Lecture Note on Lie Algebra and Representation Theory

Anthony Hong[1]

November 3, 2024

[1]Thanks to Professor Precup for her teaching Math547 Lie Algebra and Representation Theory.

Contents

Chapter 1

Lie Algebras: An Introduction

1.1 Definitions and Examples

Definition 1.1.1. *Let F be a field.* ***Lie Algebra*** *is a vector space L over F together with a bilinear map $[\cdot, \cdot] : L \times L \to L$ called* ***bracket*** *such that*

1. $[x, x] = 0$*;*
2. $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ ***(Jacobi identity)***

The ***dimension*** *of the Lie algebra is the dimension of the vector space (we assume finite dimensionality).*

Proposition 1.1.2. *Lie bracket is skew-symmetric:* $[x, y] = -[y, x]$.

Proof. $0 = [x + y, x + y] = [x, y] + [y, x]$ due to bilinearity and $[x, x] = 0$. ■

Example 1.1.3.

1. T_eG, or L, where G is a Lie group.
2. **Abelian Lie algebra**: L any F-vector space and set $[x, y] = 0$ for any $x, y \in L$.
3. **General linear algebra**: let V be a vector space. Define $L = \mathfrak{gl}(V) = \{x : V \to V \mid x \text{ a linear transformation}\}$, where $[x, y] = x \circ y - y \circ x = xy - yx$.

♣

Definition 1.1.4. *A linear map $\varphi : L \to L'$ between Lie algebras is a* ***homomorphism*** *if*

$$\varphi([x, y]) = [\varphi(x), \varphi(y)].$$

The homomorphism φ is an ***isomorphism*** *if φ is a bijection.*

A subspace $K \subseteq L$ is a ***(Lie) subalgebra*** *if $[x, y] \in K$ for any $x, y \in K$.*

Example 1.1.5. Let V be a F-vector space. For B_n, C_n, and D_n, let $\mathrm{char}(F) \neq 2$.

1. Any subalgebra of Lie algebra $\mathfrak{gl}(V)$ is called a **linear Lie algebra**.
2. Type A_n (**Special linear Lie algebra**): Suppose $\dim(V) = n + 1$. Let $\mathfrak{sl}_{n+1}(F) = \mathfrak{sl}(V) = \{x \in \mathfrak{gl}(V) \mid \mathrm{tr}(x) = 0\}$. Recall that the trace of an endomorphism is the trace of its matrix, which is independent of choice of basis. $\mathfrak{sl}_{n+1}(F)$ is a Lie subalgebra of $\mathfrak{gl}(V)$ because $\mathrm{tr}(x + y) = \mathrm{tr}(x) + \mathrm{tr}(y)$ and $\mathrm{tr}(xy) = \mathrm{tr}(yx)$.

3. Type C_n (**Symplectic Lie algebra**): Suppose $\dim V = 2n$ with basis $(w_1, \cdots, w_{2n})$. Define a nondegenerate skew-symmetric bilinear form Ω by

$$\Omega(u,v) = [-u-]\begin{bmatrix} O & I_n \\ -I_n & O \end{bmatrix}\begin{bmatrix} | \\ v \\ | \end{bmatrix}$$

It is nondegenerate in the sense that the map $\widetilde{\Omega} : V \to V^*; \Omega(u)(v) = \Omega(u)(v)$ is bijective, i.e., $U = \{u \in V | \Omega(u,v) = 0 \ \forall v \in V\}$ is a zero subspace of V. Set

$$\mathfrak{sp}(V) = \mathfrak{sp}_{2n}(F) = \{x \in \mathfrak{gl}(V) | \Omega(x(u), v) = -\Omega(u, x(v))\}.$$

It is a subalgebra because

$$\begin{aligned}\Omega([x,y](u),v) &= \Omega(x(y(u)) - y(x(u)), v) \\ &= \Omega(x(y(u)),v) - \Omega(y(x(u)),v) \\ &= -\Omega(y(u),x(v)) + \Omega(x(u),y(v)) \\ &= \Omega(u, y(x(v))) - \Omega(u, x(y(v))) \\ &= \Omega(u,[y,x](v)) \\ &= -\Omega(u,[x,y](v))\end{aligned}$$

If we denote $S = \begin{bmatrix} O & I_n \\ -I_n & O \end{bmatrix}$, then in matrix terms, the condition for $X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ (where $A, B, C, D \in \mathfrak{gl}(n, F)$) to be symplectic is that $SX = -X^t S$, i.e., $B^t = B$, $C^t = C$, and $A^t = -D$.

4. Two other families: type B_n and type D_n of **orthogonal Lie algebras** (one for odd dimension and the other for even dimension; see [1] p.3).

5. For an n-dimensional vector space over F, we can fix a basis to see $\mathfrak{gl}(V) \cong \mathfrak{gl}_n(F)$.

   We have **upper triangulars**:

$$\mathfrak{b} = \mathfrak{b}_n(F) = \{x \in \mathfrak{gl}_n(F) \mid x_{ij} = 0 \ \forall i > j\},$$

   and **strictly upper triangulars**:

$$\mathfrak{u} = \mathfrak{u}_n(F) = \{x \in \mathfrak{gl}_n(F) \mid x_{ij} = 0 \ \forall i \geqslant j\},$$

   and **abelian diagonal subalgebra**:

$$\mathfrak{t} = \mathfrak{t}_n(F) = \{x \in \mathfrak{gl}_n(F) \mid x_{ij} = 0 \ \forall i \neq j\}.$$

   It is trivial to check these the brackets are closed for them. Also note that $\mathfrak{b} = \mathfrak{t} \oplus \mathfrak{u}$ (vector space direct sum).

♣

Definition 1.1.6. *Let K be a field, and let A be a vector space over K equipped with an additional binary operation from $A \times A$ to A, denoted here by $\cdot$ (that is, if x and y are any two elements of A, then $x \cdot y$ is an element of A that is called the product of x and y). Then A is an* ***algebra over*** *K if the following identities hold for all elements x, y, z in A, and all elements (often called scalars) a and b in K:*

1. *Right distributivity: $(x + y) \cdot z = x \cdot z + y \cdot z$,*
2. *Left distributivity: $z \cdot (x + y) = z \cdot x + z \cdot y$,*
3. *Compatibility with scalars: $(ax) \cdot (by) = (ab)(x \cdot y)$.*

Remark: Every Lie algebra is an algebra with $x \cdot y = [x, y]$.

