

COURSE NOTES

---

# Lie Theory

MAT7109

---

**Instructor:**

Prof. Efim ZELMANOV

*Teaching Assistant:*

Binhe Huang

Southern University of Science and Technology

Spring, 2026

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Examples of Lie algebras</b>                      | <b>4</b>  |
| §1.1     | From associative algebras . . . . .                  | 4         |
| §1.2     | From involutions . . . . .                           | 5         |
| §1.3     | From derivations of linear algebras . . . . .        | 6         |
| §1.4     | From brackets . . . . .                              | 6         |
| §1.5     | From Lie groups . . . . .                            | 7         |
| <b>2</b> | <b>Engel's Theorem</b>                               | <b>8</b>  |
| §2.1     | A stronger version of Engel's Theorem . . . . .      | 8         |
| §2.2     | Three types of Engel's Theorem . . . . .             | 9         |
| <b>3</b> | <b>Lie's Theorem</b>                                 | <b>11</b> |
| §3.1     | Two equivalent statements of Lie's Theorem . . . . . | 11        |
| §3.2     | A stronger version of Lie's Theorem . . . . .        | 12        |
| <b>4</b> | <b>Jordan-Chevalley Decomposition</b>                | <b>14</b> |
| §4.1     | Solvable ideals and radicals . . . . .               | 14        |
| §4.2     | Review of Jordan decomposition . . . . .             | 14        |
| §4.3     | Compatibility with adjoint representation . . . . .  | 16        |
| <b>5</b> | <b>Cartan's Criteria</b>                             | <b>17</b> |
| §5.1     | Replicas . . . . .                                   | 17        |
| §5.2     | Cartan's Criterion for solvability . . . . .         | 18        |
| §5.3     | Cartan's Criterion for semisimplicity . . . . .      | 19        |
| <b>6</b> | <b>Consequences of Cartan's Criteria</b>             | <b>20</b> |
| §6.1     | Decomposition of semisimple Lie algebras . . . . .   | 20        |
| §6.2     | Derivations . . . . .                                | 21        |
| §6.3     | Abstract Jordan decomposition . . . . .              | 22        |
| <b>7</b> | <b>Casimir Elements of <math>L</math>-Modules</b>    | <b>23</b> |
| §7.1     | Tensor products of modules . . . . .                 | 23        |
| §7.2     | Schur's Lemma . . . . .                              | 23        |
| §7.3     | Casimir operators . . . . .                          | 24        |
| <b>8</b> | <b>Complete Reducibility</b>                         | <b>26</b> |
| §8.1     | Completely reducible modules . . . . .               | 26        |

|           |  |           |
|-----------|--|-----------|
| §8.2      | Weyl's Theorem . . . . .   | 27        |
| §8.3      | Preservation of Jordan decomposition . . . . .                     | 28        |
| <b>9</b>  | <b>Representations of <math>\mathfrak{sl}_2(\mathbb{F})</math></b> | <b>30</b> |
| <b>10</b> | <b>Root Space Decomposition</b>                                    | <b>32</b> |
| §10.1     | Maximal toral subalgebras . . . . .                                | 32        |
| §10.2     | The $\mathfrak{sl}_2$ copies . . . . .                             | 34        |
| §10.3     | Integral properties . . . . .                                      | 34        |
| §10.4     | The inner product . . . . .  | 35        |
| <b>11</b> | <b>Abstract Root Systems I</b>                                     | <b>37</b> |
| §11.1     | Irreducible root systems . . . . .                                 | 37        |
| §11.2     | Angles between roots . . . . .                                     | 38        |
| §11.3     | Simple roots . . . . .   | 38        |
| <b>12</b> | <b>Abstract Root Systems II</b>                                    | <b>41</b> |
| §12.1     | Weyl chambers and Weyl groups . . . . .                            | 41        |
| §12.2     | Simply transitive actions of Weyl groups . . . . .                 | 42        |
| §12.3     | Classification of root systems of rank 2 . . . . .                 | 44        |
| <b>13</b> | <b>Classification of Root Systems</b>                              | <b>46</b> |
| §13.1     | Normal systems and Coxeter graphs . . . . .                        | 46        |
| §13.2     | Classification of associated Coxeter graphs . . . . .              | 47        |
| §13.3     | Dynkin diagrams . . . . .  | 50        |
| <b>14</b> | <b>Summary</b>   | <b>53</b> |
| <b>15</b> | <b>Generators and Relations of Associative Algebras</b>            | <b>54</b> |
| §15.1     | Free associative algebras . . . . .                                | 54        |
| §15.2     | Gröbner–Shirshov Theorem . . . . .                                 | 55        |
| §15.3     | Proof of Gröbner–Shirshov Theorem . . . . .                        | 57        |
| <b>16</b> | <b>Applications of Gröbner–Shirshov Theorem</b>                    | <b>60</b> |
| §16.1     | Universal enveloping algebras . . . . .                            | 60        |
| §16.2     | Poincaré–Birkhoff–Witt Theorem . . . . .                           | 61        |
| <b>17</b> | <b>Generators and Relations of <math>L</math></b>                  | <b>63</b> |
| §17.1     | Free Lie algebras . . . . .  | 63        |
| §17.2     | Relations satisfied by $L$ . . . . .                               | 64        |
| §17.3     | Auxiliary Lie algebras . . . . .                                   | 66        |
| <b>18</b> | <b>Existence Theorem</b>   | <b>67</b> |
| §18.1     | Revisiting Serre's relations . . . . .                             | 67        |
| §18.2     | Braid group automorphisms . . . . .                                | 68        |
| §18.3     | Serre's Theorem . . . . .  | 69        |



# Chapter 1

## Examples of Lie algebras

Lie algebras were originally introduced by Sophus Lie in the 1870s to study the concept of continuous transformation groups. The book *Lie Algebras* by Nathan Jacobson, which was published in 1962 is the earliest textbook for Lie algebras, and it was the first time that the terminology “Lie algebra” appeared. Before that, people called them infinitesimal Heisenberg groups.

**Definition 1.1.** *A Lie algebra is **simple** if it is not abelian and has no nontrivial ideals.*

Where do Lie algebras come from?

### §1.1 From associative algebras

Denote by  $\mathbb{F}$  the ground field. Let  $A$  be an associative algebra.  $A^{(-)}$  is a Lie algebra with the bracket

$$[a, b] = ab - ba \quad \text{for all } a, b \in A.$$

**Theorem 1.2** (Wedderburn Theorem). *Suppose  $\mathbb{F}$  is algebraically closed.*

- Any simple finite dim. associative algebra is isomorphic to  $M_n(\mathbb{F})$ .*
- For any finite dim. associative algebra  $A$ , there exists a unique nilpotent ideal  $N \triangleleft A$  such that  $A/N \simeq M_{n_1}(\mathbb{F}) \oplus \dots \oplus M_{n_k}(\mathbb{F})$*

**Example 1.3.**  $\mathfrak{sl}_n$  is simple if  $\text{char } \mathbb{F} = 0$  or  $\text{char } \mathbb{F} \nmid n$ ; otherwise  $\mathfrak{sl}_n$  is not simple. In the case of  $\text{char } \mathbb{F} \mid n$ ,  $\mathbb{F}I_n$  is a 1-dim ideal.

## §1.2 From involutions

**Definition 1.4.** Let  $A$  be any linear algebra. A linear transformation  $*$  :  $A \rightarrow A$  is called an **involution** if for any  $a, b \in A$ ,

1.  $(a^*)^* = a$ ;
2.  $(ab)^* = b^*a^*$

**Example 1.5.** 1. Transposition  $(a_{ij}) \mapsto (a_{ji})$  is an involution.

2. Symplectic involution for quaternions:

$$\mathbb{H} = \{a + bi + cj + dk : a, b, c, d \in \mathbb{R}\}$$

with the relations

$$i^2 = j^2 = k^2 = -1, \quad ij = k, \quad jk = i, \quad ki = j.$$

The involution is defined by  $a + bi + cj + dk \mapsto a - bi - cj - dk$ . On the other hand, quaternions can be represented by matrices

$$\begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix},$$

where  $a, b, c, d \in \mathbb{R}$  and  $i$  is the square root of unity. Then the involution has the form

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \mapsto \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} = \begin{pmatrix} \delta & -\gamma \\ -\beta & \alpha \end{pmatrix}^T.$$

Given an involution of an associative algebra  $A$ , we have

$$K = \{a \in A : a^* = -a\} \subset A^{(-)}.$$

Then  $K$  is a Lie subalgebra.

**Remark 1.6.** Over an algebraically closed field  $\mathbb{F}$ , for  $\text{Mat}_n(\mathbb{F})$  (or equivalently a f.d. associative simple algebra), the classification of involutions (up to conjugation) is as follows:

1. If  $n$  is odd, any involution is isomorphic to transposition.
2. If  $n$  is even, any involution is isomorphic to transposition or symplectic involution, where symplectic involutions are defined as

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \mapsto \begin{pmatrix} \delta^T & -\beta^T \\ -\gamma^T & \alpha^T \end{pmatrix} = \begin{pmatrix} \delta & -\gamma \\ -\beta & \alpha \end{pmatrix}^T,$$

where  $\alpha, \beta, \gamma, \delta$  are  $\frac{n}{2} \times \frac{n}{2}$  matrices.

## §1.3 From derivations of linear algebras

**Definition 1.7.** Let  $A$  be an associative algebra. A linear map  $\alpha : A \rightarrow A$  is called a **derivation** if for any  $a, b \in A$ ,  $\alpha(ab) = \alpha(a)b + a\alpha(b)$ .

The commutator of two derivations is again a derivation, i.e.,  $\text{Der } A \subset \text{End}_{\mathbb{F}}(A)^{(-)}$ .

**Example 1.8.** Let  $W_n = \text{Der } \mathbb{F}[x_1, \dots, x_n]$ , where

$$\text{Der } \mathbb{F}[x_1, \dots, x_n] = \left\{ \sum_{i=1}^n f_i \frac{\partial}{\partial x_i} : f_i \in \mathbb{F}[x_1, \dots, x_n] \right\}.$$

Note that  $W_n$  is an infinite dim Lie algebra (also simple).

**Example 1.9.** Let  $A$  be an associative algebra. Fix  $a \in A$ . Then  $\text{ad } a : x \mapsto [a, x]$  is a derivation. More generally, let  $L$  be a Lie algebra. Then  $\text{ad } L \triangleleft \text{Der } L$  because  $[d, \text{ad } x] = \text{ad}(d(x))$  for any  $d \in \text{Der } L$  and  $x \in L$ . The derivations in  $\text{ad } L$  are called *inner derivations*.

**Remark 1.10.** The appearance of derivations is natural in the sense that for any  $d \in \text{Der } A$ , if  $\exp d := \text{Id} + d + \frac{d^2}{2!} + \dots$  makes sense, then  $\exp d$  is an automorphism with inverse  $\exp(-d)$ . This is the case, for example, when  $A$  is finite-dimensional or when  $d$  acts locally nilpotently.

## §1.4 From brackets

Many algebraic structures contain brackets. Some of them are Lie brackets.

**Example 1.11.** Let  $A$  be an associative commutative algebra. A bracket  $[ , ] : A \times A \rightarrow A$  is called a **Poisson bracket** if

1.  $(A, [ , ]) is a Lie algebra;$
2.  $[a^2, b] = 2a[a, b]$  for any  $a, b \in A$ .

**Remark 1.12.** In general, we require  $[ac, b] = a[b, c] + [a, b]c$ . But in the case of commutative Poisson algebras, this condition is the same as condition 2.

**Example 1.13.** 1. (Poisson bracket) Let  $B = \mathbb{F}(p_1, \dots, p_n; q_1, \dots, q_n)$ . For any  $f, g \in B$ , define

$$[f, g] = \sum_{i=1}^n \left( \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} \right).$$

2. (Lie bracket but not Poisson) Let  $B = \mathbb{F}[t]$ . For any  $f, g \in B$ , define  $[f, g] = f'g - fg'$ .

## §1.5 From Lie groups

The tangent space at the identity element of a Lie group is a Lie algebra.

# Chapter 2

## Engel's Theorem

### §2.1 A stronger version of Engel's Theorem

**Definition 2.1.** An associative algebra  $A$  is called an **enveloping algebra** of a Lie algebra  $L$  if  $L \subset A^{(-)}$  and  $A = \langle L \rangle$ .

**Theorem 2.2.** Let  $A$  be a finite dim enveloping algebra of  $L$ . If  $\exists m \geq 1$  such that  $\forall a \in L, a^m = 0$ , then  $A$  is nilpotent.

We will prove a stronger version.

**Definition 2.3.** Let  $L$  be a Lie algebra. We say a subset  $S \subset L$  is a **Lie subset** if  $\forall a, b \in S, [a, b] \in S$ .

**Theorem 2.4.** Let  $A$  be a finite dim enveloping algebra of  $L$  and let  $S \subset L$  be a Lie subset of  $L$  with  $L = \text{span } S$ . If  $\exists m \geq 1$  s.t.  $\forall a \in S, a^m = 0$ , then  $A$  is nilpotent.

*Proof.* Denote  $\dim A = n$ . By Zorn's Lemma<sup>1</sup>, we can choose a maximal Lie subset  $S_1 \subset S$  s.t. the associative algebra  $\langle S_1 \rangle$  is nilpotent.<sup>2</sup> We claim that  $S_1 = S$ . Otherwise, there exists  $s \in S \setminus S_1$  such that  $[S_1, s] \subset S_1$ .<sup>3</sup> Now let  $S_2 = S_1 \cup \{s\}$ , a Lie subset of  $S$ . We will deduce a contradiction by showing the nilpotency of  $\langle S_2 \rangle$ . Then our claim follows. Hence, we conclude that  $A = \langle L \rangle = \langle S \rangle$  is nilpotent.

<sup>1</sup>Zorn's lemma is the following statement: Let  $(P, \leq)$  be a partially ordered set. If every chain in  $P$  has an upper bound in  $P$ , then  $P$  contains a maximal element.

<sup>2</sup>First, the set  $P = \{S' \subset S : \langle S' \rangle \text{ is nilpotent}\}$  is nonempty by our assumption. For an arbitrary chain

$$Q_1 \subset Q_2 \subset \dots$$

in  $P$ , we need to prove the union of all  $Q_i$ 's, denoted by  $Q$ , is also in  $P$ . Let  $a_1 a_2 \dots a_l$  be a monomial in  $\langle Q \rangle$ . There exists  $N \in \mathbb{N}$  such that  $a_i \in Q_N$  for all  $i$ . Note that for any  $i$  we have  $Q_i^n = 0$ . Thus  $a_1 a_2 \dots a_l = 0$  as long as  $l \geq n$ . Then  $Q^n = 0$  as required.

<sup>3</sup>If  $\exists x_1 \in S_1$  s.t.  $[x_1, s] \notin S_1$ , then take  $s_1 = [x_1, s]$ . If there still exists  $x_2 \in S_1$  s.t.  $[x_2, s_1] \notin S_1$ , then take  $s_2 = [x_2, s_1]$ . Iterating, we have  $s_i = [x_i, [x_{i-1}, \dots, [x_1, s]]] \in (\text{ad } S_1)^i(s)$ . Because  $\langle S_1 \rangle$  is nilpotent,  $(\text{ad } S_1)^{2n-1} = 0$ . Thus, after finite steps, the procedure ends, i.e.,  $\exists k$  s.t.  $[S_1, s_k] \subset S_1$  as desired.

Let us prove that  $\langle S_2 \rangle$  is nilpotent. Let  $x$  be an arbitrary nonzero monomial in  $\langle S_2 \rangle$ . Then  $x$  can be written as

$$x = s_{11}s_{12} \dots s_{1m_1}s^{k_1}s_{21}s_{22} \dots s_{2m_2}s^{k_2} \dots,$$

where  $s_{ij} \in S_1$ ,  $m_i < n$  and  $k_i < m$ . Since  $[S_1, s] \subset S_1$ , we can move all  $s$  to the right of  $s_{ij}$  without changing the number of elements in  $S_1$ . For example,  $s^{k_1}s_{21} = s^{k_1-1}[s, s_{21}] + s^{k_1-1}s_{21}s$ , in which each term on both sides has one element in  $S_1$ . Thus,  $x$  has at most  $n$   $s_{ij}$  in the expression. Thus,  $x$  has length at most  $n - 1 + n(m - 1) = mn$ . As a consequence,  $S_2^{mn} = 0$ .  $\square$

## §2.2 Three types of Engel's Theorem

**Definition 2.5.** Let  $L$  be a Lie algebra and  $M$  be a vector space over  $\mathbb{F}$ . Define an operation  $L \times M \rightarrow M; (a, m) \mapsto a \cdot m$ . The following are equivalent:

1.  $L + M$  is a Lie algebra if we define  $a \cdot m = -m \cdot a$  and  $m_1 \cdot m_2 = 0$  for any  $a \in L$  and  $m_1, m_2 \in M$ .
2.  $[a, b] \cdot m = a \cdot (b \cdot m) - b \cdot (a \cdot m)$  for any  $a, b \in L$  and  $m \in M$ .
3.  $L \rightarrow (\text{End}_{\mathbb{F}} M)^{(-)}$  is a Lie algebra homomorphism.

We say  $M$  is a **module of  $L$**  if it satisfies any of these conditions.

**Definition 2.6.** Let  $L$  be a Lie algebra. Define a sequence of ideals recursively:

$$L^0 = L, \quad L^{k+1} = [L, L^k] \text{ for } k \geq 0.$$

If there exists  $n \geq 0$  such that  $L^n = 0$ , then  $L$  is called **nilpotent**.

There are three versions of Engel's Theorem:

**Theorem 2.7** (Engel's Theorem). Let  $L$  be a finite dim Lie algebra. If  $\forall a \in L$ ,  $\text{ad } a$  is nilpotent, then  $L$  is nilpotent.

**Theorem 2.8.** Let  $L$  be a Lie algebra and  $M$  be a finite dim  $L$ -module. If any element in  $L$  acts nilpotently, then  $L$  acts on  $M$  nilpotently.

**Theorem 2.9.** Let  $L$  be a Lie algebra and  $M$  be a finite dim  $L$ -module. Let  $S \subset L$  be a Lie subset with  $\text{span } S = L$ . If any  $a \in S$  acts nilpotently, then  $L$  acts nilpotently.

*Proof of Theorem 2.7.* Let  $L$  have dimension  $n$ . Then  $\forall a \in L$ ,  $(\text{ad } a)^n = 0$ . Consider the subalgebra  $\langle \text{ad } L \rangle \subset \text{End}_{\mathbb{F}} M$  generated by  $\text{ad } L$ . This theorem follows from Theorem 2.2.  $\square$

**Summary.** We have the following relations:

*Theorem 2.4*  $\implies$  *Theorem 2.9*  $\implies$  *Theorem 2.8*  $\implies$  *Theorem 2.7*.

1. *Theorem 2.7 is actually a special case of Theorem 2.8 by taking  $M$  as the adjoint representation.*
2. *Theorem 2.9 is a stronger version of Theorem 2.8.*
3. *Theorem 2.9 follows directly from Theorem 2.4 by regarding  $\varphi(S)$  as a Lie subset of the Lie algebra  $\varphi(L)$ .*

**Remark 2.10.** *Theorem 2.8 and Theorem 2.9 do not require  $L$  to be finite dim. But Theorem 2.7 requires that because it views  $L$  as an  $L$ -module.*

**Corollary 2.11.** *Let  $L$  be a Lie algebra and let  $M$  be a finite dim  $L$ -module, i.e., a finite dim representation. If any element in  $L$  acts nilpotently, then there exists a basis of  $M$  such that  $L$  consists of strictly upper triangular matrices; that is,*

$$L \subseteq \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

*Proof.* Let  $\varphi$  be the representation. By Theorem 2.8,  $\varphi(L)$  is nilpotent. Let  $n$  be such that  $\varphi(L)^n \neq 0$  and  $\varphi(L)^{n+1} = 0$ . Take a nonzero vector  $v_1$  in  $\varphi(L)^n(M)$ . Then repeating the procedure for  $M/\mathbb{C}v_1$  (not necessarily an  $L$ -module), on which  $\varphi(L)$  still acts nilpotently gives  $v_2 \in M \setminus (\mathbb{C}v_1)$ . Iterating, we get a basis  $v_1, v_2, \dots, v_{\dim M}$  as desired.  $\square$

# Chapter 3

## Lie's Theorem

**Lemma 3.1.** *Let  $\text{char } \mathbb{F} = 0$  or  $\text{char } \mathbb{F} > n$ . Let  $a, b$  be two  $n \times n$  matrices s.t.  $[a, [a, b]] = 0$ . Then  $[a, b]$  is nilpotent.*

*Proof.* Since  $\text{ad } a$  is a derivation, we have

$$(\text{ad } a)^i(b^j) = \begin{cases} 0 & \text{if } j < i, \\ n![a, b]^n & \text{if } j = i. \end{cases}$$

Let  $f = x^n + \beta_{n-1}x^{n-1} + \dots + \beta_1x + \beta_0$  be an annihilating polynomial of  $b$ . Applying  $(\text{ad } a)^n$  to  $f(b) = 0$ , we have

$$(\text{ad } a)^n(f(b)) = n![a, b]^n = 0.$$

Since  $n! \neq 0$ ,  $[a, b]^n = 0$ . □

### §3.1 Two equivalent statements of Lie's Theorem

**Definition 3.2.** *Let  $L$  be a Lie algebra. Define a sequence of ideals recursively:*

$$L^{(0)} = L, \quad L^{(k+1)} = [L^{(k)}, L^{(k)}] \text{ for } k \geq 0.$$

*If there exists  $n \geq 0$  such that  $L^{(n)} = 0$ , then  $L$  is called **solvable**.*

**Theorem 3.3** (Lie's Theorem). *Let  $L$  be a solvable Lie algebra over  $\mathbb{F}$  and  $M$  be a finite dim  $L$ -module. Let  $\text{char } \mathbb{F} = 0$  or  $\text{char } \mathbb{F} > \dim_{\mathbb{F}} M$  and  $F$  be algebraically closed. Then there exists a basis of  $M$  such that*

$$L \subset \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}.$$

Let us show an equivalent theorem.

**Theorem 3.4.** *If  $L$  is a solvable Lie algebra and  $M$  is an irreducible finite dim  $L$ -module, then  $M$  is 1-dim.*

*Proof of Theorem 3.4.* As we are mainly concerned with the image of  $L \rightarrow (\text{End}_{\mathbb{F}} M)^{(-)}$ , we may assume  $L \subset (\text{End}_{\mathbb{F}} M)^{(-)}$ . Since  $L$  is solvable, we have

$$L \supsetneq L^{(1)} \supsetneq L^{(2)} \supsetneq \dots \supsetneq L^{(s)} \supsetneq L^{(s+1)} = 0, \quad \text{for some } s < \dim M.$$

If  $s = 0$ , i.e.,  $L$  is abelian, then  $M$  contains a common eigenvector  $v$  for all elements of  $L$ .<sup>1</sup> Since  $\mathbb{C}v$  is  $L$ -invariant, by irreducibility of  $M$ , we get  $M = \mathbb{C}v$ . Then we are done.

If  $s > 0$ , then for any  $a \in L^{(s)}$  and  $b \in L$ ,  $[a, [a, b]] \in [L^{(s)}, L^{(s)}] = L^{(s+1)} = 0$ .<sup>2</sup> By Lemma 3.1, we have  $[a, b]^n = 0$ . Consider a Lie subset  $S = \{[a, b] : a \in L^{(s)}, b \in L\}$ , all of whose elements are nilpotent. By Theorem 2.4,  $\text{span } S = [L^{(s)}, L]$  is nilpotent. Note that  $[L^{(s)}, L]$  is an ideal of  $L$ .<sup>3</sup> Then  $[L^{(s)}, L]M$  is a submodule of  $M$ . By the irreducibility of  $M$  and nilpotency of  $[L^{(s)}, L]$ , this submodule must be 0, that is,  $[L^{(s)}, L] = 0$ . Then for any  $x, y \in L^{(s-1)}$ ,  $[x, [x, y]] = 0$ . By Lemma 3.1,  $[x, y]$  is nilpotent. Applying Theorem 2.4 again, we have  $L^{(s)}$  is nilpotent. Then, by the irreducibility of  $M$ , we can conclude that  $L^{(s)} = 0$  which is a contradiction.  $\square$

Now we show the equivalence between Theorem 3.3 and Theorem 3.4.

*Proof.* Theorem 3.3  $\Rightarrow$  Theorem 3.4 is obvious. Conversely, let  $M_1$  be a minimal submodule. Then consider  $M/M_1$  which has dimension  $\dim M - 1$ . Then it follows by induction.  $\square$

**Example 3.5.** *Let  $\text{char } \mathbb{F} = 2$ . Then  $\mathfrak{gl}_2(\mathbb{F})$  is solvable:*

$$L^{(1)} = \begin{pmatrix} \alpha & \gamma \\ \beta & \alpha \end{pmatrix}, \quad L^{(2)} = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}, \quad L^{(3)} = 0.$$

## §3.2 A stronger version of Lie's Theorem

In the following, we introduce a version of Lie's Theorem which does not require  $\mathbb{F}$  to be algebraically closed.

