *We consider the action $GL_r(\mathbb{R}) \curvearrowright Sym^k(\mathbb{R}^r)$, where $Sym^k(\mathbb{R}^r)$ is defined by*

$$Sym^k(\mathbb{R}^r) := \otimes^k \mathbb{R}^r \Big/ \left\langle \begin{matrix} v_{m_1} \otimes \cdots \otimes v_{m_i} \otimes v_{m_j} \otimes \cdots \otimes v_{m_k} \\ -v_{m_1} \otimes \cdots \otimes v_{m_j} \otimes v_{m_i} \otimes \cdots \otimes v_{m_k} \end{matrix} \right\rangle .$$

The action on it is well-defined since $gN \subset N$ with $N = \{v_{m_1} \otimes \cdots \otimes v_{m_i} \otimes v_{m_j} \otimes \cdots \otimes v_{m_k} - v_{m_1} \otimes \cdots \otimes v_{m_j} \otimes v_{m_i} \otimes \cdots \otimes v_{m_k}\}$.

Example 1.11. *We consider the action $GL_r(\mathbb{R}) \curvearrowright \Lambda^k(\mathbb{R}^r)$, which is called the $k-$exterior of $\mathbb{R}^r$, where $\Lambda^k(\mathbb{R}^r)$ is defined by*

$$\Lambda^k(\mathbb{R}^r) := \otimes^k \mathbb{R}^r \Big/ \left\langle \begin{matrix} v_{m_1} \otimes \cdots \otimes v_{m_i} \otimes v_{m_j} \otimes \cdots \otimes v_{m_k} \\ +v_{m_1} \otimes \cdots \otimes v_{m_j} \otimes v_{m_i} \otimes \cdots \otimes v_{m_k} \end{matrix} \right\rangle .$$

The action is also well-defined.

Remark 1.12. When $k > r$, $\Lambda^k(\mathbb{R}^r) = \{0\}$. When $k = r$, $\dim \Lambda^r(\mathbb{R}^r) = 1$. Suppose $\Lambda^r(\mathbb{R}^r) = \mathbb{R} \cdot (e_1 \wedge e_2 \cdots \wedge e_r)$ where we denote $e^\sharp = e_1 \wedge e_2 \cdots \wedge e_r$. Then the action $GL_r(\mathbb{R}) \curvearrowright \Lambda^r(\mathbb{R}^r)$ is uniquely given by $ge^\sharp = (\det g)e^\sharp$.

We examine this property by showing the following example.

Example 1.13. *Let $r = 2$, then $G_{e^\sharp} = \{g \in GL_2(\mathbb{R}) : ge^\sharp = e^\sharp\}$, where $e^\sharp = e_1 \wedge e_2$ with $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then*

$$ge^\sharp = ge_1 \wedge ge_2 = (ae_1 + ce_2) \wedge (be_1 + de_2) = (ad - bc)e_1 \wedge e_2 = (\det g)e^\sharp.$$

Hence, $G_{e^\sharp} = SL_2(\mathbb{R})$.

Example 1.14. *It is easy to see $Sym^2(\mathbb{R}^2) = span\{e_1 \otimes e_2, e_1 \otimes e_1, e_2 \otimes e_2\}$, whose dimension is 3. Take $v = e_1 \otimes e_1 + e_2 \otimes e_2$, then we can check*

$$\begin{aligned} G_v =& \{g \in GL_2(\mathbb{R}) : gv = v\} \\ =& \{g \in GL_2(\mathbb{R}) : (ae_1 + ce_2) \otimes (ae_1 + ce_2) + (be_1 + de_2) \otimes (be_1 + de_2) = v\} \\ =& \{g \in GL_2(\mathbb{R}) : a^2 + b^2 = 1, c^2 + d^2 = 1, ac + bd = 0\} \\ =& \{g \in GL_2(\mathbb{R}) : g \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} g^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\} = O_2(\mathbb{R}), \end{aligned}$$

which is the orthogonal group.

Example 1.15. *In familiar language, let $B : k^n \times k^n \to k$ be a non-degenerate bilinear form, where $k = \mathbb{R}, \mathbb{C}$. Note that the non-degeneracy is equivalent to* $\det M_B \neq 0$, *where M_B is the representation matrix of B, that is, $B(v, w) = v^T M_B w$. If B is symmetric, that is, $B(x, y) = B(y, x)$, then*

$$O_B(k) = \{g \in GL_2(k) : B(gv, gw) = B(v, w), \forall v, w \in \mathbb{R}^2\} = \{g \in GL_2(k) : g^T M_B g = M_B\}.$$

If $M_B = I$, then we get the orthogonal group $O_n(k)$. If $k = \mathbb{R}$ with signature $(p, l), p + l = n$, that is, $M_B = \begin{pmatrix} I_p & 0 \\ 0 & -I_l \end{pmatrix}$, we write $O_B(\mathbb{R}) = O_{p,l}(\mathbb{R})$. Note that all the symmetric bilinear forms on $\mathbb{R}$ are classified by all pairs (p, l) with $p + l = n$, but the classification of bilinear forms on $\mathbb{Q}_p$ is difficult, see [14].

Example 1.16. *If B is anti-symmetric, that is, $B(x, y) = -B(y, x)$, then $n \in 2\mathbb{Z}$ and we can choose a suitable basis such that $M_B = \begin{pmatrix} 0 & I_{\frac{n}{2}} \\ -I_{\frac{n}{2}} & 0 \end{pmatrix}$. The symplectic group $Sp_n(\mathbb{R}) = \{g \in GL_n(\mathbb{R}) : g^T M_B g = M_B\}$. When $n = 2$, we have $Sp_2(\mathbb{R}) = SL_2(\mathbb{R})$.*

Remark 1.17. The bilinear forms in the two examples above are related to the notions introduced before. In fact, $Sym^2((\mathbb{R}^n)^*) = span\{f_i \otimes f_j : i \leq j\}$ are the symmetric bilinear forms and $\Lambda^2((\mathbb{R}^n)^*) = span\{f_i \wedge f_j : i < j\}$ are the anti-symmetric bilinear forms.

From these examples, we find many classical lienar groups are some kind of stabilizer group G_v, which is a Lie subgroup of $GL_r(k)$.

Now we try to make an extension to quotients(cosets).

1.2.2. *Second Extension.* We consider the action $GL_r(\mathbb{R}) \circlearrowright GL_r(\mathbb{R})/H$, where H is a subgroup.

If $H = \{1\}$, then it is a trivial case. There are many ways to define the action, such as the conjugation, that is, $GL_r(\mathbb{R}) \circlearrowright GL_r(\mathbb{R})$, $g \cdot x = gxg^{-1}$.

Here comes a non-trivial example.

Example 1.18. *Let*

$$B_r(\mathbb{R}) = \left\{ \begin{pmatrix} a_1 & * & \cdots & * \\ 0 & a_2 & \cdots & \vdots \\ 0 & 0 & \ddots & * \\ 0 & \cdots & 0 & a_r \end{pmatrix} \right\}$$

be the Borel subgroup. We want to construct some space V such that $B_r(\mathbb{R}) = G_v$ for some $v \in V(\mathbb{R}^r)$.

We define

$$CFlag(\mathbb{R}^r) = \{(V_0, V_1, \cdots, V_r) : V_i \subset \mathbb{R}^r, V_i \subset V_{i+1}, \dim V_i = i\}$$

consisting of chains of subspaces, which is a set. The action $GL_r(\mathbb{R}) \circlearrowright CFlag(\mathbb{R}^r)$ is defined naturally by $(V_0, \cdots, V_r) \mapsto (gV_0, \cdots, gV_r)$. Let $C_0 = (0, \mathbb{R}, \mathbb{R}^2, \cdots, \mathbb{R}^r)$, then one can observe that the action is transitive since $\forall C \in CFlag(\mathbb{R}^r)$, there exists $g \in GL_r(\mathbb{R})$ such that $gC_0 = C$ by a direct computation. Moreover, we have

$$G_{C_0} = \{g : gC_0 = C_0\} = B_r(\mathbb{R})$$

since $ge_1 = e_1$ implies the first column of g is in the form of $(a_1, 0, \cdots, 0)^t$, then by induction we can check that $G_{C_0} = B_r(\mathbb{R})$.

And we have

$$N = \left\{ \begin{pmatrix} 1 & * & \cdots & * \\ 0 & 1 & \cdots & \vdots \\ 0 & 0 & \ddots & * \\ 0 & \cdots & 0 & 1 \end{pmatrix} \right\} = \{g \in B_r(\mathbb{R}) : g : \mathbb{R}^{i+1}/\mathbb{R}^i \to \mathbb{R}^{i+1}/\mathbb{R}^i \textit{ is identity } \}.$$

Hence, we have

$$CFlag(\mathbb{R}^r) = GL_r(\mathbb{R})/B_r(\mathbb{R})$$

as cosets where $C = gC_0 \mapsto [g] = gB_r(\mathbb{R}) = gG_{C_0}$ is an isomorphism between these two vector spaces.

For more discussion, see [11, Example 21.22].

Definition 1.19. *Let G be a Lie group. A G-homogeneous space is a topological space X together with a continuous map (aka. action)*

$$\varphi : G \times X \to X, g \cdot x := \varphi(g, x)$$

such that

(1) $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$,
(2) for all $x \in X$, $\varphi_x : G \to X$, $g \mapsto g \cdot x$ is surjective, that is, the action only has one orbit, that is, the action is transitive.

The Homogeneous Space Characterization Theorem [11, Theorem 21.18] tells us

$$F : G/G_x \to X, F(gG_x) = g \cdot x$$

is a diffeomorphism if we assume φ is smooth in the definition of G-homogeneous space. Here, we can only conclude that $G/G_x \to X$ is a homeomorphism.

Example 1.20. *Note that $CFlag(\mathbb{R}^r)$ is endowed with a transitive continuous action by $GL_r(\mathbb{R})$ thanks to the discussion before. So $CFlag(\mathbb{R}^r)$ is a $GL_r(\mathbb{R})$-homogeneous space.*

Example 1.21. *Let $\mathcal{H} = \{z = x + iy \in \mathbb{C} : y > 0\}$ be the Poincare half space. Then we can define the action $SL_2(\mathbb{R}) \subsetneq \mathcal{H}$ by*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}.$$

One can check that

(1) This action is transitive.
(2) The stablizer group

$$Stab_{SL_2(\mathbb{R})}(i) = \{g \in SL_2(\mathbb{R}) : gi = i\} = \{(\begin{matrix} a & b \\ -b & a \end{matrix}) \in SL_2(\mathbb{R})\} = SO_2(\mathbb{R}).$$

1.3. **Notations and Definitions: Tangent space and Lie bracket.** We introduce some notations before discussing exponential maps. Let $\varphi = (\varphi_1, \cdots, \varphi_m) : k^n \to k^m$, then for $a = (a_1, \cdots, a_n) \in k^n$, we denote $D(\varphi)(a) = \left(\frac{\partial \varphi_i}{\partial x_j}\right)(a)_{1 \le i \le m, 1 \le j \le n}$, which is a $m \times n$ matrix. For any $v = (v_1, \cdots, v_n)^T \in k^n$, $D_v(\varphi)(a) = (D\varphi(a)) \cdot v$, which is the directional derivative.

Definition 1.22 (Tangent space)**.** *We define*

$$T_xG = \{(c,v) : v \in \mathbb{R}^n \text{ and } (U,c) \text{ is a chart around } p,$$
$$\text{where } c : U \to c(U) \subset k^n \text{ is an homeomorphism}\}/\sim,$$

where $(c_1, v_1) \sim (c_2, v_2)$ *is the composition map* $c_2 \circ c_1^{-1} : c_1(U_1 \cap U_2) \to c_2(U_1 \cap U_2)$ *satisfies* $v_2 = \left(D(c_2 \circ c_1^{-1})(c_1(p))\right)^T(v_1)$.

Remark 1.23. The definition is equivalent to the usual definition. Take (U_1, c_1) with $(x^1, \cdots, x^n)$ and (U_2, c_2) with $(y^1, \cdots, y^n)$, then suppose $v_1 = (a^1, \cdots, a^n)^T$, $v_2 = (b^1, \cdots, b^n)^T$, that is to say, for all $f \in C^\infty(M)$, $v_1(f) = a_i \partial_{x_i} f$ and $v_2(f) = b_j \partial_{y_j} f$ Hence, $(c_1, v_1) \sim (c_2, v_2)$ is equivalent to $a_i \partial_{x_i} f = b_j \partial_{y_j} f$. By Lebniz rule,

$$\begin{pmatrix} \partial_{x_1} f \\ \vdots \\ \partial_{x_n} f \end{pmatrix} = D\left(c_2 \circ c_1^{-1}\right)(c_1(p)) \begin{pmatrix} \partial_{y_1} f \\ \vdots \\ \partial_{y_n} f \end{pmatrix},$$

by multiplying $(a_1, \cdots, a_n)$ to the left of both sides and note $a_i \partial_{x_i} f = b_j \partial_{y_j} f$, this is equivalent to $v_2 = \left(D(c_2 \circ c_1^{-1})(c_1(p))\right)^T(v_1)$.

In the following discussion, we fix a chart (U, c) with $c : x \in U \to k^n$ for simplicity. Given c, we get a bijection

$$k^n \xrightarrow{\theta_c} T_xG, \quad v \mapsto [(c, v)].$$

If $f : G \to H$ is a Lie group homomorphism, then you get

$$T(f)_x : T_xG \to T_{f(x)}H$$

with

$$\theta_{c_1}(v) \to \theta_{c_2}\left(D_v\left(c_2 \circ f \circ c_1^{-1}\right)(c_1(x))\right),$$

where $v \in k^n$.

For simplicity, it is OK to assume $\theta_c = id$, that is, we identify $T_xG = k^n$.

Let $\gamma : \mathbb{R} \to G$ be a Lie group homomorphism such that $\gamma(0) = e$, then

$$T(\gamma) : T_0(\mathbb{R}) = \mathbb{R} \to T_e(G), \quad v \mapsto D_v(c \circ \gamma)(0) = (f_1'(0), \cdots, f_n'(0)) \cdot v,$$

where $c \circ \gamma = (f_1, \cdots, f_n)$.

Example 1.24. *Here are some key examples.*

(1) Consider the Lie group homomorphism $Ad_g : G \to G$, $x \mapsto gxg^{-1}$, *you get, with a slight abuse of notation,* $Ad_g := T(Ad_g)_e : T_e(G) \to T_e(G)$.

(2) Moreover, as the construction above, you get $Ad : G \to Hom(T_e(G)) = GL(T_e(G)) \approx GL_r(k)$, $g \mapsto Ad_g$, *which is a Lie group homomorphism. Note that we are allowed to switch notations between* $GL(V)$ *and* $GL_r(k)$, *between* $End(V)$ *and* $M_r(k)$.

(3) Now we take one step further. You get $ad = T(Ad)_e : T_e(G) \to T_e(GL(T_e(G))) = End(T_e(G))$. *Here we view the identity map as* $GL_r(k) \hookrightarrow M_r(k) = T_e(GL_r(k))$, *and we switch notations to get* $T_e(GL(T_e(G))) = End(T_e(G))$.

Definition 1.25 (Lie bracket)**.** *For a Lie group G, we define*

$$[\cdot,\cdot] : T_e(G) \times T_e(G) \to T_e(G)$$

with $[X, Y] := ad(X)(Y)$, which is well-defined since $ad(X) \in End(T_e(G))$.

1.4. **The exponential map for matrices.** Formally, we define

$$\exp : M_r(k) \to GL_r(k), \quad A \mapsto \exp(A) := \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

And we denote the operator norm of $A \in M_r(k)$ by $|A| = \sup_{|v|=1}\{|Av|\} = \sup_{v\neq 0}\left\{\frac{|Av|}{|v|}\right\}$.

Lemma 1.26. *(1) $|AB| \leq |A||B|$.*

(2) $\exp$ is absolutely convergent on any set $\{A : |A| < p\}$, $p \in \mathbb{R}_+$.

(3) The function $\log(1 - A) := \sum_{n=1}^{\infty} \frac{A^n}{n}$ is absolutely convergent on $\{A : |A| < \varepsilon\}$ for all $0 \leq \varepsilon < 1$.

Proof. The first inequality is immediate if we write $|AB(v)| = |B(v)| \cdot \frac{|AB(v)|}{|B(v)|}$. And the other two follows from the first inequality immediately. □

Proposition 1.27. *Here we present that exp and $\log$ are "locally invertible".*

(1) There exists an open set $U_{I_r} \subset GL_r(k)$ such that $\exp\circ\log A = A$ for $A \in U_{I_r}$.

(2) There exists an open set $V_0 \subset M_r(k)$ such that $\log\circ\exp A = A$ for $A \in V_0$.

Proof. We only present the proof of the first property. It suffices to prove for $k = \mathbb{C}$ and we want to show $\exp\circ\log A = A$ for $|A - I_r| < 1$.

Take $A \in GL_r(\mathbb{C})$, if $A = C \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_r \end{pmatrix} C^{-1}$, then $A - I_r = C \begin{pmatrix} z_1 - 1 & & \\ & \ddots & \\ & & z_r - 1 \end{pmatrix} C^{-1}$.

Here $\lambda = z_i - 1$ are eigenvalues, so there exists v such that $\lambda = \frac{|Av|}{|v|}$, which implies $|z_i - 1| \leq |A - I_r| < 1$, then $\log A = -\sum_{n=1}^{\infty} \frac{(I-A)^n}{n} = C \begin{pmatrix} \log(z_1) & & \\ & \ddots & \\ & & \log(z_r) \end{pmatrix} C^{-1}$.

So we have $\exp\circ\log A = A$.

For general A, we use density argument. Since $\exp, \log$ are continuous, it suffices to show that diagonalizable matrices are dense. Let

$$\mathcal{S} = \{A : A \text{ has distinct eigenvalues}\} \subset \{\text{all diagonalizable matrices}\} \subset M_r(\mathbb{C}),$$

so it suffices to prove $\mathcal{S}$ is dense in $M_r(\mathbb{C})$. The discriminant is $Disc(A) := \Pi_{i<j}(r_i - r_j)^2$, where r_i is the root of $\det(xI - A)$. Note that $Disc(A)$ is a polynomial in $\{\alpha_n\}$, which in fact a polynomial in a_{ij}'s. Since $S = Disc^{-1}(\mathbb{C} \setminus \{0\})$ and $Disc$ is a polynomial, so the inverse image S is dense. For the discriminant part, see [13, Section 2.7] or [4] for more details. □

Example 1.28. *Let $G = GL_n(k)$ with $T_e(G) = M_n(k)$. For $X, Y \in T_e(G)$, set $\gamma(t) = \exp(tX)$, then $\gamma(0) = e = I_r$, $\gamma'(0) = X$. For any $Z \in T_e(G)$, we have $Ad(Z)(Y) =$*

$T(Ad_Z)_e(Y) = D_Y(Ad_Z)(e) = (Ad_Z(\gamma_Y(s)))'\,|_{s=0} = (Z\gamma_Y(s)Z^{-1})'_{s=0}$, so by definition of $ad = T(Ad)_e$, we have

$$ad(X)(Y) = \left((\gamma_X(t)\gamma_Y(s)\gamma_X(t)^{-1})'|_{s=0}\right)|_{t=0} = XY - YX.$$

Or we can compute directly that

$$\begin{aligned}[X,Y] := ad(X)(Y) &= (T(Ad)_e(X))(Y) = (Ad\circ\gamma(t))'\,|_{t=0}(Y) = ((Ad\circ\gamma(t))(Y))'\,|_{t=0}\\ &= (Ad(\gamma(t))(Y))'\,|_{t=0} = \left(T(Ad_{\gamma(t)})_e(Y)\right)'|_{t=0} = \left(D_Y(Ad_{\gamma(t)})(e)\right)'|_{t=0}\\ &= \left(\left(Ad_{\gamma(t)}(\gamma_Y(s))\right)'|_{s=0}\right)'|_{t=0} = \left(\gamma(t)Y\gamma(t)^{-1}\right)'|_{t=0} = XY - YX,\end{aligned}$$

where $\gamma_Y(s) = \exp(sY)$.

Recall that $T(f)_e([X,Y]) = [T(f)_e(X), T(f)_e(Y)]$, then we have the following lemma.

Lemma 1.29. *For $ad = T(Ad)_e : T_e(G) \to End(T_e(G))$, we have $ad[A,B] = [adA, adB]$.*

Proposition 1.30. *For a Lie group G, $A, X, Y, Z \in T_e(G)$, we have*
(1) $[A,A] = 0$;
(2) the Jacobi identity $[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0$.

Proof. We calculate

$$\begin{aligned}[A,A] = ad(A)(A) = (Ad(\gamma(t))(A))'\,|_{t=0} &= \frac{\partial\left(Ad(\gamma(t))\frac{\partial\gamma_A(s)}{\partial s}\right)}{\partial t}|_{s=0,t=0}\\ &= \frac{\partial^2\left(\gamma_A(t)\gamma_A(s)\gamma_A(t)^{-1}\right)}{\partial t\partial s}|_{s=0,t=0} = \frac{\partial^2\gamma_A(s)}{\partial t\partial s}|_{s=0,t=0} = 0.\end{aligned}$$

For the Jacobi identity, we verify

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

□

1.5. **One parameter subgroup and The exponential map for a Lie group.**

Theorem 1.31. *Let $Hom(\mathbb{R}, G)$ denote all the Lie group homomorphism between $\mathbb{R}$ and G. Suppose the map $Hom(\mathbb{R}, G) \to T_e(G)$ given by $\gamma \mapsto \gamma'(0)$ is a bijection. And for $A \in T_e(G) = k^n$, we denote the inverse by γ_A.*

Proof. For $A \in T_e(G) = k^n$, one has the vector field $\mathcal{F}_A$ satisfing $\mathcal{F}_A(g) := T(R_g)_e(A)$ with $R_g : G \to G$ denote the right multiplication $x \mapsto x \cdot g$.

The ODE theories tells us there exists a unique differentiable map $\gamma : (-\varepsilon, \varepsilon) \to G$ such that

$$\gamma(0) = e, (c\circ\gamma)'(t) = \mathcal{F}_A(\gamma(t)).$$

And one can show that $\gamma(s+t) = \gamma(s)\gamma(t)$ for $s, t, s+t \in (-\varepsilon, \varepsilon)$ by the uniqueness of this differentiable map. Indeed, fix s, the two map $t \mapsto c(\gamma(t+s))$ and $t \mapsto c(\gamma(t)\gamma(s))$ are both solutions for $\mathcal{F}$ around $\gamma(s)$, so by uniqueness, the result follows.

Now, since $(-\varepsilon, \varepsilon)$ generates $\mathbb{R}$ by the obvious decomposition for all $T \in \mathbb{R}$ that $T = nt + s$ with $n \in \mathbb{Z}, t, s \in (-\varepsilon, \varepsilon)$. So there exists an extension $\gamma_A : \mathbb{R} \to G$ of γ. □

Corollary 1.32. *Suppose $A \in T_e(G)$, $a, b \neq 0$, then $\gamma_{aA}(b) = \gamma_A(ab)$.*

Definition 1.33. *The exponential map* $\exp : T_eG \to G$ *is defined by* $\exp(A) = \gamma_A(1)$.

1.6. **Properties of the exponential map.** We followed [6, Section 8.3] in this part.

Lemma 1.34. *Suppose φ is a Lie group homomorphism, then the diagram*

$$\begin{tikzcd} T_eG \arrow[r, "T(\varphi)_e"] \arrow[d, "\exp"'] & T_eH \arrow[d, "\exp"] \\ G \arrow[r, "\varphi"] & H \end{tikzcd}$$

commutes.

Proof. We claim

$$\gamma_{(T(\varphi)_e)(X)}(t) = \varphi \circ \gamma_X(t). \tag{1.1}$$

This is because both are local solutions to

$$\begin{cases} \beta \in Hom(\mathbb{R}, H), \\ \beta(0) = e_H, \\ (c \circ \beta)'(0) = T(\varphi)_e(X). \end{cases}$$

However, this ODE admits a unique local solution, so we have verified (1.1) locally(in a neighbourhood (lying in a local chart) of e). Then since $\varphi \circ \gamma_X([0,1])$ and $\gamma_{(T(\varphi)_e)(X)}([0,1])$ are both compact, we can choose a finite local charts covering, then apply the local uniqueness result to each chart. Hence, (1.1) holds for $t \in [0,1]$. In particular, this lemma follows by taking $t = 1$ in (1.1). ☐

Theorem 1.35. *If G is connected, then the homomorphism $\varphi : G \to H$ is determined by $T(\varphi)_e$.*

Proof. **Step 1:** Note that $T_e(G)$ is a Lie group as well. For the map

$$\exp : T_e(G) \to G,$$

we claim $T(\exp)_0 = id$. (Smoothness of exp is nontrivial, c.f. [11].) Indeed,

$$T_e(G) \ni A \overset{T(\exp)_0}{\mapsto} D_A(c \circ \exp)(0),$$

with

$$D_A(c \circ \exp)(0) = (c \circ \exp(tA))'|_{t=0} = (c \circ \gamma_{tA}(1))'|_{t=0} = (c \circ \gamma_A(t))'|_{t=0} = A.$$

Now, since $\det(T(\exp)_0)(0) = 1 \neq 0$, by the inverse function theorem, exp is a local diffeomorphism near 0. Thus, the image of exp contains an open set $U \subset G$ such that $e \in U$, and $\exp : N \to U$ is a diffeomorphism where $0 \in N$.

Step 2: Now we claim $\langle U \rangle = G$, where $\langle U \rangle$ denotes the group generated by the elements in U. Firstly, for all $g \in \langle U \rangle$, we have $g \cdot U$ is an open set containing g. Moreover, $g \cdot U \subset \langle U \rangle$, so $\langle U \rangle$ is open in G.

Secondly, there exists $\{g_k\}_{k \in I}$ such that $\{g_k \cdot \langle U \rangle\}_{k \in I}$ is a family of disjoint cosets with $g_0 = e$. Then $\langle U \rangle = G \setminus \cup_{k \in I, k \neq 0} (g_k \cdot \langle U \rangle)$, which is closed in G.

Step 3: By the connectedness of G, we proved $\langle U\rangle = G$. Then φ is uniquely determined by $\varphi|_U$. Hence, by Lemma 1.34, $\varphi \circ \exp_G = \exp_H \circ T(\varphi)_e$, we know for $g \in U$, we know $\varphi(g) = \exp_H \circ T(\varphi)_e(\exp_G^{-1}(g))$, which is well-defined since $\exp_G : N \to U$ is a diffeomorphism thanks to Step 1. Hence, φ is uniquely determined by $T(\varphi)_e$. $\square$

2. Basic concepts of Lie algebras

We will introduce some classical examples in a somewhat intuitively way in this section.