Definition 1.1.7. *We say $\delta \in \mathfrak{gl}(A)$ is a* ***derivation*** *for algebra A with bilinear operator $(\cdot,\cdot)$ if $\delta(a,b) = (a, \delta(b)) + (\delta(a), b)$ for every $a, b \in A$.*

Lemma 1.1.8. *Let A be an F-algebra. Then*

$$\operatorname{Der}(A) := \{\delta \in \mathfrak{gl}(A) \mid \delta \text{ is a derivation}\}$$

is a Lie algebra.

Proof. We note that we already have a Lie algebra structure $[x, y] = xy - yx$ for $\mathfrak{gl}(A)$. Thus, we want to show that the above is a Lie subalgebra. We want: $[\delta, \tau] \in \operatorname{Der}(A)\ \forall \delta, \tau \in \operatorname{Der}(A)$. Observe that

$$\begin{aligned}
[\delta,\tau](a\cdot b) &= (\delta\tau - \tau\delta)(a \cdot b) \\
&= \delta(a\cdot\tau(b) + \tau(a)\cdot b) - \tau(a\cdot\delta(b) + \delta(a)\cdot b) \\
&= a\cdot\delta\tau(b) + \underline{\delta(a)\cdot\tau(b) + \tau(a)\cdot\delta(b)} + \delta\tau(a)\cdot b \\
&\quad - \Big(a\cdot\tau\delta(b) + \underline{\tau(a)\cdot\delta(b) + \delta(a)\cdot\tau(b)} + \tau\delta(a)\cdot b\Big) \\
&= a(\delta\tau(b) - \tau\delta(b)) + (\delta\tau(a) - \tau\delta(a))b \\
&= a\cdot[\delta,\tau](b) + [\delta,\tau](a)\cdot b.
\end{aligned}$$

■

Key fact: For each $x \in L$ a Lie algebra with bracket $[\cdot,\cdot]$, the map $\operatorname{ad} x : L \to L$; $(\operatorname{ad} x)y = [x, y]$ is a derivation.

Proof: We check that

$$\begin{aligned}
(\operatorname{ad} x)([y,z]) &= [x,[y,z]] \\
&= -[y,[z,x]] - [z,[x,y]] \quad \text{(Jacobi identity)} \\
&= [y,[x,z]] + [[x,y],z] \\
&= [y,(\operatorname{ad} x)(z)] + [(\operatorname{ad} x)(y), z].
\end{aligned}$$

Key fact: The map $\operatorname{ad} : L \to \operatorname{Der}(L) \subseteq \mathfrak{gl}(L)$; $x \mapsto \operatorname{ad} x$ is a Lie algebra homomorphism, i.e., $\operatorname{ad}([x, y]) = [\operatorname{ad} x, \operatorname{ad} y]$. This is the **adjoint representation**.

1.2 Ideals and Homomorphisms

A subspace $I \subseteq L$ is an **ideal** if $\forall x \in L, y \in I$, we have $[x, y] \in I$. An ideal is of course a Lie subalgebra.

Example:

1. 0, zero ideal; L also an ideal.
2. The **center** $Z(L) = \{z \in L \mid [z, x] = 0\ \forall x \in L\}$.
3. The **derived subalgebra** $[L, L] = F\{[x, y] \mid x, y \in L\}$ (the *span* of the set $\{[x, y] \mid x, y \in L\}$).
4. Suppose $\varphi : L \to L'$ is a homomorphism. Then $\ker(\varphi)$ is an ideal.

Note:

1. L abelian $\iff L = Z(L) \iff [L, L] = 0$.
2. Given an ideal $I \subseteq L$, L/I is a Lie algebra with $[x + I, y + I] = [x, y] + I$, and, as usual, there is a canonical homomorphism $L \to L/I$, $x \mapsto x + I$.

Proposition 1.2.1 (Isomorphism Theorems).

(a) *(First Isomorphism Theorem) If $\varphi : L \to L'$ is a homomorphism of Lie algebras, then $L/\ker(\varphi) \cong \operatorname{Im}\varphi$. If I is any ideal of L included in $\ker(\varphi)$, there exists a unique homomorphism $\psi : L/I \to L'$ making the following diagram commute (π is the canonical map):*

$$\begin{array}{ccc} L & \xrightarrow{\ \varphi\ } & L' \\ {\scriptstyle\pi}\big\downarrow & \nearrow_{\psi} & \\ L/I & & \end{array}$$

(b) *(Second Isomorphism Theorem)If I and J are ideals of L such that $I \subseteq J$, then J/I is an ideal of L/I and $(L/I)/(J/I)$ is naturally isomorphic to L/J.*

(c) *(Third Isomorphism theorem) If I, J are ideals of L, there is a natural isomorphism between $(I + J)/J$ and $I/(I \cap J)$.*

(d) *(Fourth Isomorphism Theorem) Let I be an ideal of L. Then the canonical projection map $\varphi : L \to L/I, \varphi(x) = x + I$ induces a 1-1 correspondence $\Phi : J \mapsto \varphi(J) = J/I$ between ideals of L that contain I and ideals of L/I:*

$$\begin{aligned} \mathcal{I} = \{\text{ideals of } L \text{ that contain } I\} &\longleftrightarrow \mathcal{I}' = \{\text{ideals of } L/I\} \\ \text{an ideal } J \text{ of } L \text{ that contains } I \ &\longrightarrow \text{its image } \varphi(J) = J/I \text{ in } L/I \\ \text{its inverse image } \varphi^{-1}(\mathcal{J}) \text{ in } L &\longleftarrow \ \text{an ideal } \mathcal{J} \text{ of } L/I \end{aligned}$$

Moreover, if we denote J/I by J^*, then:

- *For $J_1, J_2 \in \mathcal{I}$, $J_1 \subseteq J_2$ if and only if $J_1^* \subseteq J_2^*$, and then $\dim(J_2/J_1) = \dim(J_2^*/J_1^*)$;*
- *For $J_1, J_2 \in \mathcal{I}$, J_1 is an ideal of J_2 if and only if J_1^* is an ideal of J_2^*, and then $J_2/J_1 \cong J_2^*/J_1^*$.*

For later use we mention a couple of related notions, analogous to those which arise in group theory. The **normalizer** of a subalgebra (or just subspace) K of L is defined by $N_L(K) = \{x \in L \mid [x, K] \subset K\}$. By the Jacobi identity, $N_L(K)$ is a subalgebra of L; it may be described verbally as the largest subalgebra of L which includes K as an ideal (in case K is a subalgebra to begin with). If $K = N_L(K)$, we call K **self-normalizing**; some important examples of this behavior will emerge later. The **centralizer** of a subset X of L is $C_L(X) = \{x \in L \mid [xX] = 0\}$. Again by the Jacobi identity, $C_L(X)$ is a subalgebra of L. For example, $C_L(L) = Z(L)$.