**Theorem 3.6.** *Let  $L$  be a solvable Lie algebra over  $\mathbb{F}$  and  $M$  be a finite dim  $L$ -module. Let  $\text{char } \mathbb{F} = 0$  or  $\text{char } \mathbb{F} > \dim_{\mathbb{F}} M$ , Then  $[L, L]$  acts nilpotently.*

<sup>1</sup>Commutative operators have the same eigenspaces.

<sup>2</sup> $L^{(i)}$  are ideals of  $L$ .

<sup>3</sup>Since  $L^{(s)}$  is an ideal of  $L$ ,  $[[L^{(s)}, L], L] \subset [L^{(s)}, L]$ .

*Proof.* Denote by  $\overline{\mathbb{F}}$  the algebraic closure of  $\mathbb{F}$ . Consider  $\overline{L} = L \otimes_{\mathbb{F}} \overline{\mathbb{F}}$  and  $\overline{M} = M \otimes_{\mathbb{F}} \overline{\mathbb{F}}$ . Then  $\overline{M}$  is naturally an  $\overline{L}$ -module. Then by Theorem 3.3,  $\overline{L}$  can be represented by upper triangular matrices over  $\overline{\mathbb{F}}$ . Thus,  $[\overline{L}, \overline{L}]$  consisting of strictly upper triangular matrices acts on  $\overline{M}$  nilpotently. Since the action of  $\overline{L}$  is compatible with the original action,  $[L, L]$  also acts nilpotently.  $\square$

Under the assumption that the field is algebraically closed, we can deduce Theorem 3.4 by Theorem 3.6: Since  $[L, L]$  acts nilpotently and  $M$  is irreducible,  $[L, L] = 0$ . Thus,  $L$  is abelian. Then the same argument of the  $s = 0$  case in the proof of Theorem 3.4 completes the proof.

**Summary.** *We have the following relations:*

$$\text{Theorem 3.6} \xrightarrow{\text{alg. closed}} \text{Theorem 3.3} \iff \text{Theorem 3.4.}$$

## Chapter 4

# Jordan-Chevalley Decomposition

### §4.1 Solvable ideals and radicals

Let us start with two easy lemmas.

**Lemma 4.1.** *Let  $I \triangleleft L$ , and both  $I$  and  $L/I$  are solvable. Then  $L$  is solvable.*

*Proof.* Let  $I \supset I^1 \supset \dots \supset I^{(m)} = 0$  and  $L/I \supset (L/I)^{(1)} \supset \dots \supset (L/I)^{(n)} = 0$ . Then

$$L^{(n+m)} = (L^{(n)})^{(m)} \subset I^{(m)} = 0.$$

□

**Lemma 4.2.** *Let  $I, J$  be solvable ideals of  $L$ . Then  $I + J$  is solvable.*

*Proof.*  $(I + J)/I \cong J/(I \cap J)$ . Since  $J/(I \cap J)$  and  $I$  are both solvable, by Lemma 4.1,  $I + J$  is solvable. □

**Definition 4.3.** *Let  $\text{Rad } L$  be the maximal solvable ideal of the Lie algebra  $L$ .*

In particular, if  $L$  is finite dim,  $\text{Rad } L$  exists uniquely because of Lemma 4.2. Note that  $L/\text{Rad } L$  has no solvable ideals.

**Definition 4.4.** *A linear transformation  $\varphi$  is called **semisimple** if it satisfies*

1.  $\exists$  a basis in which the matrix of  $\varphi$  is diagonal.
2.  $(\varphi - \alpha)^2 v = 0$  if and only if  $(\varphi - \alpha)v = 0$ .

The two conditions above are equivalent.

### §4.2 Review of Jordan decomposition

Let  $\mathbb{F}$  be algebraically closed,  $V$  be a finite dim  $\mathbb{F}$ -vector space and  $\varphi : V \rightarrow V$  be a linear transformation.

Denote  $V_\alpha = \{\text{root vectors w.r.t. } \alpha\} \cup \{0\} = \ker(\varphi - \alpha)^{\dim V}$ .

**Theorem 4.5.**  $V = V_{\alpha_1} \oplus \dots \oplus V_{\alpha_r}$ , where  $\alpha_1, \dots, \alpha_r$  are distinct eigenvalues of  $\varphi$ .

*Proof.* There exists a polynomial  $f(t) = (t - \alpha_1)^{n_1} \dots (t - \alpha_r)^{n_r}$  such that  $f(\varphi) = 0$ . Denote by

$$g_i = \frac{f}{(t - \alpha_i)^{n_i}}, \quad \text{for } 1 \leq i \leq r.$$

Since they do not have common roots,  $\gcd(g_i : 1 \leq i \leq r) = 1$ . Thus, there exist  $h_i \in \mathbb{F}[t]$  such that  $\sum_{i=1}^r h_i g_i = 1$ . Then for any  $v \in V$ ,  $v = \sum_i h_i(\varphi) g_i(\varphi) v$ , where  $h_i(\varphi) g_i(\varphi) v \in V_{\alpha_i}$ .

Next, we show the linear independence of  $V_{\alpha_i}$ . Let  $v_i \in V_i$  such that  $v_1 + v_2 + \dots + v_r = 0$ . Suppose there exists  $v_i \neq 0$ . Then we have  $(\varphi - \alpha_i)^{n_i} v_i = 0$  and  $g_i v_i = 0^1$ . Since  $(t - \alpha_i)^{n_i}$  and  $g_i$  are coprime,  $v_i = 0$ , which is a contradiction.  $\square$

**Theorem 4.6.**  $\varphi$  can be decomposed as  $\varphi = \varphi_s + \varphi_n$ , where  $\varphi_s$  is semisimple and  $\varphi_n$  is nilpotent. Moreover, there exists a polynomial  $P(t) \in \mathbb{F}[t]$  with zero constant term such that  $\varphi_s = P(\varphi)$ . In particular,  $\varphi_n = (1 - P)(\varphi)$ .

*Proof.* Let  $V = V_{\alpha_1} \oplus \dots \oplus V_{\alpha_r}$  as in Theorem 4.5. Define  $\varphi_s$  to act on  $V_{\alpha_i}$  by multiplication by  $\alpha_i$ . Then  $\varphi - \varphi_s$  is naturally nilpotent.<sup>2</sup> By *Chinese Remainder Theorem*, there exists a polynomial  $P(t) \in \mathbb{F}[t]$  such that

$$P(t) \equiv \alpha_i \pmod{(t - \alpha_i)^{n_i}} \quad \text{for all } i.$$

Furthermore, we can require  $P(t)$  has zero constant term: if  $\exists \alpha_i = 0$ , then we are done; if not, we just add the equation  $P(t) \equiv 0 \pmod{t}$ . After adding, CRT still holds.<sup>3</sup>  $\square$

**Theorem 4.7.** If we further require  $[\varphi_s, \varphi_n] = 0$ , the decomposition in Theorem 4.6 is unique. We call it **Jordan decomposition**.

*Proof.* It is obvious that the decomposition in Theorem 4.6 satisfies the requirement. We use the same notations as in Theorem 4.6. Suppose that we have another decomposition  $\varphi = \varphi'_s + \varphi'_n$  satisfying  $[\varphi'_s, \varphi'_n] = 0$ . Then  $\varphi'_s$  and  $\varphi'_n$  both commute with  $\varphi$ . Since  $\varphi_s$  and  $\varphi_n$  are polynomials of  $\varphi$ , they also commute with  $\varphi_s$  and  $\varphi_n$ . Thus,  $\varphi_s - \varphi'_s$  is still semisimple and  $\varphi_n - \varphi'_n$  is still nilpotent. As a consequence,  $\varphi_s - \varphi'_s = \varphi_n - \varphi'_n$  is both semisimple and nilpotent, which implies  $\varphi_s = \varphi'_s$  and  $\varphi_n = \varphi'_n$ .  $\square$

<sup>1</sup> $g_i$  kills  $v_j$  for all  $j \neq i$ .

<sup>2</sup>Check on each  $V_{\alpha_i}$ .

<sup>3</sup> $t$  has no common roots with  $(t - \alpha_i)^{n_i}$ 's.

### §4.3 Compatibility with adjoint representation

We continue to use notations as in §4.2. In particular,  $\varphi : V \rightarrow V$  is a linear transformation in  $\text{End}_{\mathbb{F}} V = \mathfrak{gl}(V)$ .

**Theorem 4.8.** *If  $\varphi = \varphi_s + \varphi_n$  is the Jordan decomposition of  $\varphi$ , then  $\text{ad } \varphi = \text{ad } \varphi_s + \text{ad } \varphi_n$  is the Jordan decomposition. In particular,  $\text{ad } \varphi_s$  and  $\text{ad } \varphi_n$  can be expressed by polynomials of  $\text{ad } \varphi$  with zero constant terms.*

*Proof.* 1.  $\text{ad } \varphi_s$  is semisimple, because the action on the basis element  $E_{ij}$  satisfies

$$\left[ \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_r \end{pmatrix}, E_{ij} \right] = (\alpha_i - \alpha_j)E_{ij}.$$

2.  $\text{ad } \varphi_n$  is nilpotent because  $(\text{ad } \varphi_n)^{\dim V} = 0$ .

3.  $[\text{ad } \varphi_s, \text{ad } \varphi_n] = \text{ad}[\varphi_s, \varphi_n] = 0$  since  $\text{ad}$  is a Lie algebra homomorphism. □

# Chapter 5

## Cartan's Criteria

In this chapter, we always assume that  $\mathbb{F}$  has characteristic 0 and  $V$  is a finite dim vector space over  $\mathbb{F}$ .

### §5.1 Replicas

**Definition 5.1.** Let  $\varphi$  be a linear transformation on  $V$ . Suppose there is a basis of  $V$  such that

$$\varphi = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}.$$

We call a linear transformation  $\psi$  a **replica** of  $\varphi$  if

$$\psi = \begin{pmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_n \end{pmatrix} \quad \text{and} \quad \alpha_i = \alpha_j \Rightarrow \beta_i = \beta_j.$$

Then we have two easy properties of replicas.

**Proposition 5.2.** If  $\psi$  is a replica of  $\varphi$ , then  $\exists f \in \mathbb{F}[t]$  s.t.  $\psi = f(\varphi)$ .

*Proof.* By Chinese Remainder Theorem, we have a solution  $f$  for the system of equations:

$$f(t) \equiv \beta_i \pmod{t - \alpha_i}, \quad \text{for all } i.$$

Then  $f$  is as desired. □

**Proposition 5.3.** Let  $\varphi$  be a semisimple linear transformation defined as in Definition 5.1 and  $\pi \in \text{Hom}_{\mathbb{Q}}(\mathbb{F}, \mathbb{Q})$  be a  $\mathbb{Q}$ -linear function. Let  $\psi = \pi(\varphi) = \text{diag}(\pi(\alpha_1), \dots, \pi(\alpha_n))$ , then  $\text{ad } \psi$  is a replica of  $\text{ad } \varphi$ .

*Proof.* In the proof of Theorem 4.8, we have proved that  $\text{ad } \psi$  and  $\text{ad } \varphi$  are semisimple with eigenvalues  $\alpha_i - \alpha_j$ 's and  $\pi(\alpha_i) - \pi(\alpha_j)$ 's, respectively. Moreover, if  $\alpha_i - \alpha_j = \alpha_p - \alpha_q$ , then  $\pi(\alpha_i) - \pi(\alpha_j) = \pi(\alpha_p) - \pi(\alpha_q)$ .  $\square$

Then we have an immediate corollary by Proposition 5.2.

**Corollary 5.4.** *Let  $\varphi$  and  $\psi$  be as in Proposition 5.3. Then  $\exists f \in \mathbb{F}[t]$  s.t.  $\text{ad } \psi = f(\text{ad } \varphi)$ .*

## §5.2 Cartan's Criterion for solvability

Recall that  $V$  is finite dimensional.

**Lemma 5.5** (Nilpotency criterion). *Let  $A \subseteq B \subseteq \mathfrak{gl}(V)$  be subspaces. Denote by  $M = \{\varphi \in \mathfrak{gl}(V) : [\varphi, B] \subseteq A\}$ . If  $\varphi \in M$  and  $\text{tr}(\varphi M) = 0$ , then  $\varphi$  is nilpotent.*

**Remark 5.6.** *Why  $M$  is better than  $L$ ? Let  $\varphi$  and  $\psi$  be as in Proposition 5.3. Then*

$$\varphi \in M \iff \text{ad } \varphi : B \rightarrow A \xrightarrow{\text{Cor 5.4}} \text{ad } \psi : B \rightarrow A \iff \psi \in M.$$

*Proof of Lemma 5.5.* Let  $\varphi = \varphi_s + \varphi_n$  be the Jordan decomposition of  $\varphi$ . Take an arbitrary linear transformation  $\pi \in \text{Hom}_{\mathbb{Q}}(\mathbb{F}, \mathbb{Q})$  and let  $\psi = \pi(\varphi_s)$ . Under some basis,  $\varphi_s$  and  $\psi$  can be expressed as a matrix

$$\varphi_s = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \quad \text{and} \quad \psi = \begin{pmatrix} \pi(\alpha_1) & & \\ & \ddots & \\ & & \pi(\alpha_n) \end{pmatrix}.$$

By Remark 5.6,  $\psi \in M$ ; hence  $\text{tr}(\varphi\psi) = 0$ . This gives  $\text{tr}(\varphi_s\psi) = 0$ , because  $\psi$  commutes with  $\varphi_n$  (by Proposition 5.2) and  $\varphi_n\psi$  is nilpotent. Thus, we have  $\sum_i \alpha_i \pi(\alpha_i) = 0$ . Applying  $\pi$  again, we get a sum of rational numbers  $\sum_i \pi(\alpha_i)^2 = 0$ , which means  $\pi(\alpha_i) = 0$  for all  $i$ . By arbitrariness of  $\pi$ , we finally get  $\varphi_s = 0$ , that is,  $\varphi = \varphi_n$  is nilpotent.  $\square$

**Theorem 5.7** (Cartan's Criterion for solvability). *If a Lie algebra  $L \subseteq \mathfrak{gl}(V)$  satisfies  $\text{tr}(L[L, L]) = \{0\}$ , then  $L$  is solvable.*

*Proof.* Let  $M = \{x \in \mathfrak{gl}(V) : [x, L] \subseteq [L, L]\}$  (take  $A = [L, L]$ ,  $B = L$  in Lemma 5.5).

We will prove  $\text{tr}([a, b]M) = 0$  for any  $a, b \in L$ . Let  $x \in M$ . Note that

$$\text{tr}([a, b]x) = \text{tr}([ax, b]) - \text{tr}(a[x, b]) = 0,$$

because the elements in  $[L, L]$  have trace 0 and  $[x, b] \in [L, L]$ .

By Lemma 5.5 we can deduce that  $[a, b]$  acts nilpotently for any  $a, b \in L$ . Namely, the elements in the Lie subset  $S = \{[a, b] : a, b \in L\}$  act nilpotently. Then by Theorem 2.4, we have  $[L, L] = \text{span } S$  acts nilpotently, which implies  $L$  is solvable.  $\square$

### §5.3 Cartan's Criterion for semisimplicity

Recall that for a finite dimensional Lie algebra  $L$ ,  $\text{Rad } L$  is the largest solvable ideal and contains any solvable ideal of  $L$ . Moreover, the quotient Lie algebra  $L/\text{Rad } L$  does not contain any nonzero solvable ideals. We say a Lie algebra is semisimple if its radical is 0.

**Definition 5.8.** We call the bilinear form defined as  $(a|b) = \text{tr}(\text{ad } a \text{ ad } b)$  for all  $a, b \in L$  the **Killing form** of  $L$ .

The Killing form has associativity: Let  $x, y, z \in \mathfrak{gl}(V)$ .  $\text{tr}([x, y]z - x[y, z]) = \text{tr}([x, y]z + \text{ad } x[z, y]) = \text{tr}([xz, y]) = 0$ . In particular,  $\text{ad } L \subseteq \mathfrak{gl}(L)$ .

**Theorem 5.9** (Cartan's Criterion for semisimplicity). For a finite dimensional Lie algebra  $L$ ,  $L$  is semisimple if and only if its Killing form is nondegenerate.

*Proof.* Suppose  $L$  is semisimple. Let  $I = \{a \in L : (a|L) = 0\}$ . Since the Killing form is associative,  $I \triangleleft L$ . Note that  $\text{tr}(\text{ad } I \text{ ad } I) \subseteq (I|I) = \{0\}$ . By Theorem 5.7,  $\text{ad } I \subseteq \mathfrak{gl}(L)$  is solvable. Consider the Lie algebra homomorphism

$$\text{ad}|_I : I \rightarrow \text{ad } I; a \mapsto \text{ad } a.$$

We have  $\text{ad } I = I/\ker = I/(I \cap Z(L))$ , where  $Z(L)$  is the center of  $L$ . Since  $\text{ad } I$  and  $I \cap Z(L)$  are solvable,  $I$  is solvable.

Suppose the Killing form is nondegenerate and  $J \triangleleft L$  is a nonzero solvable ideal of  $L$ . Then  $L$  contains a nonzero abelian ideal  $I$ .<sup>1</sup> For any  $a \in I$  and  $b \in L$ , we have  $(\text{ad } a \text{ ad } b)^2 = 0$ , since  $[b, [a, [b, L]]] \in I$ . Thus,  $(a|b) = \text{tr}(\text{ad } a \text{ ad } b) = 0$ , which is a contradiction.  $\square$

---

<sup>1</sup>Let  $J^{(n)} \neq 0$  and  $J^{(n+1)} = 0$ . Then  $J^{(n)}$  is the desired ideal.

## Chapter 6

# Consequences of Cartan's Criteria

In this chapter we always assume  $L$  is finite dimensional.

### §6.1 Decomposition of semisimple Lie algebras

Let us first review linear algebra.

**Proposition 6.1.** *Let  $V$  be a vector space with a symmetric bilinear form  $(\mid)$  and  $W \subseteq V$  be a finite dimensional subspace on which  $(\mid)_W$  is nondegenerate. Denote by  $W^\perp = \{v \in V : (v|W) = \{0\}\}$ . Then  $V = W \oplus W^\perp$ .*

*Proof.* It is obvious that  $W \cap W^\perp = \{0\}$ . Let  $w_1, w_2, \dots, w_n$  be a basis of  $W$ . For any  $v \in V$ , we have a system of equations

$$(v - \alpha_1 w_1 - \alpha_2 w_2 - \dots - \alpha_n w_n | w_j) = 0, \quad \text{for } j = 1, \dots, n,$$

where  $\alpha_i$ 's are viewed as variables. Since  $\det(w_i | w_j) \neq 0$ , the system has a unique solution, also denoted by  $\alpha_i$ . Thus,  $v - \sum_i \alpha_i w_i \in W^\perp$  as desired.  $\square$

**Remark 6.2.** *If we require  $V$  is finite dimensional and  $(\mid)_V$  is nondegenerate, we can only deduce  $\dim V = \dim W + \dim W^\perp$ .*

**Proposition 6.3.** *If  $L$  is semisimple and  $(0) \neq I \triangleleft L$ , then  $(\mid)_I$  is nondegenerate.*

*Proof.* Assume that  $\exists$  nonzero  $a \in I$  s.t.  $\text{tr}(\text{ad } a \text{ ad } I) = \{0\}$ . Then

$$([a, I] | L) = (a | [I, L]) \subseteq (a | I) = \{0\},$$

which implies  $[a, I] = (0)$ . Then  $I \cap C_L(I) \neq (0)$ . For any  $b \in I \cap C_L(I)$ , we have  $([b, L] | I) = (b | [L, I]) = (b | I) = \{0\}$ . Thus,  $I \cap C_L(I)$  is an abelian ideal of  $L$ . This is a contradiction with semisimplicity of  $L$ .  $\square$

**Theorem 6.4.** *If  $L$  is semisimple, then  $L$  is a direct sum of simple ideals. Moreover, this decomposition is unique.*

*Proof.* Let  $I$  be an ideal of  $L$ . Then by Proposition 6.3 and Proposition 6.1,  $L = I \oplus I^\perp$ . Note that  $I^\perp$  is also an ideal of  $L$  by associativity of the Killing form. Then by induction on  $\dim L$ , we obtain the desired decomposition

$$L = I_1 \oplus I_2 \oplus \dots \oplus I_r,$$

where  $I_i$  are simple ideals of  $L$ .<sup>1</sup>

To show the uniqueness, let  $I$  be a simple ideal of  $L$ . Then  $[I, L] = I$  because  $Z(L) \subseteq \text{Rad}(L) = (0)$ . On the other hand,  $[I, L] = \bigoplus_i [I, I_i]$ . Since  $I$  is simple,  $[I, I_i] \triangleleft I$  is  $I$  itself or  $(0)$ . So all but one summand must be  $(0)$ . Say  $[I, I_i] = I$ . Then  $I = I_i$ .  $\square$

The following is a corollary of Proposition 6.3 and Cartan's Criterion of semisimplicity.

**Corollary 6.5.** *Every ideal of a semisimple Lie algebra  $L$  is still semisimple. Moreover, every quotient of  $L$  is still semisimple.*

*Proof.* The first statement follows immediately from Proposition 6.3 and Theorem 5.9. Let  $I \triangleleft L$  be an ideal. Then by Proposition 6.1,  $L/I \cong I^\perp$ . By the associativity of the Killing form,  $I^\perp$  is also an ideal of  $L$ , and thus semisimple.  $\square$

## §6.2 Derivations

Recall that inner derivations are derivations in  $\text{ad } L$ .

**Theorem 6.6.** *If  $L$  is semisimple, then all derivations are inner.*

*Proof.* Let  $d : L \rightarrow L$  be a derivation. Consider the Lie algebra  $\tilde{L} := L \oplus \mathbb{F}d$ , whose brackets are defined as

$$[d, a] = d(a) = -[a, d].$$

It is straightforward to check it is a well-defined Lie algebra. Note that  $\tilde{L}$  is not semisimple. Otherwise, by Theorem 6.4,  $\tilde{L} = L \oplus (1\text{-dim ideal})$ , which is a contradiction.

Then  $\text{Rad } \tilde{L} \neq 0$ . Say  $a + d \in \text{Rad } \tilde{L}$  for some  $a \in L$ . For any  $b \in L$ , since  $\text{Rad } \tilde{L}$  is an ideal,  $[a + d, b] = [a, b] + d(b) \in \text{Rad } \tilde{L} \cap L$ . Since  $\text{Rad } \tilde{L} \cap L$  is a solvable ideal in  $L$ , it must be  $\{0\}$ . Thus,  $d = -\text{ad } a$ .  $\square$

**Example 6.7.** *There are examples of outer derivations (derivations but not inner).*

1. *Let  $A$  be a commutative associative algebra. Then  $A^{(-)}$  is an abelian Lie algebra and all inner derivations are 0. Thus, all nonzero derivations are outer.*

<sup>1</sup>Actually we need to show if  $J \triangleleft I \triangleleft L$  be an ideal,  $J$  is an ideal of  $L$ . This holds because  $I^\perp$  is an ideal.

2. Let  $A = \text{Mat}(2; \mathbb{F}[X])$ . Consider the Lie algebra  $A^{(-)}$  and a derivation defined by

$$\begin{pmatrix} f(X) & g(X) \\ h(X) & u(X) \end{pmatrix}' = \begin{pmatrix} f'(X) & g'(X) \\ h'(X) & u'(X) \end{pmatrix}.$$

Note that it is NOT an inner derivation: Otherwise, this inner derivation maps all constant matrix (whose elements are constants) to 0. This means a matrix commutes with all constant matrices. It can only be zero matrix. Moreover, it is obvious that this derivation is not a zero map. Hence, we deduce a contradiction.

### §6.3 Abstract Jordan decomposition

Let  $L$  be a Lie algebra. We call an element  $x \in L$  **semisimple** if  $\text{ad } x$  is semisimple, and **nilpotent** if  $\text{ad } x$  is nilpotent.

In this section, we require the field  $\mathbb{F}$  to be algebraically closed. To show the abstract Jordan decomposition for an abstract Lie algebra, we need a lemma.

**Lemma 6.8.** *Let  $A$  be a finite dimensional  $\mathbb{F}$ -algebra. Then  $\text{Der } A$  contains the semisimple and nilpotent parts (in  $\text{End } A$ ) of all its elements.*

*Proof.* Let  $\varphi$  be an arbitrary element in  $\text{Der } A$  and  $\varphi = \varphi_s + \varphi_n$  be the Jordan decomposition in  $\text{End } A$ . It suffices to show that  $\varphi_s \in \text{Der } A$ .