2.1. Definitions and Constructions.

Definition 2.1. *A Lie algebra over k is a k-vector space $\mathfrak{g}$ endowed with a map $[-,-] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that for $X, Y, Z \in \mathfrak{g}$,*
(1) $[-,-]$ is k-bilinear;
(2) $[-,-]$ is anti-symmetric, that is, $[X,Y] = -[Y,X]$;
(3) $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$.

Then, for each Lie group G, we know that $\mathfrak{g} = T_e(G)$ with $[X,Y] = ad(X)(Y)$ is a Lie algebra.

Example 2.2. *Note that we have checked $(T_e(GL_r(k)), [X,Y] = XY - YX)$ is a Lie algebra. We denote this Lie algebra by $\mathfrak{gl}_r(k) = T_e(GL_r(k)) = M_r(k)$.*

Example 2.3. *Let "$\mathfrak{sl}_2(k)$"$= \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} : a + d = 0\} \subset \mathfrak{gl}_2(k)$, which is of dimension 3. One can check $X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ form a basis and $[H,X] = 2X, [H,Y] = (-2)Y, [X,Y] = H$. Once we knew $[A,B]$ between each basis, the Lie algebra is classified.*

We can check it is not isomorphic to the Lie algebra $n_3 = \left\{ \begin{pmatrix} 0 & x & z \\ & 0 & y \\ & & 0 \end{pmatrix} \right\}$.

Definition 2.4. *Let $\mathfrak{g}_1, \mathfrak{g}_2$ be Lie algebra over $k = \mathbb{R}$ or $\mathbb{C}$. A Lie algebra homomorphism is a k-linear map $f : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that $f([X,Y]) = [f(X), f(Y)]$.*

Moreover, if $\mathfrak{g}_2 = End_k(V)$ is a Lie algebra with $[T_1, T_2] := T_1 \circ T_2 - T_2 \circ T_1$, then such an f is called a representation of g_1 in V.

And we define

$$Rep(\mathfrak{g}) := \{(f, V) \mid f : \mathfrak{g} \to End(V) \text{ is a Lie algebra homomorphism}\}.$$

We consider the action $G \circlearrowleft V_i$, $i = 1, 2$ with $G \subset GL_r(k)$ and see from a viewpoint of group: Suppose we have Lie group homomorphisms $\rho_i : G \to GL(V_i)$, $i = 1, 2$. Then we get $\rho = \rho_1 \otimes \rho_2 : G \to GL(V_1 \otimes V_2)$. Respectively, we get $(\rho_i)_* = T(\rho_i)_e : \mathfrak{g} \to \mathfrak{gl}(V_i)$, $(\rho_1 \otimes \rho_2)_* = T(\rho_i)_e : \mathfrak{g} \to \mathfrak{gl}(V_1 \otimes V_2)$.

For $X \in \mathfrak{g}$, we have

$$(\rho_i)_*(X)(v_i) = D(\rho_i \circ \gamma_X)(0)(v_i) = ((\rho_i \circ \gamma_X(t))(v_i))' |_{t=0},$$

then this implies

$$\begin{aligned}(\rho_1\otimes\rho_2)_*(X)(v_1\otimes v_2) &= \left(\left((\rho_1\otimes\rho_2)\circ\gamma_X(t)\right)(v_1\otimes v_2)\right)'|_{t=0}\\ &= ((\rho_1\circ\gamma_X(t))(v_1))'|_{t=0}\otimes v_2 + v_1\otimes((\rho_2\circ\gamma_X(t))(v_2))'|_{t=0}\\ &= ((\rho_1)_*(X)(v_1))\otimes v_2 + v_1\otimes((\rho_2)_*(X)(v_2)),\end{aligned}\tag{2.1}$$

where we use the product rule in the second equality.

Now, put $v_1\otimes v_2\in V_1\otimes V_2$, we consider the stabilizer group $G_{v_1\otimes v_2}\subset G$ with the obvious action $G\subsetneq V_1\otimes V_2$ defined by $v\mapsto(\rho_1\otimes\rho_2)(g)(v_1\otimes v_2)$ for all $g\in G$. Let $X\in\mathfrak{g}_{v_1\otimes v_2}=T_e(G_{v_1\otimes v_2})$, then $\gamma_X(t)\subset G_{v_1\otimes v_2}$, which implies that $\left((\rho_1\otimes\rho_2)\circ\gamma_X(t)\right)(v_1\otimes v_2)=v_1\otimes v_2$ is a constant, so (2.1) is zero, that is,

$$((\rho_1)_*(X)(v_1))\otimes v_2 + v_1\otimes((\rho_2)_*(X)(v_2))=0.$$

Conversely, if the right hand side of (2.1) is zero, then $X\in\mathfrak{g}_{v_1\otimes v_2}$ by reversing the argument above.

For $\gamma_X(t)=\exp(tX)$, with $X\in\mathfrak{g}=T_e(G_v)$, $G_v\subset GL_r(k)$, we have $\gamma_X'(t)=\exp(tX)X$. Recall that $\exp(tX)$ is invertible near 0, so if $\gamma_X'(t)(v)=0$, then $Xv=0$, thus $\exp(tX)v=v$ for t small by the definition of the exponential of matrices. (Note that the convergence is in operator norm, so $\exp(tX)v:=\sum_{n=0}^{\infty}\frac{(tX)^n v}{n!}=v$.) Hence,

$$T_e(G_v)=\{X\in\mathfrak{gl}_r : Xv=0\}.\tag{2.2}$$

Now let us see some examples as the application of the procedure above.

Example 2.5. *Consider $G=GL_r\subsetneq V^*\otimes V^*=B(V)$, which consists the bilinear forms on V. Here $V_1=V_2=V^*$. Put $B=b^*=\sum\alpha_{ij}e_i^*\otimes e_j^*$, which corresponds to the matrix $M_B=[\alpha_{ij}]$.*

Now we consider $T_e(G_{b^*})$. For all $X\in T_e(G_{b^*})$, we have $\gamma_X(t)\subset G_{b^*}$, then $\gamma_X(t)b^*=b^*$, that is, for all $v,w\in V$, we have

$$\sum\alpha_{ij}e_i^*(\gamma_X(t)(v))e_j^*(\gamma_X(t)(w))=\sum\alpha_{ij}e_i^*(v)e_j^*(w)\Rightarrow B(\gamma_X(t)v,\gamma_X(t)w)=B(v,w),$$

then one can apply (2.1) to take the derivatives and let $t=0$, we get

$$B(X(v),w)+B(v,X(w))=0 \iff X^TM_B+M_BX=0.$$

Then using (2.2) with v replaced by b^* or use the same type of argument when deriving (2.2), we get $B(\exp(tX)(v),\exp(tX)w)=B(v,w)$ thanks to the bilinearity of B and the expansion of $\exp(tX)$. Finally, we have

$$T_e(G_B)=\{X\in\mathfrak{gl}_r : B(X(v),w)+B(v,X(w))=0\}.$$

Example 2.6. *From Example 1.15 and Example 2.5, we know that*

$$\mathfrak{o}_r(k)=T_e(O_r(k))=\{X\in\mathfrak{gl}_r(k) : X^T+X=0\}=\{\textit{skew symmetric matrices in } M_r\}.$$

Example 2.7. *We have*

$$\mathfrak{sp}_{2r}(k)=\left\{X=\begin{pmatrix}A & B\\ C & -A^t\end{pmatrix}\in M_r(k) : B=B^t, C=C^t\right\},$$

which is related to $M_B=\begin{pmatrix}0 & I_r\\ -I_r & 0\end{pmatrix}$.

Example 2.8. *Consider $GL_r \subsetneq \Lambda^r(V) = k \cdot (e_1 \wedge \cdots \wedge e_r)$ and denote $e_1 \wedge \cdots \wedge e_r$ by $e^\sharp$. Since $SL_r(k) = \{g \in GL_r(k) : ge^\sharp = e^\sharp\}$, we get*

$$\mathfrak{sl}_r(k) = T_e(SL_r(k)) = \{X \in \mathfrak{gl}_r(k) : Xe^\sharp = 0\}.$$

We calculate

$$Xe^\sharp = \sum_{i=1}^{r} e_1 \wedge e_2 \wedge \cdots \wedge X(e_i) \wedge \cdots e_r = Tr(X)e^\sharp,$$

so $\mathfrak{sl}_r(k) = \{X \in \mathfrak{gl}_r(k) : Tr(X) = 0\}$.

2.2. **Ideals, Quotients, Solvablity and Nilpotency.**

Definition 2.9. *An ideal $\mathfrak{h} \subset \mathfrak{g}$ is a Lie subalgebra such that $[X, Y] \in \mathfrak{h}$ for all $X \in \mathfrak{h}$, $Y \in \mathfrak{g}$.*

Definition 2.10. *The center of $\mathfrak{g}$ is defined as*

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

which is also an ideal of $\mathfrak{g}$.

We say $\mathfrak{g}$ is abelian if $Z(\mathfrak{g}) = \mathfrak{g}$.

One can easily verify that if $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then $(\mathfrak{g}/\mathfrak{h}, [-,-])$ with $[-,-] : \mathfrak{g}/\mathfrak{h} \times \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$ is a well-defined Lie algebra, which is called the quotient Lie algebra. Here, the Lie bracket $\mathfrak{g}/\mathfrak{h} \times \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$ is defined by $[\bar{X}, \bar{Y}] := \overline{[X, Y]}$.

Definition 2.11. *A Lie algebra $\mathfrak{g}$ is called simple if $\dim \mathfrak{g} \geq 2$ and all the ideals of $\mathfrak{g}$ are just 0 and $\mathfrak{g}$.*

From the definition, we see that $\mathfrak{g} = k^1$ is not considered as simple. One can check $\mathfrak{sl}_2(k)$ is simple.

Definition 2.12. *We define the upper central series as*

$$\mathfrak{g} = \mathfrak{g}^0 \supset [\mathfrak{g}^0, \mathfrak{g}^0](:= \mathfrak{g}^1) \supset [\mathfrak{g}^1, \mathfrak{g}^1](:= \mathfrak{g}^2) \supset \cdots$$

and the lower central series as

$$\mathfrak{g} = \mathfrak{g}_0 \supset [\mathfrak{g}, \mathfrak{g}_0](:= \mathfrak{g}_1) \supset [\mathfrak{g}, \mathfrak{g}_1](:= \mathfrak{g}_2) \supset \cdots .$$

Definition 2.13. *We say $\mathfrak{g}$ is solvable if $\mathfrak{g}^i = 0$ for some i and $\mathfrak{g}$ is nilpotent if $\mathfrak{g}_j = 0$ for some j.*

Example 2.14. *One can check $\mathfrak{b}_3 = \left\{ \begin{pmatrix} a & d & e \\ & b & f \\ & & c \end{pmatrix} \right\}$ is solvable and $\mathfrak{n}_3 = \left\{ \begin{pmatrix} 0 & d & e \\ & 0 & f \\ & & 0 \end{pmatrix} \right\}$ is nilpotent.*

Now we state some properties as a lemma with a sketch of proof.

Lemma 2.15. *Here are some properties for solvable and nilpotent Lie algebras.*

(1) If $\mathfrak{g}$ is solvable (or nilpotent), then any sub-Lie algebra and quotient Lie-algebra (quotient by an ideal) is also solvable (or niplotent).

(2) Suppose $\mathfrak{h}$ is an ideal. If $0 \to \mathfrak{h} \hookrightarrow \mathfrak{g} \to \mathfrak{g}/\mathfrak{h} \to 0$ is a exact sequence, and $\mathfrak{h}, \mathfrak{g}/\mathfrak{h}$ are solvable, then $\mathfrak{g}$ is also solvable.

(3) If $0 \to Z(\mathfrak{g}) \hookrightarrow \mathfrak{g} \to \mathfrak{g}/Z(\mathfrak{g}) \to 0$ is a exact sequence, and $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent, then g is also nilpotent.
(4) If $\mathfrak{g}$ is nilpotent, $\mathfrak{g} \neq 0$, then $Z(\mathfrak{g}) \neq 0$.
(5) Suppose $\mathfrak{h}_1, \mathfrak{h}_2 \subset \mathfrak{g}$ are solvable ideals, then $\mathfrak{h}_1 + \mathfrak{h}_2$ is also a solvable ideal.

Proof. The first property is direct by using definitions. Note that the first part implies that for any quotient Lie algebra $\mathfrak{g}/\mathfrak{h}$, $\mathfrak{h}$ is solvable since it is an ideal. To prove the fifth property, one should notice that $\mathfrak{h}_1/\mathfrak{h}_1 \cap \mathfrak{h}_2$ is isomorphic to $(\mathfrak{h}_1 + \mathfrak{h}_2)/\mathfrak{h}_2$, and $\mathfrak{h}_1/\mathfrak{h}_1 \cap \mathfrak{h}_2$ is solvable by the first property, then $\mathfrak{h}_1 + \mathfrak{h}_2$ is solvable thanks to the second property. The other three properties follow from the Exercise 5 in HW 3([6, Exercise 9.8]). □

3. Engel's theorem and Lie's theorem

We will rigorously develop the classical theories for Lie algebras in the remaining sections.

Intuitively speaking, Engel's theorem states that every nilpotent algebra $\mathfrak{g}$ is embedded in $\mathfrak{n}_r(k)$, where embedding means an injective Lie algebra homomorphism. The Lie's theorem states that every solvable algebra $\mathfrak{g}$ is embedded in $\mathfrak{b}_r(k)$ provided that k is algebraic closed. That is, Lie's theorem doesn't hold for $k = \mathbb{R}$ but for $k = \mathbb{C}$.

We follow [6, Chapter 9] in this section. In the following context, we only consider finite dimensional Lie algebras unless otherwise specified.

Theorem 3.1 (Engel's theorem)**.** *Suppose V is a finite dimensional vector space. Let $\mathfrak{g} \subset gl(V) = End(V)$ be a Lie subalgebra such that every $X \in \mathfrak{g}$ is a nilpotent element, that is, $X^n = 0$ for some $n = n_X \in \mathbb{N}$. Then*
(1) There exists $0 \neq v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{g}$.
(2) There exists a basis $\{e_1, \cdots, e_r\}$ of V such that $\mathfrak{g} \subset \mathfrak{n}_r$ with respect to $\{e_i\}$.

Proof. We prove the second part at first. Choose $v \in V$ using the first part of theorem, then we consider $V' = V/\{k \cdot v\}$. Since $\mathfrak{g}$ kills v, the action $\mathfrak{g} \circlearrowright V'$ is well-defined which is still nilpotent. So we can apply the same argument to $\mathfrak{g} \subset gl(V')$, to get $v_2' \in V/k \cdot v$. Inductively, we get $v_2, \cdots v_r$ by lifting v_i' to $v_i \in V$. Setting $v_1 = v$, then we get a basis $\{v_1, \cdots, v_r\}$ of V and $\mathfrak{g}v_k$ is a linear combination of $v_1, \cdots, v_{k-1}$, that is, $\mathfrak{g} \subset \mathfrak{n}_r$ with respect to $\{v_i\}$.

Now we turn to the proof of the first part.

Step 1: We claim that $ad(X) \in End(gl(V))$ is nilpotent. Since $X \in \mathfrak{g}$ is nilpotent, then by definition, there exists a sequence of subspaces $0 \subset V_1 \subset V_2 \subset \cdots \subset V_r = V$ such that $X(V_i) \subset V_{i-1}$. It is easy to check that $ad(X)^m(Y)$ is a linear combination of terms like $X^l Y X^k$ with $l + k = m$, so $ad(X)^m = 0$ for large m, say, $m > 2r + 2$.

Step 2: We know proceed by induction on dimension of $\mathfrak{g}$. When $\dim \mathfrak{g} = 1$, the theorem is obviously true.

Then we suppose the argument is true when dimension is strictly less than $\mathfrak{g}$.

Let $\mathfrak{h} \subset \mathfrak{g}$ be a maximal proper subalgebra. (The existence of the maximal proper subalgebra is trivial by Zorn's lemma since $\mathfrak{g}$ is finite dimensional.) Then we claim that $\dim \mathfrak{g}/\mathfrak{h} = 1$ and $\mathfrak{h} \subset \mathfrak{g}$ is an ideal. Note that $ad(\mathfrak{h})$ preserves $\mathfrak{h}$, so the action $ad(\mathfrak{h}) \circlearrowright \mathfrak{g}/\mathfrak{h}$ is well-defined. Moreover, it follows from Step 1 that $ad(X)$ acts nipotently on $\mathfrak{g}/\mathfrak{h}$ for all $X \in \mathfrak{h}$, that is, $ad(\mathfrak{h}) \subset gl(\mathfrak{g}/\mathfrak{h})$ is such that every element in $ad(\mathfrak{h})$ is nilpotent. Note that $\dim ad(\mathfrak{h}) \leq \dim \mathfrak{g}/\mathfrak{h} < \mathfrak{g}$, hence by induction, we know $\exists 0 \neq \bar{Y}_0 \in \mathfrak{g}/\mathfrak{h}$ such that

$ad(X)(\bar{Y_0}) = 0$ in $\mathfrak{g}/\mathfrak{h}$ for all $X \in \mathfrak{h}$. That is to say that $[X, Y_0] \in \mathfrak{h}$ for all $X \in \mathfrak{h}$, where $Y_0 \in \mathfrak{g} - \mathfrak{h}$ is the lift of $\bar{Y_0}$. Hence, $\mathfrak{h} \oplus k \cdot Y_0$ is a Lie subalgebra. By the maximality of $\mathfrak{h}$, we know $\mathfrak{h} \oplus k \cdot Y_0 = \mathfrak{g}$. Then the claim follows immediately.

Step 3: From induction hypothesis, we know that there exists $0 \neq v_1 \in V$ such that $Xv_1 = 0$ for all $X \in \mathfrak{h}$. Set

$$W = \cap_{X \in \mathfrak{h}} \ker(X),$$

then $\dim W \geq 1$ since $v_1 \in W$. Since $\mathfrak{h} \oplus k \cdot Y_0 = \mathfrak{g}$, it suffices to show that there exists $v \in W$, $Y_0(v) = 0$. For all $w \in W$, all $X \in \mathfrak{h}$, we write $[X, Y_0](w) = (XY_0)(w) - (Y_0X)(w)$. Since $X, [X, Y_0] \in \mathfrak{h}$, we know $(XY_0)(w) = 0$, which implies $Y_0(w) \in W$ for all $w \in W$, that is, $Y_0(W) \subset W$.

By the assumption in the theorem, Y_0 is nilpotent and $\{k \cdot Y_0\} \subset gl(W)$ is a Lie subalgebra, so by induction hypothesis, we know there exists $v \in W$ such that $Y_0(v) = 0$, which completes the proof. $\square$

Theorem 3.2 (Lie's theorem)**.** *Let $\mathfrak{g} \subset gl(V)$ with $V = \mathbb{C}^r$ be a solvable Lie algebra over $\mathbb{C}$. Then there exists $0 \neq v \in V$ such that for all $X \in \mathfrak{g}$, $X(v) = \lambda_X \cdot v$ for some $\lambda_X \in \mathbb{C}$.*

Before proving this theorem, we show a corollary.

Corollary 3.3. *Let $\mathfrak{g}$ be as above, then there exists a basis $\{e_1, \cdots, e_r\}$ of V such that $\mathfrak{g} \subset \mathfrak{b}_r(\mathbb{C})$ with respect to $\{e_i\}$.*

Proof. Take $e_1 = v$ from Lie's theorem. Then we get $V' = V/\{\mathbb{C} \cdot e_1\}$. Note that the action $\mathfrak{g} \subsetneq V'$ is well-defined since $\mathfrak{g}(e_1) \subset \mathbb{C} \cdot e_1$, so we have an injective map $\varphi : \mathfrak{g} \to gl(V')$. Since $\varphi(\mathfrak{g}) \subset gl(V')$ is also solvable, inductively, we have $\{e'_2, \ldots, e'_r\} \subset V'$ such that $\varphi(\mathfrak{g}) \subset \mathfrak{b}_r(V')$, that is, $\varphi(\mathfrak{g})e'_k \in \mathbb{C}e'_1 + \cdots \mathbb{C}e'_k$. Hence, we know $\mathfrak{g} \subset \mathfrak{b}_r(\mathbb{C})$ with respect to $\{e_i\}_{i=1}^r$. $\square$

To prove Lie's theorem, we need the following proposition, which is an important one.

Proposition 3.4 (Key proposition)**.** *Let $\mathfrak{h} \subset \mathfrak{g}$ be an ideal. Let $\sigma : \mathfrak{g} \to gl(V) = gl(\mathbb{C}^r)$ be a Lie algebra homomorphism and $\lambda : \mathfrak{h} \to \mathbb{C}$ be a $\mathbb{C}$-linear map. Set*

$$W = \{v \in V : \sigma(X)(v) = \lambda(X) \cdot v, \forall X \in \mathfrak{h}\},$$

then $\sigma(Y)(W) \subset W$ for all $Y \in \mathfrak{g}$, that is, W is $\sigma(\mathfrak{g})$-stable.

Proof. **Step 1:** Pick any $0 \neq w \in W$. For all $X \in \mathfrak{h}, Y \in \mathfrak{g}$, we consider

$$\sigma(X)\sigma(Y)(w) = \sigma(Y)\sigma(X)(w) + \sigma([X, Y])(w) = \lambda(X)\sigma(Y)(w) + \lambda([X, Y])(w). \tag{3.1}$$

Then it suffices to show $\lambda([X, Y]) = 0$ for all $X \in \mathfrak{h}, Y \in \mathfrak{g}$.

Step 2: Fix some $Y \in \mathfrak{g}$, $w \in W$. We consider $U = span_{\mathbb{C}}\{\sigma^k(Y)(w) : k \geq 0\} \subset V$, then we claim that for all $X \in \mathfrak{h}$, for any fixed i,

$$\sigma(X)\left(\sigma^i(Y)\right)(w) = \sum_{j=0}^{i} c_j(X)\sigma^j(Y)(w)$$

with $c_i(X) = \lambda(X)$, that is, $\sigma(X)$ is represented by a matrix where the diagonal elements are all $\lambda(X)$.

We prove this claim by induction. By definition, the claim holds for $i = 0$. Suppose the claim is true for $i-1$, then we use (3.1) to write

$$\begin{aligned}\sigma(X)\sigma^i(Y)(w) &= \sigma(Y)\sigma(X)\sigma^{i-1}(Y)(w) + \sigma([X,Y])\sigma^{i-1}(Y)(w)\\ &= \lambda(X)\sigma^i(Y)(w) + \sum_{j=0}^{i-1} c_j(X)\sigma^j(Y)(w) + \sum_{j=0}^{i-1} c_j([X,Y])\sigma^j(Y)(w),\end{aligned}$$

where we use the induction hypothesis in the last equality. Hence, the claim is true for i. Thus, the claim holds.

Step 3: Since $\{\sigma^i(Y)(w) : i = 0, \cdots, \dim U - 1\}$ form a basis of U, which follows easily from the minimal polynomial theory for matrix $\sigma(Y)$ in linear algebra. Then we get $Tr(\sigma(X)|_U) = \dim U \cdot \lambda(X)$. We replace X by $[X,Y]$, then we get $\lambda([X,Y]) = 0$ since $Tr(\sigma([X,Y]|_U)) = Tr(\sigma(X)\sigma(Y) - \sigma(Y)\sigma(X)|_U) = 0$. Now the proof is complete. $\square$

Now we can prove Lie's theorem.

Proof of Lie's theorem. **Step 1:** We claim that there exists an ideal $\mathfrak{h}$ of codimension 1, that is, $\dim \mathfrak{g}/\mathfrak{h} = 1$. Since $\mathfrak{g}$ is solvable, we know $\mathfrak{g}^1 = [\mathfrak{g}, \mathfrak{g}] \neq \mathfrak{g}$. and then $\mathfrak{g}/\mathfrak{g}^1$ is an abelian Lie algebra. Take any subspace $\Delta \subset \mathfrak{g}/\mathfrak{g}^1$ of codimension 1, since it is abelian, so Δ is an ideal. We denote $\delta : \mathfrak{g} \to \mathfrak{g}/\mathfrak{g}^1$. Let $\mathfrak{h} = \delta^{-1}(\Delta)$, then it is obviously an ideal in $\mathfrak{g}$ since any Lie bracket is a preimage of 0 in $\mathfrak{g}/\mathfrak{g}^1$.

And in general, the dimension of the preimage of a codimension 1 vector subspace by a surjective linear map $f : V \to W$ has codimension 1, so $\mathfrak{h}$ is an ideal of codimension 1. We prove this general result as follows. Let L be a subspace of W of codimension 1, there exists $y \notin L$ such that $k \cdot y + L = W$, write $f(x) = y$. For every $z \in V$, $f(z) = u + ay$ with $u \in L$. We write $u = f(u')$, then $f(z - u' - ax) = 0$, which implies $z - u' - ax \in \ker(f) \subset f^{-1}(L)$. Now we deduce that $V = f^{-1}(L) + k \cdot x$.

Step 2: Now we proceed by induction on the dimension of $\mathfrak{g}$. By induction, there exists $0 \neq v_0 \in V$ such that $X(v_0) = \lambda(X) \cdot v_0$, for all $X \in \mathfrak{h}$, where $\lambda : \mathfrak{h} \to \mathbb{C}$ is a $\mathbb{C}$-linear functional. Set

$$W := \{v \in V : X(v) = \lambda(X) \cdot v, \forall X \in \mathfrak{h}\},$$

then we know $\dim W \geq 1$.

Step 3: Pick $Y \in \mathfrak{g} - \mathfrak{h}$, then the Key proposition, Proposition 3.4, implies that $Y(W) \subset W$. Then over $\mathbb{C}$, there exists an eigenvector $w \in W$ of Y, hence $Y(w) = \lambda_Y \cdot w$. Then it is easy to see that for all $X \in \mathfrak{h}$, for all $c \in \mathbb{C}$, $(X + cY)(w) = (\lambda(X) + c\lambda_Y) \cdot w$, so λ extends to a $\mathbb{C}$-linear map $\lambda : \mathfrak{g} \to \mathbb{C}$ satisfying that $Z(w) = \lambda(Z) \cdot w$ for all $Z \in \mathfrak{g}$, which completes the proof. $\square$

Definition 3.5. *A representation $\sigma : \mathfrak{g} \to gl(V)$ over k is called irreducible if $\{0\}, V$ are the only $\sigma(\mathfrak{g})$-stable subspace of V.*

Now we introduce the radical of a Lie algebra, which is a well-defined notion thanks to Lemma 2.15 (5).