Definition 1.2.2. *A **representation** of L is a homomorphism $\varphi : L \to \mathfrak{gl}(V)$ for V an F-vector space.*

Example 1.2.3. Consider the adjoint representation $\operatorname{ad} : L \to \mathfrak{gl}(L); x \mapsto \operatorname{ad} x$. Then $\ker(\operatorname{ad}) = \{x \in L \mid \operatorname{ad} x = 0\} = Z(L)$. Thus, $L/Z(L) \cong \operatorname{ad}(L) \subseteq \mathfrak{gl}(L)$. ♣

Definition 1.2.4. *L is called **simple** if L has no ideals except for 0 and itself and if $[L, L] \neq 0$. Recall that a simple group is a group with no nontrivial normal subgroup.*

Remark: L simple thus nonabelian $\implies$ ideal $Z(L) = 0 \implies L \cong \operatorname{ad}(L) \subseteq \mathfrak{gl}(L)$.

Example 1.2.5. Let $\operatorname{char}(F) \neq 2$.

$$\mathfrak{sl}_2(F) = \{A \in \mathfrak{gl}_2(F) \mid \operatorname{tr}(A) = 0\} = \mathbb{C}\left\{ \overbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}^{x}, \overbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}^{h}, \overbrace{\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}}^{y} \right\}.$$

Note that $[h,x]=2x$, $[x,y]=h$, $[h,y]=-2y$. We write the linear transformation $\operatorname{ad} h : \mathfrak{sl}_2(F) \to \mathfrak{sl}_2(F)$ in matrix form with basis $\{x,h,y\}$:

$$\begin{aligned}(\operatorname{ad} h)(x) &= [h,x] = 2x\\ (\operatorname{ad} h)(h) &= [h,h] = 0\\ (\operatorname{ad} h)(y) &= [h,y] = -2y\end{aligned}$$

Thus, the matrix is

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

and x,h,y are eigenvectors for $\operatorname{ad} h$, corresponding to the eigenvalues $2,0,-2$. Since $\operatorname{char}(F) \neq 2$, these eigenvalues are distinct. Direct computation by bracketing will show that $Z(\mathfrak{sl}_2(\mathbb{C})) = 0$ and $[\mathfrak{sl}_2(\mathbb{C}), \mathfrak{sl}_2(\mathbb{C})] = \mathfrak{sl}_2(\mathbb{C})$.

$\mathfrak{sl}_2(F)$ is also a simple Lie algebra: suppose we a nonzero ideal $I \subseteq \mathfrak{sl}_2(\mathbb{C})$. Let $0 \neq ax+by+ch \in I$. Applying $\operatorname{ad} x$ twice, we get $-2bx \in I$, and applying $\operatorname{ad} y$ twice, we get $-2ay \in I$. Therefore, if a or b is nonzero, I contains either y or x ($\operatorname{char}(F) \neq 2$), and then, clearly, $I = L$ follows. On the other hand, if $a=b=0$, then $0 \neq ch \in I$, so $h \in I$, and again $I = L$ follows. ♣

1.3 Solvable and Nilpotent Lie Algebras

Definition 1.3.1. *The* ***derived series of*** L *is*

$$\underbrace{L}_{L^{(0)}} \supseteq \underbrace{[L,L]}_{L^{(1)}} \supseteq \underbrace{[L^{(1)},L^{(1)}]}_{L^{(2)}} \supseteq \cdots .$$

L is said to be ***solvable*** *if $L^{(m)} = 0$ for some $m \geqslant 1$.*

Example 1.3.2.

1. Let L be abelian. Then $L^{(1)} = 0$. Thus, L is solvable.
2. Recall that $\mathfrak{sl}_2(F) = [\mathfrak{sl}_2(F), \mathfrak{sl}_2(F)]$. Thus, $\mathfrak{sl}_2(F)$ is not solvable.
3. $\mathfrak{u}_3 = \left\{\begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}\right\} = F\left\{\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{E_{12}}, \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}}_{E_{23}}, \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{E_{13}}\right\} \subseteq \mathfrak{gl}_3(F)$. The dimensional of general $\mathfrak{u}_n$ is thus $\dim = 1 + \cdots + n = \frac{n(n+1)}{2}$. Note that bracketing the basis gives

$$\begin{aligned}[E_{12}, E_{23}] &= E_{13}\\ [E_{12}, E_{13}] &= 0 \\ [E_{23}, E_{13}] &= 0.\end{aligned} \tag{1.1}$$

Thus, $[\mathfrak{u}_3, \mathfrak{u}_3] = F\{E_{13}\}$. Note that for arbitrary x,y we have $[x,y] = x^1y^2[E_{12},E_{23}] + x^2y^1[E_{23},E_{12}] = (x^1y^2 - x^2y^1)E_{13}$. Then the center $Z(\mathfrak{u}_3)$ is $F\{E_{13}\}$ because $x^1y^2 - x^2y^1 = 0$ for any $y^1, y^2 \in F$ implies $x^1 = x^2 = 0$, i.e., $x = x^3E_{13}$. Then $L^{(2)} = [F\{E_{13}\}, F\{E_{13}\}] = [Z(\mathfrak{u}_3), Z(\mathfrak{u}_3)] = 0$. Hence, $\mathfrak{u}_3$ is solvable for $m = 2$.