Since  $\mathbb{F}$  is algebraically closed, we have root space decomposition of  $A$  w.r.t. the linear map  $\varphi$ , say  $A = \bigoplus_{\lambda} A^{\lambda}$ , where  $A^{\lambda}$  is the root space corresponding to eigenvalue  $\lambda$ . We assert that  $A^{\lambda} A^{\mu} \subseteq A^{\lambda+\mu}$  by means of the general formula:

$$(\varphi - (\lambda + \mu))^n(xy) = \sum_{i=0}^n \binom{n}{i} (\varphi - (\lambda + \mu))^i(x) \cdot (\varphi - (\lambda + \mu))^{n-i}(y),$$

for any  $x, y \in A$ . Then for any  $x \in A^{\lambda}$  and  $y \in A^{\mu}$ ,  $\varphi_s(xy) = (\lambda + \mu)(xy) = \varphi_s(x)y + x\varphi_s(y)$ , which implies  $\varphi_s \in \text{Der } A$  as desired.  $\square$

**Theorem 6.9.** *For any element  $x \in L$ , there exists a unique decomposition  $x = x_s + x_n$  s.t.  $x_s$  is semisimple,  $x_n$  is nilpotent and  $[x_s, x_n] = 0$ . We call it the **abstract Jordan decomposition of  $x$** .*

*Proof.* For any  $x \in L$ , we have the Jordan decomposition of  $\text{ad } x$  in  $\text{End } L$ . By Theorem 6.6 and Lemma 6.8, the semisimple and nilpotent parts of  $\text{ad } x$  are contained in  $\text{ad } L$ , i.e.,  $\exists x_s, x_n \in L$  s.t.  $\text{ad } x = \text{ad } x_s + \text{ad } x_n$  s.t.  $\text{ad } x_s$  is semisimple,  $\text{ad } x_n$  is nilpotent and  $\text{ad}[x_s, x_n] = [\text{ad } x_s, \text{ad } x_n] = 0$ . Since  $L$  is semisimple, the adjoint representation  $\text{ad} : L \rightarrow \mathfrak{gl}(L)$  is faithful, i.e., one-to-one. ( $\ker \text{ad} = Z(L) = 0$ .) Then  $x = x_s + x_n$  and  $[x_s, x_n] = 0$ . The uniqueness follows from the uniqueness of usual Jordan decomposition and faithfulness of  $\text{ad}$ .  $\square$

# Chapter 7

## Casimir Elements of $L$ -Modules

### §7.1 Tensor products of modules

**Proposition 7.1.** *If  $V$  and  $W$  are two  $L$ -modules, there is a well-defined  $L$ -module structure on  $V \otimes_{\mathbb{F}} W$  by*

$$a(v \otimes w) = av \otimes w + v \otimes aw, \quad \text{for all } v \in V, w \in W.$$

*Proof.* Define  $\varphi : V \times W \rightarrow V \otimes W$  by  $\varphi(v, w) = av \otimes w + v \otimes aw$ , which is a well-defined bilinear map. Thus, by the universal property of tensor products,  $\exists!$   $\tilde{\varphi}$  s.t. the following diagram commutes.

$$\begin{array}{ccc} V \times W & \xrightarrow{\varphi} & V \otimes W \\ & \searrow & \uparrow \tilde{\varphi} \\ & & V \otimes W \end{array}$$

Thus, the action is well-defined. □

**Definition 7.2.** *If  $V$  is an  $L$ -module, then  $V^* = \text{Hom}_{\mathbb{F}}(V, \mathbb{F})$  has an  $L$ -module structure defined by  $[a, \varphi]v = -\varphi(av)$ .*

Furthermore, we can define  $L$ -module structure on  $\text{Hom}_{\mathbb{F}}(V, W)$  by  $[a, \varphi]v = a\varphi(v) - \varphi(av)$ , where  $\varphi \in \text{Hom}_{\mathbb{F}}(V, W)$ . It is a special case of the above proposition if we regard  $\text{Hom}_{\mathbb{F}}(V, W)$  as  $V^* \otimes W$ .

### §7.2 Schur's Lemma

**Lemma 7.3.** *[Schur's Lemma] Let  $\mathbb{F}$  be an algebraically closed field and  $V$  be a finite dim vector space. Let  $\Sigma \subset \text{End}_{\mathbb{F}}(V)$ , and we assume that  $\Sigma$  acts on  $V$  irreducibly (with no nontrivial invariant space). Then the centralizer  $C(\Sigma) = \{\varphi \in \text{End}_{\mathbb{F}}(V) : \varphi a = a\varphi, \forall a \in \Sigma\}$  consists of scalar multiplications.*

*Proof.* Note that  $C(\Sigma)$  is an associative algebra of  $\text{End}_{\mathbb{F}}(V)$ . We will prove it actually is a division algebra. Let  $0 \neq c \in C(\Sigma)$ . Note that both  $\ker c$  and  $\text{im } c$  are  $\Sigma$ -invariant. Then  $\ker c = (0)$  and  $\text{im } c = V$ , which implies  $c$  is invertible. Moreover,  $c^{-1}$  satisfies  $c^{-1}\varphi = c^{-1}\varphi c c^{-1} = \varphi c^{-1}$ . Thus,  $c^{-1} \in C(\Sigma)$ .

**Lemma 7.4.** *The only finite dim division algebra over an algebraically closed field  $\mathbb{F}$  is  $\mathbb{F}$  itself.*

*Proof of Lemma 7.4.* Let  $C$  be a division algebra and  $c \in C$ . Denote  $\dim_{\mathbb{F}} C = d$ . Then the  $d+1$  elements  $1, c, \dots, c^d$  are linearly dependent.  $\exists f \in \mathbb{F}[t]$ , say  $f(t) = \alpha(t - \alpha_1) \dots (t - \alpha_r)$ , s.t.  $f(c) = \alpha(c - \alpha_1 1) \dots (c - \alpha_r 1) = 0$ . Since  $C$  is a division algebra, which is naturally an integral domain,  $\exists i$  s.t.  $c = \alpha_i 1$ . □

Then our claim follows immediately from Lemma 7.4. □

The following is a consequence of Schur's Lemma.

**Corollary 7.5.** *If  $L$  is a finite dim simple Lie algebra, any associative nondegenerate symmetric bilinear form must be a scalar multiple of the Killing form.*

*Proof.* Let  $(|\cdot)_1$  and  $(|\cdot)_2$  be two associative nondegenerate symmetric bilinear forms of  $L$ . Since  $L$  is finite dimensional and the forms are nondegenerate,  $\exists \alpha \in \text{End}_{\mathbb{F}}(L)$  s.t.

$$(u|\cdot)_1 = (\alpha(u)|\cdot)_2, \quad \text{for any } u \in L.$$

By associativity, we have for any  $u, v \in L$

$$(\alpha([u, v])|\cdot)_2 = ([u, v]|\cdot)_1 = (u|[v, \cdot])_1 = (\alpha(u)|[v, \cdot])_2 = ([\alpha(u), v], \cdot)_2.$$

It implies that  $\alpha$  commutes with  $\text{ad } v$  for any  $v \in L$ . By Lemma 7.3,  $\alpha \in C(\text{ad } L)$ . Since  $L$  is simple, namely  $\text{ad } L$  acts irreducibly,  $\alpha$  must be a scalar multiplication. □

## §7.3 Casimir operators

In this section, we always assume  $L$  is a finite dimensional semisimple Lie algebra. Let  $(|\cdot)$  be an associative nondegenerate symmetric bilinear form. Note that  $(|\cdot)$  is not necessarily the Killing form. But by Corollary 7.5, in each simple component, it is a scalar multiple of the Killing form.

Let  $e_1, e_2, \dots, e_n$  be a basis of  $L$  and  $e^1, e^2, \dots, e^n$  be the dual basis with respect to  $(|\cdot)$ , i.e.  $(e_i|e^j) = \delta_{ij}$ . Denote  $c = \sum_i e_i \otimes e^i \in L \otimes_{\mathbb{F}} L$ .

**Proposition 7.6.** *View  $L \otimes L$  as a tensor product of  $L$ -modules.  $L \cdot c = 0$ .*

*Proof.* We need the following lemma.

**Lemma 7.7.** For any  $a, b \in L$ ,  $(\sum_i e_i \otimes e^i | a \otimes b) = (a|b)$ .

*Proof.* Let  $a = \sum_i \alpha_i e^i$  and  $b = \sum_i \beta_i e_i$ . Then

$$\left( \sum_i e_i \otimes e^i | a \otimes b \right) = \sum_{i=1}^n (e_i | a)(e^i | b) = \sum_{i=1}^n \alpha_i \beta_i = (a|b).$$

□

For any  $a \in L$ , we need to show that  $a \cdot c = \sum_{i=1}^n [a, e_i] \otimes e^i + \sum_{i=1}^n e_i \otimes [a, e^i] = 0$ . Since  $(|)_{L \otimes L}$  is nondegenerate<sup>1</sup>, it suffices to show that for any  $b \otimes d \in L \otimes L$ ,  $(a \cdot c | b \otimes d) = 0$ .

$$\begin{aligned} \left( \sum_{i=1}^n [a, e_i] \otimes e^i + \sum_{i=1}^n e_i \otimes [a, e^i] \middle| b \otimes d \right) &= \sum_{i=1}^n ([a, e_i] | b)(e^i | d) + \sum_{i=1}^n (e_i | b)([a, e^i] | d) \\ &= - \sum_{i=1}^n (e_i | [a, b])(e^i | d) - \sum_{i=1}^n (e_i | b)(e^i | [a, d]) \\ &= -([a, b] | d) - (b | [a, d]) \\ &= 0. \end{aligned}$$

□

**Remark 7.8.** Note that  $c$  does not depend on the choice of  $e_1, \dots, e_n$ . If there is another basis  $f_1, \dots, f_n$ , then the bilinear form of  $\sum_i e_i \otimes e^i - \sum_i f_i \otimes f^i$  with every  $a \otimes b \in L \otimes L$  is zero. Since the bilinear form is nondegenerate, these two summations coincide. Therefore, we call this  $c$  the **Casimir element** of  $L$ .

**Remark 7.9.** Let  $A$  be an associative algebra satisfying  $L \subseteq A^{(-)}$  and  $L$  generates  $A$ , i.e.,  $A$  is an associative enveloping algebra of  $L$ . We can regard  $LL \subseteq A$  as an  $L$ -module by  $a \cdot bc := [a, bc] = [a, b]c + b[a, c] \in LL$ . Note that there exists an  $L$ -module homomorphism from  $L \otimes L$  to  $LL$  by  $b \otimes c \mapsto bc$ . We call the image of  $c$  under this homomorphism (also denoted by  $c$ ) the **Casimir operator** of  $A$ . Proposition 7.6 tells us that  $c = \sum_i e_i e^i \in Z(A)$ .

<sup>1</sup>We can check by its Gram matrix. Let  $A = (a_{ij})$  be the Gram matrix of  $(|)_V$  under the basis  $x_1, \dots, x_n$ . Then under the basis  $x_1 \otimes x_1, \dots, x_1 \otimes x_n, x_2 \otimes x_1, \dots, x_n \otimes x_n$ , the Gram matrix of  $(|)_{L \otimes L}$  is

$$\begin{pmatrix} a_{11}A & a_{12}A & \cdots & a_{1n}A \\ a_{21}A & a_{22}A & \cdots & a_{2n}A \\ \vdots & \vdots & & \vdots \\ a_{n1}A & a_{n2}A & \cdots & a_{nn}A \end{pmatrix},$$

which has determinant  $|A|^{2n} \neq 0$ .

# Chapter 8

## Complete Reducibility

### §8.1 Completely reducible modules

Let  $\mathbb{F}$  be a field (the characteristic and algebraic closeness are irrelevant). Let  $A$  be a linear algebra (not necessarily containing unity) and  $V$  be an  $A$ -module (not necessarily finite dimensional).

**Proposition 8.1.** *The following two conditions are equivalent:*

1.  $\forall$  submodule  $W \subseteq V$  is a direct summand,  $\exists$  a submodule  $W' \subseteq V$  s.t.  $V = W \oplus W'$ .
2.  $V = \bigoplus_{i \in I} V_i$ , where  $V_i$ 's are irreducible.

*Proof.*  $1 \Rightarrow 2$ . We claim that it is enough to show that  $V = \sum_{i \in I} V_i$ , for some index set  $I$ . Suppose we have such summation. Let  $I'$  be a maximal subset of  $I$  such that  $\sum_{i \in I'} V_i$  is direct. The existence of  $I'$  is due to Zorn's Lemma. Then for any  $V_j, j \in I \setminus I'$ ,  $(\bigoplus_{i \in I'} V_i) \cap V_j = (0)$  or  $V_j$  by irreducibility of  $V_j$ . By the maximality of  $I'$ , the intersection must be  $V_j$ . Thus,  $V = \bigoplus_{i \in I'} V_i$ .

Now we need to show that  $W := \sum(\text{irreducible submodules}) = V$ . If not, then  $\exists$  nonzero submodule  $W' \subseteq V$  s.t.  $V = W \oplus W'$ . Say  $w \in W'$  and  $w \neq 0$ . Denote by  $P$  the submodule generated by  $w$ . Note that we have a maximal submodule  $Q \subseteq P$ .<sup>1</sup> Suppose  $V = Q \oplus Q'$ . One can prove that  $P' := Q' \cap P$  satisfies  $P = P' \oplus Q$ . Then  $P' \cong P/Q$ . By the maximality of  $Q$ ,  $P'$  is irreducible, contrary to our assumption about  $W$ .

$2 \Rightarrow 1$ . For any  $W \subsetneq V = \bigoplus_i V_i$ , we denote by  $I'$  the maximal subset of  $I$  s.t.  $(\bigoplus_{i \in I'} V_i) \cap W = (0)$ .<sup>2</sup> We claim  $V = (\bigoplus_{i \in I'} V_i) + W$ . It suffices to show that it contains any  $V_j$ . If  $\exists j \in I$  s.t.  $V_j$  is not contained, then

$$\left( \bigoplus_{i \in I' \cup \{j\}} V_i \right) \cap W = (0).$$

<sup>1</sup>Since any proper module does not contain  $w$ , the union of them still does not contain  $w$ . Then by Zorn's Lemma, a maximal submodule  $Q \subseteq P$  exists.

<sup>2</sup>The existence is due to Zorn's Lemma.

This is a contradiction with the maximality of  $I'$ . □

**Definition 8.2.** *An  $A$ -module satisfying either of the equivalent conditions in Proposition 8.1 is called **completely reducible**.*

## §8.2 Weyl's Theorem

Compared with the generality of the last section, in this section, we assume  $L$  is a finite dimensional semisimple Lie algebra over an algebraically closed field  $\mathbb{F}$  of characteristic 0.

**Theorem 8.3** (Weyl's Theorem). *Every finite dimensional  $L$ -module is completely reducible.*

*Proof.* Let  $V$  be a finite dimensional  $L$ -module. It suffices to show that for every submodule  $W \exists$  a submodule of  $V$  as its direct sum complement. The proof is divided into three parts.

**Part I:  $W$  is irreducible and has codim 1.** Note that the trace form  $(a|b) := \text{tr}(a|_W b|_W)$  is nondegenerate symmetric and associative. Symmetry and associativity can be proved similarly to the Killing form. Suppose  $I \triangleleft L$  is an ideal s.t.  $(I|L) = \{0\}$ . Then  $(I|[I, I]) = \{0\}$ . By Cartan's Criterion (Theorem 5.7), the image of  $I$  is solvable. By Corollary 6.5, the image of  $I$  is semisimple, which is a contradiction.

Let  $e_i, 1 \leq i \leq n$  be a basis of  $L$  and  $e^i$  be its dual basis w.r.t.  $(\cdot | \cdot)_W$ . Then  $c = \sum_i e_i e^i$  satisfies  $\text{tr}(c|_W) = \sum_i (e_i | e_i) = n$ . On the other hand, by Proposition 7.6,  $c|_W$  lies in the centralizer of  $L|_W$ . By Schur's Lemma,  $c|_W = \alpha \cdot \text{Id}_W$  for some constant  $\alpha \in \mathbb{F}$ . Then  $\alpha = n / \dim W \neq 0$ .

The quotient  $V/W$  is a 1-dim  $L$ -module. Thus,  $L = [L, L]$  acts as 0 on  $V/W$ .<sup>3</sup> This implies that  $L \cdot V \subseteq W$ . In particular,  $c \cdot V = W$ . So we have  $V = W \oplus \ker c$ . Part I is completed.

**Part II:  $W$  is reducible and has codim 1.** We use induction on the dimension of  $V$ , thanks to Part I. Suppose  $W' \subset W$  is a proper submodule. Then  $W/W' \subseteq V/W'$  also has codim 1. Then by induction,  $\exists v \in V$  s.t.

$$V/W' = (\mathbb{F}v + W')/W' \oplus W/W'$$

with  $\dim(\mathbb{F}v + W')/W' = 1$ . Obviously,  $v \notin W'$  and  $v \notin W$ . By induction hypothesis again,  $\mathbb{F}v + W' = W' \oplus (1\text{-dim } L\text{-module})$ . Moreover, this 1-dim  $L$ -module is not contained in  $W$  since  $v$  does not belong to  $W$ . By a routine check, we have  $V = (1\text{-dim } L\text{-module}) \oplus W$ .

**Part III:  $W$  is arbitrary.** Recall that our goal is to find a direct sum complement of  $W \subseteq V$ . We consider  $\text{Hom}_{\mathbb{F}}(V, W)$  as a subspace of  $\text{End}_{\mathbb{F}}(V)$ . It suffices to find an  $L$ -module homomorphism  $p : V \rightarrow W$  such that  $p^2 = p$  and  $p|_W = \alpha \text{id}_W$  for some nonzero constant  $\alpha$ , because we would get  $V = W \oplus \ker p$ . Let

$$S = \{\varphi \in \text{Hom}_{\mathbb{F}}(V, W) : \varphi|_W = \alpha \cdot \text{id}_W \text{ for some } \alpha \in \mathbb{F}\}.$$

<sup>3</sup> $\forall x \in L, x$  acts as a scalar on  $V/W$ . Thus,  $[L, L]$  acts as 0.

Note that the elements in  $S$  are module homomorphisms:  $\forall \varphi \in S$  and  $a \in L$ ,  $[a, \varphi](w) = a\varphi(w) - \varphi(aw) = 0$ . Let  $S_0 = \{\varphi \in S : \varphi|_W = 0\}$ . Thus,  $S$  is an  $L$ -module because  $[L, S] = S_0 \subseteq S$ . Moreover, we notice that  $S_0$  as a submodule of  $S$  has codim 1. By Part I and Part II,  $\exists$  submodule  $S' \subseteq S$  s.t.  $S = S_0 \oplus S'$ . Then any nonzero  $\varphi \in S'$  is as desired.  $\square$

**Topic** (Comparison with Maschke's Theorem). *Let  $G$  be a finite group and  $\mathbb{F}$  be a field with  $\text{char } \mathbb{F} = 0$ . (The theorem actually holds for  $\text{char } \mathbb{F} \nmid |G|$ .)*

**Theorem 8.4** (Maschke's Theorem). *Any  $G$ -representation (not necessarily finite dim) is completely reducible.*

*Proof.* Let  $V$  be a  $G$ -representation. It suffices to show that any subrepresentation  $W \subseteq V$  has a direct sum complement. Let  $p$  be a projection from  $V$  to  $W$ . Define

$$\varphi = \frac{1}{|G|} \sum_{g \in G} gpg^{-1} \in \text{End}(V, W).$$

Then one can check  $\varphi$  is a homomorphism of  $G$ -representations. Thus,  $V = W \oplus \ker \varphi$ .  $\square$

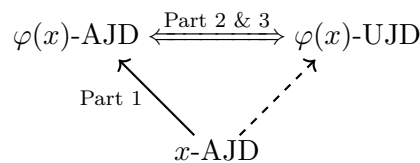
**Remark 8.5.** *We can actually prove Weyl's Theorem following the proof of Maschke's Theorem, when the Lie algebra  $L$  comes from a compact Lie group. But it would be quite complicated and involve too many details (not all Lie algebras come from compact Lie groups).*

### §8.3 Preservation of Jordan decomposition

Now we can generalize Theorem 4.8 to any finite dimensional representation.

**Theorem 8.6.** *If  $(V, \varphi)$  is a finite dimensional representation of  $L$ ,  $\varphi(x) = \varphi(x_s) + \varphi(x_n)$  is the (abstract and usual) Jordan decomposition of the linear map  $\varphi(x)$ , where  $x = x_s + x_n$  is the abstract Jordan decomposition of  $x$ .*

*Proof.* The proof is divided into three parts. The following graph illustrates the main structure.



**Part 1: The abstract Jordan decomposition of  $\varphi(x)$ .** It is obvious that  $\text{ad } \varphi(x_s)$  is semisimple and  $\text{ad } \varphi(x_n)$  is nilpotent. Moreover, we have  $[\varphi(x_s), \varphi(x_n)] = \varphi([x_s, x_n]) = 0$ ,  $\varphi(x) = \varphi(x_s) + \varphi(x_n)$  is the abstract Jordan decomposition for  $\varphi(x) \in \varphi(L)$ .

In the following, we will prove it is also the usual decomposition of  $\varphi(x)$  as a linear map in  $\text{End}(V)$ .

**Part 2:**  $\varphi(L)$  contains the semisimple and nilpotent parts of its elements. Let  $\varphi(x) = \varphi(x)_s + \varphi(x)_n$  be the usual Jordan decomposition of  $\varphi(x)$  in  $\mathfrak{gl}(V)$ . Now we consider the Lie algebra  $\varphi(L) \subseteq \mathfrak{gl}(V)$ , which is finite dimensional and semisimple, thanks to Corollary 6.5. Note that  $\text{ad}_{\mathfrak{gl}(V)} \varphi(x)_s$ , the semisimple part of  $\text{ad}_{\mathfrak{gl}(V)} \varphi(x)$ , is a polynomial with a zero constant term of  $\text{ad}_{\mathfrak{gl}(V)} \varphi(x)$ . It implies that  $\text{ad}_{\mathfrak{gl}(V)} \varphi(x)_s$  sends  $\varphi(L)$  to  $\varphi(L)$ , and so does  $\text{ad}_{\mathfrak{gl}(V)} \varphi(x)_n$ . Thus,  $\text{ad}_{\varphi(L)} \varphi(x)_s$  and  $\text{ad}_{\varphi(L)} \varphi(x)_n$  are derivations of the algebra  $\varphi(L)$ . By Lemma 6.8,  $\exists y_s, y_n \in \varphi(L)$  such that

$$\text{ad}_{\varphi(L)} y_s = \text{ad}_{\varphi(L)} \varphi(x)_s \quad \text{and} \quad \text{ad}_{\varphi(L)} y_n = \text{ad}_{\varphi(L)} \varphi(x)_n.$$

Then  $\varphi(x)_s - y_s$  commutes with any elements in  $\varphi(L)$ . By Weyl's Theorem we can write  $V = \bigoplus_i V_i$ , where  $V_i$ 's are irreducible representations. On each  $V_i$  by Schur's Lemma,  $\varphi(x)_s - y_s$  acts as a scalar multiplication. On the other hand, on each  $V_i$ , we have

$$\text{tr}(\varphi(x)_s - y_s) = \text{tr}(\varphi(x) - y_s) - \text{tr}(\varphi(x)_n) = 0,$$

where the first trace vanishes because  $\varphi(x) - y_s \in \varphi(L) = [\varphi(L), \varphi(L)]$  and the last trace vanishes because  $\varphi(x)_n$  is nilpotent. So  $\varphi(x)_s = y_s \in \varphi(L)$ , and thus  $\varphi(x)_n \in \varphi(L)$ .

**Part 3: The abstract and usual Jordan decompositions coincide.** By Theorem 4.8,  $\text{ad}_{\mathfrak{gl}(V)} \varphi(x)_s$  and  $\text{ad}_{\mathfrak{gl}(V)} \varphi(x)_n$  are semisimple and nilpotent, respectively. Since  $\varphi(L)$  is an invariant subspace of  $\mathfrak{gl}(V)$  w.r.t.  $\text{ad}_{\varphi(x)_s}$  and  $\text{ad}_{\varphi(x)_n}$ ,

$$\text{ad}_{\varphi(L)} \varphi(x)_s = (\text{ad}_{\mathfrak{gl}(V)} \varphi(x)_s)|_{\varphi(L)}$$

is semisimple, and

$$\text{ad}_{\varphi(L)} \varphi(x)_n = (\text{ad}_{\mathfrak{gl}(V)} \varphi(x)_n)|_{\varphi(L)}$$

is nilpotent. By the uniqueness of the abstract Jordan decomposition,  $\varphi(x)_s = \varphi(x_s)$  and  $\varphi(x)_n = \varphi(x_n)$ .  $\square$

## Chapter 9

# Representations of $\mathfrak{sl}_2(\mathbb{F})$

In this chapter we always assume  $\mathbb{F}$  is algebraically closed and characteristic 0. All modules in this chapter are assumed to be finite dim.

Let

$$L = \mathfrak{sl}_2(\mathbb{F}) = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} : a, b, c \in \mathbb{F} \right\}.$$

Take the standard basis

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then

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

**Definition 9.1.** Let  $V$  be an  $\mathfrak{sl}_2$ -module. A nonzero vector  $v \in V$  is called a **weight vector** of weight  $\lambda \in \mathbb{F}$  if  $hv = \lambda v$ . Denote by  $V_\lambda = \{v \in V : hv = \lambda v\}$  the corresponding weight space.

By Theorem 8.6, the action of  $h$  on every finite dim  $\mathfrak{sl}_2$ -module is semisimple. Thus every finite dim  $\mathfrak{sl}_2$ -module  $V$  is a direct sum of its weight spaces, i.e.,

$$V = \bigoplus_{\lambda \in \mathbb{F}} V_\lambda.$$

**Proposition 9.2.** Let  $v \in V_\lambda$ . Then  $e^r v \in V_{\lambda+2r}$  and  $f^r v \in V_{\lambda-2r}$  for all  $r \geq 0$ . Moreover, if  $ev = 0$ , then  $ef^r v = r(\lambda - r + 1)f^{r-1}v$  for all  $r \geq 0$ .