Definition 3.6. *For any Lie algebra $\mathfrak{g}$ over k, its maximal solvable ideal is called the radical of $\mathfrak{g}$, denoted by $Rad(\mathfrak{g})$.*

Definition 3.7. $\mathfrak{g}$ *is called semisimple if the radical is zero.*

Proposition 3.8. *We consider the exact sequence*

$$0 \to Rad(\mathfrak{g}) \hookrightarrow \mathfrak{g} \to \mathfrak{g}/Rad(\mathfrak{g}) \to 0$$

then this exact sequence splits and $\mathfrak{g}/Rad(\mathfrak{g})$ *is semisimple.*

Proof. The first part is actually due to Levi(1905), which is called the Radical splitting theorem or Levi's decomposition. See [9] or [12] for a proof. For the second part, suppose by contradiction that there is a non-trivial proper ideal I of $\mathfrak{g}/Rad(\mathfrak{g})$ and consider its preimage in $\mathfrak{g}$, which is a solvable ideal containing $Rad(\mathfrak{g})$ as a proper subset, which leads to a contradiction. □

Though the rigorous definition of $\mathfrak{g}$-module is deferred to Section 5.1, we need a notion related to $\mathfrak{g}$-module here.

Definition 3.9. *(1) Let (σ_i, V_i), $1 \leq i \leq n$ be a $\mathfrak{g}$-module, that is, $\sigma_i : \mathfrak{g} \to gl(V_i)$. Then $\otimes_{i=1}^n V_i$ is a $\mathfrak{g}$-module with action $\sigma := \otimes_{i=1}^n \sigma_i$ given by*

$$\sigma(X)(\otimes_{i=1}^n v_i) := \sum_{i=1}^n v_1 \otimes \cdots \otimes (\sigma_i(X))(v_i) \otimes \cdots \otimes v_n. \tag{3.2}$$

(2) If (σ, V) is a $\mathfrak{g}$-module, then $(\sigma^*, V^* = Hom(V, k))$ is a $\mathfrak{g}$-module with $\sigma^*(X)(f)(v) = -f(\sigma(X)(v))$ for all $X \in \mathfrak{g}, f \in V^*$.

The motivation for the second definition is as follows. Note that the action $G \subsetneq V^*$ is defined as $g(f)(v) := f(g^{-1}v)$, so if we replace $g \in G$ by $\gamma_X(t) \in G$ in the equation above and take derivative with respect to t, we will get a negative sign.

Theorem 3.10. *Suppose $\mathfrak{g}$ is a Lie algebra over $\mathbb{C}$. Every finite dimensional irreducible $\mathbb{C}$-representation of $\mathfrak{g}$ is of the form*

$$\lambda \otimes (\sigma \circ f),$$

where $\lambda : \mathfrak{g} \to \mathbb{C} = gl(\mathbb{C})$ is a one-dimensional representation, σ is an irreducible representation of $\mathfrak{g}/Rad(\mathfrak{g})$ and $f : \mathfrak{g} \twoheadrightarrow \mathfrak{g}/Rad(\mathfrak{g})$ is the quotient map.

Proof. Let $(\pi, V) \in Irr(\mathfrak{g}) := \{$all finite dimensional $\mathbb{C}$-representations of $\mathfrak{g}\}$. By Lie's theorem, there exists a $\mathbb{C}$-linear map $\lambda : Rad(\mathfrak{g}) \to \mathbb{C}$ such that

$$W = \{v \in V : \pi(X)v = \lambda(X) \cdot v, \forall X \in Rad(\mathfrak{g})\}$$

is nonzero. Since $Rad(\mathfrak{g})$ is solveble, we apply the Key Proposition, Proposition 3.4, to get $\pi(\mathfrak{g})(W) \subset W$. Since (π, V) is irreducible, so $W = V$. Hence, $Tr(\pi(X)) = \dim V \cdot \lambda(X)$ for all $X \in Rad(\mathfrak{g})$. Thus, $\lambda|_{Rad(\mathfrak{g}) \cap [\mathfrak{g},\mathfrak{g}]} = 0$.

Extend it to get a $\mathbb{C}$-linear map $\lambda : \mathfrak{g} \to \mathbb{C}$ such that $\lambda|_{[\mathfrak{g},\mathfrak{g}]} = 0$. Now we claim $\sigma := \lambda^* \otimes \pi$ is trivial on $Rad(\mathfrak{g})$. For all $X \in Rad(\mathfrak{g})$, all $c \in \mathbb{C}, v \in V$, we write by using (3.2) that

$$\sigma(X)(c \otimes v) = \lambda^*(X)(c) \otimes v + c \otimes \pi(X)(v) = -\lambda(X)c \otimes v + c \otimes (\lambda(X) \cdot v) = 0, \tag{3.3}$$

where we use the definition of W and $W = V$ in the second equation. So σ is a well-defined representation of $\mathfrak{g}/Rad(\mathfrak{g})$.

Now we check $\pi = \lambda \otimes (\sigma \circ f)$. For all $X \in \mathfrak{g}$, all $v \in V$,

$$\lambda \otimes (\sigma \circ f)(X)(c_1 \otimes c_2 \otimes v) = \lambda(X)(c_1) \otimes c_2 \otimes v - c_1 \otimes \lambda(X) c_2 \otimes v + c_1 \otimes c_2 \otimes \lambda(X) \cdot v = \pi(X)(c_1 \otimes c_2 \otimes v),$$

by using (3.3), which completes the proof. □

4. Intrinsic way to describe solvability and semisimplicity

In this section, we will introduce the Cartan's criterion named after Elie Cartan. Unless specified otherwise, $k = \mathbb{R}$ or $\mathbb{C}$. We follow [6, Appendix C] and [8, Section 4.2] in this part.

4.1. **Killing form and Cartan criterions.**

Definition 4.1. *Let V_1, V_2 be two $\mathfrak{g}$-module over k, that is, $\rho_i : \mathfrak{g} \to gl(V)$.*
A k-bilinear form $B : V_1 \times V_2 \to k$ is called $\mathfrak{g}$-invariant if

$$B(Xv_1, v_2) + B(v_1, Xv_2) = 0, \forall X \in \mathfrak{g}.$$

Since {all the bilinear forms on $V_1 \times V_2$} are isomorphic to $Hom(V_1 \otimes V_2, k)$, linear maps on $V_1 \otimes V_2$, and $Hom(V, W) \simeq V^* \otimes W$, so we can identify $B \in (V_1 \otimes V_2)^* \otimes k \simeq (V_1 \otimes V_2)^*$ with $\mathfrak{g} \subsetneq (V_1 \otimes V_2)^*$.

Consider a representation $\sigma : \mathfrak{g} \to gl(V)$ and we define $B_\sigma : \mathfrak{g} \times \mathfrak{g} \to k$ by $(X, Y) \mapsto Tr(\sigma(X) \cdot \sigma(Y))$. Then B_σ is symmetric and $ad(\mathfrak{g})$-invariant, or say, $\mathfrak{g}$-invariant with respect to the adjoint action of $\mathfrak{g}$ on $\mathfrak{g}$.

We check B_σ is indeed $ad(\mathfrak{g})$-invariant by computing

$$\begin{aligned} B_\sigma(ad(Z)X, Y) + B_\sigma(X, ad(Z)Y) &= Tr(\sigma([Z, X]) \cdot \sigma(Y) + \sigma([Z, Y]) \cdot \sigma(X)) \\ = Tr(\sigma(Z)\sigma(X)\sigma(Y) - \sigma(X)\sigma(Z)\sigma(Y) &+ \sigma(Z)\sigma(Y)\sigma(X) - \sigma(Y)\sigma(Z)\sigma(X)) = 0, \end{aligned} \tag{4.1}$$

where the last equality just follows from the property that $Tr(AB) = Tr(BA)$ for matrices.

An important example is to take $\sigma = ad$ and $V = \mathfrak{g}$.

Definition 4.2. *We call $B_\kappa := B_{ad}$ defined by*

$$B_\kappa(X, Y) = Tr(ad(X) \circ ad(Y))$$

the Cartan Killing form of $\mathfrak{g}$. See [2, Page 5] *for the reason why it is called "Cartan Killing form".*

Proposition 4.3 (Cartan)**.** *Suppose $\mathfrak{g} \subset gl(V)$ over k and we denote the canonical inclusion by $\tau : \mathfrak{g} \to gl(V)$, which is surely also a Lie algebra representation. Suppose $Tr(XY) = 0$ for all $X, Y \in \mathfrak{g}$, that is, $B_\tau(X, Y) = 0$. Then $\mathfrak{g}$ is solvable.*

Proof. **Step 0:** We may assume $k = \mathbb{C}$ since solvablity of $\mathfrak{g}$ over $\mathbb{R}$ is equivalent to the solvability of its complexification $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$. This fact is easy to check by definition.

By induction, one can check $\mathfrak{g}^{i+1} \subset ([\mathfrak{g}, \mathfrak{g}])_i$ for all i, so it suffices to show $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent. Moreover, by Engel's theorem, it suffices to show that for all $X \in [\mathfrak{g}, \mathfrak{g}]$, X is a nilpotent matrix.

Fix $X \in [\mathfrak{g}, \mathfrak{g}]$ with $\det(tI - X) = \prod_{i=1}^{r}(t - \lambda_i)$. Then it suffices to show $\lambda_i = 0$ for all i, that is,

$$\sum_{i=1}^{r} \overline{\lambda_i}\lambda_i = 0.$$

Step 1: We consider the Jordan canonical form of X, denoted still by X, with the decomposition $X = X_s + X_n$, where X_s is a diagonal matrix, then we know X_s is semisimple and X_n is nilpotent. This is the so-called Jordan-Chevalley decomposition. See [8, Section 4.2] for details. Obviously, it suffices to show $Tr(X_s\overline{X_s}) = 0$.

Step 2: Since $X \in [\mathfrak{g}, \mathfrak{g}]$, there exists $Y, Z \in \mathfrak{g}$ such that $X = [Y, Z]$. Then

$$Tr(\overline{X_s}, X) = Tr(\overline{X_s}, [Y, Z]) = -Tr([Y, \overline{X_s}], Z) = Tr([\overline{X_s}, Y], Z),$$

where the second equality can be checked analogously to (4.1) though $\overline{X_s}$ may not be in $\mathfrak{g}$ since it is just a property for matrices.

By the assumption of proposition, it suffices to show $[\overline{X_s}, Y] = 0$, then $Tr(\overline{X_s}X_s) = Tr(\overline{X_s}, X) = 0$, which completes the proof.

Step 3: We claim that $ad(\overline{X_s})(\mathfrak{g}) \subset \mathfrak{g}$. To show this, it suffices to show $ad(\overline{X_s})$ can be expressed as a polynomial of $ad(X_s)$ and $ad(X_s)$ can be expressed as a polynomial of $ad(X)$, then since $ad(X)(\mathfrak{g}) \subset \mathfrak{g}$, the result follows.

Before we prove this, we claim that $ad(X) = ad(X_s) + ad(X_n)$ is the Jordan-Chevalley decomposition. By the uniqueness of decomposition(c.f. [8, Proposition in Section 4.2]) and the commutativity $[ad(X_s), ad(X_n)] = ad[X_s, X_n] = 0$, it suffices to show $ad(X_s)$ is semisimple and $ad(X_n)$ is nilpotent. Note that $ad(X_n)$ is nilpotent thanks to Step 1 in the proof of Engel's theorem, Theorem 3.1. To prove $ad(X_s)$ is semisimple in $End(\mathfrak{g})$ (c.f. [8, Proposition in Section 4.2] for the definition), let E_{ij} be the standard basis of matrices $M_n(\mathbb{C})$, then $X = \lambda_1 E_{11} + \cdots \lambda_n E_{nn} \in gl(V) = M_n(\mathbb{C})$. Since $E_{ij}E_{kl} = \delta_{jk}E_{il}$, we know $ad(X)(E_{ij}) = XE_{ij} - E_{ij}X = (\lambda_i - \lambda_j)E_{ij}$, which implies $ad(X) \in gl(M_n(\mathbb{C})) = M_{n^2}(\mathbb{C})$ is a diagonal matrix with respect to the basis $\{E_{ij}\}_{1\le i,j\le n}$. So $ad(X_s)$ is semisimple.

Step 4: Now we know $ad(X) = ad(X_s) + ad(X_n)$ is the Jordan-Chevalley decomposition of $ad(X)$. Hence, $ad(X_s)$ can be expressed as a polynomial of $ad(X)$ by [8, Property b, Proposition, Section 4.2] thanks to the Chinese remainder theorem. We give a proof as follows. Set $\widetilde{X} = ad(X), \widetilde{X_s} = ad(X_s), \widetilde{X_n} = ad(X_n)$. Suppose $\det(tI - \widetilde{X}) = \prod_{i=1}^{l}(t - \eta_i)^{m_i}$ where all η_i's are distinct. Then $\{(t - \eta_j)^{m_j}\}$ are congruent with each other, so the Chinese remainder theorem tells us there exists a polynomial f such that $f(t) \equiv \eta_j \mod (t - \eta_j)^{m_j}$ for $j = 1, \cdots, l$. By the property that $\widetilde{X}$ is in the Jordan canonical form, we know

$$f(\widetilde{X}) = f\begin{pmatrix} X_1 & & \\ & \ddots & \\ & & X_l \end{pmatrix} = \begin{pmatrix} f(X_1) & & \\ & \ddots & \\ & & f(X_l) \end{pmatrix} = \begin{pmatrix} \eta_1 I_{m_1} & & \\ & \ddots & \\ & & \eta_l I_{m_l} \end{pmatrix} = \widetilde{X_s}.$$

Step 5: The last step is to show $ad(\overline{X_s})$ is a polynomial of $ad(X_s)$. The Lagrange interpolation theorem (c.f. [1, Example 1.2.22]) asserts that for $\{a_k\}_{k=1}^n \subset k$, $\{b_k\}_{k=1}^n \subset k$ there exists a polynomial g such that $g(a_k) = b_k$. So we know there exists a polynomial such that $g(\eta_k) = \overline{\eta_k}$, which implies

$$g(\widetilde{X_s}) = \begin{pmatrix} g(\eta_1 I_{m_1}) & & \\ & \ddots & \\ & & g(\eta_l I_{m_l}) \end{pmatrix} = \overline{\widetilde{X_s}}.$$

This completes the proof. $\square$

Theorem 4.4 (Cartan). *For any Lie algebra $\mathfrak{g}$ over k, the followings are equivalent:*
(1) $\mathfrak{g}$ is solvable;
(2) $B_\kappa(\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]) = 0$.

Proof. Again, we can assume $k = \mathbb{C}$.

Step 1: Suppose $\mathfrak{g}$ is solvable, then we apply Lie's theorem to $ad : \mathfrak{g} \to gl(\mathfrak{g})$ and we know $ad(\mathfrak{g}) \subset \mathfrak{b}(\mathfrak{g})$, which implies

$$ad([\mathfrak{g}, \mathfrak{g}]) \subset [ad(\mathfrak{g}), ad(\mathfrak{g})] \subset \mathfrak{n}(\mathfrak{g}).$$

Hence, for $X \in \mathfrak{g}$, $Y \in [\mathfrak{g}, \mathfrak{g}]$, $B_\kappa(X, Y) = Tr(ad(X) \circ ad(Y)) = 0$.

Step 2: Suppose (2) holds, then $B_\kappa(\mathfrak{g}^1, \mathfrak{g}^1) = 0$, then we consider the adjoint representation $ad : \mathfrak{g} \to gl(\mathfrak{g})$, we know by Proposition 4.3 that the image $ad(\mathfrak{g}^1)$ in $gl(\mathfrak{g})$ is solvable.

We denote the upper central series of $\mathfrak{g}$ by $\mathfrak{g}^i$. Since $ad(\mathfrak{g}^m) = 0$ for some m, we have $\mathfrak{g}^m \subset Z(\mathfrak{g})$. Hence, $\mathfrak{g}^{m+1} = 0$, which completes the proof. □

Remark 4.5. In the Step 2 of the proof above, we cannot obtain $\mathfrak{g}^1$ is solvable directly by $B_\kappa(\mathfrak{g}^1, \mathfrak{g}^1) = 0$ and Proposition 4.3 since we do not assume the canonical inclusion $\mathfrak{g} \subset gl(V)$ exists for some V but we could use the natural representation $ad : \mathfrak{g} \to gl(\mathfrak{g})$ to get the canonical inclusion $ad(\mathfrak{g}) \subset gl(\mathfrak{g})$. We should notice that Theorem 4.4 holds for any Lie algebra $\mathfrak{g}$.

Example 4.6. *Recall that* $\mathfrak{so}_2(\mathbb{R}) = \left\{ \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} : a \in \mathbb{R} \right\}$. *Though one can check that it cannot be embedded in $\mathfrak{b}_2(\mathbb{R})$, we have $[\mathfrak{g}, \mathfrak{g}] = 0$, so Theorem 4.4 implies $\mathfrak{so}_2(\mathbb{R})$ is solvable.*

Theorem 4.7 (Cartan). *Any Lie algebra $\mathfrak{g}$ over k is semisimple if and only if the Killing form $B_\kappa : \mathfrak{g} \times \mathfrak{g} \to k$ is non-degenerate.*

Proof. Set

$$S := Ker(B_\kappa) = \{X \in \mathfrak{g} : B_\kappa(X, Y) = 0, \forall Y \in \mathfrak{g}\}.$$

For all $X \in S, Y \in \mathfrak{g}$, we have $B_\kappa(ad(Z)(X), Y) = -B_\kappa(X, ad(Z)(Y)) = 0$, so $S \subset \mathfrak{g}$ is an ideal.

Step 1: Suppose $\mathfrak{g}$ is semisimple. Now we want to show that $S = 0$. By Proposition 4.3, we know that $ad(S) \subset ad(\mathfrak{g})$ is solvable since $B_\kappa(X, Y) = Tr(ad(X), ad(Y)) = 0$ for all $X, Y \in S$.

We shall follow the argument used in the Step 2 of the proof of Theorem 4.4. Since $ad(S)$ is solvable, there exists some m such that $(ad(S))^m = ad(S^m) = 0$, which implies $S^m \subset Z(\mathfrak{g})$, then $S^{m+1} = 0$. Hence, S is solvable. However, $\mathfrak{g}$ is semisimple, so $S = 0$.

Step 2: Conversely, let $S = 0$. It suffices to show that any abelian ideal $\mathfrak{a} \subset \mathfrak{g}$ is zero. Indeed, if $r := Rad(\mathfrak{g}) \neq 0$, then $r^0 \supset r^1 \supset \cdots \supset r^m \supset r^{m+1} = 0$ with $r^m \neq 0$. $r^{m+1} = [r^m, r^m] = 0$ implies r^m is a nontrivial abelian ideal in $\mathfrak{g}$.

Let $\mathfrak{a} \subset \mathfrak{g}$ be an abelian ideal. Then for all $X \in \mathfrak{a}$, $Y \in \mathfrak{g}$, we have

$$ad(X) \circ ad(Y)(\mathfrak{g}) \subset \mathfrak{a}, \quad ad(X) \circ ad(Y)(\mathfrak{a}) = 0,$$

which implies $ad(X) \circ ad(Y)$ is nilpotent. Hence, $B_\kappa(X, Y) = Tr(ad(X) \circ ad(Y)) = 0$. So $\mathfrak{a} \subset S = 0$, which completes the proof. □

Example 4.8. *Let us consider* $\mathfrak{g} := \mathfrak{sl}_2(\mathbb{R})$ *with* $\dim(\mathfrak{g}) = 3$. *Set* X, H, Y *as in Example 2.3, then* $[X, Y] = H, [H, X] = 2X, [H, Y] = -2Y$ *implies*

$$ad(X) = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad ad(H) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad ad(Y) = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

with respect to the basis $\{X, H, Y\}$. *By a direct computation, we know the representation of* B_κ *with respect to* $\{ad(X), ad(H), ad(Y)\}$ *is* $\begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}$, *so it is semisimple thanks to Theorem 4.7.*

4.2. **Relation between semisimplicity and simplicity.** Now we study the relation between "semisimple" and "simple". See [8, Section 5.2] for a reference.

Definition 4.9. *A Lie algebra* $\mathfrak{g}$ *is said to be the direct sum of ideals* $\{\mathfrak{g}_i\}_{i=1}^m$ *provided* $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_m$ *(direct sum of subspaces). This condition forces* $[\mathfrak{g}_i, \mathfrak{g}_j] \subset \mathfrak{g}_i \cap \mathfrak{g}_j = 0$ *if* $i \neq j$. *So the Lie bracket can be viewed as being defined componentwise for the external direct sum of these as vector spaces, that is,* $[X, Y] = [X_1, Y_1] + \cdots [X_m, Y_m]$ *with* $X_i, Y_i \in \mathfrak{g}_i$.

We still write $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_m$.

Theorem 4.10. *A Lie algebra* $\mathfrak{g}$ *over* k *is semisimple if and only if there exists ideals* $\{\mathfrak{g}_i\}_{i=1}^m$ *which are simple such that* $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_m$.

Proof. **"only if" part:** Suppose $\mathfrak{g}$ is not simple , then let $0 \neq \mathfrak{h} \subsetneq \mathfrak{g}$ be an ideal. Set

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

which is also an ideal of $\mathfrak{g}$. By Theorem 4.4, we know $\mathfrak{h} \cap \mathfrak{h}^\perp$ is a solvable ideal of $\mathfrak{g}$. Hence, $\mathfrak{h} \cap \mathfrak{h}^\perp = 0$.

Now choose a basis for $\mathfrak{h}$, denoted by $Y_1, \cdots, Y_k$. Denote the matrix representation of B_κ by M_B, which is invertible since B_κ is non-degenerate by Theorem 4.7. Hence, $M_B Y_1, \cdots, M_B Y_k$ are linearly independent.

Consider the map $X \mapsto (X^T M_B Y_1, \cdots, X^T M_B Y_k)$, $\mathfrak{g} \to \mathbb{R}^k$, which is surjective with kernel $\mathfrak{h}^\perp$, so $\dim \mathfrak{g} = \dim \mathfrak{h} + \dim \mathfrak{h}^\perp$. Since $\mathfrak{h} \cap \mathfrak{h}^\perp = 0$, we know $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$, which is a direct sum of two semisimple subalgebras. By induction and the fact that $\dim \mathfrak{g} < \infty$, the theorem follows.

"if" part: This part is obvious since one can check $Rad(\mathfrak{g}) = Rad(\mathfrak{g}_1) + \cdots + Rad(\mathfrak{g}_\mathfrak{m})$.

Indeed, for any ideal $\mathfrak{h} \subset \mathfrak{g}$, there exists subalgebras $\mathfrak{h}_i \subset \mathfrak{g}_i$ such that $\mathfrak{h} = \mathfrak{h}_1 + \cdots + \mathfrak{h}_m$. Since $\mathfrak{h} \supset [\mathfrak{h}, \mathfrak{g}] = [\mathfrak{h}_1, \mathfrak{g}_1] + \cdots + [\mathfrak{h}_m, \mathfrak{g}_m]$, we know $[\mathfrak{h}_i, \mathfrak{g}_i] \subset \mathfrak{h}_i$, which implies $\mathfrak{h}_i$ is an ideal of $\mathfrak{g}_i$ for all i. However, since $\mathfrak{g}_i$ is simple, we know $\mathfrak{h}_i = 0$ or $\mathfrak{g}_i$, which implies $\mathfrak{h} = \mathfrak{g}_{k_1} + \cdots + \mathfrak{g}_{k_l}$ for some index set $\{k_1, \cdots, k_l\}$. Hence, $[\mathfrak{h}, \mathfrak{h}] = \mathfrak{h}$, which cannot be solvable. Thus, $Rad(\mathfrak{g}) = 0$. □

Corollary 4.11. *Suppose* $\mathfrak{g}$ *is semisimple, then* $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ *and any ideal of* $\mathfrak{g}$ *is semisimple.*

Moreover, all homomorphism images of $\mathfrak{g}$ *are semisimple. In particular, any quotient of* $\mathfrak{g}$ *by some ideal is semisimple.*

Proof. By Theorem 4.10, we write $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_m$ with $\mathfrak{g}_i$ are simple ideals. Since $[\mathfrak{g}_i, \mathfrak{g}_i] = \mathfrak{g}_i$, we know $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.

For any ideal $\mathfrak{h} \subset \mathfrak{g}$, one can argue as in the "if part" of the proof of Theorem 4.10 to get $\mathfrak{h} = \mathfrak{g}_{k_1} + \cdots + \mathfrak{g}_{k_l}$ for some index set $\{k_1, \cdots, k_l\}$, which is semisimple thanks to Theorem 4.10.

The last part follows from the fact that the homomorphism image of a simple Lie algebra is also simple. $\square$

4.3. **Examples: Simple Lie algebras.** Now we give an example of simple Lie algebra which can be shown by computation. We will introduce more powerful tools called reductive Lie algebras in the following section to describe semisimplity.

Example 4.12. *The Lie algebra $\mathfrak{g} = \mathfrak{sl}_r(k)$ is simple. By Example 2.8, $\mathfrak{g}$ is spanned by $E_{ij}, E_{ji}, E_{ii} - E_{jj}$, $i < j$. Let $0 \neq \mathfrak{h} \subset \mathfrak{g}$ be an ideal, it suffices to show there exists $E_{ij} \in \mathfrak{h}$. This is because $[E_{ij}, E_{jk}] = E_{ik}$ for $k \neq i$, $[E_{ij}, E_{ji}] = E_{ii} - E_{jj}$ and $[[E_{ij}, E_{ji}], E_{ji}] = -2E_{ji}$, which implies $\mathfrak{g} = \mathfrak{h}$.*

Now we prove $E_{ij} \in \mathfrak{h}$ for some $i < j$. We write $[[X, E_{ij}], E_{ij}] = (-2X_{ji})E_{ij}$. If there exists $X \in \mathfrak{h}$ such that $X_{ji} \neq 0$ for some $i < j$, then $E_{ij} \in \mathfrak{h}$. Otherwise, $X = \sum_{i=1}^{r} \alpha_i E_{ii}$, then there exists $i \neq j$ such that $\alpha_i \neq \alpha_j$ since $Tr(X) = 0$. We get $[X, E_{ij}] = (\alpha_i - \alpha_j)E_{ij}$, which implies $E_{ij} \in \mathfrak{h}$.

In [17, Section 6.4 - 6.6], the author proved by computation that $\mathfrak{su}(r)$ and $\mathfrak{sp}(r)$ are simple for all r. Also, $\mathfrak{so}(r)$ is simple for $r > 4$.

5. Complete reducibility of representations

In this section, we follow [8, Section 6].

5.1. **Modules.** We first digress to introduce the definitions of modules.

Definition 5.1. *Suppose $\mathfrak{g}$ is a Lie algebra and V is a vector space. Then V endowed with the operation $\mathfrak{g} \times V \to V$ satisfying*
(1) $(c_1X + c_2Y)(v) = c_1X(v) + c_2Y(v)$;
(2) $X(c_1v + c_2w) = c_1X(v) + c_2X(w)$;
(3) $[X, Y](v) = X(Y(v)) - Y(X(v))$
is called a $\mathfrak{g}$-module.