4. In general, $\mathfrak{u}_n$ is solvable.

5. $\mathfrak{b}_3 = \left\{ \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix} \right\} \subseteq \mathfrak{gl}_3(F)$. Note that $\mathfrak{b}_3 = \mathfrak{u}_3 + \mathbb{C}\{E_{11}, E_{22}, E_{33}\}$. Recall Left-multiplying a matrix A by E_{ij} results in a matrix where the i-th row is the j-th row of A, and all other rows are zero; right-multiplying a $q \times m$ matrix A by E_{ij} results in a matrix where the j th column is the i-th column of A, and all other columns are zero. Thus,

$$\begin{gathered}\forall i, j,\ [E_{ii}, E_{jj}] = 0, \\ \forall k \neq l,\ [E_{ii}, E_{kl}] = \begin{cases} E_{kl} & , k = i \\ -E_{kl} & , l = i \\ 0 & , \text{otherwise} \end{cases}\end{gathered} \tag{1.2}$$

The remaining brackets of basis element are just (1.1). Now consider $E_{11} - E_{22}$. $E_{11} - E_{22}$ bracketing with diagonal basis elements gives 0. For off-diagonal basis elements, note that $[E_{11} - E_{22}, E_{12}] = 2E_{12}$. $[E_{11} - E_{22}, E_{13}] = E_{13}$. $[E_{11} - E_{22}, E_{23}] = -E_{23}$. Therefore, for any element in $\mathfrak{b}_3$, we already analyzed the bracket of each component of it with $E_{11} - E_{22}$, each resulting either zero or a multiple of a distinct basis element. Thus, no nontrivial element commutes with $E_{11} - E_{22}$. Thus, there are no nontrivial elements commuting with every element. Thus, $Z(\mathfrak{b}_3) = 0$. Also, $[\mathfrak{b}_3, \mathfrak{b}_3] = \mathbb{C}\{E_{12}, E_{23}, E_{13}\} = \mathfrak{u}_3$. Hence, $\mathfrak{b}_3$ is solvable.

6. In general, $\mathfrak{b}_n$ is solvable.

♣

Goal: $\text{Char}(F) = 0$ and F is algebraically closed. We will show that L is solvable $\iff$ $L/Z(L)$ is a Lie subalgebra of $\mathfrak{b}_n$ for some n.

Proposition 1.3.3. *Since an ideal is a Lie subalgebra, an ideal that is also solvable as a Lie algebra is called a solvable ideal. We have,*

1. *L is solvable $\implies$ all subalgebras and homomorphic images of L are solvable.*
2. *If $I \subseteq L$ is a solvable ideal and L/I is solvable, then L is solvable.*
3. *If $I, J \subseteq L$ are solvable ideals, then $I + J$ is solvable.*

Proof. Routine. See [1] p.11. ■

Corollary 1.3.4. $\exists!$ *maximal solvable ideal $I \subseteq L$. This maximal ideal is the* ***radical*** *of L, denoted as* $\text{Rad}(L)$.

Proof. L has finite dimension. Let $I \subseteq L$ be a solvable ideal with the largest dimension possible. Let $J \subseteq L$ be another solvable ideal. Then proposition 1.3.3 says $I + J$ is solvable. Since $I \subseteq I + J$, we see that $I = I + J$ because I has the maximum dimension. Thus, $J \subseteq I$. ■

Definition 1.3.5. *L is* ***semisimple*** *if it has no non-trivial solvable ideals, i.e.,* $\text{Rad}(L) = 0$.

Example 1.3.6. $\mathfrak{sl}_2(F)$ is semisimple (because $\mathfrak{sl}_2(F)$ is simple and is not solvable). ♣

Proposition 1.3.7. $L/\text{Rad}(L)$ *is semisimple.*

Proof. Let J be a solvable ideal of $L/\operatorname{rad}(L)$ and $\varphi : L \to L/\operatorname{rad}(L)$ the canonical projection. Then $\varphi^{-1}(J)/\text{Rad}(L) \xlongequal{\text{4th iso thm}} J$ solvable + $\text{Rad}(L)$ solvable $\overset{\text{prop.1.3.3}}{\Longrightarrow}$ $\varphi^{-1}(J)$ solvable in L. Thus, $\varphi^{-1}(L) \subseteq \operatorname{rad}(L) \implies J = \varphi(\varphi^{-1}(J)) \subseteq \varphi(\operatorname{rad}(L)) = 0$ in L. Thus, J must be zero ideal in $L/\operatorname{rad}(L)$. ■

Thus we have shown that any Lie algebra L contains a canonical solvable ideal $\operatorname{rad}(L)$ such that $L/\operatorname{rad}(L)$ is a semisimple Lie algebra. We thus have an exact sequence

$$0 \longrightarrow \operatorname{rad}(L) \longrightarrow L \longrightarrow L/\operatorname{rad}(L) \longrightarrow 0$$

so that, in some sense at least, every finite dimensional Lie algebra is ”built up” out of a semisimple Lie algebra and a solvable one. Slightly more precisely, if

$$0 \longrightarrow L_1 \longrightarrow L \longrightarrow L_2 \longrightarrow 0$$

is an exact sequence of Lie algebras, we say that L is an **extension** of L_2 by L_1. Thus the previous proposition can be rephrased as saying that any Lie algebra is an extension of the semisimple Lie algebra $L/\operatorname{rad}(L)$ by the solvable Lie algebra $\operatorname{rad}(L)$.

Proposition 1.3.8. *A finite dimensional Lie algebra L is semisimple if and only if it does not contain any non-zero abelian ideals.*

Proof. Clearly if L contains an abelian ideal it contains a solvable ideal, so that $\operatorname{rad}(L) \neq 0$. Conversely, if K is a non-zero solvable ideal in L, then the last term $K^{(m-1)}$ in the derived series of K will be a nonzero abelian ideal of L (obviously, a lie algebra $\mathfrak{g}$ is abelian $\iff [\mathfrak{g}, \mathfrak{g}] = 0$). ■

Remark 1.3.9. Simplicity implies semisimplicity. If L is simple, then $[L, L]$, which is nonzero and is an ideal, has to be the whole of L. Thus, L cannot be solvable. Therefore, $\operatorname{Rad}(L) = 0$, and L is semisimple. ♠