*Proof.* The first assertion follows directly from  $[h, e] = 2e$  and  $[h, f] = -2f$ . For the last formula, use induction on  $r$ . The case  $r = 0$  is obvious. The case  $r = 1$  is  $efv = fev + [e, f]v = \lambda v$ . Suppose the formula holds for  $r$ . Since  $hf^r v = (\lambda - 2r)f^r v$ , we have

$$ef^{r+1}v = (fe + h)f^r v = f(r(\lambda - r + 1)f^{r-1}v) + (\lambda - 2r)f^r v = (r + 1)(\lambda - r)f^r v.$$

This completes the induction.  $\square$

**Remark 9.3.** *This means  $e$  raises weights and  $f$  lowers weights.*

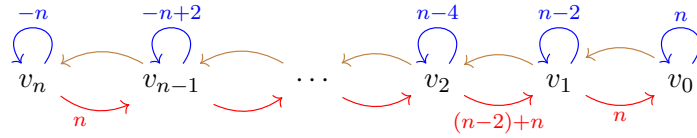
**Theorem 9.4.** *For every  $n \in \mathbb{Z}_{\geq 0}$ , there exists a unique irreducible  $\mathfrak{sl}_2$ -module  $V(n)$  of dimension  $n+1$ . It has a basis  $v_0, v_1, \dots, v_n$  such that  $hv_i = (n-2i)v_i$ ,  $fv_i = v_{i+1}$  for  $0 \leq i < n$ ,  $fv_n = 0$ ,  $ev_0 = 0$ , and  $ev_i = i(n-i+1)v_{i-1}$  for  $1 \leq i \leq n$ . Moreover, every finite dim  $\mathfrak{sl}_2$ -module is a direct sum of such modules.*

*Proof.* Let  $V$  be a nonzero irreducible  $\mathfrak{sl}_2$ -module. Since  $h$  acts semisimply,  $V$  contains a weight vector. Choose a weight vector  $v_0$  of maximal weight, say  $hv_0 = nv_0$ . Then  $ev_0 = 0$ . Put  $v_i = f^i v_0$ . By Proposition 9.2,  $ev_i = i(n-i+1)v_{i-1}$  for all  $i \geq 1$ .

Since  $V$  is finite dim, there is a minimal  $r \geq 0$  such that  $v_{r+1} = 0$ . Applying  $e$  to  $v_{r+1} = 0$  gives  $(r+1)(n-r)v_r = 0$ . Since  $v_r \neq 0$ , we get  $n = r \in \mathbb{Z}_{\geq 0}$ . Hence  $v_0, \dots, v_n$  are nonzero and have distinct weights, so they are linearly independent. Their span is a nonzero submodule of  $V$ , hence it is all of  $V$  by irreducibility. This gives the required basis and formulas.

Conversely, the formulas above define an  $\mathfrak{sl}_2$ -module. Any nonzero submodule contains a weight vector; applying  $e$  repeatedly gives a nonzero multiple of  $v_0$ , and then applying  $f$  repeatedly gives all basis vectors. Thus the module is irreducible. The last statement follows from Weyl's Theorem.  $\square$

The following graph illustrates the irreducible module  $V(n)$ .



In the graph, we use **red**, **brown** and **blue** arrows to illustrate the actions of  $e$ ,  $f$  and  $h$ , respectively.

**Theorem 9.5** ( $\mathfrak{sl}_2$ -strings). *Let  $V$  be a finite dim irreducible  $\mathfrak{sl}_2$ -module and let  $0 \neq v \in V_\lambda$ . Suppose  $p, q \in \mathbb{Z}_{\geq 0}$  are determined by  $e^p v \neq 0$ ,  $e^{p+1} v = 0$ ,  $f^q v \neq 0$  and  $f^{q+1} v = 0$ . Then  $\lambda = q - p$ .*

*Proof.* By Theorem 9.4,  $V \simeq V(n)$  for some  $n \in \mathbb{Z}_{\geq 0}$ . Hence  $V$  has a basis  $v_0, v_1, \dots, v_n$  such that  $hv_i = (n-2i)v_i$ ,  $fv_i = v_{i+1}$  and  $ev_i = i(n-i+1)v_{i-1}$ .

Since the weights  $n, n-2, \dots, -n$  are distinct, each weight space is one dimensional. Therefore, if  $0 \neq v \in V_\lambda$ , then  $v$  is a scalar multiple of some  $v_i$ . For this  $i$ , we have  $\lambda = n-2i$ . Moreover,  $e$  can act nontrivially on  $v_i$  exactly  $i$  times, while  $f$  can act nontrivially on  $v_i$  exactly  $n-i$  times. Thus  $p = i$  and  $q = n-i$ . Therefore

$$q - p = (n - i) - i = n - 2i = \lambda.$$

$\square$

# Chapter 10

## Root Space Decomposition

Throughout this chapter  $L$  denotes a finite dim semisimple Lie algebra over an algebraically closed field  $\mathbb{F}$  of characteristic 0.

### §10.1 Maximal toral subalgebras

**Definition 10.1.** Let  $L$  be a Lie algebra. A Lie subalgebra  $T \subseteq L$  is called a **toral subalgebra** if  $\forall a \in T$ ,  $a$  is semisimple (in  $L$ ).

**Lemma 10.2.** Any toral subalgebra  $T \subseteq L$  is abelian.

*Proof.* For any nonzero  $a \in T$ ,  $T$  can be decomposed as a direct sum of eigenspaces of  $a$ :

$$T = \bigoplus_{\alpha} T_{\alpha}.$$

For any  $b \in T_{\alpha}$  for some  $\alpha \neq 0$ , we have  $\text{ad}(b)L_{\beta} \subseteq L_{\alpha+\beta}$  for all  $\beta$ . Note that there exists  $k \geq 1$  s.t.  $\beta + k\alpha$  is not an eigenvalue. Then  $(\text{ad } b)^k L_{\beta} = (0)$ . Thus,  $b$  is nilpotent. But  $b \in T$  is semisimple by definition. So  $b = 0$ .  $\square$

Let  $H$  be a **maximal toral subalgebra**.<sup>1</sup> By Lemma 10.2,  $L$  can be decomposed into a direct sum of common eigenspaces of  $H$ , i.e.,

$$L = \bigoplus_{\alpha \in H^*} L_{\alpha},$$

where  $L_{\alpha} = \{x \in L : \text{ad } h(x) = [h, x] = \alpha(h)x \text{ for all } h \in H\}$ . By abuse of notation, we sometimes may denote by  $\langle \alpha, h \rangle$  or  $\langle h, \alpha \rangle$  the evaluation  $\alpha(h)$ .

Denote  $\Delta = \{\text{nonzero eigenfunctions}\} = \{\alpha \in H^* \setminus \{0\} : L_{\alpha} \neq (0)\}$ . Recall that we denote by  $( | )$  the Killing form.

**Lemma 10.3.** If  $\alpha, \beta \in \Delta \cup \{0\}$  and  $\alpha + \beta \neq 0$ , then  $(L_{\alpha} | L_{\beta}) = \{0\}$ .

<sup>1</sup>Notice that  $H$  exists because  $L$  is semisimple and finite dimensional.

*Proof.* Since  $\alpha + \beta \neq 0$ ,  $\exists$  nonzero  $h \in H$  s.t.  $\langle \alpha + \beta, h \rangle \neq 0$ . For any  $x \in L_\alpha, y \in L_\beta$ ,

$$([x, h]|y) - (x|[h, y]) = \langle \alpha + \beta, h \rangle (x|y) = 0,$$

which implies  $(L_\alpha|L_\beta) = 0$ . □

**Corollary 10.4.** *The restriction<sup>a</sup> of the Killing form on  $L_0$  is nondegenerate.*

<sup>a</sup>Note that the restriction means the form  $(\cdot|\cdot)_{L_0} = \text{tr}(\text{ad}_L \cdot \text{ad}_L \cdot)_{L_0 \times L_0}$ , instead of the Killing form of  $L_0$  (as a Lie subalgebra)  $\text{tr}(\text{ad}_{L_0} \cdot \text{ad}_{L_0} \cdot)$ . However, they coincide when  $L_0$  is an ideal, but  $L_0$  is obviously not.

*Proof.* If  $x \in L_0$  is orthogonal to  $L_0$ , then  $x$  is orthogonal to the whole  $L$ , thanks to Lemma 10.3. Since  $(\cdot|\cdot)_L$  is nondegenerate if and only if  $L$  is semisimple,  $x = 0$ . □

**Theorem 10.5.**  $H = L_0$ .

*Proof.* Thanks to Lemma 10.3,  $H \subseteq L_0$ . But the converse is much harder.

Let  $a \in L_0$  be an arbitrary element with the Jordan decomposition  $a = a_s + a_n$ . Then  $\text{ad } a_s$  and  $\text{ad } a_n$  are polynomials of  $\text{ad } a$  with zero constant terms. Thus,  $\text{ad } a_s(H) = \text{ad } a_n(H) = (0)$ , which means  $a_s$  and  $a_n$  are contained in  $L_0$ . By the semisimplicity of  $a_s$  and its commutation with  $H$ ,  $H + \mathbb{F}a_s$  still acts semisimply. Hence,  $a_s \in H$  due to the maximality of  $H$ .<sup>2</sup>

Now we need to show  $(a_n|L_0) = \{0\}$ . Once this is established, Corollary 10.4 yields  $a_n = 0$ , and consequently  $a = a_s \in H$ , which completes the argument.

Note that  $L_0$  is nilpotent. Since  $\text{ad}_{L_0} a_s = 0$ ,  $\text{ad}_{L_0} a = \text{ad}_{L_0} a_s + \text{ad}_{L_0} a_n = \text{ad}_{L_0} a_n$ . Because  $a_n$  is nilpotent,  $\text{ad}_{L_0} a$  is also nilpotent. By Engel's Theorem,  $L_0$  is nilpotent.

By Lie's theorem<sup>3</sup>,  $L_0$  can be simultaneously upper triangularized under the representation  $\text{ad}_L : L_0 \rightarrow \mathfrak{gl}(L)$ . In particular,  $\text{ad}_L a_n$  corresponds to a strictly upper triangular matrix. Thus,  $(a_n|L_0) = 0$  as required. □

Combining Corollary 10.4 and Theorem 10.5, we immediately get the following corollary.

**Corollary 10.6.** *The restriction of the Killing form on  $H$  is nondegenerate.*

**Summary.** *For a semisimple Lie algebra  $L$ , we call the decomposition*

$$L = L_0 + \sum_{0 \neq \alpha \in H^*} L_\alpha = H + \sum_{0 \neq \alpha \in H^*} L_\alpha$$

*a root decomposition of  $L$ .*

<sup>2</sup>Results of this paragraph hold for all elements in  $L_0$  since  $a$  is arbitrary.

<sup>3</sup>Nilpotency implies solvability.

## §10.2 The $\mathfrak{sl}_2$ copies

Note that Lemma 10.3 and Theorem 5.9 imply  $L_\alpha$  and  $L_{-\alpha}$  are dual spaces relative to the Killing form  $(\cdot|\cdot)$ : if  $x \in L_\alpha$  is orthogonal to  $L_{-\alpha}$ , then it is orthogonal to the whole  $L$ . Then the nondegeneracy of  $(\cdot|\cdot)$  forces  $x = 0$ . In this section, we will connect the brackets with the Killing forms.

Corollary 10.6 allows us to identify  $H$  with  $H^*$ . It will be a vital tool in the following discussion.

**Definition 10.7.** We define an isomorphism  $\nu$  from  $H$  to  $H^*$  (as vector spaces) by  $\nu(h) = (h|\cdot)$ . Namely,  $\nu(h)(h') = (h|h')$ , and  $(\nu^{-1}(\alpha)|h) = \langle \alpha, h \rangle$  for any  $\alpha \in H^*$  and  $h, h' \in H$ .

**Lemma 10.8.**  $\forall x_\alpha \in L_\alpha$  and  $y_\alpha \in L_{-\alpha}$ , for some  $\alpha \in \Delta$ ,  $[x_\alpha, y_\alpha] = (x_\alpha|y_\alpha)\nu^{-1}(\alpha)$ .

*Proof.* By the associativity of the Killing form,  $([h, x_\alpha]|y_\alpha) = (h|[x_\alpha, y_\alpha])$ . On the other hand,

$$([h, x_\alpha]|y_\alpha) = \langle \alpha, h \rangle (x_\alpha|y_\alpha) = (h|\nu^{-1}(\alpha))(x_\alpha|y_\alpha) = (h|(x_\alpha|y_\alpha)\nu^{-1}(\alpha)).$$

By the nondegeneracy of  $(\cdot|\cdot)_H$ ,  $[x_\alpha, y_\alpha] = (x_\alpha|y_\alpha)\nu^{-1}(\alpha)$ . □

**Lemma 10.9.** If  $\alpha \in \Delta$ ,  $(\alpha|\alpha) \neq 0$ .

*Proof.* Thanks to the dual relation, we can take some  $x_\alpha \in L_\alpha$  and  $y_\alpha \in L_{-\alpha}$  s.t.  $(x_\alpha|y_\alpha) \neq 0$ . Then, by Lemma 10.8,  $\mathbb{F}x_\alpha + \mathbb{F}\nu^{-1}(\alpha) + \mathbb{F}y_\alpha$  is a subalgebra of  $L$ , denoted by  $S$ . It is nilpotent since  $[\nu^{-1}(\alpha), x_\alpha] = (\alpha|\alpha)x_\alpha = 0$ . So  $\mathbb{F}\nu^{-1}(\alpha) = [S, S]$  acts nilpotently on  $L$ , which is a contradiction with the semisimplicity of  $\nu^{-1}(\alpha) \in H$ . □

**Remark 10.10.** In the next section, we will prove that  $(\alpha|\alpha) > 0$ .

**Proposition 10.11.** For any  $\alpha \in \Delta$  and  $0 \neq x_\alpha \in L_\alpha$ , we can find  $y_\alpha \in L_{-\alpha}$  s.t.  $h_\alpha = [x_\alpha, y_\alpha]$  and  $(x_\alpha, y_\alpha, h_\alpha)$  form an  $\mathfrak{sl}_2$ -triple.

*Proof.* We can find  $y_\alpha \in L_{-\alpha}$  s.t.  $(x_\alpha|y_\alpha) = \frac{2}{(\alpha|\alpha)}$ . Then  $h_\alpha = \frac{2\nu^{-1}(\alpha)}{(\alpha|\alpha)}$  and

$$[h_\alpha, x_\alpha] = \frac{2}{(\alpha|\alpha)}[\nu^{-1}(\alpha), x_\alpha] = \frac{2}{(\alpha|\alpha)} \langle \nu^{-1}(\alpha), \alpha \rangle x_\alpha = 2x_\alpha.$$

Similarly, we have  $[h_\alpha, y_\alpha] = -2y_\alpha$ . □

## §10.3 Integral properties

Denote by  $S_\alpha$  the  $\mathfrak{sl}_2$  copy in Proposition 10.11.

**Proposition 10.12.** *If  $\alpha \in \Delta$ , then  $k\alpha \in \Delta$  for  $k \in \mathbb{F}$  if and only if  $k = \pm 1$ .*

*Proof.* The subspace  $M = \bigoplus_{k \in \mathbb{F}} L_{k\alpha}$  bears a natural  $S_\alpha$ -module structure. The weight space  $L_{k\alpha}$  corresponds to the eigenspace of eigenvalue  $2k$ . Since  $L$  is finite-dimensional,  $2k \in \mathbb{Z}$ , i.e.,  $k \in \frac{1}{2}\mathbb{Z}$ . By the complete reducibility,  $M$  is the sum of all irreducible  $S_\alpha$ -submodules. Moreover, we can divide these submodules into 2 parts by the parity of weights.

Let  $M_0$  be an irreducible submodule with even weights. Then  $M_0$  contains eigenvalue 0. Namely,  $M_0$  intersects  $H$  nontrivially. Note that  $H + S_\alpha$  is a  $S_\alpha$ -submodule, which the irreducible submodule  $M_0$  intersects nontrivially. Thus,  $M_0 \subseteq H + S_\alpha$ . It implies that  $2\alpha \notin \Delta$  if  $\alpha \in \Delta$ .

Now we can deduce that submodules of  $M$  with odd weights can not exist. Otherwise,  $\alpha/2 \in \Delta$  and  $\alpha \in \Delta$ , contrary to the result above.

In conclusion, we prove that  $M = H \oplus S_\alpha = L_{-\alpha} \oplus L_0 \oplus L_\alpha$ .  $\square$

From the proof, these two corollaries follows.

**Corollary 10.13.** *If  $\alpha \in \Delta$ , then  $\dim L_\alpha = 1$*

**Corollary 10.14.** *If  $\alpha, \beta, \alpha + \beta \in \Delta$ ,  $L_{\alpha+\beta} = [L_\alpha, L_\beta]$ .*

Furthermore, if we consider the space  $\bigoplus_{i \in \mathbb{Z}} L_{\beta+i\alpha}$  for two roots  $\alpha, \beta$ , we have the following proposition.

**Proposition 10.15.** *Suppose  $\alpha, \beta \in \Delta$ ,  $\bigoplus_{i \in \mathbb{Z}} L_{\beta+i\alpha}$  is an irreducible  $S_\alpha$ -module.*

*Proof.*  $L_\beta$  as a weight space of  $S_\alpha$  has weight  $\langle \beta, h_\alpha \rangle = \frac{2(\beta|\alpha)}{(\alpha|\alpha)} \in \mathbb{Z}$ .<sup>4</sup> Moreover,  $L_{\beta+i\alpha}$  has weight  $\frac{2(\beta|\alpha)}{(\alpha|\alpha)} + 2i$ . Since each weight space has dimension 1 (by Corollary 10.13) and all weights have the same parity,  $\bigoplus_{i \in \mathbb{Z}} L_{\beta+i\alpha}$  is irreducible as an  $S_\alpha$ -module.  $\square$

## §10.4 The inner product

Since  $(\cdot | \cdot)_{H^*}$  is a nondegenerate symmetric bilinear form, we have  $H^* = \mathbb{F}\alpha \oplus \ker(\alpha|\cdot)$ . We consider the symmetry of this decomposition, i.e., the reflection  $s_\alpha$  w.r.t. the hyperplane  $\ker(\alpha|\cdot)$ . Explicitly,

$$s_\alpha : x \mapsto x - \frac{2(x|\alpha)}{(\alpha|\alpha)} \alpha.$$

**Proposition 10.16.**  $s_\alpha(\Delta) = \Delta$ .

*Proof.* For an arbitrary  $\beta \in \Delta$ , suppose that  $\beta - q\alpha, \beta - (q-1)\alpha, \dots, \beta + p\alpha$  is the  $\alpha$ -string of  $\beta$ . Then  $\frac{2(\beta|\alpha)}{(\alpha|\alpha)} = q - p$ . Thus  $-q \leq \frac{2(\beta|\alpha)}{(\alpha|\alpha)} \leq p$  and  $\beta - \frac{2(\beta|\alpha)}{(\alpha|\alpha)}\alpha$  is in the  $\alpha$ -string.  $\square$

<sup>4</sup>One can check by  $\mathfrak{sl}_2$ -theory that  $\langle \beta, h_\alpha \rangle = q - p$ , where  $\beta - q\alpha, \beta - (q-1)\alpha, \dots, \beta + p\alpha$  is the  $\alpha$ -string of  $\beta$ .

**Lemma 10.17.**  $\text{span}(\Delta) = H^*$ .

*Proof.* Otherwise  $\exists 0 \neq h \in H$  s.t.  $\langle h, \text{span}(\Delta) \rangle = 0$ . Then  $h$  commutes with  $L_\alpha$  for any  $\alpha \in \Delta$  and with  $H$ , i.e.,  $h \in Z(L)$ , contrary to the fact that  $L$  is semisimple.  $\square$

**Lemma 10.18.**  $(\alpha|\beta) \in \mathbb{Q}$ , for all  $\alpha, \beta \in \Delta$ .

*Proof.* Since  $(\alpha|\beta) = \frac{(\beta|\alpha)}{(\alpha|\alpha)} \cdot (\alpha|\alpha)$  and  $\frac{2(\beta|\alpha)}{(\alpha|\alpha)} \in \mathbb{Z}$ , it suffices to prove  $(\alpha|\alpha) \in \mathbb{Q}$  for all  $\alpha \in \Delta$ . Note that

$$(\alpha|\alpha) = (\nu^{-1}(\alpha)|\nu^{-1}(\alpha)) = \text{tr}(\nu^{-1}(\alpha)^2) = \sum_{\beta \in \Delta} (\beta|\alpha)^2 = \sum (\text{rat. num.})^2 (\alpha|\alpha)^2.$$

Thus  $(\alpha|\alpha) = \frac{1}{\sum (\text{rat. num.})^2}$  is still rational.  $\square$

From the proof, we immediately obtain the following corollary.

**Corollary 10.19.**  $(\alpha|\alpha) > 0$  for all  $\alpha \in \Delta$ .

By Lemma 10.17, we can choose  $\alpha_1, \alpha_2, \dots, \alpha_n \in \Delta$  as a basis of  $H^*$ . Then every  $\beta \in \Delta$ ,  $\beta = \sum_{1 \leq i \leq n} k_i \alpha_i$  for some  $k_i \in \mathbb{F}$ .

**Proposition 10.20.** All  $k_i \in \mathbb{Q}$ .

*Proof.* We can regard  $k_i$  as  $n$  unknowns in the equation system  $(\beta|\alpha_j) = \sum_i k_i (\alpha_i|\alpha_j)$  for all  $1 \leq j \leq n$ . By Lemma 10.18, the coefficient matrix and  $(\beta|\alpha_j)$ 's are rational. Thus, the solution  $k_i$ 's are rational.  $\square$

**Proposition 10.21.**  $(|\cdot) : \mathbb{Q}\Delta \times \mathbb{Q}\Delta \rightarrow \mathbb{Q}$  is positive definite.

*Proof.* Let  $\lambda = \sum_i k_i \alpha_i \in \mathbb{Q}\Delta$ .  $(\lambda|\lambda) = (\nu^{-1}\lambda|\nu^{-1}\lambda) = \text{tr}(\text{ad}(\nu^{-1}\lambda)^2) = \sum_{\alpha \in \Delta} (\lambda|\alpha)^2 \geq 0$  and the equation holds if and only if  $\lambda = 0$ .  $\square$

Let  $E = \mathbb{Q}\Delta \otimes_{\mathbb{Q}} \mathbb{R}$ . With the inner product  $(|\cdot)$ ,  $E \cong \mathbb{R}^n$  is a Euclidean space.

**Theorem 10.22.** In this Euclidean space  $E$ ,  $\Delta$  has the following properties:

1.  $\Delta \subseteq E \setminus \{0\}$ ,  $|\Delta| < \infty$  and  $\text{span}_{\mathbb{R}}(\Delta) = E$ ;
2.  $\mathbb{R}\alpha \cap \Delta = \{\alpha, -\alpha\}$ ;
3.  $\forall \alpha \in \Delta$ ,  $s_\alpha(\Delta) = \Delta$ ;
4.  $\forall \alpha, \beta \in \Delta$ ,  $\frac{2(\beta|\alpha)}{(\alpha|\alpha)} \in \mathbb{Z}$ ;

# Chapter 11

## Abstract Root Systems I

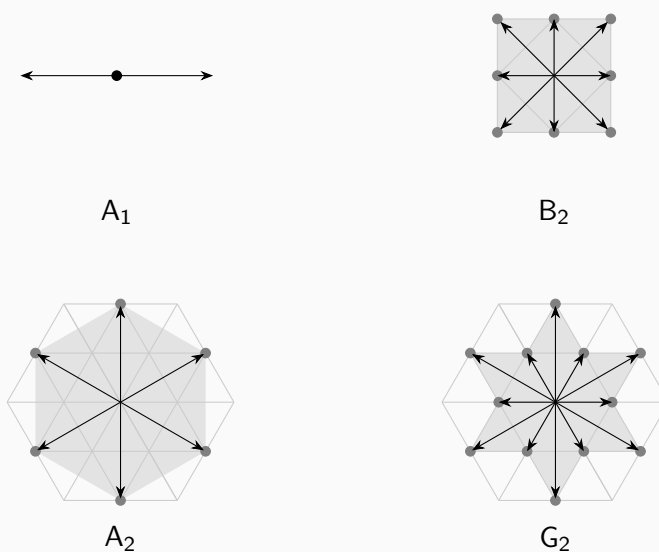
Throughout this chapter we are concerned with a fixed Euclidean space  $E$ , which is a finite dim vector space over  $\mathbb{R}$  with a positive definite inner product  $(\cdot | \cdot)$ . In the Euclidean space  $E$ , we still have the reflection map  $s_\alpha$  with respect to some nonzero vector  $\alpha$ .

### §11.1 Irreducible root systems

**Definition 11.1.** A subset  $\Delta$  of the Euclidean space  $E$  is called a **root system** in  $E$  if  $\Delta$  satisfies the conditions in Theorem 10.22.