Obviously, if $\sigma : \mathfrak{g} \to gl(V)$ is a representation of $\mathfrak{g}$, then one can view V as a $\mathfrak{g}$-module via the operation $(X, v) \mapsto \sigma(X)v$, that is, $X \cdot v := \sigma(X)v$. Naturally, we say V is irreducible if it has precisely two $\mathfrak{g}$-submodules V and $\{0\}$. This matches the definition of irreducible representation. In this view, a submodule is related to a $\sigma(\mathfrak{g})$-stable subspace.

A homomorphism of $\mathfrak{g}$-modules is a linear map $f : V \to W$ satisfying $f(X(v)) = X(f(v))$.

If V, W are two $\mathfrak{g}$-modules, then we can view $V \otimes W$ as a $\mathfrak{g}$-module thanks to Definition 3.9. Since $V^* \otimes W \simeq Hom(V, W)$, $\mathfrak{g}$ acts naturally on the space $Hom(V, W)$ by the rule $(X\phi)(u) = X(\phi(u)) - \phi(X(u))$ for $u \in V, \phi \in Hom(V, W)$. Note that for $f \in V^*, w \in W$, the isomorphism $V^* \otimes W \simeq Hom(V, W)$ maps $f \otimes w$ to the element $\phi : v \mapsto f(v)w$ in $Hom(V, W)$. Then the action $\mathfrak{g} \subsetneq V^* \otimes W$ $X(f \otimes w) = Xf \otimes w + f \otimes Xw$, corresponds to the action $\mathfrak{g} \subsetneq Hom(V, W)$

$$u \mapsto (Xf)(u)w + f(u)Xw = f(u)Xw - f(Xu)w = X(f(u)w) - f(Xu)w = X(\phi(u)) - \phi(Xu).$$

Hence, the action $\mathfrak{g} \subsetneq Hom(V, W)$ arises from the isomorphism $V^* \otimes W \simeq Hom(V, W)$ naturally.

Note that for $\phi \in Hom(V,W)$, $X\phi = 0$ is equivalent to $X(\phi(v)) = \phi(X(v))$ for all $v \in V$, it follows that

$$\begin{aligned} Hom_{\mathfrak{g}}(V,W) :=& \{\phi : V \to W \text{ is a } \mathfrak{g} - \text{module homomorphism}\} \\ =& \{\phi \in Hom_k(V,W)(= Hom(V,W)) : X\phi = 0, \forall X \in \mathfrak{g}\}. \end{aligned}$$

5.2. **Weyl's theorem on complete reducibility of $\sigma : \mathfrak{g} \to gl(V)$.** Now we discuss the semisimplicity of representations.

Definition 5.2. *A representation $\sigma : \mathfrak{g} \to gl(V)$ over k is called semisimple if for all $\sigma(\mathfrak{g})$-stable subspace $W \subset V$, there is a $\sigma(\mathfrak{g})$-stable complement $W' \subset V$ such that $V = W \oplus W'$. Or we say W, W' are $\mathfrak{g}$-stable via σ.*

Suppose $\mathfrak{g}$ is semisimple, then we know from Theorem 4.10 that $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$, where $\mathfrak{g}_i$ are simple . By definition, $\mathfrak{h} \subset \mathfrak{g}$ is an $ad(\mathfrak{g})$-subspace if and only if $\mathfrak{h}$ is an ideal. On the other hand, the ideals of $\mathfrak{g}$ is of the form $\mathfrak{g}_{l_1} \oplus \cdots \oplus \mathfrak{g}_{l_m}$ by Corollary 4.11. Hence, the representation $ad : \mathfrak{g} \to gl(\mathfrak{g})$ is semisimple provided that $\mathfrak{g}$ is semisimple.

And it is trivial that if any representation of $\mathfrak{g}$ is semisimple, then $ad(\mathfrak{g})$ is semisimple. So now it is natural to discuss the relation between the semisimplicity of $\mathfrak{g}$ and the condition that "any representation of $\mathfrak{g}$ is semisimple" and "$ad(\mathfrak{g})$ is semisimple". This question is answered by Weyl's theorem.

In the previous subsection, we proved that a semisimple Lie algebra has nondegenerate Killing form B_κ in Theorem 4.7. Here we generalize it to B_σ.

Let $\mathfrak{g}$ be a semisimple Lie algebra and $\sigma : \mathfrak{g} \to gl(V)$ is a faithful(injective) representation, then we claim B_σ is non-degenerate. The argument is similar to the proof of Theorem 4.7. Since $\mathfrak{g}$ is semisimple, it suffices to show $\ker(B_\sigma)$ is a solvable ideal. By Proposition 4.3, it suffices to show $\ker(B_\sigma)$ is an ideal. One can check this by a direct computation.

This leads to the following definition of the Casimir element naturally.

Definition 5.3. *Suppose $\mathfrak{g}$ is a semisimple Lie algebra. For any representation $\sigma : \mathfrak{g} \to gl(V)$, choose two bases $\{X_i\}, \{Y_j\} \subset \mathfrak{g}$ such that $B_\sigma(X_i, Y_j) = \delta_{ij}$. we say*

$$c_{\mathfrak{g}} := \sum_{i=1}^{r} \sigma(X_i)\sigma(Y_i) \in gl(V),$$

is the Casimir element of $\mathfrak{g}$ with respect to the representation σ, where $r = \dim \mathfrak{g}$.

One can check it is independent of the choice of basis. Note that $c_{\mathfrak{g}}$ may not lie in $\mathfrak{g}$.

Lemma 5.4. *We have $[c_{\mathfrak{g}}, \sigma(\mathfrak{g})] = 0$ and $Tr(c_{\mathfrak{g}}) = r$.*

Proof. For $X \in \mathfrak{g}$, let $[X, X_i] = \sum_{j=1}^{r} a_{ij} X_j$, $[X, Y_i] = \sum_{j=1}^{r} b_{ij} Y_j$, then

$$a_{ij} = B_\sigma([X, X_i], Y_j) = -B_\sigma(X_i, [X, Y_j]) = -b_{ji}.$$

Hence, we compute

$$\begin{aligned} [\sigma(X), c_{\mathfrak{g}}] = \sum_{i=1}^{r} [\sigma(X), \sigma(X_i)\sigma(Y_i)] &= \sum_{i=1}^{r} [\sigma(X), \sigma(X_i)]\sigma(Y_i) + \sigma(X_i)[\sigma(X), \sigma(Y_i)] \\ &= \sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij}\sigma(X_j)\sigma(Y_i) + b_{ij}\sigma(X_i)\sigma(Y_j) = 0. \end{aligned}$$

The second assertion is direct since $Tr(c_{\mathfrak{g}}) = \sum_{i=1}^{r} Tr(\sigma(X_i)\sigma(Y_i)) = r$. $\square$

Example 5.5. *Recall the classical example $\mathfrak{sl}_2(k)$ with the basis $\{X, H, Y\}$ satisfying that $Tr(XY) = 1$, $Tr(HH) = 2$, $Tr(XX) = Tr(YY) = 0$. Hence, $c_g = Tr(XY) + Tr(YX) + Tr(HH) = \begin{pmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}$, which is not in $\mathfrak{sl}_2(k)$. However, it is in $Z(\mathfrak{sl}_2(k))$.*

Lemma 5.6 (Schur's Lemma)**.** *If $\sigma : \mathfrak{g} \to gl(V)$ is an irreducible representation over k and $M \in gl(V)$ is such that $M\sigma(X) = \sigma(X)M$ for all $X \in \mathfrak{g}$. Then $M = 0$ or $M \in GL(V)$.*

Moreover, if $k = \mathbb{C}$, then $M = \lambda I_n$, where $n = \dim V$.

Proof. Note that $\ker M$ is $\sigma(\mathfrak{g})$-stable, then $\ker(M) = 0$ or V since σ is irreducible. Here, the dimension of V is finite and $\ker(M) = 0$ imply $M \in GL(V)$.

Now suppose $k = \mathbb{C}$, then M has at least one nontrivial eigenspace unless $M = 0$. Since any eigenspace of M is $\sigma(\mathfrak{g})$-stable, thus each eigenspace is exactly equal to V, which implies $M = \lambda I_n$. $\square$

Now we introduce the Weyl's theorem. The original proof used the "unitary trick". But we will present a different one here. The main idea is induction.

Lemma 5.7. *Any representation $\tau : \mathfrak{g} \to gl_1(k)$ is trivial provided that $\mathfrak{g}$ is semisimple.*

Proof. From Corollary 4.11, we know $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$. However, $\tau([\mathfrak{g}, \mathfrak{g}]) = 0$, which implies the map is trivial. $\square$

Theorem 5.8 (Weyl's theorem)**.** *Let $\mathfrak{g}$ be a semisimple Lie algebra over $k = \mathbb{R}, \mathbb{C}$ and $\sigma : \mathfrak{g} \to gl(V)$ is a representation. If $W \subset V$ is $\sigma(\mathfrak{g})$-stable, then there exists $\sigma(\mathfrak{g})$-stable subspace $W' \subset V$ such that $V = W \oplus W'$.*

Proof. **Step 1:** We start with a special case. If $W \subset V$ is an irreducible $\mathfrak{g}$-submodule and $\dim V/W = 1$, then $V = W \oplus \ker(c_{\mathfrak{g}})$ as $\mathfrak{g}$-modules.

We claim that the Casimir element $c_{\mathfrak{g}}$ acts on W invertibly. Suppose not, then by Lemma 5.4, we know $c_{\mathfrak{g}} \in gl(V)$ and it commutes with $\sigma(\mathfrak{g})$. Moreover, we have W is an irreducible submodule, so we can apply Schur's lemma, Lemma 5.6. Hence, we know $c_{\mathfrak{g}}$ acts on W either invertibly or trivially.

Now suppose $c_{\mathfrak{g}}$ acts on W trivially by contradiction. From Lemma 5.7, we know $\mathfrak{g}$ acts on V/W trivially, that is $\sigma(\mathfrak{g})(V) \subset W$. This action is well-defined since W is $\sigma(\mathfrak{g})$-stable. Then $c_{\mathfrak{g}}(V) \subset W$, that is, $c_{\mathfrak{g}}$ acts on V/W trivially. Hence, $Tr_V(c_{\mathfrak{g}}) = 0$, which contradicts to the second assertion in Lemma 5.4. Thus, the claim is true, that is, $c_{\mathfrak{g}}$ acts on W invertibly. In particular, $\ker(c_{\mathfrak{g}}) \cap W = \{0\}$.

Pick $v \in V - W$ and set $w_0 = c_{\mathfrak{g}}(v)$. Since $c_{\mathfrak{g}}$ acts on W invertibly, there exists $(c_{\mathfrak{g}})^{-1}(w_0) \in W$, then $v - (c_{\mathfrak{g}})^{-1}(w_0) \in \ker(c_{\mathfrak{g}})$, implies $V = W + \ker(c_{\mathfrak{g}}) = W \oplus \ker(c_{\mathfrak{g}})$. Since $[c_{\mathfrak{g}}, \mathfrak{g}] = 0$ by Lemma 5.4, $\ker(c_{\mathfrak{g}})$ is also $\mathfrak{g}$-stable. This completes the proof.

Step 2: Now we consider another special case. If $W \subset V$ is a $\mathfrak{g}$-submodule and $\dim V/W = 1$, then there exists a $\mathfrak{g}$-module W' such that $V = W \oplus W'$.

From Step 1, we assume with loss of generality that W is not irreducible, then one can pick a $\sigma(\mathfrak{g})$-stable submodule $0 \neq Z \subsetneq W$. Then $W/Z \subset V/Z$ is of codimension 1. By induction, since $\dim W/Z < \dim W$, we know there exists a submodule A of V/Z such that $V/Z = A \oplus W/Z$ with $\dim A = \dim W/Z - \dim V/Z = \text{codim} W/V = 1$. (The induction hypothesis is just the step 1 since it reduces to the irreducible case when dimension goes down.) We denote the canonical map as $f : V \to V/Z$ and let $Y = f^{-1}(A)$. Now by induction, since $A = Y/Z$ has dimension 1, we know there exists $\mathfrak{g}$-submodule U such that $Y = Z \oplus U$. It follows that $V = W \oplus U$ since $\dim W + \dim U = \dim W + \dim Y - \dim Z = \dim W + 1 = \dim V$ and $W \cap U = \{0\}$.

Step 3: Now for general submodule $W \subset V$, let

$$\varphi : Hom(V, W) \to Hom(W, W), \quad f \mapsto f|_W.$$

Consider $k \cdot id_W \subset Hom(W, W)$, for all $f \in \varphi^{-1}(k \cdot id_W)$, we have $f|_W = c \cdot id_W$ for some $c \in k$. Then for $w \in W$, $(Xf)(w) = X(f(w)) - f(Xw) = X(cw) - c(Xw) = 0$, that is, $\mathfrak{g}$ maps $\varphi^{-1}(k \cdot id_W)$ to $\varphi^{-1}(0 \cdot id_W)$, and both $\varphi^{-1}(k \cdot id_W)$ and $\varphi^{-1}(0 \cdot id_W)$ are $\mathfrak{g}$-modules.

Moreover, $\varphi^{-1}(k \cdot id_W)/\varphi^{-1}(0 \cdot id_W)$ has dimension one since for all $f \in \varphi^{-1}(k \cdot id_W)$, it is determined modulo $\varphi^{-1}(0 \cdot id_W)$ by the scalar $f|_W$, hence there exists a one dimensional subspace of $\varphi^{-1}(k \cdot id_W)$, say $\mathcal{W}$ such that $\varphi^{-1}(k \cdot id_W) = \varphi^{-1}(0 \cdot id_W) \oplus \mathcal{W}$ thanks to the claim in Step 2. Let $0 \neq f \in \mathcal{W}$, then $\mathcal{W} = \text{span}\,\{f\}$. One may assume $f|_W = id_W$ by multiplying a constant. From Lemma 5.7, we know that $\mathfrak{g}$ acts trivially on $\mathcal{W}$, and hence on f, that is, $0 = (Xf)(v) = X(f(v)) - f(Xv)$ for all $v \in V, X \in \mathfrak{g}$. This is to say f is a $\mathfrak{g}$-module homomorphism. Therefore, $\ker f$ is a $\mathfrak{g}$-submodule of V. Since f maps V into W and acts as id_W on W, $\dim \ker f = \dim V - \dim W$, so we conclude that $V = W \oplus \ker f$, which completes the proof. $\square$

As an application of Weyl's theorem, we can show the following theorem which is closely related to the Jordan-Chevalley decomposition used in the proof of Proposition 4.3.

Theorem 5.9. *Let $\mathfrak{g} \subset gl_r(\mathbb{C})$ be a semisimple linear Lie algebra. Then the usual Jordan decomposition (Jordan-Chevalley decomposition) of $X \in \mathfrak{g} \subset gl(V)$ in $gl(V)$, say $X = X_s + X_n$, satisfies $X_s, X_n \in \mathfrak{g}$.*

Proof. **Step 1:** From Lemma 4.11, $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$, we have $\mathfrak{g} \subset sl_r(\mathbb{C})$ since $Tr(AB - BA) = 0$. Since X_n is nilpotent, then $X_n \in sl_r(\mathbb{C})$, and so is X_s.

Step 2: For any $\mathfrak{g}$-submodule $W \subset \mathbb{C}^r = V$, we set

$$S_W := \{Y \in gl_r(\mathbb{C}) : Y(W) \subset W, Tr(Y|_W) = 0\}.$$

Then $\mathfrak{g} \subset S_W$ since W is a $\mathfrak{g}$-module and $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$. Since X_s and X_n are polynomials of X, we know $X_s(W) \subset W$ and $X_n(W) \subset W$. Moreover, since X_n is nilpotent, $Tr(X_n|_W) = 0$, thus $Tr(X_s|_W)$ is also zero. Hence, $X_n, X_s \in S_W$.

Step 3: Set

$$\mathfrak{g}' = N_{gl_r(\mathbb{C})}(\mathfrak{g}) \cap \left(\bigcap_{W \subset V, \text{submodule}} S_W \right),$$

where $N_{gl_r(\mathbb{C})}(\mathfrak{g}) := \{Y \in gl_r(\mathbb{C}) : [Y, \mathfrak{g}] \subset \mathfrak{g}\}$, then we have $\mathfrak{g} \subset \mathfrak{g}'$. Since $ad(Y_s)$ and $ad(Y_n)$ are polynomials of $ad(Y)$, and $ad(Y)\mathfrak{g} \subset \mathfrak{g}$ for all $Y \in N_{gl_r(\mathbb{C})}(\mathfrak{g})$, so $Y_s, Y_n \in N_{gl_r(\mathbb{C})}(\mathfrak{g})$. It follows from Step 2 and Step 3 that $X_s, X_n \in \mathfrak{g}'$ for all $X \in \mathfrak{g}$.

Step 4: We view $\mathfrak{g}'$ as a $\mathfrak{g}$-module. Then by Weyl's theorem, there exists a $\mathfrak{g}$-module $\mathfrak{g}''$ such that $\mathfrak{g} = \mathfrak{g}' \oplus \mathfrak{g}''$. Then it suffices to show $\mathfrak{g}'' = 0$. For $X \in \mathfrak{g}$, $Y'' \in \mathfrak{g}'' \subset \mathfrak{g}$, $[X, Y''] \in \mathfrak{g} \cap \mathfrak{g}'' = \{0\}$.

Let $W \subset V$ be an irreducible $\mathfrak{g}$-module. We claim that for all $Y \in \mathfrak{g}''$, we have $Y|_W = 0$. Since $[Y, \mathfrak{g}] = 0$, we know by Schur's lemma that Y acts on W as a scalar, that is, $Y|_W = \lambda \cdot id_W$. On the other hand, $Tr(Y|_W) = 0$ since $Y \in S_W$, so $Y|_W = 0$.

Since V can be written as a direct sum of irreducible $\mathfrak{g}$-submodules thanks to Weyl's Theorem, so in fact $Y = 0$. Hence, $\mathfrak{g}'' = 0$, which completes the proof. □

As a corollary of this theorem (Theorem 5.9), we can define the abstract Jordan decomposition for any semisimple Lie algebra.

Definition 5.10. *Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$. We have the canonical embedding $ad : \mathfrak{g} \to gl(\mathfrak{g})$. Then for any $X \in \mathfrak{g}$, we define*

$$X_s^{ad} := ad^{-1}((adX)_s), \quad X_n^{ad} := ad^{-1}((adX)_n).$$

This definition is well-defined since $(adX)_s, (adX)_n \in ad(\mathfrak{g})$ thanks to Theorem 5.9.

Remark 5.11. If $\mathfrak{g} \subset gl_r(\mathbb{C})$ and $X \in \mathfrak{g}$, then we have $X_s = X_s^{ad}, X_n = X_n^{ad}$ thanks to the result in the Step 3 in the proof of Proposition 4.3.

Hence, from now on, we denote $X_s := X_s^{ad}, X_n := X_n^{ad}$ for all $X \in \mathfrak{g}$, where $\mathfrak{g}$ is a general semisimple Lie algebra.

Corollary 5.12. *If $\phi : \mathfrak{g} \to gl(V)$ is a representation of the semisimple Lie algebra $\mathfrak{g}$, then if $X \in \mathfrak{g}$ and $X = X_s + X_n$ is its abstract Jordan decomposition, $\phi(X) = \phi(X_s) + \phi(X_n)$ is the Jordan decomposition of the operator $\phi(X)$.*

Proof. From Theorem 4.10, we know $\mathfrak{g} = \mathfrak{g}^1 \oplus \cdots \oplus \mathfrak{g}^m$, where g_i are simple. We can write $X = X^1 + \cdots + X^m$ with $X_i \in \mathfrak{g}^i$. Each X^i has its own Jordan decomposition $X^i = X_s^i + X_n^i$, where $X_s^i, X_n^i \in \mathfrak{g}^i$. By the mutual commutativity of the $\mathfrak{g}^i$, we therefore have that $X = \sum_{k=1}^m X_s^k + \sum_{k=1}^m X_n^k$ is the Jordan decomposition of X.

The representation ϕ is faithful on the subalgebra $\mathfrak{g}' = \sum_{i=1}^{m'} \mathfrak{g}^i$ and has kernel $\mathfrak{g}'' = \sum_{i=m'+1}^{n} \mathfrak{g}^i$. Put $X = X' + X''$, where $X' = X^1 + \cdots + X^{m'}$ and $X'' = X^{m'+1} + \cdots + X^n$.

Now $\phi|_{\mathfrak{g}'}$ is a faithful representation, so by Theorem 5.9, we have

$$\phi(X) = \phi(X') = \phi(X_s') + \phi(X_n') = \phi(X_s) + \phi(X_n),$$

which completes the proof. □

Proposition 5.13. *Let $\mathfrak{g}, \mathfrak{g}^\sharp$ be semisimple Lie algebras over $\mathbb{C}$ and $\varphi : \mathfrak{g} \to \mathfrak{g}^\sharp$ be a Lie algebra homomorphism. Then we have*

$$\varphi(X)_s = \varphi(X_s), \quad \varphi(X)_n = \varphi(X_n).$$

Proof. From Theorem 4.10, we know $\mathfrak{g} = \mathfrak{g}^1 \oplus \cdots \oplus \mathfrak{g}^m$ and $\mathfrak{g}^\sharp = \mathfrak{g}^{\sharp 1} \oplus \cdots \oplus \mathfrak{g}^{\sharp n}$, where $\mathfrak{g}^i$ and $\mathfrak{g}^{\sharp j}$ are simple.

Let $X \in \mathfrak{g}$ and we have the decomposition $X = X^1 + \cdots + X^m$, which implies that $ad(X) = ad(X^1) + \cdots + ad(X^m)$ Moreover, $X_s = \oplus_i X_s^i$ and $X_n = \oplus_i X_n^i$.

From the second part of Corollary 4.11, we know for any $1 \le i \le m$, there exists j so that $\varphi : \mathfrak{g}^i \to \mathfrak{g}^{\sharp j}$ is an isomorphism. Hence, it reduces to to prove that for an isomorphism $\tau : \mathfrak{g} \to \mathfrak{g}'$, $\tau(X_s) = X_s'$ and $\tau(X_n) = X_n'$ provided that $\tau(X) = X'$.

Note that the diagram

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{ad} & ad(\mathfrak{g}) \\ \tau\downarrow & & \downarrow\tau^* \\ \mathfrak{g}' & \xrightarrow{ad'} & ad'(\mathfrak{g}') \end{array}$$

commutes. By definition of X_s, in order to prove $\tau(X_s) = X'_s$, it suffices to prove $ad' \circ \tau\,(X_s) = (ad'(\tau(X)))_s$. Since $ad' \circ \tau$ is a representation of $\mathfrak{g}$, by the previous corollary, we have $ad' \circ \tau(X_s) = (ad'(\tau(X)))_s$, which completes the proof.

□

To close this section, we recall that at the beginning of this subsection, we showed that $ad : \mathfrak{g} \to gl(\mathfrak{g})$ is semisimple provided that $\mathfrak{g}$ is semisimple. Actually, it relates to a useful definition.

Definition 5.14. *We say $\mathfrak{g}$ is reductive if $ad : \mathfrak{g} \to gl(\mathfrak{g})$ is semisimple.*

For more equivalent characterizations, see [6, Exercise 9.25] and [3, Section 6.4, Proposition 5].

6. Irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$

In this section, we follow [8, Section 7]. As usual, we denote $X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and we have $[H, X] = 2X, [H, Y] = (-2)Y, [X, Y] = H$.

Let $\varphi : \mathfrak{sl}_2 \to gl(V)$ be an irreducible representation, then $\varphi(H) = \varphi(H_s) = \varphi(H)_s$ is diagonalizable thanks to [8, Corollary 6.4], that is, $V = \oplus_{\alpha\in\mathbb{C}} V_\alpha$, where

$$V_\alpha := \{v \in V : \varphi(H)v = \alpha v\}$$

is the eigenspace corresponds to α and the sum is a finite sum over eigenvalues.

Now we compute the action of X and Y on V_α. For $v \in V_\alpha$, we have

$$H(X(v)) = X(H(v)) + [H, X](v) = \alpha(X(v)) + 2X(v) = (\alpha + 2)X(v).$$

Hence, $X : V_\alpha \to V_{\alpha+2}$. Analogously, $Y : V_\alpha \to V_{\alpha-2}$.

Lemma 6.1. *If $V \in Irr(\mathfrak{sl}_2(\mathbb{C}))$, then there exists $\alpha \in \mathbb{C}$ such that $V = \oplus_{i=1}^k V_{\alpha_0+2i}$ for some $k \in \mathbb{N}_{\geq 0}$ with $V_{\alpha_0+2k} \neq \{0\}$.*

Proof. For all eigenvalues α_l of H, we observe that $\oplus_{i=-\infty}^{\infty} V_{\alpha_l+2i}$ is a $\mathfrak{g}$-stable subspace of V. However, since V is irreducible, we know $\oplus_{i=-\infty}^{\infty} V_{\alpha_l+2i} = V$ for all l. Hence, for any l, m, there exists $j \in \mathbb{Z}$ such that $\alpha_l + 2j = \alpha_m$ and since the number of eigenvalues is finite, so there exists $k \in \mathbb{N}_{\geq 0}$ such that $V = \oplus_{i=1}^k V_{\alpha_0+2i}$. □

Remark 6.2. Until now, we cannot expect $\dim V_{\alpha_0+2i} \neq 0$ for all $1 \leq i \leq k$. However, we will see this is indeed true soon.

Set $n = \alpha_0 + 2k$, then pick any $v \in V_n$, we have $X(v) = 0$. Consider the set

$$S := \{v, Y(v), \cdots, Y^i(v), \cdots\} \subset V,$$

then we claim the following properties hold.

Lemma 6.3. *We have*

$$HY^i(v) = (n-2i)Y^i(v), \quad XY^i(v) = i(n-i+1)Y^{i-1}(v).$$

In particular, S spans V.

Proof. We compute

$$HY^i(v) = HYY^{i-1}(v) = Y\left(HY^{i-1}(v)\right) - 2YY^{i-1}(v),$$

then by induction on i and $Hv = nv$ we know the first assertion holds. On the other hand, the relation

$$XY^i(v) = YXY^{i-1}(v) + HY^{i-1}(v) = Y\left(XY^{i-1}(v)\right) + (n-2i+2)Y^{i-1}(v),$$

also allow us to derive the second assertion by induction.

Hence, span S is $\mathfrak{g}$-stable, so span $S = V$. □

Corollary 6.4. *For all $\alpha \in \mathbb{C}$, $\dim V_\alpha \leq 1$.*

Proof. Let $n = \alpha_0 + 2k$. Since $Y^i(v) \in V_{n-2i}$ for all i, and span $S = V$, we know V_{n-2i} is spanned by $Y^i(v)$. Hence, $\dim V_{n-2i} \leq 1$ and it is non-zero if $Y^i(v) \neq 0$. □

Corollary 6.5. *Let $n = \alpha_0 + 2k \in \mathbb{N}_{\geq 0}$. Then* $\dim V = n+1$.