Definition 1.3.10. *The **lower central series** (or **descending central series**) of L is:*

$$\underbrace{L}_{L^0} \supseteq \underbrace{[L, L^0]}_{L^1} \supseteq \underbrace{[L, L^1]}_{L^2} \supseteq \cdots.$$

where

$$\forall i \geqslant 2, \quad L^i := [L, L^{i-1}].$$

*A Lie algebra L is **nilpotent** if $L^m = 0$ for some $m \geqslant 1$.*

Remark 1.3.11. $L^1 = L^{(1)}$, thus it's easy to see by induction that $L^{(i)} \subseteq L^i$, so all nilpotent Lie algebras are solvable. ♠

Example 1.3.12. Continue with example 1.3.2,

1. L abelian $\implies$ L nilpotent.
2. $\mathfrak{u}_3 \supseteq [\mathfrak{u}_3, \mathfrak{u}_3] = F\{E_{13}\} \supseteq [\mathfrak{u}_3, F\{E_{13}\}] \overset{(1.1)}{=\!=} 0 \implies \mathfrak{u}_3$ is nilpotent.
3. For the same reason, $\mathfrak{u}_n$ is nilpotent.
4. $\mathfrak{b}_3 \supseteq [\mathfrak{b}_3, \mathfrak{b}_3] = \mathfrak{u}_3 \supseteq [\mathfrak{b}_3, \mathfrak{u}_3] \overset{(1.2)}{=\!=} \mathfrak{u}_3 \supseteq \mathfrak{u}_3 \supseteq \cdots \implies \mathfrak{b}_3$ is not nilpotent.
5. For the same reason, $\mathfrak{b}_n$ is not nilpotent.

♣

Proposition 1.3.13. *Let L be a Lie algebra.*

(a) If L is nilpotent, then so are all subalgebras and homomorphic images of L.

(b) If $L/Z(L)$ is nilpotent, then so is L.

(c) If L is nilpotent and nonzero, then $Z(L) \neq 0$.

Proof. See [1] p.12. ■

Remark 1.3.14. The condition for L to be nilpotent can be rephrased as follows:

For some m (depending only on L), $(\operatorname{ad} x_1 \operatorname{ad} x_2 \cdots \operatorname{ad} x_m)(y) = 0$ for all $x_i, y \in L$.

In particular, this condition implies $(\operatorname{ad} x)^m = 0$ for all $x \in L$. Now if L is any Lie algebra, and $x \in L$, we call x **ad-nilpotent** if $\operatorname{ad} x$ is a **nilpotent**, i.e. $(\operatorname{ad} x)^k = 0$ for some $k > 0$. Thus, we see ♠

Remark 1.3.15. If L is nilpotent, then all elements of L are ad-nilpotent. ♠

It is a pleasant surprise to find that the converse is also true.

Theorem 1.3.16 (Engel)**.** *If all elements of L are ad-nilpotent, then L is nilpotent.*

Lemma 1.3.17. *Let $x \in \mathfrak{gl}(V)$ be nilpotent ($x^r = 0$ for sme $r > 0$), then $\operatorname{ad} x$ is also nilpotent.*

Proof. $(\operatorname{ad} x)(y) = [x, y] = xy - yx$ and

$$(\operatorname{ad} x)^2 y = [x, [x, y]] = [x, xy - yx] = x^2 y - 2xyx + yx^2$$

By induction, we will have

$$(\operatorname{ad} x)^m (y) = \sum_{k=0}^{m} c_k x^k y x^{m-k}$$

Thus, the r such that $x^r = 0$ gives rise to some $m = 2r$ such that each term contains x^r and thus $(\operatorname{ad} x)^m = 0$. ■

Remark 1.3.18 (Converse is not true)**.** A word of warning: It is easy for a matrix to be ad-nilpotent in $\mathfrak{gl}(n, F)$ without being nilpotent. (The identity matrix is an example.) It should be kept in mind two contrasting types of nilpotent linear Lie algebras: $\mathfrak{b}(n, F)$ and $\mathfrak{u}(n, F)$. ♠

Engel's Theorem will be deduced from the following result, which is of interest in its own right.

Recall that a single nilpotent linear transformation always has at least one nonzero eigenvector, corresponding to its unique eigenvalue 0: $Ax = \lambda x \implies A^r x = \lambda^r x = 0 \implies \lambda^r = 0, \lambda = 0$. To show it has a nonzero eigenvector, i.e., $\exists 0 \neq x$ s.t. $Ax = 0$ is to show it is singular, but $A^r = 0 \implies \det(A)^r = 0 \implies \det(A) = 0$ (entries of the matrix are from field and thus integral domain without zero divisor, i.e., $ab = 0 \implies a = 0$ or $b = 0$).

This is just the case $\dim L = 1$ of the following theorem.

Theorem 1.3.19. *Let L be a subalgebra of $\mathfrak{gl}(V)$, V finite dimensional. If every element of L is nilpotent and $V \neq 0$, then $\exists 0 \neq w \in V$ such that $x(w) = 0 \; \forall x \in L$, or $L.w = 0$.*

Proof. Induction on $\dim L$.

$\dim L = 1$: $L = F\{x\}$. We have just shown this above.

$\dim L > 1$: Let $K \subseteq L$ be a maximal proper subalgebra.

claim: $\dim K = \dim L - 1$ and K is an ideal of L.
proof: let $\overline{L} = L/K$ and consider $\varphi : K \to \mathfrak{gl}(\overline{L})$; $y \mapsto \operatorname{ad}(y)$. That is, for $y \in K$, we have $\varphi(y)(x + K) = [y, x] + K$. Now, φ is a homomorphism, i.e., $\varphi([y, z]) = [\varphi(y), \varphi(z)]$, by Jacobi identity. Thus $\varphi(K)$ is a Lie subalgebra of $\mathfrak{gl}(\overline{L})$. Also, $\dim \varphi(K) \leqslant \dim(K) < \dim(L)$. Furthermore, the Lemma implies that every element of $\varphi(K)$ is nilpotent. Thus we can apply induction hypothesis to $\varphi(K) \subseteq \mathfrak{gl}(\overline{L})$. Then $\exists K \neq (z + K) \in \overline{L}$ (i.e., $z \notin K$) such that $\varphi(K).(z + K) = 0$. This implies $\forall y \in K$, $\operatorname{ad}(y)(z + K) = [y, z] + K = 0$ (i.e., $[y, z] \in K$). Thus, we have $z \in K$ such that $\forall y \in K$, $[y, z] \in K$. Thus, K is properly contained in the normalizer $N_L(K)$, which is also a subalgebra. Thus, by maximality, it has to be the case that $N_L(K) = L$. That is, $\forall x \in L$, $[x, K] \subset K \implies K$ is an ideal.
We then show $\dim K = \dim L - 1$ and K. If $\dim L/K$ were greater than one, then the inverse image in