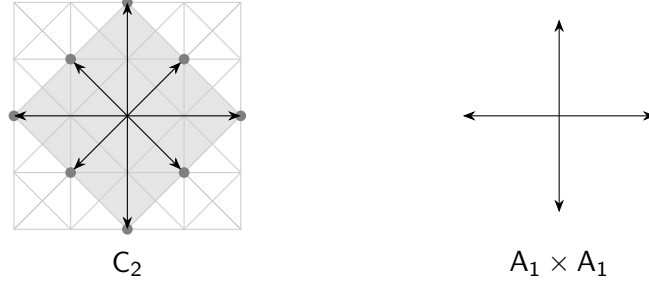
**Definition 11.2.** If a root system  $\Delta$  can be decomposed as a disjoint union  $\Delta_1 \sqcup \Delta_2$  satisfying  $\Delta_i \neq \emptyset$ ,  $(\Delta_1 | \Delta_2) = \{0\}$  and  $\text{span}(\Delta_1) \oplus \text{span}(\Delta_2) = E$  then  $\Delta$  is called **reducible**; otherwise it is called **irreducible**.

**Example 11.3.** There are some examples of root systems. Theorem 12.16 will prove that they are all irreducible root systems (up to isomorphism<sup>a</sup>) in  $\mathbb{R}$  and  $\mathbb{R}^2$ .



<sup>a</sup>Later we will give the definition.

**Remark 11.4.** 1. Note that although the root system of type  $C_2$  seems different, it is isomorphic to the root system of type  $B_2$ . Furthermore, they are dual to each other.



2. Type  $A_1 \times A_1$  can also be embedded in  $\mathbb{R}^2$ , but it is not irreducible.

## §11.2 Angles between roots

Let  $\alpha, \beta \in \Delta$  and  $\theta$  be the angle between them. We have

$$\frac{2(\beta|\alpha)}{(\alpha|\alpha)} \cdot \frac{2(\alpha|\beta)}{(\beta|\beta)} = 4 \frac{\|\beta\| \|\alpha\| \cos \theta \cdot \|\alpha\| \|\beta\| \cos \theta}{\|\alpha\|^2 \cdot \|\beta\|^2} = 4(\cos \theta)^2. \quad (\star)$$

Since  $0 \leq (\cos \theta)^2 \leq 1$  and the left-hand side of  $(\star)$  is a nonnegative integer,  $4(\cos \theta)^2 = 0, 1, 2, 3, 4$ . Using high-school knowledge, we have all possible values of  $\theta$ .

|               |       |                       |                       |                  |                 |                 |                      |                      |   |
|---------------|-------|-----------------------|-----------------------|------------------|-----------------|-----------------|----------------------|----------------------|---|
| $\cos \theta$ | -1    | $-\frac{\sqrt{3}}{2}$ | $-\frac{\sqrt{2}}{2}$ | $-\frac{1}{2}$   | 0               | $\frac{1}{2}$   | $\frac{\sqrt{2}}{2}$ | $\frac{\sqrt{3}}{2}$ | 1 |
| $\theta$      | $\pi$ | $\frac{5\pi}{6}$      | $\frac{3\pi}{4}$      | $\frac{2\pi}{3}$ | $\frac{\pi}{2}$ | $\frac{\pi}{3}$ | $\frac{\pi}{4}$      | $\frac{\pi}{6}$      | 0 |

Now we consider the case of  $(\beta|\alpha) \neq 0$  and  $\beta \neq \pm\alpha$ . In the product at left-hand side of  $(\star)$ , at least one term is 1. WLOG, say  $\frac{2(\beta|\alpha)}{(\alpha|\alpha)} = 1$ . Then

$$\frac{\|\beta\|}{\|\alpha\|} = \frac{2\|\beta\|\|\alpha\| \cos \theta}{\|\alpha\|^2} \cdot \frac{1}{2 \cos \theta} = \frac{1}{2 \cos \theta}.$$

Combining with the condition (2) in Theorem 10.22, we can determine the ratio of the lengths of two roots from the angle  $\theta$  between them, except when  $\theta = \frac{\pi}{2}$ .

## §11.3 Simple roots

**Definition 11.5.** A subset  $B = \{\beta_1, \beta_2, \dots, \beta_n\} \subset \Delta$  is called a **base** if

1.  $B$  is a basis of  $E$ ;
2.  $\forall \alpha \in \Delta$ ,  $\alpha = \sum_i k_i \beta_i$  and either all  $k_i \in \mathbb{Z}_{\geq 0}$  or all  $k_i \in \mathbb{Z}_{< 0}$

Suppose  $B$  is a base of  $\Delta$ . Then  $\Delta = \Delta_+ \sqcup \Delta_-$  by the definition of  $B$ . We say  $(\widehat{\beta_i, \beta_j})$  is **obtuse** if  $(\widehat{\beta_i, \beta_j}) \leq 0$ .

**Proposition 11.6.**  $\forall \beta_1 \neq \beta_2 \in B, (\beta_1|\beta_2) \leq 0$ .

*Proof.* Suppose  $(\beta_1|\beta_2) > 0$ . Then at least one of  $\frac{2(\beta|\alpha)}{(\alpha|\alpha)}$  and  $\frac{2(\alpha|\beta)}{(\beta|\beta)}$  is 1. Say  $\frac{2(\beta|\alpha)}{(\alpha|\alpha)} = 1$ . Then by condition (3) in Theorem 10.22,  $s_{\beta_2}(\beta_1) = \beta_1 - \frac{2(\beta|\alpha)}{(\alpha|\alpha)}\beta_2 = \beta_1 - \beta_2 \in \Delta$ . WLOG, we assume  $\beta_1 - \beta_2 = \sum k_i \beta_i$  with all  $k_i \geq 0$ . But since  $B$  is a basis of  $E$ , the expression is unique, which implies  $k_1 = 1, k_2 = -1$ , contrary to the assumption.  $\square$

Now we focus on the existence of the base. Consider the hyperplanes  $\alpha^\perp, \forall \alpha \in \Delta$ . We have  $\bigcup_{\alpha \in \Delta} \alpha^\perp \neq E$ .<sup>1</sup>

**Definition 11.7.** We call  $\gamma \in E$  **regular** if  $\gamma \in E \setminus (\bigcup_{\alpha \in \Delta} \alpha^\perp)$ , i.e.,  $(\gamma|\alpha) \neq 0$  for all  $\alpha \in \Delta$ .

Then for a regular vector  $\gamma$ , we obtain  $\Delta = \Delta_+ \sqcup \Delta_-$ , where  $\Delta_+ = \{\alpha \in \Delta : (\gamma|\alpha) > 0\}$  and  $\Delta_- = -\Delta_+$ . We call the decomposition a **regular decomposition** of  $\Delta$ . Denote by

$$B(\gamma) = \{\beta \in \Delta_+ | \beta \text{ is not a sum of elements in } \Delta_+\}.$$

We are going to show this  $B$  is a base.

**Remark 11.8.**  $B(\gamma)$  has the following properties:

1.  $B(\gamma)$  **is not empty**.  $\{\beta \in \Delta_+ : (\beta|\gamma) \text{ is minimal among } (\Delta_+|\gamma)\} \subset B(\gamma)$  because if  $\beta = \alpha_1 + \alpha_2 + \dots + \alpha_m$  with  $\alpha_i \in \Delta_+$  then  $(\beta|\gamma) > (\alpha_i|\gamma)$ .
2.  $B(\gamma)$  **satisfies Definition 11.5(2)**. It follows directly from the definition.
3.  $B(\gamma)$  **spans**  $\Delta$ . It follows directly from the construction.
4.  $B(\gamma)$  **satisfies Proposition 11.6**. Suppose  $\exists (\beta_1|\beta_2) > 0$ . By the same argument in the proof of Proposition 11.6,  $\beta_1 - \beta_2 \in \Delta$ . Then either  $\beta_1 - \beta_2 \in \Delta_+$  or  $\beta_2 - \beta_1 \in \Delta_-$ . In the first case, we deduce a contradiction that  $\beta_1 = (\beta_1 - \beta_2) + \beta_2$  is decomposed; In the second case, we have that  $\beta_2 = (\beta_2 - \beta_1) + \beta_1$  is decomposed.

Now it only remains to show  $B(\gamma)$  is linearly independent which can be proved by the following lemma.

**Lemma 11.9.** Let  $P = \{\xi_1, \xi_2, \dots, \xi_r\}$  be a finite subset of nonzero vectors in  $E$  satisfying

1.  $\forall i \neq j, (\xi_i|\xi_j) \leq 0$ ;
2.  $\exists 0 \neq \gamma \in E$  s.t.  $(\gamma|P) > 0$ .

Then  $P$  is linearly independent.

<sup>1</sup>Over an infinite field, a finite union of proper subspaces can not cover the whole space.

*Proof.* Let  $\emptyset \neq I \subset \{1, 2, \dots, r\}$  and  $\sum_{i \in I} k_i \xi_i = 0$  for some  $k_i \in \mathbb{R} \setminus \{0\}$ . Then  $(\gamma | \sum_i k_i \xi_i) = \sum_i k_i (\gamma | \xi_i) = 0$ , where  $(\gamma | \xi_i) > 0$ . We can divide  $I$  into two parts by the signs of  $k_i$ 's. Say  $k_i > 0$  if  $i \in I_1 \subset I$  and  $k_j < 0$  if  $j \in I_2 = I \setminus I_1$ . Then  $\sum_{i \in I_1} k_i \xi_i = \sum_{j \in I_2} -k_j \xi_j$  and

$$0 < \left( \sum_{i \in I_1} k_i \xi_i \middle| \sum_{i \in I_1} k_i \xi_i \right) = \left( \sum_{i \in I_1} k_i \xi_i \middle| \sum_{j \in I_2} -k_j \xi_j \right) \leq 0.$$

□

**Summary.** We can obtain this one-to-one correspondence for any root system  $\Delta$ :

$$\{\text{Bases}\} \xleftrightarrow{1-1} \{\text{Regular decompositions}\}$$

by

$$\begin{aligned} & B \mapsto \{\pm\text{-combination of } B\} \\ \left\{ \begin{array}{l} \text{all indecomposable} \\ \text{roots in } \Delta_+ \end{array} \right\} & \leftrightarrow \Delta_+ \sqcup \Delta_- \end{aligned}$$

# Chapter 12

## Abstract Root Systems II

### §12.1 Weyl chambers and Weyl groups

Note that hyperplanes  $\alpha^\perp$ ,  $\alpha \in \Delta$  divide  $E$ .

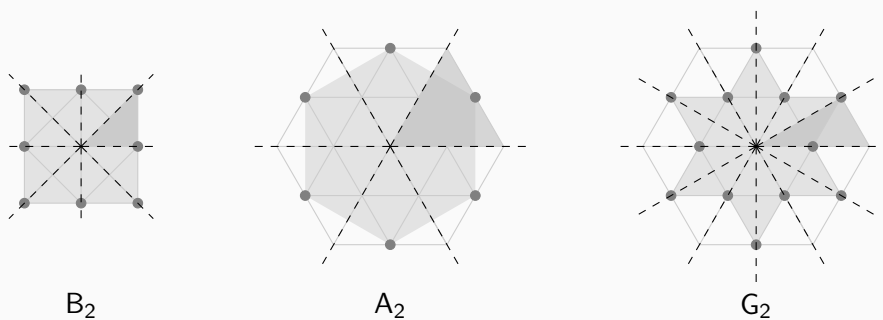
**Definition 12.1.** A *Weyl chamber* of the root system  $\Delta$  is a nonempty subset

$$\left\{ \gamma \in E \setminus \bigcup \alpha^\perp : (\gamma|\beta) < 0 \ \forall \beta \in \Delta' \text{ and } (\gamma|\beta) < 0 \text{ otherwise} \right\}$$

for some subset  $\Delta' \subset \Delta$ .

Note that every root system  $\Delta$  has finitely many Weyl chambers because  $\Delta' \in 2^\Delta$ .

**Example 12.2.** In the figures below, dashed lines indicate hyperplanes and the darker region illustrates a Weyl chamber. To simplify the figures, we denote each root by a gray point instead of an arrow.



**Definition 12.3.** The *Weyl group* of  $\Delta$  is a subgroup of  $GL(E)$  generated by all  $s_\alpha$ ,  $\alpha \in \Delta$ .

By definition  $\forall w \in W$ ,  $w(\Delta) = \Delta$ , because it holds for every generator. Then we have  $W \hookrightarrow S_\Delta$ . Hence,  $|W| < \infty$ .

**Lemma 12.4.**  $\forall w \in W$ ,  $(w(\beta)|w(\gamma)) = (\beta|\gamma)$  for any  $\beta, \gamma \in E$ .

*Proof.* It suffices to show this for its generator  $s_\alpha$ , for all  $\alpha \in \Delta$ . By direct calculation, we

have

$$(s_\alpha(\beta)|s_\alpha(\gamma)) = (\beta|\gamma) - \frac{2(\beta|\alpha)}{(\alpha|\alpha)}(\alpha|\gamma) - \frac{2(\gamma|\alpha)}{(\alpha|\alpha)}(\alpha|\beta) + \frac{2(\beta|\alpha)}{(\alpha|\alpha)} \frac{2(\gamma|\alpha)}{(\alpha|\alpha)}(\alpha|\alpha) = (\beta|\gamma).$$

□

**Lemma 12.5.** *Let  $B$  be a base. Then  $s_\alpha$ ,  $\alpha \in B$ , permutes all  $\Delta_+ \setminus \{\alpha\}$ .*

*Proof.* Let  $\gamma \in \Delta_+ \setminus \{\alpha\}$ . Then  $\gamma = k_1\beta_1 + k_2\beta_2 + \dots + k_t\beta_t$ , for some  $\beta_i \in B$  and  $k_i > 0$ . By definition,  $s_\alpha(\gamma) = k_1\beta_1 + \dots + k_t\beta_t + \square\alpha$ , for some  $\square \in \mathbb{Z}$ . Since the coefs of  $\beta_i$ 's, other than  $\alpha$ , are unchanged,  $s_\alpha(\gamma) \in \Delta_+$ . □

We denote by  $\rho = \frac{1}{2} \sum_{\alpha > 0} \alpha$ , which turns out to be a very important element in  $E$ .

**Corollary 12.6.** *Let  $\beta \in B$ .  $s_\beta(\rho) = \rho - \beta$ .*

**Definition 12.7.** *Let  $B$  be a base of  $\Delta$ . We call*

$$C(B) := \{\gamma \in E : (\gamma|\beta) > 0, \forall \beta \in B\}$$

*the **fundamental Weyl chamber** relative to  $B$ .*

Since  $B$  is a basis of  $E$ ,  $C(B)$  is indeed a nonempty set, i.e.,  $C(B)$  is a Weyl chamber. Moreover, any regular element  $\gamma$  is contained in precisely one Weyl chamber. Combining with the summary in §11.3, we obtain one-to-one correspondences:

$$\{\text{Bases}\} \xleftrightarrow{1-1} \{\text{Regular decompositions}\} \xleftrightarrow{1-1} \{\text{Weyl chambers}\}.$$

## §12.2 Simply transitive actions of Weyl groups

Let  $B$  be a base. Denote by  $W_B = \langle s_\beta : \beta \in B \rangle$ .

**Proposition 12.8.** *Given two bases  $B$  and  $B'$  of  $\Delta$ ,  $\exists w \in W_B$  s.t.  $w(B') = B$ .<sup>a</sup>*

<sup>a</sup>In Proposition 12.12 we will prove this  $w$  is unique.

*Proof.* Choose a regular  $\gamma'$  s.t.  $(B'|\gamma') > 0$ . We are looking for some  $w \in W_B$  s.t.  $(B|w(\gamma')) > 0$ . Then by Lemma 12.4 we have  $(w(B')|w(\gamma')) = (B'|\gamma') > 0$ , which implies  $B = w(B')$  by the 1-1 correspondence between bases and Weyl chambers.

Consider  $w \in W_B$  s.t.  $(\rho|w(\gamma'))$  is maximal. Then  $\forall \beta \in B$ ,  $(\rho|s_\beta w(\gamma')) \leq (\rho|w(\gamma'))$ . On the other hand, by Corollary 12.6,

$$(\rho|s_\beta w(\gamma')) = (s_\beta(\rho)|w(\gamma')) = (\rho - \beta|w(\gamma')).$$

Thus  $(\beta|w(\gamma')) \geq 0$ . Since  $\gamma'$  is regular,  $(B|w(\gamma')) > 0$  as desired. □

**Proposition 12.9.**  $\forall$  root can be included in some base.

*Proof.* Our goal is to find a regular  $\gamma$  for any  $\alpha \in \Delta$  s.t. the base  $B$  associated to  $\gamma$  contains  $\alpha$ . Thanks to Remark 11.8(1), it reduces to finding a regular  $\gamma$  s.t.  $(\alpha|\gamma)$  is minimal among  $(\Delta_+|\gamma)$ .

$\forall \beta \neq \pm\alpha, \beta^\perp \neq \alpha^\perp$ . Then  $\alpha^\perp \notin \bigcup_{\beta \neq \pm\alpha} \beta^\perp$ .  $\exists 0 \neq \gamma' \in \alpha^\perp$  such that  $(\gamma'|\alpha) = 0$  and  $(\gamma'|\beta) \neq 0$  for any  $\beta \neq \pm\alpha$ . Choose  $\gamma$  very close to  $\gamma'$  and  $(\gamma|\alpha) > 0$ , where “very close” means  $0 < (\gamma|\alpha) < |(\gamma|\beta)|$  for all  $\beta \neq \pm\alpha$ . This  $\gamma$  is as desired.  $\square$

Combining Proposition 12.8 and Proposition 12.9, we have the following conclusion.

**Corollary 12.10.** Given a base  $B, \forall \alpha \in \Delta, \exists w \in W_B$  s.t.  $w(\alpha) \in B$ . Namely,  $\Delta = W_B B$ .

Now we are ready to prove  $W = W_B$ .

**Proposition 12.11.**  $W = W_B$ .

*Proof.* It is straightforward to show that  $s_w \alpha = w s_\alpha w^{-1}$  for all  $\alpha \in \Delta$  and  $w \in W$ . Thanks to Corollary 12.10,  $\forall \alpha \in \Delta, \exists w \in W_B$  s.t.  $\alpha = w\beta$  for some  $\beta \in B$ . Thus  $s_\alpha = w s_\beta w^{-1} \in W_B$ .  $\square$

It is natural to consider the uniqueness of  $w$  in Proposition 12.8. Namely, does the Weyl group act simply transitively on bases? The answer is yes.

**Proposition 12.12.**  $\forall w \in W_B, w(B) = B$ , then  $w = \text{id}$ .

*Proof.* Suppose  $w \neq \text{id}$  and denote  $w = s_{\beta_{i_1}} s_{\beta_{i_2}} \dots s_{\beta_{i_m}}$ , where  $\beta_{i_j} \in B$ . WLOG, we assume this expression is reduced, i.e.,  $m$  is minimal. We claim that  $w(\beta_m) < 0$ . This contradiction forces  $w = \text{id}$ .

Suppose  $w(\beta_m) > 0$ , then  $s_{\beta_{i_1}} \dots s_{\beta_{i_{m-1}}} \beta_m < 0$ . Consider the sequence of roots:

$$0 < \beta_m, \quad s_{\beta_{i_{m-1}}} \beta_m, \quad \dots, \quad s_{\beta_{i_1}} \dots s_{\beta_{i_{m-1}}} \beta_m < 0.$$

Then there exists  $r$  s.t.  $s_{\beta_{i_r}} \dots s_{\beta_{i_{m-1}}} \beta_m > 0$  and  $s_{\beta_{i_{r+1}}} \dots s_{\beta_{i_{m-1}}} \beta_m < 0$ . Let  $u = s_{\beta_{i_{r+1}}} \dots s_{\beta_{i_{m-1}}}$ . Then  $u(\beta_m) > 0$  and  $s_{\beta_{i_r}} u(\beta_m) < 0$ . By Lemma 12.5,  $u(\beta_m) = \beta_{i_r}$ . Then  $s_{\beta_{i_r}} = u s_{\beta_m} u^{-1}$  and

$$s_{\beta_{i_r}} s_{\beta_{i_{r+1}}} \dots s_{\beta_{i_m}} = u s_{\beta_m} s_{\beta_m} = s_{\beta_{i_{r+1}}} \dots s_{\beta_{i_{m-1}}},$$

contrary to the fact that  $w$  is reduced.  $\square$

**Summary.** Corollary 12.10 and Proposition 12.11 show that any base of a root system together with its inner product encodes the whole structure. Based on this, we are able to characterize a root system by a matrix, that is, so-called a Cartan matrix.

**Definition 12.13.** Given a root system  $\Delta$  with a base  $B = \{\alpha_1, \dots, \alpha_n\}$ , we call the matrix

$$C = \left( \frac{2(\alpha_i | \alpha_j)}{(\alpha_i | \alpha_i)} \right)$$

the **Cartan matrix** of  $\Delta$ .<sup>a</sup>

<sup>a</sup>The definition depends on the ordering of simple roots, but this is not essential. We say the Cartan matrices are equivalent if they are equal up to a reordering.

**Remark 12.14.** Since  $( | )$  is positive definite, the Cartan matrix is nondegenerate.

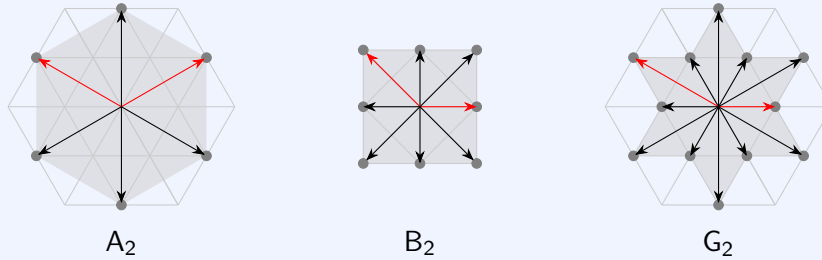
**Definition 12.15.** We say two root systems  $\Delta \subseteq E$  and  $\Delta' \subseteq E'$  are isomorphic if there exists an isomorphism  $\psi : E \rightarrow E'$  mapping  $\Delta \rightarrow \Delta'$  and satisfying  $\langle \psi(\alpha), \psi(\beta) \rangle = \langle \alpha, \beta \rangle$ .

By definition, we can easily see that two root systems are isomorphic if and only if they have equivalent Cartan matrices.

### §12.3 Classification of root systems of rank 2

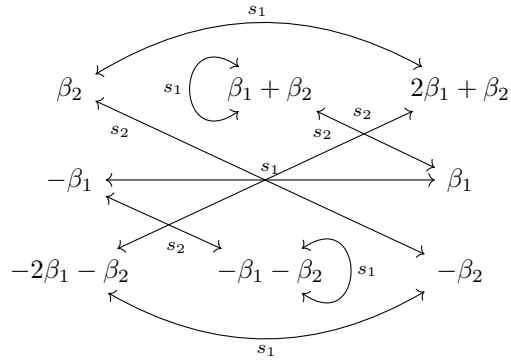
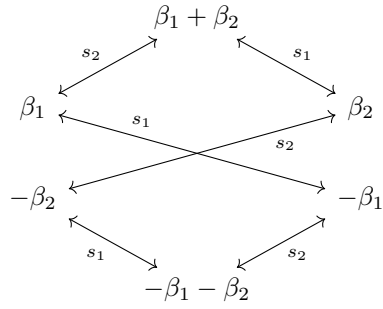
We are now in the position to classify the irreducible root system of rank 2. (Rank one is plain.)

**Theorem 12.16.** All irreducible root systems of rank 2 are as follows (up to isomorphism). Red arrows indicate the simple roots.

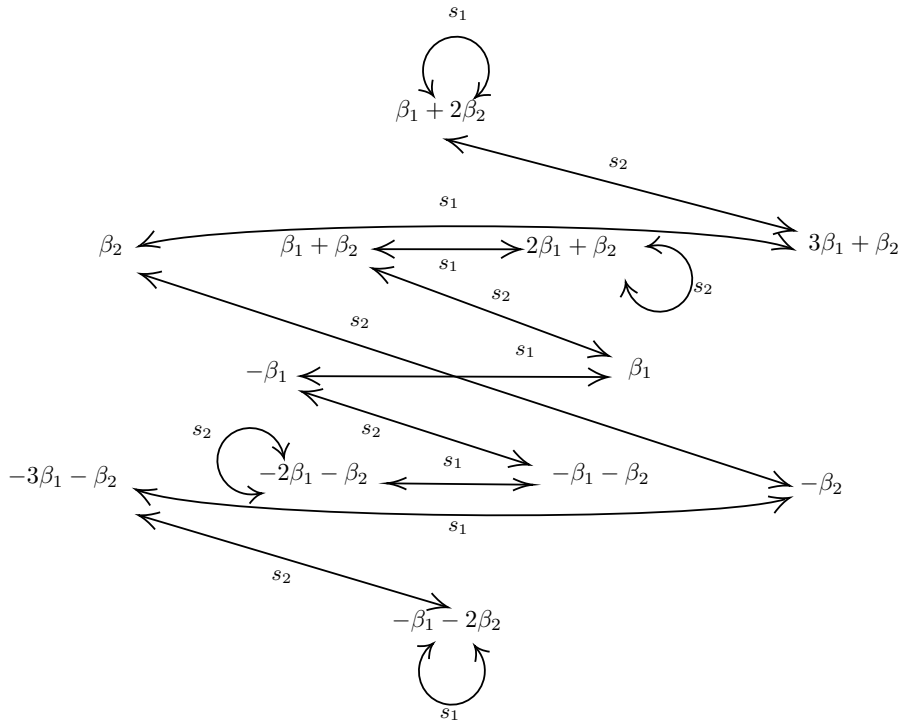


*Proof.* For  $\Delta$  of rank two, every base  $B$  has two vectors  $\beta_1, \beta_2$ . WLOG, we assume  $\|\beta_1\| \leq \|\beta_2\|$ . Thanks to Proposition 11.6 and the discussion in §11.2, the angle  $\theta = \widehat{(\beta_1, \beta_2)} = \frac{2\pi}{3}, \frac{3\pi}{4}, \frac{5\pi}{6}$  and  $\frac{\|\beta_2\|}{\|\beta_1\|} = 1, \sqrt{2}, \sqrt{3}$ , respectively. To simplify the notation, we denote by  $s_i = s_{\beta_i}$ .