Proof. Consider $m = \min\{i \in \mathbb{N} \geq 1 : Y^i(v) = 0\}$, we have $0 = XY^m(v) = m(n-m-1)Y^{m-1}(v)$ and $Y^{m-1}(v) \neq 0$. Hence, $n = m-1$ and $\dim V = m = n+1$. □

Then we have the following main theorem, which can be found in [10, Theorem 4.59].

Theorem 6.6. *There exists a bijection $Irr(\mathfrak{sl}_2(\mathbb{C}))/\sim \to \mathbb{N}_{\geq 0}$ that maps V to n, where $V = V_{-n} \oplus V_{-n+2} \oplus \cdots V_n$ and n is the largest eigenvalue of H. More precisely, we have these two claims:*

(1) For any $n \geq 0$, let $V^{(n)}$ be the finite-dimensional vector space with basis $v^0, v^1, \cdots, v^n$. Define the action of $\mathfrak{sl}_2(\mathbb{C})$ by

$$\begin{aligned} Hv^i &= (n-2i)v^i; \\ Yv^i &= v^{i+1}, i < n; \quad Yv^n = 0; \\ Xv^i &= i(n-i+1)v^{i-1}, i > 0; \quad Xv^0 = 0. \end{aligned} \tag{6.1}$$

Then $V^{(n)}$ is an irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$. We call it the irreducible representation with highest weight n.

(2) For $n \neq m$, representation $V^{(n)}$, $V^{(m)}$ are non-isomorphic.

(3) Every finite dimensional irreducible representation of $\mathfrak{sl}_2(\mathbb{C})$ is isomorphic to one of representations $V^{(n)}$.

In particular, all eigenvalues of V are all integers and each occurs along with its negative (an equal number of times).

Proof. (1) Notice that if $n = 0$, then it is the trivial representation. For $n > 0$, in order to check that $V^{(n)}$ is indeed a representation of $\mathfrak{sl}_2(\mathbb{C})$, it suffices to check that

$$Hv^i = [X,Y]v^i, \quad 2Xv^i = [H,X]v^i, \quad -2Ye^i = [H,Y]e^i.$$

Now we prove irreduciblility. Suppose $W \subset V^{(n)}$ is a submodule, then pick $0 \neq w \in W$. Suppose $w = a_l v^l + a_{l+1} v^{l+1} + \cdots + a_n v^n$, where $a_l \neq 0, l \geq 0$. Then $X^n Y^{n-l} w = c a_l v^0$, which implies $v^0 \in W$. Hence, $W = V^{(n)}$, as required.

(2) For $n \neq m$, $\dim V^{(n)} \neq \dim V^{(m)}$. Therefore they are not isomorphic.

(3) The injectivity part follows from the discussion above.

$\square$

As an easy corollary, we can see the dimension map $Irr(\mathfrak{sl}_2(\mathbb{C}))/\sim \longrightarrow \mathbb{N}_{\geq 1}$ is a bijection such that $V \mapsto \dim V = n + 1$.

Corollary 6.7. *For any representation $\sigma : \mathfrak{sl}_2 \to gl(V)$, we have*

$$V = \oplus_{n \geq 0} m(\sigma, V^{(n)}) V^{(n)},$$

where $V^{(n)}$ is the unique irreducible representation of dimension $n+1$ as in Theorem 6.6, $\sigma_n := \{v \in V : \sigma(H)v = nv\}$. Then we have $\sum_{n \geq 0} m(\sigma, V^{(n)}) = \dim \sigma_0 + \dim \sigma_1$.

Proof. Using Weyl's thoerem, we can decompose any representation σ into irreducible ones in the form $V = \oplus_{n \geq 0} m(\sigma, V^{(n)}) V^{(n)}$. Then notice that each irreducible $\mathfrak{g}$-module has a unique occurence of either the weight 0 or else the weight 1 (but not both) by applying Theorem 6.6 to each irreducible subrepresentation (submodule), thus $\sum_{n \geq 0} m(\sigma, V^{(n)}) = \dim \sigma_0 + \dim \sigma_1$. $\square$

Now we discuss the concrete realization of $V^{(n)}$, which is the irreducible $n+1$-dim representation of $\mathfrak{sl}_2(\mathbb{C})$. We have $V^{(0)} : \mathfrak{sl}_2(\mathbb{C}) \to gl(\mathbb{C})$ is trivial by (6.1). For $V^{(1)}$, we know from (6.1) that $Hv^1 = -v^1, Hv^0 = v^0, Yv^0 = v^1, Yv^1 = 0$.

7. Cartan Subalgebras

In this section, we follow [16, Chapter 3].

7.1. **Cartan subalgebra, Regular element, Zariski topology.** It is natural to ask whether we can give a classification of $Irr(\mathfrak{g})$ for general semisimple Lie algebra $\mathfrak{g}$. Note that when $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$, what plays an important role is $H(a) = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}$. Let $X(b) = \begin{pmatrix} & b \\ & \end{pmatrix}$ and $Y(c) = \begin{pmatrix} & \\ c & \end{pmatrix}$. Then we have $ad(H(a))(X(b)) = 2a \cdot X(b)$, $ad(H(a))(Y(c)) = -2a \cdot Y(c)$ and $ad(H(a))(H(a')) = 0 \cdot H(a')$. It is more or less analogous to the simultaneous diagonalization of matrices. Thus, we want to get an analogue of span $\{H(a)\}$ for general semisimple Lie algebras.

Definition 7.1. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$. A subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is called a Cartan subalgebra if*

(1) $\mathfrak{h}$ is nilpotent;

(2) $\mathfrak{h} = N_{\mathfrak{g}}(\mathfrak{h})$, where $N_{\mathfrak{g}}(\mathfrak{h}) := \{X \in \mathfrak{g} : [X, \mathfrak{h}] \subset \mathfrak{h}\}$ is the normalizer.

Example 7.2. *(1) Suppose $\mathfrak{g}$ is nilpotent, then $\mathfrak{h} = \mathfrak{g}$ is a Cartan subalgebra.*

(2) Suppose $\mathfrak{g} = \mathfrak{b}_3(\mathbb{C})$, then $\mathfrak{h}_3 := \left\{ \begin{pmatrix} a & & \\ & b & \\ & & c \end{pmatrix} : a, b, c \in \mathbb{C} \right\}$ is a Cartan subalgebra.

(3) *Suppose* $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$, *then* $\{diag(a_1, \dots, a_n) : a_i \in \mathbb{C}, \sum a_i = 0\} \subset \mathfrak{g}$ *is a Cartan subalgebra.*

For $X \in \mathfrak{g} \neq \{0\}$, consider the non-degenerate representation $ad : \mathfrak{g} \to gl(\mathfrak{g})$. Let $P_X(t) := \det(tI_n - ad(X)) = \sum_{i=0}^n a_i(X)T^i$ with $n = \dim \mathfrak{g}$. Note $a_0(X) = 0$ since $ad(X)(X) = 0$, that is, 0 is an eigenvalue of $ad(X)$.

Definition 7.3. *We define the rank of* $\mathfrak{g}$ *as*

$$\text{rank } \mathfrak{g} = \min\{i : a_i(X) \neq 0 \text{ for some } X \in \mathfrak{g}\}.$$

Moreover, an element $X \in \mathfrak{g}$ *is called regular if* $a_{\text{rank } \mathfrak{g}}(X) \neq 0$.

Note that this definition is independent of the choice of basis of $\mathfrak{g}$ since different representations of $ad(X) \in gl(\mathfrak{g})$ are similar to each other and similar matrices have the same characteristic polynomial.

Remark 7.4. Obviously, $a_n = 1$. If rank $\mathfrak{g} = n$, then for all $X \in \mathfrak{g}$, $ad(X)$ is nilpotent. By Engel's theorem, we know $\mathfrak{g}$ is nilpotent if and only if every $X \in \mathfrak{g}$, $ad(X)$ is nilpotent (HW04 Problem 1). (Sketch of proof: $\Rightarrow$: $ad(X)^m(Y) \in \mathfrak{g}^m$; $\Leftarrow$: Since $ad(\mathfrak{g}) \subset \mathfrak{n}_r$, we have $ad(\mathfrak{g}_r) = 0$, which implies $\mathfrak{g}_{r+1} = 0$, so $\mathfrak{g}$ is nilpotent.) This tells us

$$\text{rank } \mathfrak{g} = n \iff \mathfrak{g} \text{ is nilpotent.}$$

Example 7.5. *If* $\mathfrak{g} = \mathfrak{sl}_2$, *we can calculate that for* $X = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$, $P_X(T) = T^3 - 4(a^2+bc)T$, *so rank* $\mathfrak{g} = 1$. *And* X *is regular if and only if* $a^2 + bc = \det(X) \neq 0$.

Consider

$$\mathfrak{g}_r = \{X \in \mathfrak{g} : X \text{ is regular}\} = \{X \in \mathfrak{g} : a_{\text{rank } \mathfrak{g}}(X) \neq 0\} \subset \mathfrak{g},$$

we introduce a new notion.

Definition 7.6 (Zariski topology)**.** *Let* $V \simeq \mathbb{C}^n$ *be the* n*-dimensional vector space over* $\mathbb{C}$. *Denote by* $\mathcal{A}$ *the algebra of complex polynomials in* n*-variables, viewed as functions over* V. *For any* $S \subset \mathcal{A}$, *define*

$$V(S) = \{v \in V : f(v) = 0 \text{ for every } f \in S\}.$$

Then

$$\mathcal{T}_{Zar} := \{V - V(S) : S \in \mathcal{A}\}$$

is a well-defined topology on V, *called the Zariski topology.*

One can easily check that the Zariski topology is coarser than the usual Euclidean topology on V.

Lemma 7.7. $\mathfrak{g}_r \subset \mathfrak{g}$ *is open, dense and connnected with respect to the Zariski topology on* $\mathfrak{g}$.

Proof. Openness is true since it follows immediately by taking $S = \{a_{\text{rank } \mathfrak{g}}\}$ in the definition of the Zariski topology.

For density, suppose not, then $U := \{X \in \mathfrak{g} : a_{\text{rank } \mathfrak{g}}(X) = 0\}$ contains an open set $\widetilde{U}$ with respect to the Zariski topology on $\mathfrak{g}$. By definition, $\widetilde{U}$ is also open in the Euclidean topology. Hence, $a_{\text{rank } \mathfrak{g}} = 0$ as a complex polynomial since it vanishes on a non-empty Euclidean open set of $\mathfrak{g}$.

For connectedness, suppose by contradiction that there exists $X, Y \in \mathfrak{g}_r$ lying in different components, then there exists two disjoint open sets $A, B \subset \mathfrak{g}_r$ such that $A \cup B = \mathfrak{g}_r$ and $X \in A, Y \in B$. Then we consider the straight complex line joining X to Y in $\mathfrak{g}$, denoted by $L(X,Y)$. Since $L(X,Y) \cap \mathfrak{g}_r = (L(X,Y) \cap A) \cup (L(X,Y) \cap B)$, we know $L(X,Y) \cap \mathfrak{g}_r$ is disconnected since $L(X,Y) \cap A$ and $L(X,Y) \cap B$ are both non-empty.

However, since $L(X,Y)$ can be parametrized by one variable t, we know each closed set in the Zariski topology only meets $L(X,Y)$ at finitely many points or contains the whole line $L(X,Y)$,

Hence, for closed sets A, B, we find $L(X,Y) \cap A$ and $L(X,Y) \cap B$ only contains finitely many points.

On the other hand, for the open set $\mathfrak{g}_r$, $L(X,Y) \cap \mathfrak{g}_r$ is just the line $L(X,Y)$ with at most finitely many points removed from it, which contradicts to $L(X,Y) \cap \mathfrak{g}_r = (L(X,Y) \cap A) \cup (L(X,Y) \cap B)$.

This completes the proof. □

Definition 7.8. *For any $X \in \mathfrak{g}$, $\lambda \in \mathbb{C}$, we say*

$$\mathfrak{g}_X^\lambda = \{Y \in \mathfrak{g} : (ad(X) - \lambda)^n(Y) = 0 \text{ for some } n \geq 1\}$$

is the nilspace of $ad(X)$ with respect to λ.

Note that $\dim \mathfrak{g}_X^\lambda$ is the multiplicity of 0 as an eigenvalue of $ad(X)$, that is, $\dim \mathfrak{g}_X^\lambda = \min\{i : a_i(X) \neq 0\}$. So $\dim \mathfrak{g}_X^\lambda \geq \operatorname{rank} \mathfrak{g}$ and the equality holds if and only if X is regular.

Lemma 7.9. *Fix $X \in \mathfrak{g}$, then*

(1) $\mathfrak{g} = \oplus_{\lambda \in \mathbb{C}} \mathfrak{g}_X^\lambda$;
(2) $[\mathfrak{g}_X^\lambda, \mathfrak{g}_X^\mu] \subset \mathfrak{g}_X^{\lambda+\mu}$ for all $\lambda, \mu \in \mathbb{C}$;
(3) $\mathfrak{g}_X^0 \subset \mathfrak{g}$ is a Lie subalgebra.

Proof. The first assertion follows immediately from the classical Jordan form theory in linear algebra. The third assertion is an easy corollary of the second. For the second one, it suffices to show

$$(ad(X) - \lambda - \mu)^n[Y, Z] = \sum_{i=0}^{n} \binom{n}{i} \left[(ad(X) - \lambda)^i(Y), (ad(X) - \mu)^{n-i}(Z)\right] \tag{7.1}$$

for $Y \in \mathfrak{g}_X^\lambda$, $Z \in \mathfrak{g}_X^\mu$. We prove this by induction. For $n = 1$, by Jacobi identity, we have

$$(ad(X) - \lambda - \mu)[Y, Z] = [(ad(X) - \lambda)Y, Z] + [Y, (ad(X) - \mu)Z].$$

Then it follows from induction by applying the fundamental formula $\binom{n}{i} = \binom{n-1}{i-1} + \binom{n-1}{i}$. □

Theorem 7.10. *If $X \in \mathfrak{g}$ is regular, then $\mathfrak{g}_X^0 \subset \mathfrak{g}$ is a Cartan subalgebra.*

Proof. **Step 1:** We claim $\mathfrak{g}_X^0$ is nilpotent. By Engel's theorem, it suffices to show for all $Y \in \mathfrak{g}_X^0$, the restriction $ad(Y)|_{\mathfrak{g}_X^0} : \mathfrak{g}_X^0 \to \mathfrak{g}_X^0$ is nilpotent.

Let $\widetilde{ad}(Y) : \mathfrak{g}/\mathfrak{g}_X^0 \to \mathfrak{g}/\mathfrak{g}_X^0$ be a vector space homomorphism induced by $ad(Y)$, then we put

$$U = \{Y \in \mathfrak{g}_X^0 : ad(Y)|_{\mathfrak{g}_X^0} : \mathfrak{g}_X^0 \to \mathfrak{g}_X^0 \text{ is not nilpotent}\} = \{Y \in \mathfrak{g}_X^0 : \left(ad(Y)|_{\mathfrak{g}_X^0}\right)^{\dim \mathfrak{g}_X^0} \neq 0\},$$

$$V = \{Y \in \mathfrak{g}_X^0 : \widetilde{ad}(Y) \text{ is invertible}\} = \{Y \in \mathfrak{g}_X^0 : \det\left(\widetilde{ad}(Y)\right) \neq 0\}.$$

It follows from the second equalities in both equations above that $U, V \in \mathfrak{g}$ are both open with respect to the Zariski topology by noting the relation between U, V and the zeros of certain polynomials respectively.

Moreover, we claim that $X \in V$. It follows from Lemma 7.9 (1) and (2) that for any $Y \in \mathfrak{g}$ such that $ad(X)(Y) \in \mathfrak{g}_X^0$, we have $Y \in \mathfrak{g}_X^0$, and hence $\ker \widetilde{ad}(X) = 0$. In particular, $V \neq \varnothing$.

By the same type of argument as in the proof of Lemma 7.7, we show V is dense. Note that V^c is a set of solutions of a polynomial equation, which implies its interior is empty, so V is dense.

Now, it suffices to show $U \cap V = \varnothing$ to conclude $U = \varnothing$. Suppose not, choose $Y \in U \cap V$, then $ad(Y)|_{\mathfrak{g}_X^0}$ has 0 as an eigenvalue with multiplicity strictly less than $\dim \mathfrak{g}_X^0 = \operatorname{rank} \mathfrak{g}$. On the other hand, 0 is not an eigenvalue of $\widetilde{ad}(Y) : \mathfrak{g}/\mathfrak{g}_X^0 \to \mathfrak{g}/\mathfrak{g}_X^0$, thus it follows from these two results that the multiplicity of 0 as an eigenvalue of $ad(Y) : \mathfrak{g} \to \mathfrak{g}$ is strictly less than rank $\mathfrak{g}$. However, by the definition of rank, we know

$$\text{the multiplicity of 0 as an eigenvalue of } ad(Y) \geq \operatorname{rank} \mathfrak{g},$$

which is a contradiction.

Step 2: Now we show $\mathfrak{g}_X^0 = N_{\mathfrak{g}}(\mathfrak{g}_X^0)$. For all $Z \in N_{\mathfrak{g}}(\mathfrak{g}_X^0)$, we have $[Z, \mathfrak{g}_X^0] \subset \mathfrak{g}_X^0$. In particular, $ad(X)(Z) \in \mathfrak{g}_X^0$. By definition, there exists p such that $ad^p(X)ad(X)(Z) = 0 \in \mathfrak{g}_X^0$. Therfore, $Z \in \mathfrak{g}_X^0$, which completes the proof. ☐

7.2. **Conjugacy of Cartan Subalgebras.** Since we have $\mathfrak{g} \xrightarrow{ad} gl(\mathfrak{g}) = End(\mathfrak{g}) \xrightarrow{\exp} GL(\mathfrak{g})$, we define

$$G^0 := \langle \exp \circ ad(\mathfrak{g}) \rangle \subset GL(\mathfrak{g})$$

as the subgroup generated by $\exp \circ ad(\mathfrak{g})$.

In fact, G^0 acts transitively on {all Cartan subalgebras of $\mathfrak{g}$}. This fact can be found at [16].

In the following three propositions, we use the following notations. Let $\mathfrak{h} \subset \mathfrak{g}$ be a Cartan subalgebra. For $X \in \mathfrak{h}$, we write $ad(X)|_{\mathfrak{h}} : \mathfrak{h} \to \mathfrak{h}$ and $\widetilde{ad}(X) : \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$.

Proposition 7.11. *Let $\mathfrak{h} \subset \mathfrak{g}$ be a Cartan subalgebra. Then the set*

$$V_{\mathfrak{h}} = \{X \in \mathfrak{h} : \widetilde{ad}(X) \textit{ is invertible}\} \neq \varnothing.$$

Proof. **Step 1:** In particular, $\mathfrak{h}$ is nilpotent, then $ad(\mathfrak{h})$ is nilpotent by definition. Thus $\widetilde{ad}(\mathfrak{h}) \subset gl(\mathfrak{g}/\mathfrak{h})$ is nilpotent and hence solvable. Now it follows from Lie's theorem that there exists a complete flag

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

of $\widetilde{ad}(\mathfrak{h})$-modules such that $\dim D_{i+1}/D_i = 1$ and the action $\widetilde{ad}(\mathfrak{h}) \subsetneq D_i/D_{i-1}$ is given by

$$X \cdot Z_i := \widetilde{ad}(X)(Z_i) = \alpha_i(X) Z_i \mod D_{i-1}$$

for all $Z_i \in D_i$, $X \in \mathfrak{h}$, where $\alpha_i : \mathfrak{h} \to \mathbb{C}$. Hence, the eigenvalues of $\widetilde{ad}(X)$ are $\alpha_1(X), \cdots, \alpha_n(X)$. Then, it suffices to show α_i is not identically 0 for any i.

Step 2: In the following steps, we suppose by contradiction that there exists some k such that α_k is identically zero and $\alpha_1, \dots, \alpha_{k-1}$ are not. Then for $i = 1, \dots, k-1$, $\ker(\alpha_i)$ are of codimension 1 in $\mathfrak{h}$, which are hyperplanes in $\mathfrak{h}$, then $\bigcup_{i=1}^{k-1} \ker(\alpha_i)$ is a proper subset of $\mathfrak{h}$. Hence, there exists $X_0 \in \mathfrak{h}$ such that $\alpha_i(X_0) \neq 0$ for all $i = 1, \cdots, k-1$.

Then $\widetilde{ad}(X_0) : D_{k-1} \to D_{k-1}$ is invertible and $\widetilde{ad}(X_0) : D_k \to D_k$ has 0 as an eigenvalue of multiplicity 1.

Let

$$D' = \{Y \in D_k : \left(\widetilde{ad}(X_0)|_{D_k}\right)^n (Y) = 0 \text{ for some } n > 0\}$$

be the nilspace of $\widetilde{ad}(X_0)|_{D_k}$, which is of dimension 1 and disjoint with D_{k-1}. Hence, $D_k = D_{k-1} \oplus D'$. Moreover, we know D' is the nilspace of $\widetilde{ad}(X_0)|_{D_k}$.

Step 3: We claim that for all $Z \in D'$, we have $\widetilde{ad}(X)(Z) = 0$ for all $X \in \mathfrak{h}$. This is obvious valid for $X = X_0$ by the definition of D'. Now one can check by induction and the fact $\widetilde{ad}(X_0)(Z) = 0$ that

$$\widetilde{ad}(X_0)^n \left(\widetilde{ad}(X)(Z)\right) = \widetilde{ad}\left(ad(X_0)^n(X)\right)(Z).$$

On the other hand, since $\mathfrak{h}$ is nilpotent, we know $ad(X_0)^n(X) = 0$ for n sufficiently large. Hence, $\widetilde{ad}(X_0)^n \left(\widetilde{ad}(X)(Z)\right) = 0$, which implies $\widetilde{ad}(X)(Z) \in D'$.

But $\widetilde{ad}(X)(D_k) \subset D_{k-1}$, so $\widetilde{ad}(X)(Z) = 0$. Thus this claim is true.

Step 4: It follows from the claim that $[\mathfrak{h}, Z] \subset \mathfrak{h}$, that is, $Z \in N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$ since $\mathfrak{h}$ is a Cartan subalgebra. However, $Z \notin \mathfrak{h}$, so this leads to a contradiction. ☐

As an easy corollary, it follows from the definiton of $V_{\mathfrak{h}}$ that $V_{\mathfrak{h}}$ is a Zariski open set in $\mathfrak{h}$ since the invertibility is related to the fact that the determinant polynomial has no zeros.

Proposition 7.12. *The set $W_{\mathfrak{h}} := G^0 \cdot V_{\mathfrak{h}} \subset \mathfrak{g}$ is open in $\mathfrak{g}$.*

Proof. We follow the idea of the proof of [6, Theorem D.22].

Step 1: We do some preparation work at first. For $\alpha \in Hom(\mathfrak{h}, \mathbb{C})$, we define

$$\mathfrak{g}_{\mathfrak{h}}^{\alpha} := \{Y \in \mathfrak{g} : \text{for every } X \in \mathfrak{h}, \left(ad(X) - \alpha(X)\right)^n (Y) = 0 \text{ for some } n = n_X\}.$$

Then we have the following facts

(1) $\mathfrak{g} = \oplus_{\alpha \in Hom(\mathfrak{h}, \mathbb{C})} \mathfrak{g}_{\mathfrak{h}}^{\alpha}$;
(2) $[\mathfrak{g}_{\mathfrak{h}}^{\alpha}, \mathfrak{g}_{\mathfrak{h}}^{\beta}] \subset \mathfrak{g}_{\mathfrak{h}}^{\alpha+\beta}$;
(3) $\mathfrak{g}_{\mathfrak{h}}^{0} = \mathfrak{h}$,

where the first fact can be found in [5, Chapter VII, Section 1.3, Proposition 9], and we can deduce the second one by (7.1), the third one by $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$.

Step 2: In fact, we can prove a more stronger argument. Consider the map

$$F : \mathfrak{g}_{\mathfrak{h}}^{\alpha_1} \times \cdots \times \mathfrak{g}_{\mathfrak{h}}^{\alpha_l} \times \mathfrak{h} \to \mathfrak{g}, \quad F(X_{\alpha_1}, \cdots, X_{\alpha_l}, Y) = \left(\exp \circ ad(X_{\alpha_1})\right) \circ \cdots \circ \left(\exp \circ ad(X_{\alpha_l})\right)(Y).$$

Note that

$$\left\{ \prod_{\alpha \in Hom(\mathfrak{h}, \mathbb{C})} \left(\exp \circ ad(X_\alpha)\right) : X_\alpha \in \mathfrak{g}_{\mathfrak{h}}^{\alpha} \right\}$$

is a subset of G^0. Then it suffices to prove $F(\mathfrak{g}_{\mathfrak{h}}^{\alpha_1} \times \cdots \times \mathfrak{g}_{\mathfrak{h}}^{\alpha_l} \times V_{\mathfrak{h}})$ contains a Zariski open set in $\mathfrak{g}$.

Step 3: From the property (2) in Step 1, we know $ad(X_{\alpha_j})$ is nilpotent for all $X_{\alpha_j} \in \mathfrak{g}_{\mathfrak{h}}^{\alpha_j}$ since $ad(X_{\alpha_j})^k(\mathfrak{g}_{\mathfrak{h}}^{\beta})$ lies in $\mathfrak{g}_{\mathfrak{h}}^{k\alpha+\beta}$, which is a trivial set for sufficiently large k. Hence, $\exp \circ ad(X_\alpha)$ is a polynomial in $ad(X_\alpha)$.

Since F is a polynomial mapping from one complex vector space to another of the same dimension thanks to the property (1) in Step 1 and $V_{\mathfrak{h}}$ is a nonempty Zariski open set in $\mathfrak{h}$ thanks to Proposition 7.11, so it suffices to show the derivative $F_*|_p$ is invertible at some point p by the following fact in Algebraic Geometry: if $F : \mathbb{C}^N \to \mathbb{C}^N$ is a polynomial whose for derivative $F_*|_p$ is invertible at some point p, then for any nonempty Zariski open set $U \subset \mathbb{C}^N$, $F(U)$ contains a nonempty Zariski open set. For more details, refer to [5, Chapter VII, Appendix I, Section 2, Corollary] or [7, Chapter III, Proposition 10.4, Corollary 10.7].

Step 4: Now it suffices to show the differential map is surjective.