L of a one dimensional subalgebra of L/K (which always exists) would be a proper subalgebra properly containing K, which is absurd; therefore, K has codimension one. This allows us to write $L = K + F\{z\}$ for any $z \in L - K$.//

Now we show L with all ele. nilp has some $w \in V$ s.t. $L.w = 0$. Consider the subspace of V killed by K, i.e., $W = \{v \in V | y(v) = 0 \; \forall y \in K\} \subseteq V$. Induction hypothesis on K ($\dim K = \dim L - 1 < \dim L$ and all elements of L thus K nilpotent) $\implies W \neq 0$. Let $x \in L$ and $y \in K$ and $w \in W$. Then

$$(yx)(w) = \underbrace{(xy)(w)}_{=0} - \underbrace{[x,y]}_{\in K}(w) = 0 \implies \forall x \in L \text{ and } w \in W, \; xw \in W$$

That is, W is an L-stable subspace of V. In particular, for that $z \in L - K$, we have $z(W) \subseteq W$. Thus, $z|_W$ is also nilpotent. Then $\exists 0 \neq w \in W$, s.t., $z(w) = 0$. Given $x \in L = K \oplus F\{z\}$, we can write $x = y + az$ for some $y \in K$, $a \in F$. Now,

$$x(w) = y(w) + az(w) = 0$$

■

proof of Engel's theorem. We want to show that if every $x \in L$ is ad-nilp then L is nilp. We proceed by induction on dimension of L.

If the dimension is 1, then ad is trivial and $\mathrm{ad}(L)$ would trivially be nilp. Let $\dim(L) > 1$. Apply above theorem to $\mathrm{ad}(L) \subseteq \mathfrak{gl}(L)$ to get some nonzero $x \in L$ such that $[L, x] \neq 0$, which implies $x \in Z(L)$. Recall from example 1.2.3 that $\mathrm{ad}(L) = L/Z(L)$. Now, $\dim \mathrm{ad}(L) = \dim L/Z(L) < \dim(L)$. The induction hypothesis gives $\mathrm{ad}(L)$ nilpotent. Proposition 1.3.3 (2) then implies L is nilpotent. ■

Definition 1.3.20. *Let V be a finite-dimensional vector space over a field F with $\dim V = n$. A* ***flag*** *in V is a chain of subspaces:*

$$0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_n = V$$

such that $\dim V_i = i$ for each $i = 0, 1, \ldots, n$. An endomorphism $x \in \mathrm{End}(V)$ is said to ***stabilize*** *or* ***preserve*** *the flag if*

$$x(V_i) \subseteq V_i \quad \text{for all } i.$$

Corollary 1.3.21. *Let L be a subalgebra of $\mathfrak{gl}(V)$, V finite dimensional. If every element of L is nilpotent and $V \neq 0$, then there exists a flag (V_i) in V stable under the action of L, i.e., for all i, $x(V_i) \subseteq V_{i-1}$ for all $x \in L$. In other words, there exists a basis of V relative to which the matrices of L are all in $\mathfrak{u}(n, F)$.*

Proof. Begin with any nonzero $v \in V$ killed by L, whose existence is assured by above theorem. Set $V_1 = Fv$. Let $W = V/V_1$, and observe that the induced action of L on W is also by nilpotent endomorphisms. By induction on $\dim V$, W has a flag stabilized by L, whose inverse image in V does the trick.

The action on W induced by $x : V \to V$ is $x'(a + V_1) = x(a) + V_1$.

If the Lie algebra L stabilizes a flag, then in a basis adapted to this flag (i.e., where the basis vectors span the successive subspaces V_i in the flag), the action of any element $x \in L$ will map each basis vector to a linear combination of basis vectors that correspond to smaller subspaces in the flag. This ensures that the matrix representation of x in this basis will be upper triangular. Conversely, if the elements of L are represented by upper triangular matrices in some basis, then these matrices stabilize the subspaces spanned by the first i basis vectors. This gives a flag that is preserved by all elements of L, where each subspace is spanned by a certain number of basis vectors, corresponding to the positions of non-zero entries in the upper triangular matrix. ■

Lemma 1.3.22. *Let L be nilpotent, and let K be an ideal of L. If $K \neq 0$, then $K \cap Z(L) \neq 0$ (In particular, $Z(L) \neq 0$).*

Proof. L acts on K via the adjoint representation, so above theorem yields nonzero $x \in K$ killed by L, i.e., $[Lx] = 0$, so $x \in K \cap Z(L)$. ■

Chapter 2

Semisimple Lie Algebras

Assume $\operatorname{char} F = 0$ and F algebraically closed.

2.1 Structure of Solvable Lie Algebras

2.1.1 Lie's theorem

Theorem 1.3.19 asserts the existence of a common eigenvector for a Lie algebra consisting of nilpotent endomorphisms. The next theorem is of similar nature.

Theorem 2.1.1 (Lie's Theorem). *Suppose $L \subseteq \mathfrak{gl}(V)$ is a solvable subalgebra. $\dim(V) < \infty$ and $V \neq 0$. Then V contains a common nonzero eigenvector for all $x \in L$.*

Proof. Use induction on $\dim L$, the case $\dim L = 0$ being trivial. When $\dim L = 1$, the eigenvector exists for any matrix with entries in algebraically closed field. We attempt to imitate the proof of Theorem 1.3.19. The idea is
(1) to locate an ideal K of codimension one,
(2) to show by induction that common eigenvectors exist for K,
(3) to verify that L stabilizes a space consisting of such eigenvectors, and
(4) to find in that space an eigenvector for a single $z \in L$ satisfying $L = K + F\{z\}$.

Step (1) is easy. Since L is solvable, of positive dimension, L properly includes $[L, L]$. $L/[L, L]$ being abelian, any subspace is automatically an ideal. Using Fourth isomorphism theorem to take a subspace K' of codimension one, then its inverse image $K = \varphi^{-1}(K')$ is an ideal of codimension one (easy to compute) in L, containing $[L, L]$.