1.  $\theta = \frac{2\pi}{3}$ . By the following action table,  $\Delta = \{\pm\beta_1, \pm\beta_2, \pm(\beta_1 + \beta_2)\}$  is of type  $A_2$ .
2.  $\theta = \frac{3\pi}{4}$ . By the following action table,  $\Delta = \{\pm\beta_1, \pm\beta_2, \pm(\beta_1 + \beta_2), \pm(2\beta_1 + \beta_2)\}$  is of type  $B_2$ .



3.  $\theta = \frac{5\pi}{6}$ . By the following action table,  $\Delta = \{\pm\beta_1, \pm\beta_2, \pm(\beta_1 + \beta_2), \pm(2\beta_1 + \beta_2), \pm(3\beta_1 + \beta_2), \pm(\beta_1 + 2\beta_2)\}$  is of type  $G_2$ .



□

# Chapter 13

## Classification of Root Systems

### §13.1 Normal systems and Coxeter graphs

**Definition 13.1.** In a Euclidean space  $E$ , a **normal system** consists of roots  $v_1, \dots, v_k$  satisfying

1.  $v_1, \dots, v_k$  are linearly independent.
2.  $\|v_i\| = 1$ .
3.  $(v_i|v_j) = 0, -\frac{1}{2}, -\frac{\sqrt{2}}{2}, -\frac{\sqrt{3}}{2}$ .

**Example 13.2.** Let  $B = \{\beta_1, \dots, \beta_n\}$  be a base of a root system  $\Delta$ . Then  $\left\{ \frac{\beta_1}{\|\beta_1\|}, \frac{\beta_2}{\|\beta_2\|}, \dots, \frac{\beta_n}{\|\beta_n\|} \right\}$  is a normal system.

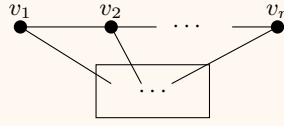
**Definition 13.3.** Given a normal system  $\{v_1, \dots, v_n\}$ , the associated **Coxeter graph** has the normal system as its vertex set, with  $4(v_i|v_j)$  edges joining  $v_i$  and  $v_j$ .

**Remark 13.4.** Let  $\Gamma$  be the Coxeter graph  $\Gamma$  associated to the normal system in Example 13.2. We call  $\Gamma$  the Coxeter graph of  $B$ . Since the Weyl group  $W$  preserves the inner product and acts transitively on bases (thanks to Proposition 12.8), Coxeter graphs of two bases of  $\Delta$  are isomorphic. Therefore, we also call  $\Gamma$  the Coxeter graph of  $\Delta$ .

**Proposition 13.5.** A normal system has the following properties:

1. A subset of a normal system is a normal system.
2. Let  $v_1, \dots, v_k$  be a normal system. Then  $\#\{\text{connected pairs}\} < k$
3. The associated Coxeter graph has no cycles.
4. For any  $v_i$ ,  $\#\{\text{edges adjacent to } v_i\} < 4$ .
5. (Contractibility) If  $v_1, \dots, v_r$  satisfies  $(v_i|v_j) = -\frac{1}{2}\delta_{i+1,j}$ , i.e., the associated

Coxeter graph is as follows, then replacing  $v_1, \dots, v_r$  by  $v_1 + v_2 + \dots + v_r$  we obtain a normal system.

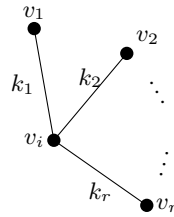


*Proof.* The first property is obvious.

2. Let  $v = v_1 + v_2 + \dots + v_k$ . Then  $0 < (v|v) = \sum_{i=1}^k (v_i|v_i) + 2 \sum_{i < j} (v_i|v_j) \leq k - \#\{\text{connected pairs}\}$ .

3. It follows directly from 2.

4. Let  $v_j$ ,  $1 \leq j \leq r$  be all roots adjacent to  $v_i$  with  $k_j$  edges, respectively. Then  $v_j$ 's are pairwise orthogonal, otherwise we have a cycle. By orthonormal decomposition,  $v_i = (v_1|v_i)v_1 + (v_2|v_i)v_2 + \dots + (v_r|v_i)v_r + v'$ , where  $(v'|v_j) = 0$  and  $v' \neq 0$ . Taking the inner product with  $v_i$ , we have  $(v_1|v_i)^2 + \dots + (v_r|v_i)^2 = 1 - (v'|v') < 1$ . Our claim follows from the fact that  $(v_i|v_j)^2 = \frac{k_j}{4}$ .



5. Since any  $v_j$ , where  $j \notin \{1, \dots, r\}$ , is connected to at most one vector in  $\{v_1, \dots, v_r\}$  (otherwise a cycle would arise), we only need to consider the length of  $v_1 + \dots + v_r$ . Moreover, we have  $(v_1 + \dots + v_r|v_1 + \dots + v_r) = k + 2 \sum_{i=1}^{r-1} (v_i|v_{i+1}) = 1$ .  $\square$

## §13.2 Classification of associated Coxeter graphs

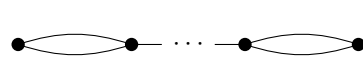
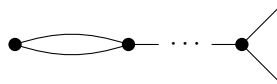
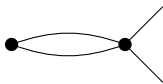
Throughout this section, we denote by  $\Gamma$  an irreducible Coxeter graph associated to normal systems. Thanks to Proposition 13.5,  $\Gamma$  can not contain a  $k$ -edge with  $k \geq 4$ .

$\Gamma$  contains a triple edge. It is easy to show that  $\Gamma$  could only contain two vertices.

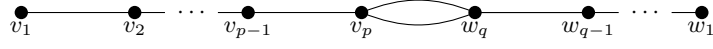
**Proposition 13.6.** *If  $\Gamma$  contains a triple edge,  $\Gamma$  is isomorphic to the following graph.*



$\Gamma$  contains a double edge. By contractibility,  $\Gamma$  can not contain a 3-branch points or another 2-edge, i.e., the following graphs are NOT allowed.



Hence,  $\Gamma$  can only be a double edge with simple chains attached to both of its vertices. Let  $\{v_1, \dots, v_p\}$  and  $\{w_1, \dots, w_q\}$  be the two simple chains (where  $p, q \geq 1$ ), with  $v_p$  connected to  $w_q$  by a double edge.



Let  $v = v_1 + 2v_2 + \dots + pv_p$  and  $w = w_1 + 2w_2 + \dots + qw_q$ . Then we have

$$(v|v) = 1^2 + 2^2 + \dots + p^2 - (1 \cdot 2 + 2 \cdot 3 + \dots + (p-1)p) = p^2 - \frac{(p-1)p}{2} = \frac{p(p+1)}{2},$$

$$(w|w) = \frac{(q-1)q}{2}, \quad (v|w) = pq(v_p|w_q) = -\frac{\sqrt{2}}{2}pq.$$

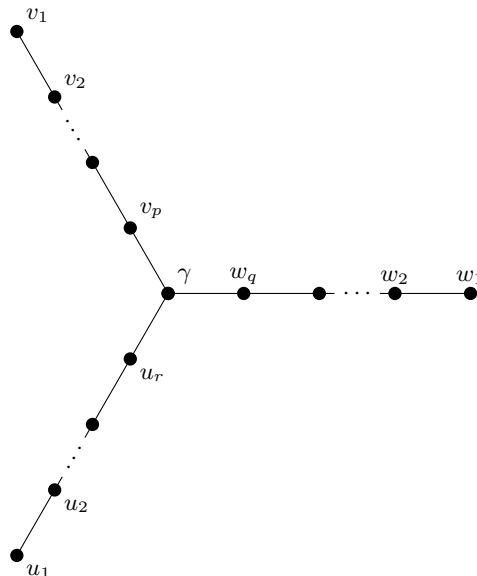
By Cauchy-Schwarz inequality,  $(v|w)^2 < (v|v)(w|w)$ , and together with the above equations, gives  $(p-1)(q-1) < 2$ . Hence,  $\Gamma$  can occur only in two cases:  $p = 1$  or  $q = 1$ ;  $p = q = 2$ .

**Proposition 13.7.** *If  $\Gamma$  contains a 2-edge,  $\Gamma$  is isomorphic to one of the following graphs.*



$\Gamma$  contains only 1-edges. We consider the branch points. If all points are 1- or 2-branch points, then  $\Gamma$  is a simple chain. By Proposition 13.5(4),  $\Gamma$  can not contain a  $k$ -branch point with  $k \geq 4$ . Combined with Proposition 13.5(5),  $\Gamma$  can not contain more than one 3-branch point (otherwise, after contraction we would obtain a graph with a 4-branch point.) Therefore, it remains to consider the graph with exactly one 3-branch point.

Let  $\gamma$  be the 3-branch point. Let  $\{v_1, \dots, v_p\}$ ,  $\{w_1, \dots, w_q\}$  and  $\{u_1, \dots, u_r\}$  be the three simple chains (where  $p, q, r \geq 1$ ), with  $v_p$ ,  $w_q$  and  $u_r$  each connected to  $\gamma$ .



Let  $v = \sum_{i=1}^p iv_i$ ,  $w = \sum_{i=1}^q iw_i$  and  $u = \sum_{i=1}^r iu_i$ . Then  $(\gamma|v) = p(\gamma|v_p) = -\frac{1}{2}p$  and  $(v|v) = \frac{p(p+1)}{2}$ ; analogous formulas hold for  $w$  and  $u$ . Since  $v$ ,  $w$  and  $u$  are pairwise orthogonal, we have

$$\left(\gamma \left| \frac{v}{\|v\|} \right.\right)^2 + \left(\gamma \left| \frac{w}{\|w\|} \right.\right)^2 + \left(\gamma \left| \frac{u}{\|u\|} \right.\right)^2 = \frac{p}{2(p+1)} + \frac{q}{2(q+1)} + \frac{r}{2(r+1)} < 1.$$

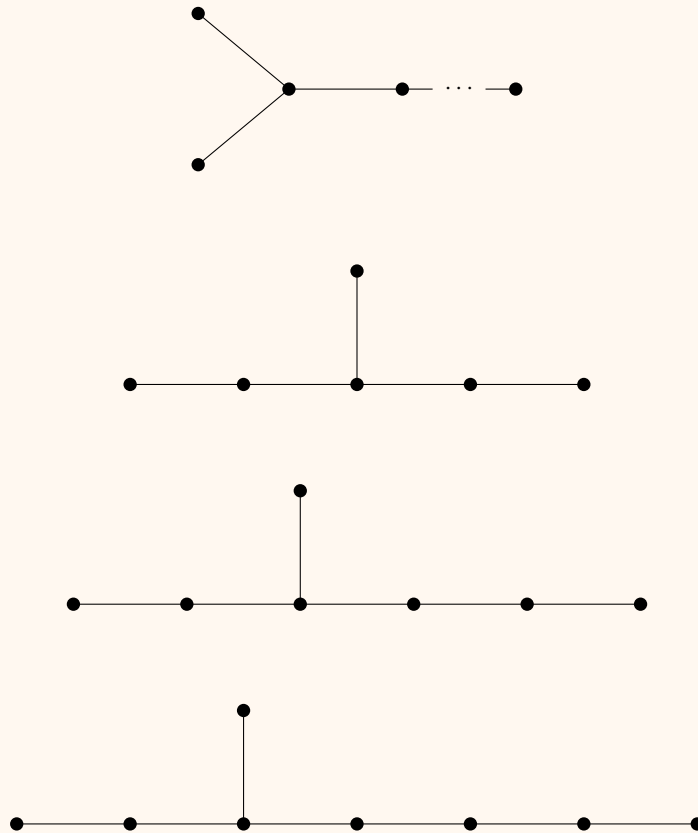
Simplifying the above expression, we obtain

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

It is easy to see that at least one of  $p, q, r$  is 1, say  $p = 1$ . Then we have the inequality  $\frac{1}{q+1} + \frac{1}{r+1} > \frac{1}{2}$ . WLOG, we assume  $q \leq r$ .

1. If  $q = 1$ , then  $r$  is arbitrary;
2. If  $q = 2$ , then  $r < 5$ , i.e.,  $r = 2, 3, 4$ ;
3. If  $q \geq 3$ , then  $r < 3 \leq q$ , which contradicts our assumption that  $q \leq r$ .

**Proposition 13.8.** *If  $\Gamma$  contains only 1-edges,  $\Gamma$  is isomorphic to one of the following graphs.*

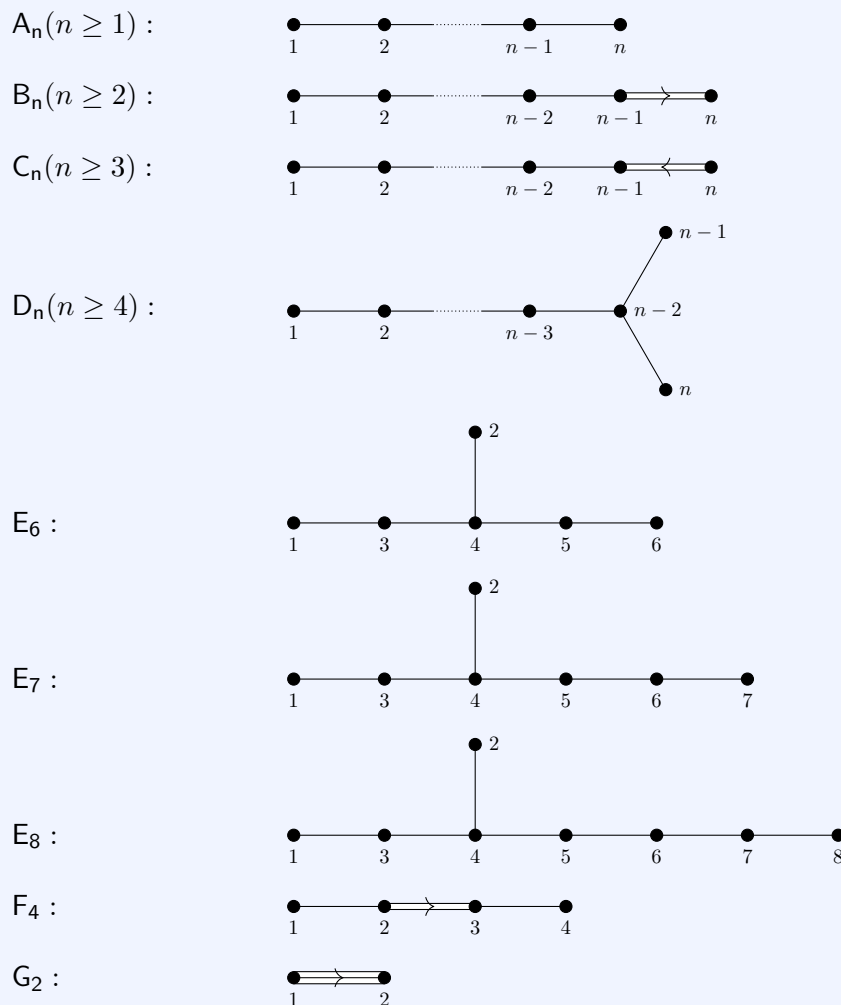


**Remark 13.9.** *So far we have not proved each graph in Proposition 13.6, 13.7 and 13.8 does belong to a root system. It will be proved in the next section.*

## §13.3 Dynkin diagrams

**Definition 13.10.** Let  $\Gamma$  be the Coxeter graph of  $\Delta$ . Whenever a double or triple edge occurs, we add an arrow pointing to the shorter of the two roots; we call the resulting figure the **Dynkin diagram** of  $\Delta$ .

**Theorem 13.11.** If  $B$  is a base of an irreducible root system, its Dynkin diagram must be one of the following:



Moreover, each graph above is a Dynkin diagram of some root system  $\Delta$ .

*Proof.* The first claim follows directly from Proposition 13.6, 13.7 and 13.8. For the last assertion, we need to construct a suitable root system for each type.

**Idea:** Construct  $\Delta = \{\text{elements of certain lengths in a lattice } \Lambda \text{ of } E\}$ .

Let  $\omega_1, \dots, \omega_l$  be an orthonormal basis of  $E$ , where  $l = \dim E$  is taken as required.

$A_n$  : Let  $\Lambda = \{\sum_{i=1}^{n+1} k_i \omega_i : \sum k_i = 0, k_i \in \mathbb{Z}\}$ . Let  $\Delta = \{\alpha \in \Lambda : (\alpha|\alpha) = 2\}$ . Then  $\Delta = \{\omega_i - \omega_j : 1 \leq i \neq j \leq n+1\}$ . Note that  $B = \{\alpha_i := \omega_i - \omega_{i+1} : 1 \leq i \leq n\}$  is a base.

Moreover, all base elements have the same length, and for  $i \neq j$ ,

$$\left( \frac{\alpha_i}{\|\alpha_i\|} \middle| \frac{\alpha_j}{\|\alpha_j\|} \right) = -\delta_{\{|i-j| \leq 1\}} \frac{1}{2},$$

as desired.

$B_n$  : Let  $\Lambda = \sum_{i=1}^n \mathbb{Z}\omega_i$  and  $\Delta = \{\alpha \in \Lambda : (\alpha|\alpha) = 1, 2\} = \{\pm\omega_i, \pm\omega_i \pm \omega_j : i \neq j\}$ . Note that  $B = \{\alpha_i := \omega_i - \omega_{i+1}, \alpha_n := \omega_n : 1 \leq i \leq n-1\}$  is a base. Moreover, for  $1 \leq i \leq n-1$ ,  $\|\alpha_i\| = \sqrt{2} > \|\alpha_n\| = 1$  and

$$\begin{aligned} \left( \frac{\alpha_i}{\|\alpha_i\|} \middle| \frac{\alpha_j}{\|\alpha_j\|} \right) &= -\delta_{\{|i-j| \leq 1\}} \frac{1}{2}, \\ \left( \frac{\alpha_i}{\|\alpha_i\|} \middle| \frac{\alpha_n}{\|\alpha_n\|} \right) &= -\delta_{i,n-1} \frac{\sqrt{2}}{2}, \end{aligned}$$

$C_n$  : Let  $\Lambda = \sum_{i=1}^n \mathbb{Z}\omega_i$  and  $\Delta = \{\alpha \in \Lambda : (\alpha|\alpha) = 2, 4\} = \{\pm\omega_i \pm \omega_j, \pm 2\omega_i : i \neq j\}$ . Note that  $B = \{\alpha_i := \omega_i - \omega_{i+1}, \alpha_n := 2\omega_n : 1 \leq i \leq n-1\}$  is a base. Moreover, for  $1 \leq i \leq n-1$ ,  $\|\alpha_i\| = \sqrt{2} < \|\alpha_n\| = 2$  and

$$\begin{aligned} \left( \frac{\alpha_i}{\|\alpha_i\|} \middle| \frac{\alpha_j}{\|\alpha_j\|} \right) &= -\delta_{\{|i-j| \leq 1\}} \frac{1}{2}, \\ \left( \frac{\alpha_i}{\|\alpha_i\|} \middle| \frac{\alpha_n}{\|\alpha_n\|} \right) &= -\delta_{i,n-1} \frac{\sqrt{2}}{2}, \end{aligned}$$

$D_n$  : Let  $\Lambda = \sum_{i=1}^n \mathbb{Z}\omega_i$ . Let  $\Delta = \{\pm\omega_i \pm \omega_j : 1 \leq i \neq j \leq n\} = \{\alpha \in \Lambda : (\alpha|\alpha) = 2\}$ . Note that  $B = \{\alpha_i := \omega_i - \omega_{i+1} : 1 \leq i \leq n-1\} \cup \{\alpha_n := \omega_{n-1} + \omega_n\}$  is a base. Moreover, all base elements have the same length, and for  $1 \leq i \neq j \leq n-1$ ,

$$\begin{aligned} \left( \frac{\alpha_i}{\|\alpha_i\|} \middle| \frac{\alpha_j}{\|\alpha_j\|} \right) &= -\delta_{\{|i-j| \leq 1\}} \frac{1}{2}, \\ \left( \frac{\alpha_i}{\|\alpha_i\|} \middle| \frac{\alpha_n}{\|\alpha_n\|} \right) &= -\delta_{i,n-2} \frac{1}{2}, \end{aligned}$$

as desired.

$F_4$  : Let  $\Lambda = \sum_{i=1}^4 \mathbb{Z}\omega_i + \mathbb{Z}(\sum_{i=1}^4 \omega_i)/2$  and

$$\Delta = \{\alpha \in \Lambda : (\alpha|\alpha) = 1, 2\} = \{\pm\omega_i, \pm\omega_i \pm \omega_j, \frac{1}{2}(\pm\omega_1 \pm \omega_2 \pm \omega_3 \pm \omega_4) : i \neq j\}.$$

Note that  $B = \{\omega_1 - \omega_2, \omega_2 - \omega_3, \omega_3, -(\omega_1 + \omega_2 + \omega_3 + \omega_4)\}$  is a base. In this case,

$$\Delta_+ = \{\omega_1, \omega_2, \omega_3, -\omega_4, \omega_i - \omega_j, \frac{1}{2}(\pm\omega_1 \pm \omega_2 \pm \omega_3 - \pm\omega_4) : i < j\}.$$

One can check  $B$  satisfies the Dynkin diagram of type  $F_4$ .

$G_2$  : Let  $F$  be the hyperplane orthogonal to the vector  $\omega_1 + \omega_2 + \omega_3$  in  $E = \mathbb{R}^3$ . Let  $\Lambda = \sum_{i=1}^3 \mathbb{Z}\omega_i \cap F$  and  $\Delta = \{\alpha \in \Lambda : (\alpha|\alpha) = 2, 6\}$ . So  $\Delta = \pm\{\omega_1 - \omega_2, \omega_2 - \omega_3, \omega_1 - \omega_3, 2\omega_1 - \omega_2 - \omega_3, 2\omega_2 - \omega_1 - \omega_3, 2\omega_3 - \omega_1 - \omega_2\}$ . We take  $B = \{\omega_1 - \omega_2, -2\omega_1 + \omega_2 + \omega_3\}$ . One can check that it satisfies the Dynkin diagram of type  $G_2$ .

$E_6, E_7, E_8$  can be found in Humphreys' book (12.1). □

**Remark 13.12.**  $\text{Der } \mathbb{O}_{\mathbb{F}}$  is of type  $G_2$ , where  $\mathbb{O}_{\mathbb{F}}$  is the octonion algebra over  $\mathbb{F}$ .



## Chapter 15

# Generators and Relations of Associative Algebras

In this chapter,  $\mathbb{F}$  is just a field (without any additional conditions).

### §15.1 Free associative algebras

Let  $X = \{x_i\}_{i \in I}$  be a set, called an **alphabet**, whose elements are called **letters**. A **word** over  $X$  is either a finite sequence of letters from  $X$  or the empty word 1, which is regarded as having length 0.

Denote by  $X^*$  the set of all words over  $X$ . Then  $X^*$  is equipped with a natural multiplication by adjunction

**Definition 15.1.** We call  $\mathbb{F}\langle X \rangle = \{\sum_i \alpha_i w_i : \alpha_i \in \mathbb{F} \text{ and } \alpha_i \text{ are almost all zero}\}$  a **free associative algebra on  $X$  over  $\mathbb{F}$** , where the operations extend linearly from the semigroup structure.

Clearly,  $X^*$  is a basis of  $\mathbb{F}\langle X \rangle$ .

For an  $\mathbb{F}$ -algebra  $A$ , suppose  $\{a_i\}_{i \in I}$  is a generating set, i.e. every element of  $A$  can be represented (not necessarily uniquely) as a finite linear combination of products of the  $a_i$ 's.

Choose a set  $X = \{x_i\}_{i \in I}$  to be an alphabet indexed by the same set  $I$ , so that we have a map

$$\varphi : X \rightarrow A, \quad x_i \mapsto a_i, \quad i \in I.$$

Then we obtain an algebra homomorphism  $\bar{\varphi} : \mathbb{F}\langle X \rangle \rightarrow A$ . Let  $J = \ker \bar{\varphi}$ . Then  $J$  is an ideal of  $\mathbb{F}\langle X \rangle$ . Since  $\text{Im } \bar{\varphi} = A$ , we have  $A \cong \mathbb{F}\langle X \rangle / J$ .