Consider the differential map at any $p \in \mathfrak{g}_{\mathfrak{h}}^{\alpha_1} \times \cdots \times \mathfrak{g}_{\mathfrak{h}}^{\alpha_l} \times \mathfrak{h}$, with the identification of the tangent spaces and the vector spaces themselves:

$$F_*|_p : \mathfrak{g}_{\mathfrak{h}}^{\alpha_1} \times \cdots \times \mathfrak{g}_{\mathfrak{h}}^{\alpha_l} \times \mathfrak{h} \to \mathfrak{g}.$$

Since the differential of the curve $F(p+(0, \cdots, Y, \cdots, 0, H))$ at $t = 0$ is $[Y, H] = -ad(H)(Y)$ and the differential of the curve $F(p + (0, \ldots, 0, H))$ at $t = 0$ is H, we know $Im(\widetilde{ad}(V_{\mathfrak{h}})) \subset Im(F_*|_p)$ and $\mathfrak{h} \subset Im(F_*|_p)$. By Proposition 7.11, we know that $\dim Im(\widetilde{ad}(V_{\mathfrak{h}})) + \dim \mathfrak{h} = \dim \mathfrak{g}$ so $Im(\widetilde{ad}(V_{\mathfrak{h}})) + \mathfrak{h} = \mathfrak{g}$, which implies $Im(F_*|_p) \supset \mathfrak{g}$. This completes the proof. □

The following proposition is more or less a converse of Theorem 7.10.

Proposition 7.13. *Let $\mathfrak{h} \subset \mathfrak{g}$ be a Cartan subalgebra, then there exists a regular $X \in \mathfrak{g}$ such that $\mathfrak{h} = \mathfrak{g}_X^0$.*

Proof. Since $W_{\mathfrak{h}} \subset \mathfrak{g}$ is open and $\mathfrak{g}_r$ is open and dense, there exists $a \in G^0, X \in V_{\mathfrak{h}} \subset \mathfrak{h}$ such that $a \cdot X \in W_{\mathfrak{h}} \cap \mathfrak{g}_r$. We choose a basis of $\mathfrak{g}$, namely $\{v_i\}$ and recall that the definition of regular elements, Definition 7.3, is independent of the choice of basis. Then for all $Y \in \mathfrak{g}$, $ad(Y)$ has a representation matrix $M_1(Y)$ with respect to $\{v_i\}$. Since $a \in GL(\mathfrak{g})$, we know $\{a \cdot v_i\}$ is also a basis of $\mathfrak{g}$, and then $ad(Y)$ has another representation matrix $M_2(Y)$ such that M_1, M_2 are similar. Note that $M_2(a \cdot X) = M_1(X)$, so X is also regular.

But $X \in V_{\mathfrak{h}}$ implies $ad(X) : \mathfrak{h} \to \mathfrak{h}$ is nilpotent and $\widetilde{ad}(X) : \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$ is invertible, so $\mathfrak{h} = \mathfrak{g}_X^0$, which completes the proof. □

7.3. **The semisimple case.** In this subsection, we suppose $\mathfrak{g}$ is semisimple and $\mathfrak{h} \subset \mathfrak{g}$ is a Cartan subalgebra over $\mathbb{C}$.

Theorem 7.14. *Suppose $\mathfrak{g}$ is semisimple. Let $\mathfrak{h} \subset \mathfrak{g}$ be a Cartan subalgebra over $\mathbb{C}$, then*
(1) $B_\kappa|_{\mathfrak{h}\times\mathfrak{h}}$ is non-degenerate;
(2) $\mathfrak{h}$ is abelian;
(3) $\mathfrak{h} = C_{\mathfrak{g}}(\mathfrak{h}) := \{Y \in \mathfrak{g} : [Y, X] = 0, \forall X \in \mathfrak{h}\}$;
(4) $\forall X \in \mathfrak{h}$, X is semisimple.

Proof. (1) By Proposition 7.13, we pick a regular element X such that $\mathfrak{h} = \mathfrak{g}_X^0$. Then $\mathfrak{g} = \mathfrak{g}_X^0 \oplus \left(\oplus_{\lambda\neq 0}\mathfrak{g}_X^\lambda\right)$. We claim that if $\lambda + \mu \neq 0$, then $B_\kappa(\mathfrak{g}_X^\lambda, \mathfrak{g}_X^\mu) = 0$. For $Y \in \mathfrak{g}_X^\lambda$,

$Z \in \mathfrak{g}_X^{\mu}$, we use adjoint to get $B_\kappa(ad(X)Y, Z) + B_\kappa(Y, ad(X)Z) = 0$, which implies $(\lambda + \mu)B_\kappa(Y, Z) = 0$.

Hence, we have a decomposition of $\mathfrak{g}$ into mutually orthogonal subspaces with respect to B_κ that $\mathfrak{g} = \mathfrak{g}_X^0 \oplus \left(\oplus_{\lambda \neq 0} \left(\mathfrak{g}_X^{\lambda} \oplus \mathfrak{g}_X^{-\lambda}\right)\right)$. Since B_κ is non-degenerate on $\mathfrak{g}$ by Theorem 4.7, then it is also non-degenerate on each of these subspaces by the orthogonality. In particular, $B_\kappa|_{\mathfrak{h}\times\mathfrak{h}}$ is non-degenerate.

(2) Since $\mathfrak{h}$ is nilpotent, Theorem 4.4 implies $Tr(ad(X) \circ ad(Y)) = 0$ for $X \in \mathfrak{h}$, $Y \in [\mathfrak{h}, \mathfrak{h}]$, that is, $[\mathfrak{h}, \mathfrak{h}] \perp \mathfrak{h}$ with respect to $B_\kappa|_{\mathfrak{h}\times\mathfrak{h}}$. Since $B_\kappa|_{\mathfrak{h}\times\mathfrak{h}}$ is non-degenerate, $[\mathfrak{h}, \mathfrak{h}] = 0$.

(3) Since $\mathfrak{h}$ is a Cartan subalgebra, we have $\mathfrak{h} \subset C_{\mathfrak{g}}(\mathfrak{h}) \subset N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$.

(4) Let $X = X_s + X_n \in \mathfrak{h}$ is the Jordan decomposition, then it suffices to show $X_n = 0$. For all $Y \in \mathfrak{h}$, we have $[Y, X] = 0$, then $[Y, X_s] = [Y, X_n] = 0$. So $X_s, X_n \in C_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$.

Since $[Y, X_n] = 0$, we know $ad(Y) \circ ad(X_n) = ad(X_n) \circ ad(Y)$, so $ad(Y) \circ ad(X_n)$ is nilpotent by the fact that X_n is nilpotent. Hence, $B_\kappa(Y, X_n) = 0$, which implies $X_n \in \mathfrak{h}^{\perp}$, that is, $X_n = 0$.

□

Corollary 7.15. *Let $\mathfrak{h}, \mathfrak{g}$ be as above, then we have*

(1) $\mathfrak{h} \subset \mathfrak{g}$ is a maximal abelian subalgebra;

(2) Every regular element of $\mathfrak{g}$ is semisimple.

Proof. (1) It follows from Theorem 7.14 (3) immediately.

(2) It follows from Theorem 7.10 and Theorem 7.14(4) immediately.

□

8. Root Systems

In this section, we follow [16, Chapter V].

8.1. Root systems over $k = \mathbb{R}, \mathbb{C}$.

In this part, we follow [16, Chapter V, Section 1-4, 17]

Definition 8.1. *Suppose V is a vector space over k and $\alpha \in V \setminus \{0\}$. A symmetry with α is an element $s_\alpha \in GL_k(V)$ such that*

(1) $s_\alpha(\alpha) = -\alpha$;

(2) $H_\alpha := \{v \in V : s_\alpha(v) = v\}$ is a hyperplane of V, that is, a subspace of codimension 1.

Obviously, $\alpha \notin H_\alpha$, so H_α is a complement for the line $\mathbb{R}\alpha$ spanned by α. The symmetry s_α is completely determined by the choice of α and H_α.

Moreover, for $\alpha \in V \setminus \{0\}$, we have the following bijection

$$\{\text{symmetries with } \alpha \text{ in } V\} \leftrightarrow \{\alpha^* \in V^* : \langle \alpha^*, \alpha \rangle = 2\}.$$

The bijection is described as follows.

For any symmetry with α, denoted by s_α, suppose α^* is the unique element in V^* such that $\langle \alpha^*, \alpha \rangle = 2$ and $\ker \alpha^* = H_\alpha$. Then s_α is given by

$$s_\alpha(x) = x - \langle \alpha^*, x \rangle \alpha,$$

that is, $s_\alpha = id - \alpha^* \otimes \alpha$. On the other hand, for $\alpha^* \in V^* \setminus \{0\}$ such that $\langle \alpha^*, \alpha \rangle = 2$, $s_\alpha = id - \alpha^* \otimes \alpha$ is a symmetry with α with $H_\alpha = \ker \alpha^*$ is as desired since $\dim \ker \alpha^* = \dim V - \dim Im(\alpha^*) = \dim V - 1$.

Definition 8.2. *A finite set $R \subset V$ is said to be a root system in V if the following conditions are satisfied:*
(1) span $R = V$, $0 \notin R$;
(2) for any $\alpha \in R$, there exists a symmetry with α, namely s_α, such that $s_\alpha(R) \subset R$;
(3) for any $\alpha, \beta \in R$, $s_\alpha(\beta) - \beta \in \mathbb{Z} \cdot \alpha$, that is, $\langle \alpha^*, \beta \rangle \in \mathbb{Z}$, where α^* is as above.
The rank of R is defined as the dimension of V.

Note that the second condition $s_\alpha(R) \subset R$ implies that $s_\alpha(R) = R$ since $s_\alpha^2 = id$. Moreover, such symmetry with α in the second condition is actually unique by the following lemma.

Lemma 8.3. *Let $\alpha \in V \setminus \{0\}$, and let R be a finite subset of V which spans V. There is at most one symmetry with α which leaves R invariant.*

Proof. Let s_α, s'_α be two symmetries with α having the properties in Defintion 8.2. Then $u = s_\alpha \circ s'_\alpha$ satisfies

$$u(R) = R, \quad u(\alpha) = \alpha, \quad u : V/k \cdot \alpha \to V/k \cdot \alpha \text{ is the identity map,}$$

where the third assertion follows from the third property in Definition 8.2. Hence, all the eigenvalues of u are 1. Since $u^n(R)$ is a permutation of R for all $n \in \mathbb{N}$, then there exists some $n \geq 0$, $u^n = id$ in $GL_k(V)$. So u is diagonalizable and hence $u = 1$. Since $s_\alpha^2 = id$, we know $s_\alpha = s'_\alpha$. □

Definition 8.4. *We say a root system R is reduced if $R \cap k\alpha = \pm\alpha$.*

If a root system R is not reduced, then it contains $\alpha, t\alpha$ with $0 < |t| < 1$, then by the third condition in Definition 8.2, $2t \in \mathbb{Z}$, which implies $t = \frac{1}{2}$. Then the roots proportional to α are simply $-\alpha, -\frac{1}{2}\alpha, \frac{1}{2}\alpha, \alpha$.

Remark 8.5. The reduced roots systems are those which arise in the theory of semisimple Lie algebras (or algebraic groups) over an algebraically closed field; they are the only ones we shall need. Nonreduced systems occur when one no longer assumes that the base field is algebraically closed.

Definition 8.6. *We say $W := \langle s_\alpha : \alpha \in R \rangle \subset GL_k(V)$ is the Weyl group of R, which is the group generated by all symmetries with α associated with R.*

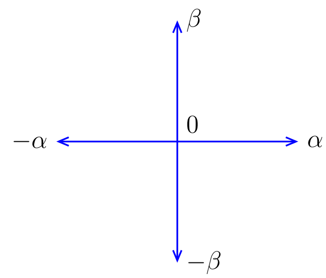
Since R spans V and all $s_\alpha \in W$ leaves R invariant, the Weyl group W can be identified with a subgroup of the group of all permutations of R, which is hence a finite group.

Example 8.7. *Here are some examples for root systems.*

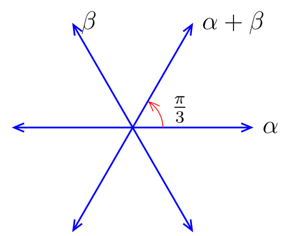
(1) For $V = \mathbb{R}$, $R = \{\pm\alpha\}$ is a root system. In this example, s_α is the usual reflection. The Weyl group is S_2.

$$-\alpha \xleftrightarrow{\quad 0 \quad} \alpha$$

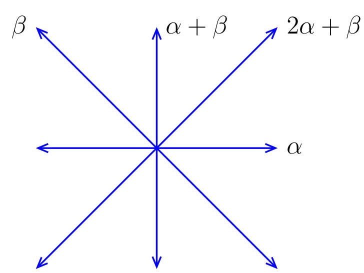
(2) For $V = \mathbb{R}^2$, $R = \{\pm\alpha, \pm\beta\}$ is a root system. In this example, s_α, s_β are all the usual reflections. The Weyl group is $S_2 \times S_2$.



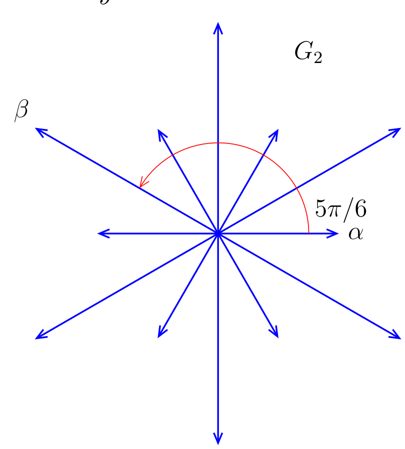
(3) For $V = \mathbb{R}^2$, $R = \{\pm\alpha, \pm\beta, \pm(\alpha+\beta)\}$ is a root system. In this example, s_γ is the reflection along the line perpendicular to γ. The Weyl group is D_6, which is the dihedral group of order 6.



(4) For $V = \mathbb{R}^2$, $R = \{\pm\alpha, \pm\beta, \pm(\alpha+\beta), \pm(2\alpha+\beta)\}$ is a root system. In this example, s_γ can be defined using (8.2), where (γ_1, γ_2) in (8.1) is the Euclidean inner product for $\gamma_1, \gamma_2 \in R$. One can check s_γ is the reflection along the line $\gamma^\perp$. The Weyl group is D_8, which is the dihedral group of order 8.



(5) For $V = \mathbb{R}^2$, there is another root system.



8.2. **Root systems over $\mathbb{R}$.** From now on, the ground field $k = \mathbb{R}$. We follow [16, Chapter V, Section 5-7] in this part.

In the previous definition of root systems, one need to verify some conditions related to V^*. Now we seek for a way to define the root systems in a more intrisic way.

Proposition 8.8. *Let R be a root system in V. There is a positive definite symmetric bilinear form $(-,-)$ on V which is invariant under the Weyl group W of R.*

Proof. We choose an arbitrary positive definite symmetric bilinear form on V, namely $B(X,Y)$. Then we define

$$(X,Y) := \sum_{s_\gamma \in W} B(s_\gamma(X), s_\gamma(Y)) \tag{8.1}$$

which is a well-defined positive definite W-invariant symmetric bilinear form on V since W is a finite group. $\square$

Note that $\alpha' = 2\frac{\alpha}{(\alpha,\alpha)}$ is an element in V^* defined by $\langle \alpha', x\rangle = 2\frac{(x,\alpha)}{(\alpha,\alpha)}$, then $\langle \alpha', \alpha\rangle = 2$. Moreover, for all $h \in H$, by the W-invariance of $(-,-)$, we have $(h,\alpha) = (s_\alpha(h), s_\alpha(\alpha)) = (h,-\alpha) = -(h,\alpha)$. which implies $H \perp \mathbb{R}\alpha$ with respect to $(-,-)$, that is, $H = \ker \alpha'$. On the other hand, we know from the previous discussion that there exists a unique $\alpha^* \in V^*$ such that $s_\alpha = id - \alpha^* \otimes \alpha$ and $H = \ker \alpha^*$. Hence, $\alpha^* = \alpha'$ in V^*. So we deduce from this that we have

$$s_\alpha(x) = x - 2\frac{(\alpha,x)}{(\alpha,\alpha)}\alpha, \quad \forall x \in V. \tag{8.2}$$

Now, the third condition in Definition 8.2 can be interpreted as

$$2\frac{(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}, \quad \forall \alpha,\beta \in R.$$

Now we discuss the relation of any two roots α, β in R. Put

$$n(\beta,\alpha) = \langle \alpha^*, \beta\rangle = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z},$$

which is equivalent to $n(\beta,\alpha) = 2\frac{|\beta|}{|\alpha|}\cos\theta \in \mathbb{Z}$. It follows that $n(\beta,\alpha)\cdot n(\alpha,\beta) = 4\cos^2\theta \in \mathbb{Z}$, which implies $\cos^2\theta \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$. More precisely, if α and β are not proportional to each other, the only possibilities are as follows:

$\cos\theta$	θ	$n(\beta,\alpha)$	$n(\alpha,\beta)$	$\lvert\beta\rvert/\lvert\alpha\rvert$
0	$\frac{\pi}{2}, \frac{3\pi}{2}$	0	0	
$\frac{1}{2}$	$\frac{\pi}{3}$	1	1	1
$-\frac{1}{2}$	$\frac{2\pi}{3}$	-1	-1	1
$\frac{\sqrt{2}}{2}$	$\frac{\pi}{4}$	2	1	$\sqrt{2}$
$-\frac{\sqrt{2}}{2}$	$\frac{3\pi}{4}$	-2	-1	$\sqrt{2}$
$\frac{\sqrt{3}}{2}$	$\frac{\pi}{6}$	3	1	$\sqrt{3}$
$-\frac{\sqrt{3}}{2}$	$\frac{5\pi}{6}$	-3	-1	$\sqrt{3}$

Proposition 8.9. *Let $\alpha, \beta \in R$ be two non-proportional roots. If $n(\beta,\alpha) > 0$, that is, $(\alpha,\beta) > 0$, they form an acute angle, then $\alpha - \beta \in R$.*

Proof. The above table shows that $n(\alpha,\beta) = 1$ or $n(\beta,\alpha) = 1$. Without loss of generality, we assume $n(\alpha,\beta) = 1$, then $\alpha - \beta = \alpha - n(\alpha,\beta)\beta = s_\beta(\alpha) \in R$, which completes the proof. $\square$

8.3. **Bases of a root system and Dynkin diagrams.**

Definition 8.10. *A subset $S \subset R$ is a base for R if*

(1) S is a basis for the vector space V;

(2) Each $\beta \in R$ can be written as a linear combination

$$\beta = \sum_{\alpha \in S} m_\alpha \alpha,$$

where either $m_\alpha \in \mathbb{Z}_{\geq 0}$ for all α or $m_\alpha \in \mathbb{Z}_{\leq 0}$ for all α.

Now we shall prove that there indeed exists a base. The proof is a direct construction with steps below:

(1) Choose $t \in V^*$ such that $\langle t, \alpha \rangle \neq 0$ for all $\alpha \in R$. (From the discussion above, it is also equivalent to pick $v \in V$ such that $(v, \alpha) \neq 0$ for all $\alpha \in R$.)

(2) Let $R_t^+ = R \cap t^{-1}(\mathbb{R}_{>0})$, $R_t^- = R \cap t^{-1}(\mathbb{R}_{<0})$. Since $t^{-1}(\{0\}) = \varnothing$, we get $R = R_t^+ \cup R_t^-$ and $V = t^{-1}(\mathbb{R}_{>0}) \cup t^{-1}(\mathbb{R}_{<0})$. Moreover, one can observe that $R_t^+ = -R_t^-$.

(3) We say an element $\alpha \in R_t^+$ is decomposable if there exists $\beta, \gamma \in R_t^+$ such that $\alpha = \beta + \gamma$; otherwise, α is said to be indecomposable. Let

$$S_t = \{\alpha \in R_t^+ : \alpha \text{ is indecomposable}\}.$$

In the following theorem, we prove that S_t is a base for R.

Theorem 8.11. *Let S_t be defined as above, then S_t is a base for R.*

Proof. **Step 1:** Let

$I = \{\alpha \in R_t^+ : \alpha$ is not a linear combination of elements of S_t with non-negative coefficients$\}$

and we want to show $I = \varnothing$. Suppose not, then there exists an element $\alpha \in I$ with $\langle t, \alpha \rangle$ minimal since I is a finite set. If α is indecomposable, then α itself is in S_t, which contradicts $\alpha \in I$. Hence, α is decomposable. Then there exists $\beta, \gamma \in R_t^+$ such that $\alpha = \beta + \gamma$. Since $\langle t, \alpha \rangle = \langle t, \beta \rangle + \langle t, \gamma \rangle$ and all these three terms are strictly bigger then 0, so by the minimality of α, we know $\beta \notin I$ and $\gamma \notin I$. Hence, $\alpha \notin I$, which leads to a contradiction.

Step 2: We claim that $(\alpha, \beta) \leq 0$ for all $\alpha, \beta \in S_t$ such that $\alpha \neq \beta$. Since α, β are indecomposable, they are nonproportional; otherwise, we have $\alpha = \frac{1}{2}\beta$ by the third condition in Definition 8.2 and then $\beta = \alpha + \alpha$, which contradicts that β is indecomposable.

Now we suppose by contradiction that $(\alpha, \beta) > 0$ and shall apply Proposition 8.9 to get $\gamma = \alpha - \beta \in R$. If $\gamma \in R_t^+$, then α is decomposable since $\alpha = \beta + \gamma$. If $\gamma \in R_t^-$, then $-\gamma \in R_t^+$ and hence β is decomposable since $\beta = \alpha + (-\gamma)$. Both lead to a contradiction.

Step 3: Since $V = \operatorname{span} R = \operatorname{span} R_t^+ = \operatorname{span} S_t$, it suffices to show the linearly independence of S_t.

Suppose $\sum_{\alpha \in S_t} c_\alpha \alpha = 0$ and let $I = \{i : c_{\alpha_i} > 0\}$ and $J = \{j : c_{\alpha_j} < 0\}$. It suffices to show $I = J = \varnothing$.

Put

$$\lambda := \sum_{i \in I} c_{\alpha_i} \alpha_i = \sum_{j \in J} (-c_{\alpha_j}) \alpha_j,$$

Since

$$0 \leq (\lambda, \lambda) = \sum_{i \in I, j \in J} c_{\alpha_i} (-c_{\alpha_j}) (\alpha_i, \alpha_j) \leq 0,$$

where the last inequality follows from Step 2, we know $\lambda = 0$. However,

$$0 = \langle t, \lambda \rangle = \sum_{i \in I} c_{\alpha_j} \langle t, \alpha_i \rangle,$$

$$0 = \langle t, \lambda \rangle = \sum_{j \in I} (-c_{\alpha_j}) \langle t, \alpha_j \rangle,$$

with $\langle t, \alpha_i \rangle > 0$ and $\langle t, \alpha_j \rangle < 0$, so $I = J = \varnothing$, which completes the proof. □

Theorem 8.12. *Conversely, if S is a base for R, and if $t \in V^*$ is such that $\langle t, \alpha \rangle > 0$ for all $\alpha \in S$, then $S = S_t$.*

See [16, Section 8] for a proof.

Definition 8.13. *Let $S \subset R$ is a base for R. We say the elements in*

$$R_S^+ := \{\beta \in R : \text{ there exists some } c_\alpha \geq 0 \text{ such that } \beta = \sum_{\alpha \in S} c_\alpha \alpha\},$$

are the positive roots with respect to S. And we can define R_S^- with the obvious modification and the elements in R_S^- are said to be the negative roots with respect to S.

The elements in S are also called the simple roots of R.

Definition 8.14. *Let S is a base for R. The Cartan matrix with respect to S is the matrix*

$$Car(R) := (n(\alpha, \beta))_{\alpha, \beta \in S},$$

where $n(\alpha, \beta) = \langle \beta^*, \alpha \rangle$.

Remark 8.15. By the second step in the proof of Theorem 8.11 and the statement of Theorem 8.12, we know $n(\alpha, \beta) \leq 0$ unless $\alpha = \beta$. Hence, $n(\alpha, \beta) \in \{0, -1, -2, -3\}$ if $\alpha \neq \beta$.

Definition 8.16. *The Coxeter graph for R (with respect to S is a graph with vertices denoting the elements in S and two distinct elements $\alpha \neq \beta \in S$ are connected by $n(\alpha, \beta) \times n(\beta, \alpha)$ many edges.*

Definition 8.17. *The Dynkin diagram $Dyn(R)$ is a partially directed graph whose underlying graph is the Coxeter graph and if $(\alpha, \alpha) > (\beta, \beta)$, then there is an arrow from α to β.*

Example 8.18. *We take the root systems in Example 8.7 as new examples.*

(1) *The Cartan matrix for (3) in Example 8.7 with respect to $S = \{\alpha, \beta\}$ is $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ and the Dynkin diagram is* .
(2) *The Cartan matrix for (4) in Example 8.7 with respect to $S = \{\alpha, \beta\}$ is $\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}$ and the Dynkin diagram is* .
(3) *The Cartan matrix for (5) in Example 8.7 with respect to $S = \{\beta, \alpha\}$ is $\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$ and the Dynkin diagram is* .

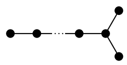
(4) The Cartan matrix for (2) in Example 8.7 with respect to $S = \{\alpha, \beta\}$ *is* $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ *and the Dynkin diagram is* $\bullet \quad \bullet$.

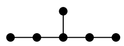
(5) All the Dynkin diagrams of the reduced simple root systems can be classified as follows:

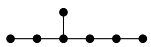
A_r :

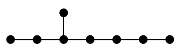
B_r :

C_r :

D_r : 

E_6 : 

E_7 : 

E_8 : 

F_4 :

G_2 :

Definition 8.19. *A root system* $R \subset V$ *is called irreducible if there is no* $V = V_1 \oplus V_2$ *such that* $V_1 \neq 0, V_2 \neq 0$ *and* $R \subset V_1 \cup V_2$.

9. Root systems and Semisimple Lie algebras

Note that we discuss the semisimple Lie algebras with the ground field $\mathbb{C}$ but discuss the root systems with the ground field $\mathbb{R}$ above, so it is natural to ask what are the relations between root systems over $\mathbb{C}$ and those over $\mathbb{C}$. Throughout this section, $\mathfrak{g}$ denotes a complex semisimple Lie algebra.

9.1. Root systems over $k = \mathbb{C}$.

We shall follow [16, Chapter V, Section 17].

Proposition 9.1. *Let* $R \subset V_0$ *be a root system over* $\mathbb{R}$*, then* $R \subset V_0 \otimes_{\mathbb{R}} \mathbb{C}$ *is a root system with*

$$s_\alpha^{\mathbb{C}}(v \otimes z) := z s_\alpha(v), \quad \forall \alpha \in R,$$

being all the symmetries with α *of* $V_0 \otimes \mathbb{C}$*, the* $\mathbb{C}$*-linear extension of* s_α.