For step (2), to apply induction to K, we verify that K has a dimension lower than L and that K is solvable: if $K = 0$, then L is abelian of dimension 1 and an eigenvector for a basis vector of L finishes the proof; so for $K \neq 0$, it is solvable as a subalgebra of solvable L (using Proposition 1.3.3 (1)). Induction gives an common eigenvector v so that $y \in K$, $y(v) = \lambda(y)v$ for some linear function $\lambda : K \to F$. Fix this λ, and denote by W_λ the subspace

$$\{w \in V \mid y(w) = \lambda(y)w, \text{ for all } y \in K\}; \text{ so } v \in W_\lambda,\ W_\lambda \neq 0.$$

Step (3) consists in showing that L leaves W_λ invariant. Assuming for the moment that this is done, proceed to step (4): Write $L = K + F\{z\}$ for a $z \in L - K$, and use the fact that F is algebraically closed to find an eigenvector $v_0 \in W_\lambda$ of z now acting on W_λ (for some eigenvalue of z). Then v_0 is obviously a common eigenvector for L (and λ can be extended to a linear function on L such that $x(v_0) = \lambda(x)v_0$, $x \in L$).

It remains to show that L stabilizes W_λ. Let $w \in W_\lambda, x \in L$. To test whether or not $x(w)$ lies in W_λ, we must take arbitrary $y \in K$ and examine $yx(w) = xy(w) - [x,y](w) = \lambda(y)x(w) - \lambda([x,y])w$. Thus we have to prove that $\lambda([x,y]) = 0$. For this, fix $w \in W_\lambda, x \in L$. Let $n > 0$ be the smallest integer for which $w, x(w), \cdots, x^n(w)$ are linearly dependent. Let W_i be the subspace of V spanned by $w, x(w), \ldots, x^{i-1}(w)$. (set $W_0 = 0$), so $\dim W_n = n, W_n = W_{n+i} (i \geqslant 0)$ and x maps W_n into W_n. It is easy to check that each $y \in K$ leaves each W_i invariant. Relative to the basis $w, x.w, \ldots, x^{n-1}.w$ of W_n, we claim that $y \in K$ is represented by an upper triangular matrix whose diagonal entries equal $\lambda(y)$. This follows immediately from the congruence:

$$(*) \qquad yx^i.w \equiv \lambda(y)x^i.w \,(\operatorname{mod} W_i)$$

which we prove by induction on i, the case $i = 0$ being obvious. Write $yx^i(w) = yxx^{i-1}(w) = xyx^{i-1}(w) - [x,y]x^{i-1}(w)$. By induction, $yx^{i-1}(w) = \lambda(y)x^{i-1}(w) + w' \,(w' \in W_{i-1})$; since x maps W_{i-1} into W_i (by construction), $(*)$ therefore holds for all i.

According to our description of the way in which $y \in K$ acts on W_n, $\operatorname{tr}_{W_n}(y) = n\lambda(y)$. In particular, this is true for elements of K of the special form $[x,y]$ (x as above, y in K). But x, y both stabilize W_n, so $[x,y]$ acts on W_n as the commutator of two endomorphisms of W_n; its trace is therefore 0. We conclude that $n\lambda([x,y]) = 0$. Since char $F = 0$, this forces $\lambda([x,y]) = 0$, as required. ■

Corollary 2.1.2. *Let L be a solvable subalgebra of $\mathfrak{gl}(V)$, $\dim V = n < \infty$. Then L solvable $\iff$ $L \subseteq \mathfrak{b}$.*

Proof. $\implies$: $\mathfrak{b}$ is solvable.

$\impliedby$: this is the same as saying L stabilizes some flag in V. We do this by induction on $\dim(V) = k$. Let its basis be $\{v_1, \cdots, v_k\}$. Let $V' := V/F\{v_1, \cdots, v_k\}$. Note that the induced $L' \subseteq \mathfrak{gl}(V')$ is solvable. Due to Lie's theorem, $\exists 0 \neq v_{k+1} \in V'$ s.t. $\forall x \in L'$, $x(v_{k+1}) \subseteq F\{v_{k+1}\}$. Then pull this v_{k+1} back to V to form another basis for vector space of dimension $\kappa + 1$. ■

More generally, let L be any solvable Lie algebra, $\phi : L \to \mathfrak{gl}(V)$ a finite dimensional representation of L. Then $\phi(L)$ is solvable, by Proposition 1.3.3, hence stabilizes a flag (Corollary above). For example, if ϕ is the adjoint representation, a flag of subspaces stable under L is just a chain of ideals of L, each of codimension one in the next. We can also translate this to upper triangular matrices. We have:

Corollary 2.1.3. *A Lie algebra L is solvable $\iff$ there exists a chain of ideals of L, $0 = L_0 \subset L_1 \subset \ldots \subset L_n = L$, such that $\dim L_i = i \iff L/Z(L) \cong \operatorname{ad}(L)$ is solvable in $\mathfrak{b}$.*

Corollary 2.1.4. *A Lie algebra L is solvable $\iff$ $[L,L]$ is nilpotent.*

Proof. $\impliedby$: $[L,L]$ is nilpotent $\overset{\text{Fact1.3.11}}{\implies}$ $[L,L]$ solvable. Since $L/[L,L]$ is abelian and thus solvable, proposition 1.3.3 concludes.

$\implies$: we show $x \in [L,L]$ implies that $d_L x$ is nilpotent. Find a flag of ideals as in Corollary above. Relative to a basis $(x_1, \ldots, x_n)$ of L for which $(x_1, \ldots, x_i)$ spans L_i, the matrices of ad L lie in $\mathfrak{b}(n,F)$. Therefore the matrices of $[\operatorname{ad} L, \operatorname{ad} L] = \operatorname{ad}_L[L,L]$ lie in $\mathfrak{u}(n,F)$, the derived algebra of $\mathfrak{b}(n,F)$. It follows that $\operatorname{ad}_L x$ is nilpotent for $x \in [L,L]$; a fortiori $\operatorname{ad}_{[LL]} x$ is nilpotent, so $[L,L]$ is nilpotent by Engel's Theorem. ■

2.1.2 Jordan-Chevalley Decomposition

Read Steven Roman's *Advanced Linear Algebra* for Jordan canonical form in a more general setting.