Let  $R \subseteq J$  be a subset generating  $J$  as an ideal, i.e.,  $J$  is the smallest ideal of  $\mathbb{F}\langle X \rangle$  containing  $R$ . Explicitly,

$$J = \left\{ \sum a_i r_i b_i \mid a_i, b_i \in \mathbb{F}\langle X \rangle, r_i \in R \right\}.$$

We say that  $\mathbb{F}\langle X \rangle / J$  is the algebra presented by generators  $X$  and relations  $R$ , and denote it by  $\mathbb{F}\langle X | R = 0 \rangle$  or  $\langle X | R = 0 \rangle$  when the field is clear. Since  $A \cong \mathbb{F}\langle X \rangle / J$ , we call

$\mathbb{F}\langle X | R = 0 \rangle$  a presentation of  $A$ . We say a presentation is finite if  $\exists$  finite sets  $X$  and  $R$  s.t.  $A = \langle X | R = 0 \rangle$

## §15.2 Gröbner–Shirshov Theorem

Even given a finite presentation  $\mathbb{F}\langle X | R = 0 \rangle$  of an associative  $\mathbb{F}$ -algebra, it is difficult to determine whether two elements  $a, b \in \mathbb{F}\langle X \rangle$  are equal modulo the ideal  $J = \text{id}_{\mathbb{F}\langle X \rangle}(R)$ . Indeed, there exist finite presentations that defy any algorithm.

We now describe a class of finite presentations for which such an algorithm does exist. Although this method does not always work, it is often effective.

In this section, we assume  $X$  is equipped with a total order and satisfies **the descending chain condition**, i.e.,  $\nexists$  infinite strictly descending chain  $x_1 > x_2 > \dots$

**Definition 15.2.** For any  $f \in K\langle X \rangle$ ,  $f = a_1 w_1 + \dots + a_k w_k$ , where  $w_i \in X^*$ ,  $a_i \in \mathbb{F} \setminus \{0\}$ . Let  $w_j$  be the maximal word with respect to the length-lexicographic order.<sup>a</sup> Then call  $w_j$  the **leading monomial** of  $f$ , denoted by  $\bar{f}$ . Denote by  $f = \bar{f} + \{f\}$ .

<sup>a</sup>The length-lexi order means  $w_i < w_j$  if the length of  $w_i$  is less than that of  $w_j$ , or they have the same length and the first distinct letter of  $w_i$  is less than the one of  $w_j$ .

**Remark 15.3.** For  $A = \langle X | R \rangle$ , if  $f \in R$ , then

$$\bar{f} = w_j = \sum_{i \neq j} -\frac{a_i}{a_j} w_i.$$

Thus  $\bar{f}$  can be written as a linear combination of smaller words in  $A$ .

**Definition 15.4.** A word  $w \in X^*$  is **reducible** if it contains some  $\bar{f}$ ,  $f \in R$ , as a subword, i.e.

$$w = w' \bar{f} w'', \quad w', w'' \in X^*.$$

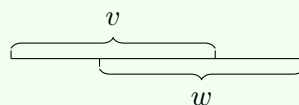
Otherwise,  $w$  is called **irreducible**.

Denote the set of all irreducible words w.r.t.  $R$  by  $\text{Irr}(R)$ .

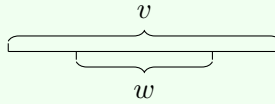
**Proposition 15.5.**  $\text{Irr}(R)$  spans  $A$ .

**Definition 15.6.** Given words  $v$  and  $w \in X^*$ , we say  $v, w$  **admit a composition** if

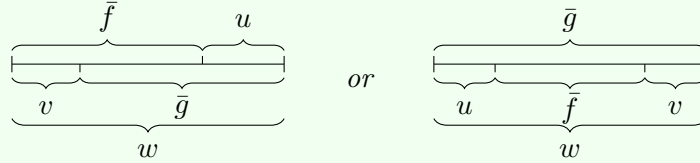
1. the end of one word is the beginning of the other;



2. one of these words is a subword of the other.



**Definition 15.7.** Let  $f, g \in \mathbb{F}\langle X \rangle$ . The coefficients of  $\bar{f}, \bar{g}$  are equal to 1. Suppose that  $\bar{f}, \bar{g}$  admit a composition, i.e.



The element  $(f, g)_w = fu - vg$  (or  $ufv - g$ ) is called the **composition** of  $f$  and  $g$  w.r.t. the word  $w$ .

**Theorem 15.8** (Gröbner–Shirshov Theorem). Let  $A = \mathbb{F}\langle X | R = 0 \rangle$ .  $\text{Irr}(R)$  is a basis of  $A$  if and only if for any two relations  $f, g \in R$  that admit a composition, all their compositions reduce to 0.

We say  $R$  is **closed under composition** if  $R$  satisfies the necessary condition in the theorem.

**Example 15.9.** Let us consider the algebra  $A = \langle x, y, z | [x, y] = z, [z, x] = 2x, [z, y] = -2y \rangle$ . Define the order by  $x < y < z$ . The relation set  $R$  contains

$$\begin{aligned} &yx - xy + z, \\ &zx - xz - 2x, \\ &zy - yz + 2y. \end{aligned}$$

The only relations admitting a composition are  $yx - xy + z$  and  $zy - yz + 2y$  w.r.t. the word  $zyx$ . The composition can be reduced to 0:

$$\begin{aligned} (zy - yz + 2y)x - z(yx - xy + z) &= -yzy + 2yx + zxy - z^2 \\ &\rightarrow -y(xz + 2x) + 2(xy - z) + (xz + 2x)y - z^2 \\ &= -yzy - 2yx + 2xy - 2z + xzy + 2xy - z^2 \\ &\rightarrow 0 \end{aligned}$$

**Example 15.10.** Let us consider the algebra

$$A = \langle x, y, z | xy + yx = z^2, yz + zy = x^2, zx + xz = y^2 \rangle.$$

Define the order by  $x < y < z$ . The relation set  $R$  contains

$$\begin{aligned} z^2 - xy - yx, \\ zy + yz - x^2, \\ zx + xz - y^2. \end{aligned}$$

The relations  $z^2 - xy - yx$  and  $zy + yz - x^2$  admit a composition w.r.t. the word  $z^2y$ . The composition is

$$\begin{aligned} (z^2 - xy - yx)y - z(zy + yz - x^2) \\ &= -xy^2 - yxy - \color{blue}{zyz} + \color{blue}{zxx} \\ &\rightarrow -xy^2 - yxy - (x^2 - yz)z + (y^2 - xz)x \\ &= -xy^2 - yxy - x^2z + yz^2 + y^2x - xzx \\ &\rightarrow -xy^2 - yxy - x^2z + y(xy + yx) + y^2x - x(y^2 - xz) \\ &= -xy^2 - yxy - x^2z + yxy + y^2x + y^2x - xy^2 + x^2z \\ &= 2y^2x - 2xy^2. \end{aligned}$$

This means  $\text{Irr}(R)$  is NOT a basis of this algebra if  $\text{char } \mathbb{F} \neq 2$ .

**Remark 15.11.** Suppose  $A = \mathbb{F}\langle X | R = 0 \rangle$  is an associative algebra with  $R$  not closed under composition, e.g., Example 15.10. Let  $R = R_0$  and  $R_1$  be the union of  $R_0$  and all reduced compositions of elements of  $R_0$ . Note that by definition, the compositions of elements in  $R$  are still contained in  $\text{id}(R)$ . Thus adding these elements into relation set  $R$  keeps the algebra unchanged. Namely,  $\mathbb{F}\langle X | R_0 = 0 \rangle = \mathbb{F}\langle X | R_1 = 0 \rangle$ .

Iterating this process, we obtain a sequence of subsets of  $\text{id}(R)$ :

$$R = R_0 \subseteq R_1 \subseteq R_2 \subseteq \dots \quad \text{and} \quad \mathbb{F}\langle X | R_i = 0 \rangle = \mathbb{F}\langle X | R_{i+1} = 0 \rangle.$$

If  $A$  is a finitely generated commutative algebra, this process terminates after finitely many steps, i.e., the sequence stabilizes. Let  $R_\infty = \bigcup_i R_i$ . Then by Theorem 15.8,  $\text{Irr}(R_\infty)$  is a basis of  $A$ . We call  $R_\infty$  the **Gröbner-Shirshov basis** of  $A$  and we call this algorithm for commutative algebra **Buchberger's theorem** or **Buchberger's algorithm**.

If  $A$  is graded associative algebra, say  $A = \bigoplus_{i \in \mathbb{N}} A_i$ . Let  $A_{\leq N}$  be the graded subspaces  $\bigoplus_{0 \leq i \leq N} A_i$ . Given  $N$ , we can find a basis of  $A_{\leq N}$  by this process as well.

### §15.3 Proof of Gröbner–Shirshov Theorem

**Sufficiency.** If there exists one reduction not 0, then it is a nontrivial linear combination of irreducible words. Since  $f, g \in R$ ,  $(f, g)_w = 0$  in  $A$ . Thus, this linear combination = 0, contrary to the linear independence.

**Necessity.** We need a lemma.

**Lemma 15.12.** *If  $R$  is closed under composition, then for all  $f \in \text{id}(R) \setminus \{0\}$ , the leading monomial  $\bar{f}$  is reducible.*

With this lemma, we can prove that if  $g$  is a nontrivial linear combination of irreducibles, then  $\bar{g}$  is irreducible. Hence  $g \notin \text{id}(R)$ , i.e.  $g \neq 0$  in  $A$ . Thus all irreducibles in  $A$  are linearly independent. Together with Proposition 15.5, we are done.

We now prove Lemma 15.12. Take it easy! The proof is almost straightforward. :)

*Proof of Lemma 15.12.* Write  $f \in \text{id}(R) \setminus \{0\}$  as  $\sum_i \alpha_i u_i r_i v_i$ , where  $\alpha_i \in \mathbb{F} \setminus \{0\}$ ,  $u_i, v_i \in X^*$ ,  $r_i \in R \setminus \{0\}$ . Note that  $\overline{u_i r_i v_i} = u_i \bar{r}_i v_i$  ( $u_i, v_i$  are monomials).

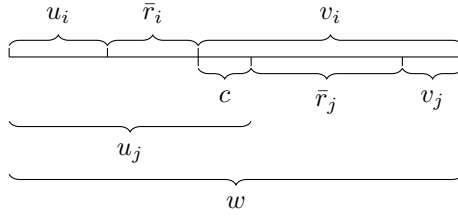
Let  $w = \max\{\overline{u_i r_i v_i} : i\}$ . Let  $\#w$  denote the number of occurrences of  $w$  among the terms  $u_i \bar{r}_i v_i$ . If  $\#w = 1$ , then  $\bar{f} = w$ , which is reducible. We will prove by induction on the pair  $(w, \#w)$ , ordered lexicographically.

Suppose that  $w = \overline{u_i r_i v_i} = \overline{u_j r_j v_j}$  for distinct  $i, j$ . Then

$$\alpha_i u_i r_i v_i + \alpha_j u_j r_j v_j = (\alpha_i + \alpha_j) u_i r_i v_i - \alpha_j (u_i r_i v_i - u_j r_j v_j).$$

In the following, we will prove  $u_i r_i v_i - u_j r_j v_j$  is a linear combination of lower terms (in the sense of length-lexi order). Then we can complete the proof by induction on  $\#w$ . We divide the remaining proof into three cases.

*Case 1:  $\bar{r}_i$  and  $\bar{r}_j$  do not overlap in  $w$ .* Write  $w$  as follows.



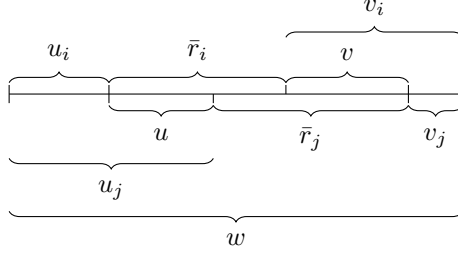
Case 1

Then we have

$$\begin{aligned} u_i r_i v_i - u_j r_j v_j &= u_i (\bar{r}_i + \{r_i\}) v_i - u_j (\bar{r}_j + \{r_j\}) v_j \\ &= u_i \{r_i\} v_i - u_j \{r_j\} v_j \\ &= u_i \{r_i\} c \bar{r}_j v_j - u_j \bar{r}_i c \{r_j\} v_j. \end{aligned}$$

The maximal summands of both  $u_i \{r_i\} c \bar{r}_j v_j$  and  $u_j \bar{r}_i c \{r_j\} v_j$  have length-lexi order strictly less than  $w$ .

*Case 2:  $\bar{r}_i$  and  $\bar{r}_j$  overlap, but neither is a subword of the other.* Write  $w$  as follows.



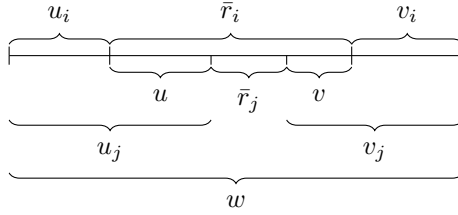
Case 2

Then we have

$$\begin{aligned}
 u_i r_i v_i - u_j r_j v_j &= u_i r_i v v_j - u_i u r_j v_j \\
 &= u_i (r_i v - u r_j) v_j \\
 &= u_i (r_i, r_j)_{u_i \bar{r}_i} v_j
 \end{aligned}$$

Since  $(r_i, r_j)_{u_i \bar{r}_i}$  can be reduced to 0,  $(r_i, r_j)_{u_i \bar{r}_i} = \sum_k \alpha'_k u'_k r'_k v'_k$  for some  $r'_k \in R$  and  $u'_k r'_k v'_k < w$ .

*Case 3: one of  $\bar{r}_i$  and  $\bar{r}_j$  is a subword of the other.* Write  $w$  as follows.



Case 3

We have  $u_i r_i v_i - u_j r_j v_j = u_i r_i v_i - u_i u r_j v v_i = -u_i (r_j, r_i)_{\bar{r}_i} v_i$ . Since  $(r_j, r_i)_{\bar{r}_i}$  can be reduced to 0,  $(r_j, r_i)_{\bar{r}_i} = \sum_k \alpha'_k u'_k r'_k v'_k$  for some  $r'_k \in R$  and  $u'_k r'_k v'_k < w$ .

□

## Chapter 16

# Applications of Gröbner–Shirshov Theorem

In this chapter we assume that  $\mathbb{F}$  is an arbitrary field (unless otherwise stated) and  $L$  is a finite dimensional Lie algebra over  $\mathbb{F}$ .

### §16.1 Universal enveloping algebras

**Definition 16.1.** For a Lie algebra  $L$ , the **universal enveloping algebra** of  $L$  is a pair  $(U, \iota)$ , where  $U$  is an associative algebra with 1 and a Lie homomorphism  $\iota : L \rightarrow U^{(-)}$  satisfying the universal property: for any associative algebra  $A$  with a Lie homomorphism  $\varphi : L \rightarrow A^{(-)}$ , there exists a unique associative algebra homomorphism  $\tilde{\varphi} : U \rightarrow A$  such that  $\tilde{\varphi} \circ \iota = \varphi$ .

$$\begin{array}{ccc} L & \xrightarrow{\iota} & U \\ & \searrow \varphi & \downarrow \tilde{\varphi} \\ & & A \end{array}$$

The uniqueness of universal enveloping algebras is easy to show, and the existence is proved by the following construction.

Let  $\{e_i : i \in I\}$  be a basis of  $L$ . Denote by  $\gamma_{ij}^k$  the structural constants, i.e.,

$$[e_i, e_j] = \sum_{k \in I} \gamma_{ij}^k e_k.$$

Let  $X = \{x_i : i \in I\}$  be an alphabet. Consider the associative algebra

$$U(L) := \mathbb{F}\langle X \mid x_i x_j - x_j x_i = \sum_{k \in I} \gamma_{ij}^k x_k \rangle$$

and the Lie algebra homomorphism  $\iota : L \rightarrow U(L); e_i \mapsto x_i$ . One can easily check that  $(U(L), \iota)$  is the universal enveloping algebra.

**Remark 16.2.** Let  $V$  be an  $L$ -module by  $\rho : L \rightarrow \text{End}(V)$ . Then by definition, we have an associative algebra homomorphism  $\tilde{\rho} : U(L) \rightarrow \text{End}(V)$  satisfying the universal property, which means  $V$  is a  $U(L)$ -module.

Furthermore, it is a one-to-one correspondence between  $L$ -modules and  $U(L)$ -modules by the PBW theorem, which will be proved in the next section.

## §16.2 Poincaré–Birkhoff–Witt Theorem

Using the same notations as §16.1, we have the following famous theorem.

**Theorem 16.3** (Poincaré–Birkhoff–Witt Theorem). *If  $I$  is a totally ordered set with the descending chain condition,<sup>a</sup> then  $U(L)$  has a basis*

$$x_{i_1}x_{i_2}\dots x_{i_n}, \quad \text{for various } i_1 \leq i_2 \leq \dots \leq i_n, \quad n \in \mathbb{N}.$$

<sup>a</sup>If  $I$  is finite, this holds naturally.

*Proof.* Denote by  $R = \{x_i x_j - x_j x_i - \sum_{k \in I} \gamma_{ij}^k x_k : i \neq j\}$ . Define a total order on  $X$  by  $x_j < x_i$  if  $j < i$ . Note that  $x_i x_j$  for any  $j < i$  is reducible. Thus

$$\text{Irr}(R) = \{x_{i_1}x_{i_2}\dots x_{i_n} \mid \text{for various } i_1 \leq i_2 \leq \dots \leq i_n, \quad n \in \mathbb{N}\}.$$

Thanks to Theorem 15.8, it suffices to show that  $R$  is closed under composition.

In order to simplify the calculation, we denote  $\sum_{k \in I} \gamma_{ij}^k x_k$  by  $\{x_i, x_j\}$ . The only relations admitting a composition in  $R$  are

$$f = x_i x_j - x_j x_i - \{x_i, x_j\}, \quad g = x_j x_l - x_l x_j - \{x_j, x_l\},$$

for all  $l < j < i$ . Let  $w = x_i x_j x_l$ . We have

$$\begin{aligned} (f, g)_w &= -x_j x_i x_l - \{x_i, x_j\} x_l + x_i x_l x_j + x_i \{x_j, x_l\} \\ &\rightarrow -x_j (x_l x_i - \{x_i, x_l\}) - \{x_i, x_j\} x_l + (x_l x_i + \{x_i, x_l\}) x_j + x_i \{x_j, x_l\} \\ &= -x_j x_l x_i - x_j \{x_i, x_l\} - \{x_i, x_j\} x_l + x_l x_i x_j + \{x_i, x_l\} x_j + x_i \{x_j, x_l\} \\ &\rightarrow -(x_l x_j + \{x_j, x_l\}) x_i - x_j \{x_i, x_l\} - \{x_i, x_j\} x_l + x_l (x_j x_i + \{x_i, x_j\}) \\ &\quad + \{x_i, x_l\} x_j + x_i \{x_j, x_l\} \\ &= -\{x_j, x_l\} x_i + x_i \{x_j, x_l\} - x_j \{x_i, x_l\} + \{x_i, x_l\} x_j - \{x_i, x_j\} x_l + x_l \{x_i, x_j\} \\ &= \{x_i, \{x_j, x_l\}\} + \{x_j, \{x_l, x_i\}\} + \{x_l, \{x_i, x_j\}\} \\ &\rightarrow 0. \end{aligned}$$

The last reduction is due to Jacobi identity. □

**Corollary 16.4.** *Let  $L$  be a Lie algebra.*

1. *The map  $\iota : L \rightarrow U(L)$  given by  $\iota(e_i) = x_i$  is an injection.*

2. If  $L'$  is a subalgebra of  $L$ , then  $U(L') \subset U(L)$

# Chapter 17

## Generators and Relations of $L$

In this chapter, we first introduce free Lie algebras in §17.1. This construction does not depend on any special assumption on the field  $\mathbb{F}$ . Afterwards, we use free Lie algebras to construct semisimple Lie algebras associated with root systems. In the remaining sections, unless otherwise stated, we assume that  $\mathbb{F}$  is an algebraically closed field with  $\text{char } \mathbb{F} = 0$ .

### §17.1 Free Lie algebras

Let  $X = \{x_1, x_2, \dots, x_n\}$  be an alphabet.

**Definition 17.1.** *The free Lie algebra generated by  $X$  is a Lie algebra  $\text{Lie } \langle X \rangle$  together with a map  $\iota : X \rightarrow \text{Lie } \langle X \rangle$  satisfying the universal property:  $\forall$  Lie algebra  $L$ ,  $\forall$  map  $\varphi : X \rightarrow L$ , there is a unique Lie algebra homomorphism  $\tilde{\varphi}$  s.t.  $\tilde{\varphi} \circ \iota = \varphi$ .*

$$\begin{array}{ccc} X & \xrightarrow{\iota} & \text{Lie } \langle X \rangle \\ & \searrow \varphi & \downarrow \tilde{\varphi} \\ & & L \end{array}$$

As in the case of universal enveloping algebras, the uniqueness of free Lie algebras is obvious. Now we will show the existence by construction.

A natural first candidate is  $\mathbb{F} \langle X \rangle^{(-)}$ . However, this algebra is too large. For example,  $x_{i_1} x_{i_2} \in \mathbb{F} \langle X \rangle^{(-)}$ , but it cannot be generated by  $X$  through the Lie bracket. Therefore, we consider the Lie subalgebra generated by all commutators.

**Definition 17.2.** *We define a **commutator** in  $\mathbb{F} \langle X \rangle$  by the following conditions:*

1. Every  $x_i$  is a commutator;
2. If  $\rho', \rho''$  are commutators, then  $[\rho', \rho''] = \rho' \rho'' - \rho'' \rho'$  is a commutator.

**Proposition 17.3.**  *$\text{Lie } \langle X \rangle = \{\sum_i \alpha_i \rho_i \mid \rho_i \text{ are commutators}\}$  is the free Lie algebra generated by  $X$ .*

*Proof.* First,  $\text{Lie}\langle X \rangle$  is closed under the Lie bracket by the definition of commutators. Hence it is a Lie subalgebra of  $\mathbb{F}\langle X \rangle^{(-)}$ .

Let  $L$  be a Lie algebra and let  $\varphi : X \rightarrow L$  be a map. By Corollary 16.4, we may regard  $L$  as a Lie subalgebra of  $U(L)^{(-)}$ . Thus  $\varphi$  may also be viewed as a map from  $X$  to  $U(L)$ . By the universal property of  $\mathbb{F}\langle X \rangle$ , there exists a unique associative algebra homomorphism

$$\widehat{\varphi} : \mathbb{F}\langle X \rangle \rightarrow U(L)$$

extending  $\varphi$ .

We claim that the restriction of  $\widehat{\varphi}$  to  $\text{Lie}\langle X \rangle$  has image in  $L$ . Indeed, every element of  $\text{Lie}\langle X \rangle$  is a linear combination of commutators, and the image of a commutator under  $\widehat{\varphi}$  is the corresponding iterated Lie bracket of elements of  $L$ . Hence it lies in  $L$ . Therefore, we obtain a map

$$\tilde{\varphi} = \widehat{\varphi}|_{\text{Lie}\langle X \rangle} : \text{Lie}\langle X \rangle \rightarrow L.$$

Since  $\widehat{\varphi}$  is an associative algebra homomorphism, its restriction is compatible with commutators. Hence  $\tilde{\varphi}$  is a Lie algebra homomorphism, and clearly  $\tilde{\varphi} \circ i = \varphi$ .

Finally, the uniqueness follows because  $\text{Lie}\langle X \rangle$  is generated by  $X$  as a Lie algebra. Thus any Lie algebra homomorphism  $\text{Lie}\langle X \rangle \rightarrow L$  extending  $\varphi$  is uniquely determined by its values on  $X$ .  $\square$

For a subset  $R \subset \text{Lie}\langle X \rangle$ , we denote by  $\text{Lie}\langle X | R = 0 \rangle = \text{Lie}\langle X \rangle / \text{id}_{\text{Lie}\langle X \rangle}(R)$ .

**Proposition 17.4.**  $U(\text{Lie}\langle X | R = 0 \rangle) = \mathbb{F}\langle X | R = 0 \rangle$ .

*Proof.* Since  $\text{id}_{\text{Lie}\langle X \rangle}(R) \subset \text{id}_{\mathbb{F}\langle X \rangle}(R)$ , the assignment  $x_i + \text{id}_{\text{Lie}\langle X \rangle}(R) \mapsto x_i + \text{id}_{\mathbb{F}\langle X \rangle}(R)$  extends to a Lie algebra homomorphism  $\iota : \text{Lie}\langle X | R = 0 \rangle \rightarrow \mathbb{F}\langle X | R = 0 \rangle^{(-)}$ .

Let  $\varphi : \text{Lie}\langle X | R = 0 \rangle \rightarrow A^{(-)}$  be a Lie algebra homomorphism. We extend the assignment

$$\tilde{\varphi} : \mathbb{F}\langle X | R = 0 \rangle \rightarrow A; \quad x_i + \text{id}_{\mathbb{F}\langle X \rangle}(R) \mapsto \varphi(x_i + \text{id}_{\text{Lie}\langle X \rangle}(R))$$

as an associative algebra homomorphism.