Proof. It is easy to check by definition. □

Proposition 9.2. *Let* $R \subset V$ *over* $\mathbb{C}$ *be a root system. One can view* $V_0 = span_{\mathbb{R}} R \subset V$ *as a real subspace, then*

(1) R *is a root system in* V_0 *over* $\mathbb{R}$*;*
(2) The natural map $V_0 \otimes_{\mathbb{R}} \mathbb{C} \to V$ *is an isomorphism;*
(3) For all $\alpha \in R$, $s_\alpha \in GL_{\mathbb{C}}(V)$ *is the linear extension of the symmetry* $s_\alpha^0 \in GL_{\mathbb{R}}(V_0)$ *of* V_0.

Proof. (1) We check the three conditions in Definition 8.2 as follows:

(i) Clearly, $\text{span}_{\mathbb{R}} R = V_0$.
(ii) For $\alpha \in R$, $s_\alpha(R) = R$, then $s_\alpha(V_0) = V_0$ since $\text{span}_{\mathbb{R}} R = V_0$. Put $s_\alpha^0 := s_\alpha|_{V_0}$, then $s_\alpha^0(R) = R$. For $\beta \in R$, then $s_\alpha^0(\beta) = \beta - \langle \alpha^*, \beta \rangle \alpha$, where $\alpha^* \in V^* = Hom_{\mathbb{C}}(V, \mathbb{C})$ is associated to $\alpha \in R \subset V$. On the other hand, $\alpha_0^* := \alpha^*|_{V_0} \in Hom_{\mathbb{R}}(V_0, \mathbb{C})$ is the element associated to $\alpha \in R \subset V_0$. So $s_\alpha^0 = id - \alpha_0^* \otimes \alpha$ is a symmetry with α in V_0.

(iii) For $\alpha, \beta \in R$, $\langle \alpha_0^*, \beta \rangle = \langle \alpha^*, \beta \rangle \in \mathbb{Z}$.

(iv) In fact, from the third fact, we can see that $\alpha_0^* \in V_0^* = Hom_{\mathbb{R}}(V_0, \mathbb{R})$ since $V_0 = \mathrm{span}_{\mathbb{R}} R$.

(2) Denote the natural map by φ. Since $\mathrm{span}_{\mathbb{C}} R = V$, it is obvious that $\varphi : V_0 \otimes_{\mathbb{R}} \mathbb{C} \to V$ is surjective.

It then suffices to show $\varphi : V_0 \otimes_{\mathbb{R}} \mathbb{C} \to V$ is injective. Moreover, it suffices to show the induced map $\varphi^* : V^* \to (V_0 \otimes_{\mathbb{R}} \mathbb{C})^* = V_0^* \otimes_{\mathbb{R}} \mathbb{C}$ is surjective since for any $0 \neq x \in V_0 \otimes_{\mathbb{R}} \mathbb{C}$, there exists at least one linear functional $f \in (V_0 \otimes_{\mathbb{R}} \mathbb{C})^*$ such that $f(x) \neq 0$.

Since $R_0^* := \{\alpha_0^* = \alpha^*|_{V_0} : \alpha \in R\}$ is a root system in V_0^* by [16, Chapter V, Section 6, Proposition 2], we know $\mathrm{span}_{\mathbb{R}} R_0^* = V_0^*$. On the other hand, $\alpha_0^* = \alpha^*|_{V_0} = \varphi^*(\alpha^*)$, which implies φ^* is surjective, and this completes the proof.

(3) This follows from Proposition 9.1 and the second part of this proposition above.

□

Remark 9.3. In the proof of the second assertion above, we use a basic linear algebra fact for complexification that $(V_0 \otimes_{\mathbb{R}} \mathbb{C})^* = V_0^* \otimes_{\mathbb{R}} \mathbb{C}$. We digress a bit to recall the general form of the statement in linear algebra and present a proof for the sake of the convenience for the readers.

Suppose V is a finite dimensional $\mathbb{R}$-vector space, then the complexification of $V^* = Hom_{\mathbb{R}}(V, \mathbb{R})$ is $V^* \otimes_{\mathbb{R}} \mathbb{C} = Hom_{\mathbb{R}}(V, \mathbb{C})$. Moreover, for any $\mathbb{R}$-linear map $\varphi \in V \to \mathbb{C}$, one can extend it to a $\mathbb{C}$-linear map $\varphi^{\mathbb{C}} : V \otimes \mathbb{C} \to \mathbb{C}$ by $\varphi^{\mathbb{C}}(v \otimes z) = z\varphi(v)$. Moreover,

$$\dim_{\mathbb{C}} Hom_{\mathbb{R}}(V, \mathbb{C}) = \dim V^* \otimes_{\mathbb{R}} \mathbb{C} = \dim_{\mathbb{R}} V^* = \dim_{\mathbb{R}} V = \dim_{\mathbb{C}} Hom_{\mathbb{C}}(V \otimes \mathbb{C}, \mathbb{C}).$$

This gives an isomorphism $Hom_{\mathbb{R}}(V, \mathbb{C}) \simeq Hom_{\mathbb{C}}(V \otimes \mathbb{C}, \mathbb{C})$, that is, $V^* \otimes_{\mathbb{R}} \mathbb{C} \simeq (V \otimes_{\mathbb{R}} \mathbb{C})^*$.

9.2. **Semisimple Lie algebras revisited.** Recall that for a Cartan subalgebra $\mathfrak{h}$ of a semisimple Lie algebra $\mathfrak{g}$ over $\mathbb{C}$, there exists a decomposition $\mathfrak{g} = \mathfrak{h} \oplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}^\alpha$, where

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

and $\mathfrak{h} = \mathfrak{g}^0$. In fact, there are at most finitely many $\alpha \in \mathfrak{h}^*$ such that $\mathfrak{g}^\alpha$ is non-trivial. We denote the set of such αs except for 0 by R, then the decomposition is

$$\mathfrak{g} = \mathfrak{h} \oplus_{\alpha \in R \subset \mathfrak{h}^*} \mathfrak{g}^\alpha.$$

Note that this decomposition is similar to the one for non-semisimple Lie algebras in the Step 1 of the proof of Proposition 7.12, but the proof is rather simple. It follows directly from $ad([\mathfrak{h}, \mathfrak{h}]) = 0$, Theorem 7.14 (4) and the criterion for simultaneously diagonalizable.

The goal of this part is to show R is a root system in $\mathfrak{h}^*$.

Proposition 9.4. *(1) For $X \in \mathfrak{g}^\alpha$, $Y \in \mathfrak{g}^{-\alpha}$, $H \in \mathfrak{h}$, one has*

$$B_\kappa(H, [X, Y]) = \alpha(H) \cdot B_\kappa(X, Y).$$

(2) Since $B_\kappa|_{\mathfrak{h} \times \mathfrak{h}}$ is non-degenerate by Theorem 7.14, denote $A_\alpha \in \mathfrak{h}$ be the element corresponds to $\alpha \in \mathfrak{h}^*$ via the isomorphism $\mathfrak{h} \to \mathfrak{h}^*$ induced by B_κ, that is, $\langle \alpha, H \rangle = B_\kappa(A_\alpha, H)$ for all $H \in \mathfrak{h}$. Then we have $[X, Y] = B_\kappa(X, Y)A_\alpha$ for all $X \in \mathfrak{g}^\alpha$, $Y \in \mathfrak{g}^{-\alpha}$.

Proof. (1) Since the killing form B_κ is ad-invariant, we have

$$B_\kappa(H, ad(X)Y) = B_\kappa(ad(H)(X), Y) = B_\kappa(\alpha(H)X, Y) = \alpha(H)B_\kappa(X, Y).$$

(2) By the definition, we have $B_\kappa(H, A_\alpha) = \alpha(H)$, then
$$B_\kappa(H, [X,Y]) = B_\kappa(H, A_\alpha)B_\kappa(X,Y) = B_\kappa(H, B_\kappa(X,Y)A_\alpha),$$
for all $H \in \mathfrak{h}$, which implies $[X,Y] = B_\kappa(X,Y)A_\alpha$ since they all lie in $\mathfrak{h}$ and $B_\kappa|_{\mathfrak{h}\times\mathfrak{h}}$ is non-degenerate.

$\square$

From the proof of Theorem 7.14, we know the decomposition is of the form
$$\mathfrak{g} = \mathfrak{h} \oplus \left(\oplus_{\alpha \in R/\{\pm 1\}} \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha}\right).$$
Then these $\mathfrak{g}^\alpha$'s have the following properties:

Theorem 9.5. *(a) For all $\alpha \in R$, put $\mathfrak{h}_\alpha := [\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}] \subset \mathfrak{h}$, then $\dim \mathfrak{g}^\alpha = \dim \mathfrak{h}_\alpha = 1$. Also, there exists a unique $H_\alpha \in \mathfrak{h}_\alpha$ such that $\alpha(H_\alpha) = 2$.*
(b) For all nonzero $X_\alpha \in \mathfrak{g}^\alpha$, there exists a unique element $Y_\alpha \in \mathfrak{g}^{-\alpha}$ such that $[X_\alpha, Y_\alpha] = H_\alpha$. Moreover, $\mathfrak{sl}_2$ is isomorphic to $\mathfrak{h}_\alpha \oplus \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha}$ defined by the linear extension of the mapping of the $\mathfrak{sl}_2$-triple $H \mapsto H_\alpha, X \mapsto X_\alpha, Y \mapsto Y_\alpha$, where H, X, Y is defined as in Example 2.3.
(c) R is a reduced root system in $\mathfrak{h}^*$.
(d) If $\alpha, \beta \in R$ with $\alpha + \beta \neq 0$, then $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$.

Proof. The outline of the proof is (1)(2)(3)(4)(5) $\Rightarrow$ $(a)(b)$, (6)(7)(8) $\Rightarrow$ (c) and (9) $\Rightarrow$ (d).

(1) From Proposition 9.4, $\dim \mathfrak{h}_\alpha = 1$.
(2) Now we prove there exists a unique $H_\alpha \in \mathfrak{h}_\alpha$ such that $\alpha(H_\alpha) = 2$. Since $\dim \mathfrak{h}_\alpha = 1$, it suffices to show $\alpha(\mathfrak{h}_\alpha) \neq 0$. Suppose by contradiction that $\alpha(\mathfrak{h}_\alpha) = 0$, then pick $X_\alpha \in \mathfrak{g}^\alpha, Y_\alpha \in \mathfrak{g}^{-\alpha}, 0 \neq Z = [X_\alpha, Y_\alpha] \in \mathfrak{h}_\alpha$. We find $[Z, X_\alpha] = \alpha(Z)X_\alpha = 0$ and $[Z, Y_\alpha] = -\alpha(Z)Y_\alpha = 0$ since $\alpha(Z) = 0$. Denote the subalgebra generated by X_α, Y_α, Z by $\mathfrak{a} := \langle X_\alpha, Y_\alpha, Z\rangle \subset \mathfrak{g}$, which is a nilpotent subalgebra. Apply Lie's theorem to the faithful representation $ad : \mathfrak{a} \to gl(\mathfrak{g})$, there exists a flag D of $\mathfrak{g}$ such that $0 = D_0 \subset D_1 \subset \cdots \subset D_n = \mathfrak{g}$ such that $\dim D_i/D_{i-1} = 1$ and $ad(\mathfrak{a})(D_i) \subset D_i$. Then since $Z \in [\mathfrak{a}, \mathfrak{a}]$, we know $ad(Z)(D_i) \subset D_{i-1}$. Hence, $ad(Z)$ is nilpotent. On the other hand, $ad(Z)$ is semisimple by Theorem 7.14. Thus, $Z = 0$, which leads to a contradiction.
(3) For any $X_\alpha \in \mathfrak{g}^\alpha$, there exists $Y_\alpha \in \mathfrak{g}^{-\alpha}$ such that $[X_\alpha, Y_\alpha] \neq 0$ since $\dim \mathfrak{h}_\alpha = 1$. Then $B_\kappa(X_\alpha, Y_\alpha) \neq 0$ by Proposition 9.4. Using $\dim \mathfrak{h}_\alpha = 1$ again, we can get $[X_\alpha, Y_\alpha] = H_\alpha$ by multiplying Y_α by a suitable scalar. Hence, $[H_\alpha, X_\alpha] = \alpha(H_\alpha)X_\alpha = 2X_\alpha$ and $[H_\alpha, Y_\alpha] = -2Y_\alpha$. Then $\mathfrak{sl}_2 \simeq \langle H_\alpha, X_\alpha, Y_\alpha\rangle$.
(4) Now we prove $\dim \mathfrak{g}^\alpha = 1$. Suppose not, $\dim \mathfrak{g}^\alpha \geq 2$, then for a fixed $Y \in \mathfrak{g}^{-\alpha}$, there exists $X_\alpha \in \mathfrak{g}^\alpha$ such that $B_\kappa(X_\alpha, Y) = 0$. Then by Proposition 9.4, $[X_\alpha, Y] = 0$. Moreover, $[H_\alpha, Y] = -2Y$ by definition. Thanks to (3), we can view $\mathfrak{g}$ as an $\mathfrak{sl}_2$-module, then apply the theory of representations of $\mathfrak{sl}_2$, like Theorem 6.6, we know the highest weight is a positive number. However, $[H_\alpha, Y] = -2Y, [X_\alpha, Y] = 0$ implies the highest weight is -2, which is a contradiction.
(5) Then the element Y_α in (3) is unique since $\dim \mathfrak{g}^{-\alpha} = 1$. And we know the subalgebra $\langle H_\alpha, X_\alpha, Y_\alpha\rangle$ in (3), which is isomorphic to $\mathfrak{sl}_2$, is actually equal to $\mathfrak{h}_\alpha \oplus \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha}$. And we denote this by
$$\mathfrak{sl}_{2,\alpha} := \langle H_\alpha, X_\alpha, Y_\alpha\rangle = \mathfrak{h}_\alpha \oplus \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha}.$$

(6) Now we prove span $R = \mathfrak{h}^*$. Otherwise, there exists $H \in \mathfrak{h}$ such that $\alpha(H) = 0$ for all $\alpha \in R$. (It suffices to pick $H \in \{A_\alpha : \alpha \in R\}^\perp$ thanks to Proposition 9.4.) Then $ad(H) = 0$ on $\mathfrak{g}$, which implies $H \in Z(\mathfrak{g})$. However, the center of $\mathfrak{g}$ is trivial since $\mathfrak{g}$ is semisimple.

(7) For all $\alpha, \beta \in R$, we prove $\beta(H_\alpha) \in \mathbb{Z}$ and $\beta - \beta(H_\alpha)\alpha \in R$. Put $p = \beta(H_\alpha)$, then for all $Y_\beta \in \mathfrak{g}^\beta$, $[H_\alpha, Y_\beta] = pY_\beta$, which shows Y_β has weight p when we view $\mathfrak{g}$ as an $\mathfrak{sl}_2$-module. Hence, $p \in \mathbb{Z}$.

Let $Z = \begin{cases} ad(Y_\alpha)^p(Y_\beta), & \text{if } p \geq 0 \\ ad(X_\alpha)^{-p}(Y_\beta), & \text{if } p < 0 \end{cases}$, then $Z \neq 0$ by Theorem 6.6. Obviously, $\beta - p\alpha \in \mathfrak{h}^*$. Also, $[H_\alpha, Z] = (\beta - p\alpha)(H_\alpha)Z$ by induction implies $\beta - p\alpha \in R$.

(8) Now we prove R is reduced. Otherwise, there exists $\alpha \in R$ such that $2\alpha \in R$. Pick $Y \in \mathfrak{g}^{2\alpha}$, then $[H_\alpha, Y] = 4Y$. But $[X_\alpha, Y] = 0$ since $3\alpha \notin R$. On the other hand,

$$[H_\alpha, Y] = [[X_\alpha, Y_\alpha], Y] = -[[Y_\alpha, Y], X_\alpha] = ad(X_\alpha)\left([Y_\alpha, Y]\right) = 0$$

since $[Y_\alpha, Y] \in \mathfrak{g}^\alpha$, which is a constant multiple of X_α. Hence, $Y = 0$, which contradicts.

(9) Let α, β be two non-proportional roots, then $\alpha + \beta \neq 0$. Now we prove $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$.

Consider $E := \sum_{k \in \mathbb{Z}} \mathfrak{g}^{\beta + k\alpha}$, then as in (7), E is a $\mathfrak{sl}_{2,\alpha}$-module and the weights of E are the integers $\beta(H_\alpha) + 2k$, for all k such that $\beta + k\alpha \in R$. Each weight has multiplicity 1 by (4).

Let p, q be the greatest integer such that $\beta - p\alpha, \beta + q\alpha \in R$, respectively. Then by Theorem 6.6, we know E is an irreducible $sl_{2,\alpha}$-module of dimension $m+1$, where $m = \beta(H_\alpha) + 2q = -\beta(H_\alpha) + 2p$. And $ad(X_\alpha) : \mathfrak{g}^{\beta+k\alpha} \to \mathfrak{g}^{\beta+(k+1)\alpha}$ is an isomorphism for all $-p \leq k \leq q-1$ by the structure of irreducible representation of $\mathfrak{sl}_2$.

The result follows by taking $k = 0$.

$\square$

From the discussion above, we know that any semisimple Lie algebra corresponds to a root system R, and hence corresponds to a Dynkin diagram. Actually, the correspondence is surjective. We will not go into details of the proof but we show some examples instead.

Here we introduce a lemma, which gives another sufficient condition for identifying Cartan subalgebras. One can compare this with Theorem 7.14 and Corollary 7.15.

Lemma 9.6. *Suppose $\mathfrak{g}$ is a semisimple Lie algebra. If $\mathfrak{h} \subset \mathfrak{g}$ is maximally abelian and any $H \in \mathfrak{h}$ is semisimple, then $\mathfrak{h}$ is Cartan.*

Proof. Note that $\mathfrak{h}$ is abelian and $ad(H) \in gl(\mathfrak{g})$ is semisimple(diagonalizable) for all $H \in \mathfrak{h}$, so by the criterion for simultaneously diagonalization in linear algebra, there exists a decomposition

$$\mathfrak{g} = \mathfrak{g}_\mathfrak{h}^0 \oplus \left(\oplus_{\alpha \in \mathfrak{h}^* - \{0\}, \mathfrak{g}^\alpha \neq 0} \mathfrak{g}_\mathfrak{h}^\alpha\right),$$

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

Since $\mathfrak{h}$ is maximally abelian, we have $\mathfrak{h} = \mathfrak{g}_\mathfrak{h}^0$.

Clearly, $\mathfrak{h}$ is nilpotent. It suffices to show $\mathfrak{h} = N_\mathfrak{g}(\mathfrak{h})$. Pick $X \in N_\mathfrak{g}(\mathfrak{h})$, we write $X = c_0 H + \sum_\alpha c_\alpha X_\alpha$. Since $[X, \mathfrak{h}] = c_0[H, \mathfrak{h}] + \sum_\alpha c_\alpha [X_\alpha, \mathfrak{h}] \subset \mathfrak{h}$, for any fixed α such that $\mathfrak{g}^\alpha \neq 0$, there exists H' such that $\alpha(H') \neq 0$, then $[X_\alpha, H'] = \alpha(H')X_\alpha \neq 0$, which does not lie in $\mathfrak{h}$, which contradicts $c_0[H, \mathfrak{h}] + \sum_\alpha c_\alpha [X_\alpha, \mathfrak{h}] \subset \mathfrak{h}$ unless all c_α's vanish, that is, $X = c_0 H \in \mathfrak{h}$. $\square$

Example 9.7. *(1) One can check*

$$\mathfrak{h} = \{H = diag(a_1, a_2, a_3) : a_1 + a_2 + a_3 = 0\}$$

is maximally abelian by direct computation, thus $\mathfrak{h}$ *is a Cartan subalgebra of* $\mathfrak{sl}_3$.
Set $\alpha_i \in \mathfrak{h}^*$ *such that* $\alpha_i(H) = a_i - a_{i+1}$ *for* $i = 1, 2$. *Then*

$$R = \{\pm\alpha_1, \pm\alpha_2, \pm(\alpha_1 + \alpha_2)\}.$$

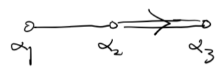
This is related to A_2 *type Dynkin diagram.*

(2) For $M = M^T$, $\mathfrak{so}_M(\mathbb{C}) = \{A_{n\times n} : A^T M + MA = 0\}$. *For* $n = 7$, $M = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$,

we have

$$\mathfrak{h} = \{H = diag(a_1, a_2, a_3, 0, -a_3, -a_2, -a_1) : a_i \in \mathbb{C}\}$$

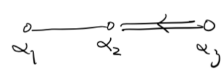
is a Cartan subalgebra. Then let $\alpha_1(H) = a_1 - a_2$, $\alpha_2(H) = a_2 - a_3$ *and* $\alpha_3(H) = a_3$, *then* $\{\alpha_1, \alpha_2, \alpha_3\}$ *is a base of R, which corresponds to a* B_3 *type Dynkin diagram*



(3) For $M = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, $\mathfrak{sp}_6(\mathbb{C}) = \{A_{6\times 6} : A^T M + MA = 0\}$. *we have*

$$\mathfrak{h} = \{H = diag(a_1, a_2, a_3, -a_3, -a_2, -a_1) : a_i \in \mathbb{C}\}$$

is a Cartan subalgebra. Then let $\alpha_1(H) = a_1 - a_2$, $\alpha_2(H) = a_2 - a_3$ *and* $\alpha_3(H) = 2a_3$, *then* $\{\alpha_1, \alpha_2, \alpha_3\}$ *is a base of R, which corresponds to a* C_3 *type Dynkin diagram*



Definition 9.8. *Let R be the root system associated with* $(\mathfrak{g}, \mathfrak{h})$, *where* $\mathfrak{h}$ *is a Cartan subalgebra of semisimple Lie algebra* $\mathfrak{g}$. *Let S be a base of R and* R^+ *be the set of positive roots with respect to S. Put*

$$\mathfrak{n} := \sum_{\alpha\in R^+} \mathfrak{g}^\alpha, \quad \mathfrak{n}^- := \sum_{\alpha\in R^-} \mathfrak{g}^\alpha,$$

then we say $\mathfrak{b} := \mathfrak{h} \oplus \mathfrak{n}$ *is the Borel subalgebra corresponding to* $\mathfrak{h}$ *and S.*

10. Universal enveloping algebras

In this section, we follow [8, Section 17].

First, we may forget temporarily all the specialized theory of Lie algebras. We recall the definition of algebras.

Definition 10.1. *An algebra A over a field k is $(A, (-,-))$ such that*

(1) A is a vector space over k;

(2) $(-,-) : A \times A \to A$ is a k-bilinear form.

Most Lie algebras are non-associative algebras. $k[x]$ and $End_k(V)$ are typical associative algebras. For an associative algebra A, we usually use the notation

$$a \cdot b := (a, b), \quad \forall a, b \in A.$$

Then an associative algebra A can generate a Lie algebra by defining

$$[a, b] := a \cdot b - b \cdot a, \quad \forall a, b \in A. \tag{10.1}$$

Definition 10.2. *Let $\mathfrak{g}$ be a Lie algebra over $k = \mathbb{R}, \mathbb{C}$, which is allowed to be infinite dimensional, contratry to our usual convention.*

A universal enveloping algebra of $\mathfrak{g}$ is a pair $(U(\mathfrak{g}), \varphi)$ which satisfies the following three properties:

(1) $U(\mathfrak{g})$ is an associative algebra with a unit 1 over k;

(2) $\varphi : \mathfrak{g} \to U(\mathfrak{g})$ is a Lie algebra homomorphism, where we adopt the Lie algebra structure of $U(\mathfrak{g})$ as in (10.1), that is, $\varphi([X, Y]) = \varphi(X) \cdot \varphi(Y) - \varphi(Y) \cdot \varphi(X)$;

(3) For any associative algebra with a unit over k and for all Lie algebra homomorphism $\alpha : \mathfrak{g} \to A$, that is, $\alpha([X, Y]) = \alpha(X) \cdot \alpha(Y) - \alpha(Y) \cdot \alpha(X)$, there exists a unique algebra homomorphism $\widetilde{\alpha} : U(\mathfrak{g}) \to A$(sending 1 to 1) such that $\alpha = \widetilde{\alpha} \circ \varphi$, that is, the following diagram commutes.

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\varphi"] \arrow[d, "\alpha"'] & U(\mathfrak{g}) \arrow[dl, "\widetilde{\alpha}"] \\ A & \end{tikzcd}$$

Proposition 10.3. *If a universal enveloping algebra of $\mathfrak{g}$ exists, denoted by $(U(\mathfrak{g}), \varphi)$, then such a pair is unique up to isomorphism of algebras.*

Proof. Suppose there exists another pair (V, φ') satisfying the same hypotheses, then by the third property in the definition above, we get two algebra homomorphisms $\beta : U(\mathfrak{g}) \to V$ and $\gamma : V \to U(\mathfrak{g})$ such that $\varphi' = \beta \circ \varphi$ and $\varphi = \gamma \circ \varphi'$. Then $\gamma \circ \beta$ is the unique dotted map making the following diagram commutes,

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\varphi"] \arrow[d, "\varphi"'] & U(\mathfrak{g}) \arrow[dl, dotted] \\ U(\mathfrak{g}) & \end{tikzcd}$$

where the uniqueness follows from the third property in Definition 10.2. However, $id_{U(\mathfrak{g})}$ is another dotted map making the diagram commutes, which implies $\gamma \circ \beta = id_{U(\mathfrak{g})}$. Similarly, $\beta \circ \gamma = id_V$. This proves the uniqueness. $\square$

Proposition 10.4. *For any Lie algebra $\mathfrak{g}$ over $k = \mathbb{R}, \mathbb{C}$, there exists a universal enveloping algebra $(U(\mathfrak{g}), \varphi)$.*

Proof. Let $T(\mathfrak{g}) = \oplus_{n \geq 0} T^n(\mathfrak{g})$ be the tensor algebra of $\mathfrak{g}$, where $T^0(\mathfrak{g}) = k$, $T^n(\mathfrak{g}) = \mathfrak{g} \otimes \cdots \otimes \mathfrak{g}$ (n copies). Let

$$I(\mathfrak{g}) := \langle [X, Y] - X \otimes Y + Y \otimes X : X, Y \in \mathfrak{g} \rangle$$

be the two sided ideal generated by $[X, Y] - X \otimes Y + Y \otimes X$.

Then $U(\mathfrak{g}) := T(\mathfrak{g})/I(\mathfrak{g})$ is a well-defined associative algebra and we denote the canonical map $\mathfrak{g} = T^1(\mathfrak{g}) \overset{i}{\hookrightarrow} T(\mathfrak{g}) \overset{\pi}{\twoheadrightarrow} U(\mathfrak{g})$ by $\varphi := \pi \circ i$. We claim that $(U(\mathfrak{g}), \varphi)$ is a universal enveloping algebra of $\mathfrak{g}$.

The first two conditions in Definition 10.2 are satisfied thanks to the definition of $I(\mathfrak{g})$. Let $\alpha : \mathfrak{g} \to A$ be as in the definition. The universal property of $T(\mathfrak{g})$ implies that there exists a unique algebra homomorphism $\alpha_1 : T(\mathfrak{g}) \to A$ that extends α such that $\alpha = \alpha_1 \circ i$. Then by the property $\alpha([X, Y]) = \alpha(X) \cdot \alpha(Y) - \alpha(Y) \cdot \alpha(X)$ and the commutativity $\alpha = \alpha_1 \circ i$, we know that $[X, Y] - X \otimes Y + Y \otimes X \in \ker \alpha_1$ for all $X, Y \in \mathfrak{g}$. Thus α_1 induces a homomorphism $\widetilde{\alpha} : U(\mathfrak{g}) \to A$ such that $\widetilde{\alpha} \circ \pi = \alpha_1$, and hence $\widetilde{\alpha} \circ \varphi = \alpha$.

The uniqueness of $\widetilde{\alpha}$ is obvious since 1 and $Im(\varphi)$ generate $U(\mathfrak{g})$.

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "i"] \arrow[rr, bend left, "\varphi"] \arrow[d, "\alpha"] & T(\mathfrak{g}) \arrow[r, "\pi"] \arrow[dl, "\alpha_1"] & U(\mathfrak{g}) \arrow[dll, "\widetilde{\alpha}"] \\
A & &
\end{tikzcd}$$

$\square$

Example 10.5. *Suppose $\mathfrak{g}$ is abelian, that is, $[-, -] = 0$ on $\mathfrak{g}$. Then*

$$U(\mathfrak{g}) = T(\mathfrak{g})/\langle X \otimes Y - Y \otimes X \rangle = Sym(\mathfrak{g})$$

is the symmetric algebra of vector space $\mathfrak{g}$.

In fact, one can check the symmetric algebra $Sym(\mathfrak{g})$ can be identified, through a canonical isomorphism, to the polynomial ring $K[B]$, where B is a basis of $\mathfrak{g}$. See [8, Section 17.1] *or* [15, Chapter III, Section 3] *for more discussions.*

Now we define a filtration as follows. Suppose $(U(\mathfrak{g}), \varphi)$ is a universal enveloping algebra of $\mathfrak{g}$. Let $U_i(\mathfrak{g}) := \{\varphi(X_1) \cdots \varphi(X_j) : j \leq i, X_k \in \mathfrak{g}\ \forall 1 \leq k \leq j\}$, then

$$U_0(\mathfrak{g}) \subset U_1(\mathfrak{g}) \subset \cdots U_i(\mathfrak{g}) \subset \cdots U(\mathfrak{g})$$

and $U_i(\mathfrak{g}) U_j(\mathfrak{g}) \subset U_{i+j}(\mathfrak{g})$. Set $gr_i U(\mathfrak{g}) := U_i(\mathfrak{g})/U_{i-1}(\mathfrak{g})$, which is a vector space. Then we say $grU(\mathfrak{g}) := \oplus_{i \geq 0} gr_i U(\mathfrak{g})$ is the graded universal enveloping algebra.

Since $U_{i-1}(\mathfrak{g}) U_j(\mathfrak{g}) \subset U_{i+j-1}(\mathfrak{g})$ and $U_i(\mathfrak{g}) U_{j-1}(\mathfrak{g}) \subset U_{i+j-1}(\mathfrak{g})$, the multimplication in $U(\mathfrak{g})$ induce a well-defined bilinear map $gr_i U(\mathfrak{g}) \times gr_j U(\mathfrak{g}) \to gr_{i+j} U(\mathfrak{g})$. Moreover, it extends to $grU(\mathfrak{g}) \times grU(\mathfrak{g}) \to grU(\mathfrak{g})$, making $grU(\mathfrak{g})$ a graded associative algebra with identity.

The composite linear map $\phi_m : T^m(\mathfrak{g}) \to U_m(\mathfrak{g}) \to gr_m U(\mathfrak{g})$ makes sense and is surjective. It yields a linear map $\phi : T(\mathfrak{g}) \to grU(\mathfrak{g})$ by combining these maps and ϕ is surjective.

Lemma 10.6. *$\phi : T(\mathfrak{g}) \to grU(\mathfrak{g})$ is an algebra homomorphism, which factors through Sym$(\mathfrak{g})$, that is, ϕ induces a homomorphism ϕ^* such that the following diagram commutes.*

$$\begin{tikzcd} T(\mathfrak{g}) \arrow[r, "\pi"] \arrow[d] \arrow[rr, bend left, "\phi"] & U(\mathfrak{g}) \arrow[r, "g"] & grU(\mathfrak{g}) \\ Sym(\mathfrak{g}) \arrow[urr, dotted, "\phi^*"] & & \end{tikzcd}$$

In particular, $grU(\mathfrak{g})$ is commutative.

Proof. By definition, $\phi(X \otimes Y) = \phi(X) \cdot \phi(Y)$, it follows that ϕ is an algebra homomorphism.

Since $\mathrm{Sym}(\mathfrak{g}) = T(\mathfrak{g})/\langle X \otimes Y - Y \otimes X \rangle$, it suffices to show $\phi(X \otimes Y - Y \otimes X) = 0$ for all $X, Y \in \mathfrak{g}$. We compute

$$\begin{aligned} &\phi(X \otimes Y - Y \otimes X) = \phi(X) \cdot \phi(Y) - \phi(Y) \cdot \phi(X) \\ =&g\left(\varphi(X) \cdot \varphi(Y) - \varphi(Y) \cdot \varphi(X)\right) = g\left(\varphi([X,Y])\right) \in U_1/U_1 = 0, \end{aligned}$$

which completes the proof.

Since $\mathrm{Sym}(\mathfrak{g})$ is commutative and ϕ^* is surjective, we know $grU(\mathfrak{g})$ is commutative. (Note that $U(\mathfrak{g})$ may not be commutative.) $\square$

Theorem 10.7 (Poincaré-Birkhoff-Witt Theorem)**.** *For any Lie algebra over $k = \mathbb{R}, \mathbb{C}$, these following are equivalent:*

(1) Sym$(\mathfrak{g}) \xrightarrow{\phi^*} grU(\mathfrak{g})$ is an isomorphism of algebras;
(2) fixing an ordered basis $\{e_1, \cdots, e_n\}$ of $\mathfrak{g}$, the set

$$\left\{\varphi(e_{i_1}) \cdot \varphi(e_{i_2}) \cdots \varphi(e_{i_{m-1}}) \cdot \varphi(e_{i_m}) : m \geq 0, i_1 \leq i_2 \leq \cdots \leq i_m\right\}$$

is a k-basis of $U(\mathfrak{g})$.

Also, they both hold.

Proof. For the proof of any one of them, see [8, Section 17.4] or [15, Chapter III]. We only prove the equivalency of these two statements.

We claim that

$$U_r(\mathfrak{g}) = \mathrm{span}\ \{\varphi(e_{i_1}) \cdot \varphi(e_{i_2}) \cdots \varphi(e_{i_{m-1}}) \cdot \varphi(e_{i_m}) : m \leq r, i_1 \leq i_2 \leq \cdots \leq i_m\}.$$

This is obvious by induction on r. We omit the proof but we can check intuitively as follows. By defintion, for any $\varphi(X_1) \cdot \varphi(X_2) \cdots \varphi(X_j) \in U_r(\mathfrak{g})$ for $j \leq r$, we can write each X_i as a linear combination of e_l's. Then for terms like $\cdots \varphi(e_{l_1}) \cdot \varphi(e_{l_2}) \cdots$ with $l_1 > l_2$, we replace it by using the fact $\varphi(e_{l_1}) \cdot \varphi(e_{l_2}) - \varphi(e_{l_2}) \cdot \varphi(e_{l_1}) = \varphi([e_{l_1}, e_{l_2}])$.

Now for $M = (i_1, \cdots, i_r)$ with $i_1 \leq \cdots \leq i_r$, we write $l(M) = r$ and $X_M = \varphi(e_{i_1}) \cdots \varphi(e_{i_r})$. Note that the first statement is equivalent to the injectivity of ϕ^* since the surjectivity is obvious. Moreover, ϕ^* is injective if and only if for any $r \geq 0$, the equation

$$\sum_{l(M)=r} c_M X_M \equiv 0, \quad \mod U_{r-1}(\mathfrak{g})$$

of $c_M \in \mathbb{C}$ has only the trivial solution. And this is equivalent to

$$\sum_{l(M)=r} c_M X_M = \sum_{l(M)<r} d_M X_M \tag{10.2}$$

has only the trivial solution for c_M's.

Obviously, the second statement of the theorem and the claim above implies (10.2) has only trivial solutions for c_M's. In fact, if (10.2) has only trivial solutions for c_M's for all r, then by induction on r, one can observe that the second statement is true.

This completes the proof of equivalency. □

11. Linear representations of semisimple Lie algebras

In this section, $\mathfrak{g}$ denotes a semisimple Lie algebra over $\mathbb{C}$ and $\mathfrak{h}$ denotes a Cartan subalgera. We follow [16, Chapter VII] in this part.

Definition 11.1. *For any $\mathfrak{g}$-module (π, V) (not necessarily finite dimensional) and $\omega \in \mathfrak{h}^*$, we set*

$$V^\omega := \{v \in V : \pi(H)(v) = \omega(H)(v), \forall H \in \mathfrak{h}\},$$

which is a vector subspace of V. We call ω a weight of V if $V^\omega \neq \{0\}$ and $v \in V^\omega$ is said to have weight ω. And we call $\dim V^\omega$ the multiplicity of ω.

Example 11.2. *For $(\pi, V) = (ad, \mathfrak{g})$, the weights are the roots of the pair $(\mathfrak{g}, \mathfrak{h})$ as in Theorem 9.5.*

Motivated from the example above, here is a generalization of Theorem 9.5 (d).

Proposition 11.3. *One has $\pi(\mathfrak{g}^\alpha)(V^\omega) \subset V^{\omega+\alpha}$ if $\omega \in \mathfrak{h}^*$, $\alpha \in R$. And then it is an easy corollary that $\oplus_{\omega \text{ weights of } V} V^\omega \subset V$ is a $\pi(\mathfrak{g})$-stable.*

Proof. For all $X \in \mathfrak{g}^\alpha$, $v \in V^\omega$, $H \in \mathfrak{h}$, we compute

$$\pi(H)\pi(X)(v) = \pi(X)\pi(H)(v) + \pi([H, X])(v) = (\omega(H) + \alpha(H))\,\pi(X)(v),$$

which implies that $\pi(X)v \in V^{\omega+\alpha}$. □

Definition 11.4. *Let V be a $\mathfrak{g}$-module. A vector $v \in V$ is called a primitive element of weight ω if it satisfies*

(1) $v \neq 0$ and has weight ω;

(2) $\mathfrak{g}^\alpha(v) = 0$ for all $\alpha \in R^+$.

Note that the second condition is equivalent to say that $\mathfrak{g}^\alpha(v) = 0$ holds for all $\alpha \in S$ thanks to the definition of bases.

For each $\alpha \in R^+$, we choose $X_\alpha \in \mathfrak{g}^\alpha$, $Y_\alpha \in \mathfrak{g}^{-\alpha}$, $H_\alpha := [X_\alpha, Y_\alpha] \in \mathfrak{h}$ as in Theorem 9.5.

Proposition 11.5. *Let V be a $\mathfrak{g}$-module and $v \in V$ be a primitive element of weight ω. Set*

$$E := \mathfrak{g}\text{-submodule of } V \text{ generated by } v,$$

that is, $E = \{X_1 \cdot (X_2 (\cdots (X_n \cdot v))) : X_i \in \mathfrak{g}, n \in \mathbb{Z}_{\geq 0}\}$. Then

(1) If $R^+ := \{\beta_1, \cdots, \beta_k\}$, then

$$E = span_{\mathbb{C}}\{Y_{\beta_1}^{m_1} \cdots Y_{\beta_k}^{m_k} v : m_i \in \mathbb{Z}_{\geq 0}\}.$$

(2) The weights of E are of the form $\omega - \sum_{i=1}^n p_i\alpha_i$, where $p_i \in \mathbb{Z}_{\geq 0}$, $S = \{\alpha_1, \cdots, \alpha_n\}$. And they have finite multiplicity.
(3) ω is a weight of E of multiplicity 1.
(4) E is an indecomposable $\mathfrak{g}$-module, that is, if there exists two submodules E_1, E_2 such that $E = E_1 \oplus E_2$, then either $E_1 = 0$ or $E_2 = 0$.

Remark 11.6. It is obvious by definition that if $\mathfrak{g}$ is irreducible, then $\mathfrak{g}$ is indecomposable. If it is a finite dimensional $\mathfrak{g}$-module, then by Weyl's theorem (Theorem 5.8), we know that these two notions are equivalent. However, this may not hold for infinite dimensional representations.

Proof. (1) Note that $\{Y_{\beta_1}, \cdots, Y_{\beta_k}\}$ is an ordered basis of the vector space $\mathfrak{n}^-$, we can obtain an ordered basis $\{Y_1, \cdots, Y_k, \widetilde{Y_1}, \cdots, \widetilde{Y_m}\}$ of the vector space $\mathfrak{g}$, where we write $Y_j := Y_{\beta_j}$ for simplicity and $\{\widetilde{Y_1}, \cdots, \widetilde{Y_m}\}$ is an ordered basis of the vector space $\mathfrak{b}$.

Thanks to the PBW theorem (Theorem 10.7), we know that

$$\{\varphi(Y_1)^{n_1} \cdots \varphi(Y_k)^{n_k} \varphi(\widetilde{Y_1})^{l_1} \cdots \varphi(\widetilde{Y_m})^{l_m} : n_i, l_j \in \mathbb{Z}_{\geq 0}\},$$

form a basis of $U(\mathfrak{g})$ and the bases of $U(\mathfrak{b})$ and $U(\mathfrak{n}^-)$ are of similar forms. Then $U(\mathfrak{g})$ is a free $U(\mathfrak{n}^-)$-module and $U(\mathfrak{g}) = U(\mathfrak{n}^-) \cdot U(\mathfrak{b})$. By the definition of v, we know that for all $B \in \mathfrak{b}$, either $B \cdot v = 0$ or $B \cdot v$ is proportional to v, so

$$E = U(\mathfrak{g}) \cdot v = U(\mathfrak{n}^-) \cdot U(\mathfrak{b}) \cdot v = U(\mathfrak{n}^-) \cdot v.$$

On the other hand, $\{\varphi(Y_1)^{m_1} \cdots \varphi(Y_k)^{m_k} : m_i \in \mathbb{Z}_{\geq 0}\}$ form a basis of $U(\mathfrak{n}^-)$, hence (1).

(2) By Proposition 11.3, $\varphi(Y_1)^{m_1} \cdots \varphi(Y_k)^{m_k} v$ has weight $\omega - \sum_{j=1}^k m_j\beta_j$. Since $\beta_j = \sum_i c_{ji}\alpha_i$ with $c_{ji} \in \mathbb{Z}_{\geq 0}$, we know all the weights of E are of the form $\omega - \sum_i p_i\alpha_i$ with $p_i \in \mathbb{Z}_{\geq 0}$.

The multiplicity of $\omega - \sum_{i=1}^n p_i\alpha_i$ is equal to

$$\sharp\left\{(m_1, \cdots, m_k) \in \mathbb{Z}_{\geq 0}^k : (m_1, \cdots, m_k)(c_{ji}) = (p_1, \cdots, p_n), (c_{ji}) \in (\mathbb{Z}_{\geq 0})_{k\times n}, k \geq n, \text{ at least one } p_l > 0\right\},$$

which is finite since $\sum_{j=1}^k m_j c_{jl} = p_l$ has only finite solutions for $m_j \in \mathbb{Z}_{\geq 0}$ if p_l is a positive integer.

(3) Since $\omega = \omega - \sum_j m_j\beta_j$ if and only if $m_j = 0$ for all j, we know from (2) that $E^\omega = \mathbb{C} \cdot v$.

(4) Suppose $E = E_1 \oplus E_2$, where E_1, E_2 are two submodules, then one can check $E^\omega = E_1^\omega \oplus E_2^\omega$. Since $\dim E^\omega = 1$ by (3), we must have $E^\omega = E_1^\omega$ or $E^\omega = E_2^\omega$, which implies that either $E_2 = 0$ or $E_1 = 0$ holds.

□

Up to now, we know nothing about the existence and uniqueness of the primitive element. In fact, if $\dim V < \infty$, then there exists a primitive element, but it is not always the case if $\dim V = \infty$. And we can show the primitive element is unique up to scalar multiplication in the following case.

Corollary 11.7. *Let V be an irreducible $\mathfrak{g}$-module containing a primitive element v of weight ω. Then*

(1) Every primitive element of V lies in $\mathbb{C} \cdot v$ and we say its weight ω is the highest weight of V.

(2) *Let V_1, V_2 be two irreducible $\mathfrak{g}$-modules with highest weight ω_1, ω_2, then V_1 and V_2 are isomorphic if and only if $\omega_1 = \omega_2$.*

Proof. Suppose $v' \in V$ is primitive of weight $\omega' \in \mathfrak{h}^*$. Since V is irreducible, we know the $\mathfrak{g}$-submodule E of V generated by v is equal to V. By applying Proposition 11.5, we have

$$\omega' = \omega - \sum_i m_i \alpha_i, \quad m_i \geq 0.$$

Similarly, exchanging the roles of v and v', we see that

$$\omega = \omega' - \sum_i m_i' \alpha_i, \quad m_i' \geq 0.$$

Thus, $\omega = \omega'$. And by Proposition 11.5 (3), we know $\dim V^\omega = 1$ and hence $v = cv'$.

It suffices to prove that if $\omega_1 = \omega_2$, then V_1 and V_2 are isomorphic. Let $v_i (i = 1, 2)$ be a primitive element of V_i of weight $\omega = \omega_1 = \omega_2$. Clearly, the $\mathfrak{g}$-module $V := V_1 \oplus V_2$ has $v = v_1 + v_2$ as a primitive element of weight ω. Let E be the $\mathfrak{g}$-submodule of V generated by v and the projection $\pi_2 : V \to V_2$ induces a $\mathfrak{g}$-module homomorphism $f_2 : E \to V_2$ as in the diagram below.

$$\begin{tikzcd}
V \arrow[r, "\pi_2"] & V_2 \\
E \arrow[u, "i"] \arrow[ur, "f_2"] &
\end{tikzcd}$$

One has $f_2(v) = v_2$. Note that v_2 generates V_2, so f_2 is surjective. Moreover, $\ker f_2 = V_1 \cap E \subset V_1$. However, $v_1 \notin \ker f_2$. Suppose not, then $v_1 \in \ker f_2 \in E$ and $v_1 + v_2 \in E$. Since both are primitive element of weight ω, v_1 is a constant multiple of $v_1 + v_2$ thanks to Proposition 11.5, which leads to a contradiction. Thus, $v_1 \notin \ker f_2$. Since V_1 is irreducible, the $\mathfrak{g}$-submodule $\ker f_2 = 0$. Thus $f_2 : E \to V_2$ is an isomorphism. Similarly, one can prove that E is isomorphic to V_1. Therefore, V_1 and V_2 are isomorphic. $\square$

Theorem 11.8. *For each $\omega \in \mathfrak{h}^*$, there exists a unique irreducible $\mathfrak{g}$-module with highest weight equal to ω.*

Proof. The uniqueness is up to isomorphism and it follows from Corollary 11.7.

Step 1: Now we consider the following diagram.

$$\begin{tikzcd}
\mathfrak{n} \arrow[r, "i"] & \mathfrak{b} \arrow[r] \arrow[dr, "\omega"] & \mathfrak{h} \arrow[d, "\omega"] \\
 & & \mathbb{C}
\end{tikzcd}$$

This gives ω is a Lie algebra homomorphism from $\mathfrak{b}$ to $\mathbb{C}$. Let L_ω be a one-dimensional $\mathfrak{b}$-module whose basis is an element v such that $Hv = \omega(H)v$ for all $H \in \mathfrak{h}$ and $X_\alpha v = 0$ for $X_\alpha \in \mathfrak{n}$. Then one can naturally view it as a $U(\mathfrak{b})$-module. By taking the tensor product of $U(\mathfrak{g})$ with the one dimensional space L_ω, we get a $U(\mathfrak{g})$-module

$$V_\omega := U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} L_\omega.$$

It is obvious that $1 \otimes v \in V_\omega$ and it follows from the defintion of v that $1 \otimes v$ is a primitive element of weight ω in V_ω. Now we get a $\mathfrak{g}$-module V_ω generated by a primitive element v of weight ω.

Step 2: Set

$$N_\omega := \langle \text{proper } \mathfrak{g}\text{-submodule of } V_\omega \rangle \subset V_\omega,$$

then the quotient module $E_\omega = V_\omega / N_\omega$ is irreducible if $E_\omega \neq 0$. Hence, it suffices to show $N_\omega \neq V_\omega$.

Step 3: Put

$$V_\omega^- := \oplus_{\pi \neq \omega, \pi \text{ is a weight}} (V_\omega)^\pi,$$

then we claim that if V' is a proper $\mathfrak{g}$-submodule of V_ω, then $V' \subset V_\omega^-$.

Since V' is stable under $\mathfrak{h}$, one has $V' = \oplus_{\pi \text{ is a weight}} V'^\pi$. By Proposition 11.5 (3), we have $\dim(V_\omega)^\omega = 1$. So if $V'^\omega \neq 0$, then it has dimension 1 and would contain $1 \otimes v$. Hence, one would have $V' = V_\omega$ since V_ω is generated by v, which leads to a contradiction. Thus $V' = \oplus_{\pi \neq \omega, \pi \text{ is a weight}} (V_\omega)^\pi$, that is ,$V' \subset V_\omega^-$.

Hence, $N_\omega \subset V_\omega^-$, which is a proper submodule of V_ω and this completes the proof. ☐

Proposition 11.9. *Let $V \neq 0$ be a finite dimensional $\mathfrak{g}$-module. Then*

(1) $V = \oplus_{\pi \text{ is a weight}} V^\pi$.
(2) V contains a primitive element.
(3) If V is generated by a primitive element, then V is irreducible.
(4) For any weight ω of V, $\pi(H_\alpha)$ is an integer for all $\alpha \in R$.

Proof. (1) By Theorem 7.14 (4), we know that the endomorphisms defined by elements in $\mathfrak{h}$ are diagonalizable and commute with each other, hence can be diagonalizable simultaneously.

(2) We apply Lie's theorem (Theorem 3.2) to the solvable Lie algebra $\mathfrak{b}$ and find that there exists $\lambda \in \mathfrak{b}^*$, $v \in V$ such that $X \cdot v = \lambda(X)v$ and $H \cdot v = \lambda(H)v$ for all $X \in \mathfrak{g}^\alpha$, $\alpha \in R^+$, $H \in \mathfrak{h}$. Moreover, from Theorem 9.5 (d), we know that $[X, H] \in \mathfrak{g}^\alpha$ and $[X, H] \cdot v = \lambda(H)\lambda(X)v - \lambda(X)\lambda(H)v = 0$, which implies $\lambda([X, H]) = 0$. Thanks to Theorem 9.5 (a), $\lambda|_{\mathfrak{g}^\alpha} = 0$ for all $\alpha \in R^+$. Hence, v is a primitive element of weight λ.

(3) This follows from Proposition 11.5 (4) and Weyl's theorem (Theorem 5.8).

(4) Note that for $\alpha \in R^+$, we know from Theorem 9.5 that one can view V as a $\mathfrak{sl}_{2,\alpha} := \langle H_\alpha, X_\alpha, Y_\alpha \rangle$-module. It follows from the representation theories of $\mathfrak{sl}_2$ that the eigenvalues of H_α on V belong to $\mathbb{Z}$, which completes the proof.

☐

Corollary 11.10. *Every finite dimensional irreducible $\mathfrak{g}$-module has a highest weight.*

Proof. This follows from Proposition 11.9 (2) and Corollary 11.7. ☐

References

1. Michael. Artin, *Algebra*, second ed., Prentice Hall, Upper Saddle River, N.J., 2011 (eng).
2. Armand Borel, *Essays in the history of Lie groups and algebraic groups*, History of Mathematics, vol. 21, American Mathematical Society, Providence, RI; London Mathematical Society, Cambridge, 2001. MR 1847105
3. Nicolas Bourbaki, *Lie groups and Lie algebras. Chapters 1–3*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 1989, Translated from the French, Reprint of the 1975 edition. MR 979493

4. ______, *Algebra. II. Chapters 4–7*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 1990, Translated from the French by P. M. Cohn and J. Howie. MR 1080964
5. ______, *Lie groups and Lie algebras. Chapters 7–9*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 2005, Translated from the 1975 and 1982 French originals by Andrew Pressley. MR 2109105
6. William Fulton and Joe Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991, A first course, Readings in Mathematics. MR 1153249
7. Robin Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics, No. 52, Springer-Verlag, New York-Heidelberg, 1977. MR 0463157
8. James E. Humphreys, *Introduction to Lie algebras and representation theory*, Graduate Texts in Mathematics, Vol. 9, Springer-Verlag, New York-Berlin, 1972. MR 0323842
9. Nathan Jacobson, *Lie algebras*, Interscience Tracts in Pure and Applied Mathematics, No. 10, Interscience Publishers (a division of John Wiley & Sons, Inc.), New York-London, 1962. MR 0143793
10. Alexander Kirillov, Jr., *An introduction to Lie groups and Lie algebras*, Cambridge Studies in Advanced Mathematics, vol. 113, Cambridge University Press, Cambridge, 2008. MR 2440737
11. John M. Lee, *Introduction to smooth manifolds*, second ed., Graduate Texts in Mathematics, vol. 218, Springer, New York, 2013. MR 2954043
12. Hannes Nilsson, *On levi decompositions in finite and infinite dimensional lie algebras*, `https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=9040553&fileOId=9040554`, 2021, Student Paper.
13. Pierre Samuel, *Algebraic theory of numbers*, Houghton Mifflin Co., Boston, Mass., 1970, Translated from the French by Allan J. Silberger. MR 0265266
14. Jean-Pierre Serre, *A course in arithmetic*, Graduate Texts in Mathematics, No. 7, Springer-Verlag, New York-Heidelberg, 1973, Translated from the French. MR 0344216
15. ______, *Lie algebras and Lie groups*, second ed., Lecture Notes in Mathematics, vol. 1500, Springer-Verlag, Berlin, 1992, 1964 lectures given at Harvard University. MR 1176100
16. ______, *Complex semisimple Lie algebras*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2001, Translated from the French by G. A. Jones, Reprint of the 1987 edition. MR 1808366
17. John Stillwell, *Naive Lie theory*, Undergraduate Texts in Mathematics, Springer, New York, 2008. MR 2431402

Zhejiang University, Hangzhou, China
Email address: `3180102677@zju.edu.cn` `math_tn@berkeley.edu`