Let us show it is well-defined. Note that  $\tilde{\varphi}(r) = 0$  for any  $r \in R$ . Let  $\rho = \sum_i \alpha_i u_i r_i v_i \in \text{id}_{\mathbb{F}\langle X \rangle}(R)$ , where  $u_i, v_i \in X^*$  and  $r_i \in R$ . Then  $\tilde{\varphi}(\rho) = \sum_i \alpha_i \tilde{\varphi}(u_i) \tilde{\varphi}(r_i) \tilde{\varphi}(v_i) = 0$ .

Moreover, it is routine to check  $\tilde{\varphi}$  satisfies the universal property.  $\square$

## §17.2 Relations satisfied by $L$

Let us review the root decomposition of semisimple Lie algebra  $L$ . Given a maximal toral subalgebra  $H$ , we have a root system  $\Delta$  such that

$$L = H + \sum_{\alpha \in \Delta} L_{\alpha}.$$

Let  $B = \{\alpha_1, \dots, \alpha_n\}$  be a base of  $\Delta$ . By Proposition 10.11, for any nonzero  $x_i \in L_{\alpha_i}$ ,

there exists  $y_i \in L_{-\alpha_i}$  such that

$$(x_i | y_i) = \frac{2}{(\alpha_i | \alpha_i)}$$

and

$$[x_i, y_i] = \frac{2\nu^{-1}(\alpha_i)}{(\alpha_i | \alpha_i)} =: \alpha_i^\vee \in \mathfrak{h}.^1$$

Recall that for  $\alpha \in \mathfrak{h}^*$  and  $h \in \mathfrak{h}$ , we defined  $\langle \alpha, h \rangle := \alpha(h)$ .  $C$  is the Cartan matrix associated to  $(\Delta, B)$  (Definition 12.13). A straightforward calculation shows that

$$c_{ij} = \frac{2(\alpha_i | \alpha_j)}{(\alpha_i | \alpha_i)} = \langle \alpha_j, \alpha_i^\vee \rangle = \begin{cases} \leq 0 & \text{if } i \neq j, \\ 2 & \text{if } i = j. \end{cases}$$

For  $x_j \in L_{\alpha_j}$  and  $y_j \in L_{-\alpha_j}$ , by the definition of root spaces we have

$$[\alpha_i^\vee, x_j] = \alpha_j(\alpha_i^\vee)x_j = c_{ij}x_j, \quad [\alpha_i^\vee, y_j] = -\alpha_j(\alpha_i^\vee)y_j = -c_{ij}y_j.$$

In conclusion, for any  $1 \leq i, j \leq n$ , the following relations hold in  $L$ :

$$\begin{aligned} [\alpha_i^\vee, \alpha_j^\vee] &= 0, \\ [\alpha_i^\vee, x_j] &= c_{ij}x_j, \\ [\alpha_i^\vee, y_j] &= -c_{ij}y_j, \\ [x_i, y_j] &= \begin{cases} \alpha_i^\vee & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \end{aligned} \tag{R'}$$

However, these are not all the relations among the elements  $x_i, y_j$  and  $\alpha_k^\vee$ . It remains to describe the commutation relations between  $x_i$  and  $x_j$ , and between  $y_i$  and  $y_j$ , for  $i \neq j$ .

Now regard  $(x_i, y_i, \alpha_i^\vee)$  as an  $\mathfrak{sl}_2$ -triple, and view  $L$  as an  $\mathfrak{sl}_2$ -module under the adjoint action. Thanks to Proposition 10.15, for  $j \neq i$ ,  $\bigoplus_{k \in \mathbb{Z}} L_{\alpha_j + k\alpha_i}$  is an irreducible  $\mathfrak{sl}_2$ -module and the element  $x_j$  is an eigenvector of  $\alpha_i^\vee$  with eigenvalue  $c_{ij}$ . By Theorem 9.5 and  $(\text{ad } y_i)(x_j) = 0$ , we obtain for any  $i \neq j$

$$(\text{ad } x_i)^{-c_{ij}+1}(x_j) = 0, \tag{(\Theta_{ij}^+)}$$

$$(\text{ad } y_i)^{-c_{ij}+1}(y_j) = 0. \tag{(\Theta_{ij}^-)}$$

These relations are called the **Serre relations**.

**Summary.** For a semisimple Lie algebra  $L$ , we can take root vectors  $x_i, y_i$  and  $\alpha_i^\vee$  for each simple root  $\alpha_i \in B$  s.t.

$$\alpha_1^\vee, \dots, \alpha_n^\vee, x_1, \dots, x_n, y_1, \dots, y_n$$

generate  $L$  and at least satisfy the relations  $(R')$ ,  $(\Theta_{ij}^+)$  and  $(\Theta_{ij}^-)$ .

<sup>1</sup>In this section,  $\alpha_i^\vee$  is not a letter.

### §17.3 Auxiliary Lie algebras

For an arbitrary root system  $\Delta$  with a base  $B = \{\alpha_1, \dots, \alpha_n\} \subset \Delta$ , we consider an alphabet

$$\tilde{X} = \{\alpha_1^\vee, \dots, \alpha_n^\vee, x_1, \dots, x_n, y_1, \dots, y_n\},$$

with the order  $\alpha_1^\vee < \dots < \alpha_n^\vee < x_1 < \dots < x_n < y_1 < \dots < y_n$ .

**Definition 17.5.** We define the *auxiliary Lie algebra*  $L' = \text{Lie} \langle \tilde{X} | R' \rangle$ , where  $[\cdot, \cdot]$  in  $(R')$  are regarded as commutators.

Thanks to Proposition 17.4,  $U(L') = \langle \tilde{X} | R' \rangle$ .

By a painful but straightforward calculation, we can check that  $R'$  is closed under composition. Using the Grobner-Shirshov Theorem, we can prove that  $U(L')$  has a basis

$$\alpha_{i_1}^\vee \dots \alpha_{i_p}^\vee x_{j_1} \dots x_{j_q} y_{k_1} \dots y_{k_s}, \quad i_1 \leq \dots \leq i_p.$$

Denote the sets  $\{x_1, \dots, x_n\}$  and  $\{y_1, \dots, y_n\}$  by  $X$  and  $Y$ , respectively. Note that every word  $x_{j_1} \dots x_{j_q}$  is irreducible. Thus the Lie subalgebra generated by  $X$  is a free Lie algebra  $\text{Lie} \langle X \rangle$ . The same argument applies to  $Y$ .

**Theorem 17.6.**  $L' = \text{Lie} \langle X \rangle + \text{span}\{\alpha_1^\vee, \dots, \alpha_n^\vee\} + \text{Lie} \langle Y \rangle$ .

*Proof.* It remains to show the right-hand side is a subalgebra of  $L'$ . Since we can easily prove that it is closed under bracket by  $\alpha_i^\vee, x_i$  and  $y_i$ , it is an ideal of  $L$  and thus a subalgebra.  $\square$

# Chapter 18

## Existence Theorem

In this chapter, we always assume  $\mathbb{F}$  is an algebraically closed field with  $\text{char } \mathbb{F} = 0$ . Our main goal is to construct a Lie algebra for an arbitrary root system.

Recall that we denote  $L' = \text{Lie} \langle \tilde{X} | R' \rangle$ . In this chapter, we denote  $L = \text{Lie} \langle \tilde{X} | R', \Theta_{ij}^{\pm} \rangle$ .

### §18.1 Revisiting Serre's relations

Let us consider the remaining relations, Serre's relations. By abuse of notation, we denote by  $\Theta_{ij}^+$  and  $\Theta_{ij}^-$  the left-hand side of  $(\Theta_{ij}^+)$  and  $(\Theta_{ij}^-)$ , respectively.

**Lemma 18.1.** *For  $i \neq j$ ,  $\text{id}_{L'} \Theta_{ij}^+ = \text{id}_{\text{Lie} \langle X \rangle} \Theta_{ij}^+$  and  $\text{id}_{L'} \Theta_{ij}^- = \text{id}_{\text{Lie} \langle Y \rangle} \Theta_{ij}^-$ .*

*Proof.* By the symmetry of  $X$  and  $Y$ , it suffices to show  $\text{id}_{\text{Lie} \langle X \rangle} \Theta_{ij}^+$  is an ideal of  $L'$ . Since  $x_i$ 's are eigenvectors of  $\text{ad } \alpha_i^{\vee}$ 's,  $\alpha_i^{\vee}$  leaves  $\text{id}_{\text{Lie} \langle X \rangle} \Theta_{ij}^+$  invariant. Hence we only need to consider the action of  $y_k$ . Since  $[\text{ad } y_k, \text{ad } x_{k'}] = -\delta_{kk'} \text{ad } \alpha_k^{\vee}$ , it suffices to check  $(\text{ad } y_k)(\Theta_{ij}^+) = 0$ .

*Case 1:*  $k \neq i, j$ .  $\text{ad } y_k$  commutes with  $\text{ad } x_i$  and annihilates  $x_j$ . Thus  $(\text{ad } y_k)(\Theta_{ij}^+) = 0$ .

*Case 2:*  $k = j$ .  $\text{ad } y_k$  still commutes with  $\text{ad } x_i$ . Then we have

$$(\text{ad } y_k)(\Theta_{ij}^+) = (\text{ad } x_i)^{-c_{ij}+1} \text{ad } y_j(x_j) = -(\text{ad } x_i)^{-c_{ij}+1} \alpha_j^{\vee}.$$

If  $c_{ij} < 0$ ,  $-c_{ij} + 1 \geq 2$  and then  $(\text{ad } x_i)^{-c_{ij}+1} \alpha_j^{\vee} = 0$ . If  $c_{ij} = 0$ ,  $(\text{ad } x_i)(\alpha_j^{\vee}) = -c_{ij} x_i = 0$ . Thus  $(\text{ad } y_k)(\Theta_{ij}^+) = 0$ .

*Case 3:*  $k = i$ . We have

$$\begin{aligned} (\text{ad } y_i)(\text{ad } x_i)^{-c_{ij}+1} x_j &= [\text{ad } y_i, \text{ad } x_i](\text{ad } x_i)^{-c_{ij}} x_j + (\text{ad } x_i)[\text{ad } y_i, \text{ad } x_i](\text{ad } x_i)^{-c_{ij}-1} x_j \\ &\quad + \dots + (\text{ad } x_i)^{-c_{ij}} [\text{ad } y_i, \text{ad } x_i] x_j \\ &= -(\text{ad } \alpha_i^{\vee})(\text{ad } x_i)^{-c_{ij}} x_j - (\text{ad } x_i)(\text{ad } \alpha_i^{\vee})(\text{ad } x_i)^{-c_{ij}-1} x_j \\ &\quad - \dots - (\text{ad } x_i)^{-c_{ij}} (\text{ad } \alpha_i^{\vee}) x_j \\ &= ((2c_{ij} - c_{ij}) + (2(c_{ij} - 1) - c_{ij}) + \dots + c_{ij}) (\text{ad } x_i)^{-c_{ij}} x_j \\ &= 0. \end{aligned}$$

□

**Summary.** Write  $H = \text{span}\{\alpha_1^\vee, \dots, \alpha_n^\vee\}$ . We have

$$L = \text{Lie}\langle \tilde{X}|R', \Theta_{ij}^\pm \rangle = \text{Lie}\langle X|\Theta_{ij}^+, i \neq j \rangle + H + \text{Lie}\langle Y|\Theta_{ij}^-, i \neq j \rangle.$$

By relations in  $(R')$ ,  $H$  is a toral subalgebra. Since  $x_i$ 's and  $y_i$ 's are eigenvectors of  $H$ ,  $L = H + \sum_{\gamma \in H^*} L_\gamma$ . Moreover, it is easy to see that given  $x_{\beta_i} \in L_{\beta_i}$ ,  $[x_{\beta_1}, x_{\beta_2}]$  has the root  $\beta_1 + \beta_2$  if the bracket is nonzero.

Identifying  $H^*$  and  $E$ , the Euclidean space associated to  $\Delta$ , by  $\langle \alpha_j, \alpha_i^\vee \rangle = c_{ij}$ , we obtain

$$\text{Lie}\langle X|\Theta_{ij}^+, i \neq j \rangle = \sum_{\gamma \in \text{NB} \setminus \{0\}} L_\gamma, \quad \text{Lie}\langle Y|\Theta_{ij}^-, i \neq j \rangle = \sum_{\gamma \in -\text{NB} \setminus \{0\}} L_\gamma, \quad H = L_0.$$

**Proposition 18.2.** For a simple root  $\alpha_i$ , we have

$$\dim L_{k\alpha_i} = \begin{cases} 1 & \text{if } k = \pm 1; \\ 0 & \text{if } k \neq 0, \pm 1. \end{cases}$$

*Proof.* From the prospective of the height of root, it is easy to see that  $L_{\alpha_i} = \mathbb{F}x_i$  and  $L_{-\alpha_i} = \mathbb{F}y_i$  and  $L_{k\alpha_i} = 0$  if  $k \neq 0, \pm 1$ . □

## §18.2 Braid group automorphisms

Recall that for a locally nilpotent operator  $\varphi$ ,  $\exp \varphi$  is well-defined (Remark 1.10). Moreover, if  $\varphi$  is a derivation, to check it is locally nilpotent, it suffices to verify that it is locally nilpotent on a set of generators (Leibniz rule!). Furthermore, if  $\varphi$  is a nilpotent derivation,  $\exp \varphi$  is an automorphism:

$$\begin{aligned} \exp \varphi(ab) &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=0}^k \binom{k}{i} \varphi^i(a) \varphi^{k-i}(b) \\ &= \sum_{k=0}^{\infty} \sum_{i=0}^k \frac{1}{i!(k-i)!} \varphi^i(a) \varphi^{k-i}(b) \\ &= \sum_{i,j \geq 0} \frac{1}{i!j!} \varphi^i(a) \varphi^j(b) \\ &= \left( \sum_{i=0}^{\infty} \frac{\varphi^i(a)}{i!} \right) \left( \sum_{j=0}^{\infty} \frac{\varphi^j(b)}{j!} \right) \\ &= \exp \varphi(a) \exp \varphi(b). \end{aligned}$$

By  $(\Theta_{ij}^+)$  and  $(\text{ad } x_i)^3(H) = 0$ ,  $\text{ad } x_i$  is a locally nilpotent derivation on  $L$ . Hence

$\exp \operatorname{ad} x_i$  is a well-defined algebra automorphism with inverse  $\exp(-\operatorname{ad} x_i)$ . The same argument applies to  $\operatorname{ad} y_i$ .

Set  $r_i := \exp \operatorname{ad} x_i \exp(-\operatorname{ad} y_i) \exp \operatorname{ad} x_i \in \operatorname{Aut}(L)$ .

**Definition 18.3.** We call an automorphism  $\varphi \in \operatorname{Aut}(L)$  a **braid group automorphism**<sup>a</sup> if  $\varphi$  is generated by  $r_i$ 's.

<sup>a</sup>The name comes from the fact that the subgroup generated by  $r_i$ 's can be verified to be a braid group.

**Lemma 18.4.** We write  $s_i$  for the reflection associated to the simple root  $\alpha_i$  in Weyl group  $W$ .<sup>a</sup>

1.  $\langle w(h), w(\alpha) \rangle = \langle h, \alpha \rangle$  for any  $w \in W$ ,  $h \in H$  and  $\alpha \in H^*$ ;
2.  $r_i|_H = s_i$  for any  $i$ ;
3.  $r_i(L_\alpha) \subseteq L_{s_i(\alpha)}$ .

<sup>a</sup>The action of  $W$  on  $H$  is defined by  $s_i(h) = h - \langle h, \alpha_i \rangle \alpha_i^\vee$ .

*Proof.* It suffices to show  $\langle s_i(h), s_i(\alpha) \rangle = \langle h, \alpha \rangle$  for each  $i$ .

$$\langle s_i(h), s_i(\alpha) \rangle = \langle h - \langle h, \alpha_i \rangle \alpha_i^\vee, \alpha - \langle \alpha, \alpha_i^\vee \rangle \alpha_i \rangle = \langle h, \alpha \rangle.$$

The second claim is given by a painful but straightforward calculation, i.e.,  $r_i(h) = h - \langle h, \alpha_i \rangle \alpha_i^\vee$ .

Let us prove (3). Since  $r_i$  is an automorphism, for any  $a_\alpha \in L_\alpha$ , we have

$$[h, r_i(a_\alpha)] = r_i([r_i^{-1}(h), a_\alpha]) = \langle s_i^{-1}(h), \alpha \rangle r_i(a_\alpha) = \langle h, s_i(\alpha) \rangle r_i(a_\alpha).$$

□

## §18.3 Serre's Theorem

If  $w = s_{i_1} \dots s_{i_k}$  we write  $r_w = r_{i_1} \dots r_{i_k}$ .<sup>1</sup>

**Proposition 18.5.** The set of roots of  $L$ ,  $\{\gamma \in H^* \setminus \{0\} : L_\gamma \neq 0\}$ , coincides with  $\Delta$ . Moreover, for all  $\gamma \in \Delta$ ,  $\dim L_\gamma = 1$  and thus  $\dim L < \infty$ .

*Proof.* If  $\gamma$  is collinear to some root, i.e.,  $\gamma = k\alpha$  for some  $\alpha \in \Delta$  with  $k \neq 0$ , then by Corollary 12.10  $\exists w \in W$  s.t.  $\alpha_i = w(\alpha)$  for some simple root  $\alpha_i$ . By Lemma 18.4 and Proposition 18.2,  $r_w(L_\gamma) \subseteq L_{k\alpha_i}$ . Note that  $r_w$  is an automorphism. Hence  $\dim L_\alpha = 1$  if  $k = \pm 1$ , and 0 otherwise.

<sup>1</sup>Here one should be careful:  $w$  may have multiple expressions, so this notation might not be well-defined. However, this automorphism is indeed well-defined, and this subtlety does not affect our following discussion.

If  $\gamma$  is not collinear to any root, we will prove that  $\exists w \in W$  s.t.  $w(\gamma)$  is mixed, i.e.,  $\gamma$  is a linear combination of  $\alpha_i$ 's with both positive and negative coeffs. Since  $\gamma$  is not collinear to any root,  $\gamma^\perp \not\subseteq \bigcup_{\alpha \in \Delta} \alpha^\perp$ . Thus  $\exists \beta \in \gamma^\perp$  s.t.  $\beta \notin \bigcup_{\alpha \in \Delta} \alpha^\perp$ , that is,  $\beta$  is regular. By the 1-1 correspondence between bases and Weyl chambers,  $\exists w \in W$ , s.t.  $(w(\beta)|w(\gamma)) = 0$  but  $(w(\beta)|\alpha_i) > 0$ . Thus  $w(\gamma)$  is mixed (otherwise  $(w(\gamma)|w(\beta)) > 0$ ). But by the decomposition in Summary of §18.1,  $w(\gamma)$  is not a root, neither is  $\gamma$ .  $\square$

**Lemma 18.6.** *If  $I \subset L$  is an ideal, then  $I = (I \cap H) + \sum_{\alpha \in \Delta} (I \cap L_\alpha)$ .<sup>a</sup>*

<sup>a</sup>It is valid for any  $L$ -module which decomposes as a direct sum of eigenspaces of a finite dimensional commutative subalgebra. The proof is exactly the same.

*Proof.* Let  $a = a_{\gamma_1} + \dots + a_{\gamma_r} \in I$ , where  $a_{\gamma_i} \in L_{\gamma_i}$  and  $\alpha_i \neq \alpha_j$  if  $i \neq j$ . Then the hyperplanes of  $\gamma_i$  and  $\gamma_i - \gamma_j$  are proper subspaces of  $H^*$ . Thus there exists  $h \in H$ , s.t.  $\gamma_i(h)$ 's are nonzero and distinct. Then iteratively applying  $\text{ad } h$  to  $a$ , we will get a system of equations which has a Vandermonde matrix as its coef matrix. Thus each component  $a_{\gamma_i} \in I$ .  $\square$

As a consequence,  $H$  is a maximal toral subalgebra.

**Proposition 18.7.**  *$L$  is semisimple.*

*Proof.* To show  $L$  is semisimple, it suffices to show any nonzero ideal  $I$  can not be abelian. Suppose  $I \subset L$  is a nonzero abelian ideal. Let  $a = a_{\gamma_1} + \dots + a_{\gamma_r} \in I$ , where  $a_{\gamma_i} \in L_{\gamma_i}$  and  $\alpha_i \neq \alpha_j$  if  $i \neq j$ . Thanks to Lemma 18.6, each component  $a_{\gamma_i} \in I$ .

It is easy to see that  $H \cap I = 0$ . Suppose  $a_\alpha \in I$  for some nonzero  $a_\alpha \in L_\alpha$ . Then by symmetry, there exists  $b_\alpha \in L_{-\alpha}$  s.t.  $0 \neq [a_\alpha, b_\alpha] \in L_0 = H$ . Otherwise, by a braid group action, we obtain  $[x_i, y_i] = 0$  for some simple root  $\alpha_i$  contrary to the relation in  $(R')$ . But in this case  $H \cap I \neq 0$ , which is absurd.  $\square$

**Summary.** *Combining Proposition 18.5 and Proposition 18.7, we obtain Serre's Theorem.*

**Theorem 18.8** (Serre's Theorem). *Given a root system  $\Delta$  with a base  $B$ , let  $L = \text{Lie} \langle \tilde{X} | R', \Theta_{i_j}^\pm \rangle$ . Then  $L$  is a finite dimensional semisimple Lie algebra with maximal toral subalgebra spanned by  $\alpha_i^\vee$ 's and with corresponding root system  $\Delta$ .*

# Afterword / 课程后记



Prof. Efim Zelmanov teaching Lie Theory, March 16, 2026.

## 中文版

Zelmanov 教授上课时只带一杯咖啡。除此之外，好像什么都不需要。所有的定理和证明似乎都早已在他的脑海中。他的板书流畅而优美，讲解过程中几乎没有停顿，也很少出现卡壳或“挂黑板”的情况。我常常惊叹于他对基础知识的掌握程度：他仿佛能够从李代数的定义出发，不借助任何参考资料，一路自然地讲到 Serre 定理。

他曾在课上说：“我只会教那些我不需要准备的内容。因为如果我为了上课去翻某本教科书，把定理和证明背下来，再复述给你们听，那是作弊。最终，我会忘记这个证明，你们也会忘记。这没有任何意义。”这句话给我留下了很深的印象。它让我意识到，真正的理解并不是记住某个证明的每一步，而是能够把它放回整个理论的脉络中，知道它为什么自然、为什么必要。

Zelmanov 教授的课并不追求很快的进度。相反，他常常在一节课的开始，花很长时间回顾“where we stand”。有时候，这样的回顾会占去半节课，甚至更久；有时候，一个重要定理的证明会被重新讲一遍。起初，我常常觉得这样的回顾没有必要，甚至会拖慢课程进度。但他曾作过一个比喻：学习就像往水里加糖。你可以不断往水中加糖，但到了某个点以后，糖就不再溶解了。

现在回想起来，确实如此。很多知识我曾经以为自己已经学过，但其实并没有真正理解；如果让我重新证明其中的一些结果，我未必能够完整地写出来。这门课让我意识到，数学学

习并不只是不断向前推进，更重要的是让已经学过的内容真正“溶解”在自己的理解之中。

也许多年以后，我会忘记某些具体的证明，忘记某些技术性的细节，但我仍会记得这门课的节奏，记得黑板上一行行自然生长出的公式，记得一位数学家手拿一杯咖啡走进教室，却仿佛带来了整个理论世界的样子。

## English version

Professor Zelmanov came to class with only a cup of coffee. Apart from that, he seemed to need nothing else. All the theorems and proofs seemed already to be present in his mind. His writing on the blackboard was fluent and elegant; throughout his lectures, he almost never paused, nor did he often get stuck in front of the board. I was often amazed by the depth of his command of the fundamentals: it felt as if he could begin from the definition of a Lie algebra and, without relying on any reference materials, naturally proceed all the way to the Serre's Theorem.

He once said in class, "I only teach what I do not need to prepare. If I have to open a textbook, memorize a theorem and its proof, and then repeat them to you, that is cheating. In the end, I will forget the proof, and so will you. It has no meaning." These words left a deep impression on me. They made me realize that true understanding is not merely about remembering every step of a proof, but about being able to place it back into the broader structure of the theory and to know why it is natural and why it is necessary.

Professor Zelmanov's lectures did not pursue a fast pace. On the contrary, at the beginning of a class, he often spent a long time reviewing "where we stand". Sometimes such a review would take half a lecture, or even longer; sometimes the proof of an important theorem would be explained all over again. At first, I often felt that such reviews were unnecessary and that they slowed down the progress of the course. But he once offered a metaphor: learning is like adding sugar to water. You can keep adding sugar to the water, but after a certain point, it will no longer dissolve.

Looking back now, this is indeed true. Much of the knowledge I once thought I had learned, I had not truly understood. If I were asked to reproduce some of the results, I might not be able to write down complete proofs. This course made me realize that learning mathematics is not only about constantly moving forward; more importantly, it is about allowing what has already been learned to truly "dissolve" into one's own understanding.

Perhaps many years from now, I will have forgotten some specific proofs and some technical details. But I will still remember the rhythm of this course, the lines of formulas that seemed to grow naturally across the blackboard, and the image of a mathematician walking into the classroom with a cup of coffee in hand, as if he had brought with him an entire theoretical world.