Let F be an algebraically closed field. A general matrix in **Jordan canonical form** looks like

$$\begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_r \end{pmatrix}$$

where each A_i is a Jordan block matrix $J_t(\lambda)$ for some $t \in \mathbb{N}$ and $\lambda \in F$:

$$J_t(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}_{t\times t}$$

Our linear algebra says that every endomorphism x can be written into a Jordan canonical form with a certain basis. That is, the matrix M of x is of the form $M = SJS^{-1}$ for some invertible matrix S and its unique Jordan canonical form J. Note that $J = D + N$ where $D = \operatorname{diag}(\underbrace{\lambda_1, \cdots, \lambda_1}, \cdots, \underbrace{\lambda_r, \cdots, \lambda_r})$ and D is a matrix consisting of shift matrices as blocks. The shift matrices are simply of the form $J_t(\lambda) - \operatorname{diag}(\lambda, \cdots, \lambda)$. Thus, x can be written as a sum of a diagonal matrix and a nilpotent matrix which commute. We can make this decomposition more precise.

The **minimal polynomial** of $x \in \mathfrak{gl}(V) \cong \mathfrak{gl}(n, F)$ is the monic polynomial $p(t) \in F(t)$ with minimal degree such that $p(x) = 0$.

Lemma 2.1.5. *Let $x \in \mathfrak{gl}(n, F)$ and F an algebraically closed field. The following are equivalent:*

(1) There exists invertible g such that gxg^{-1} is a diagonal matrix, i.e., x is diagonalizable.

(2) There exists a basis $\{v_i\}$ for F^n consisting of eigenvectors for x.

(3) The minimal polynomial has distinct roots, i.e., $f(t) = \prod_i (t - \lambda_i)$ with $\lambda_i \neq \lambda_j$ for $i \neq j$.

Proof. See [3] Corollary B.1.2 for $F = \mathbb{C}$. ■

We call x **semisimple** if it satisfies any of the above three conditions.

Remark 2.1.6. We remark that two commuting semisimple endomorphisms $x, y \in \mathfrak{gl}(V)$ can be simultaneously diagonalized; therefore, their sum or difference is again semisimple. The sum or difference of two commuting nilpotent endomorphisms is nilpotent as well.

It is an exercise to show that their (additive) Jordan-Chevalley decomposition is indeed additive: $(x + y)_n = x_n + y_n$, $(x + y)_s = x_s + y_s$.

Also, if x is semisimple and maps a subspace W of V into itself, then obviously the restriction of x to W is semisimple. ♠

Proposition 2.1.7. *Let V be a finite dimensional vector space over F, $x \in \mathfrak{gl}(V)$.*

(a) There exist unique $x_s, x_n \in$ End V satisfying the conditions: $x = x_s + x_n$, x_s is semisimple, x_n is nilpotent, x_s and x_n commute.

(b) There exist polynomials $p(t), q(t) \in F(t)$, without constant term, i.e, $p(0) = q(0) = 0$, such that $p(x) = x_s, q(x) = x_n$. In particular, x_s and x_n commute with any endomorphism commuting with x.

The decomposition $x = x_s + x_n$ is called the (additive) ***Jordan-Chevalley decomposition*** *of x, or just the* ***Jordan decomposition****; x_s, x_n are called (respectively) the* ***semisimple part*** *and the* ***nilpotent part*** *of x.*

Example 2.1.8.

$$x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$x^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{bmatrix} \quad \begin{aligned} p(t) &= \frac{1}{2}t^2 \\ q(t) &= t - \frac{1}{2}t^2 \end{aligned}$$

♣

Corollary 2.1.9 (Corollary of second part of the proposition). *(a) If $y \in \mathfrak{gl}(V)$ such that $[x, y] = 0$, then $[x_s, y] = [x_n, y] = 0$.*

(b) If $A \subset B \subset V$ are subspaces, and x maps B into A, then x_s and x_n also map B into A.

Proof of the proposition. Let $a_1, \ldots, a_k$ (with multiplicities $m_1, \ldots, m_k$) be the distinct eigenvalues of x, so the characteristic polynomial is $\Pi\left(t-a_i\right)^{m_i}$. If $V_i = \operatorname{Ker}\left(x - a_i.1\right)^{m_i}$, then V is the direct sum of the subspaces $V_1, \ldots, V_k$, each stable under x. On V_i, x clearly has characteristic polynomial $\left(t-a_i\right)^{m_i}$. Now apply the Chinese Remainder Theorem (for the ring $F(t)$) to locate a polynomial $p(t)$ satisfying the congruences, with pairwise relatively prime moduli: $p(t) \equiv a_i\left(\bmod \left(t-a_i\right)^{m_i}\right), p(t) \equiv 0(\bmod t)$. (Notice that the last congruence is superfluous if 0 is an eigenvalue of x, while otherwise t is relatively prime to the other moduli.) Set $q(t) = t - p(t)$. Evidently each of $p(t), q(t)$ has zero constant term, since $p(t) \equiv 0(\bmod t)$.

Set $x_s = p(x), x_n = q(x)$. Since they are polynomials in x, x_s and x_n commute with each other, as well as with all endomorphisms which commute with x. They also stabilize all subspaces of V stabilized by x, in particular the V_i. The congruence $p(t) \equiv a_i\left(\bmod \left(t-a_i\right)^{m_i}\right)$ shows that the restriction of $x_s - a_i.1$ to V_i is zero for all i, hence that x_s acts diagonally on V_i with single eigenvalue a_i. By definition, $x_n = x - x_s$, which makes it clear that x_n is nilpotent.

It remains only to prove the uniqueness assertion in (a). Let $x = s + n$ be another such decomposition, so we have $x_s - s = n - x_n$. Because of (b), all endomorphisms in sight commute. Sums of commuting semisimple (resp. nilpotent) endomorphisms are again semisimple (resp. nilpotent), whereas only 0 can be both semisimple and nilpotent. This forces $s = x_s, n = x_n$ ∎

Remark 2.1.10. In proving the Engel's theorem, we have shown the lemma that

$$x \in \mathfrak{gl}(V) \text{ nilpotent implies } \operatorname{ad} x \text{ nilpotent.}$$

It can also be shown that

$$x \in \mathfrak{gl}(V) \text{ semisimple implies } \operatorname{ad} x \text{ semisimple.}$$

Since x_s is semisimple, we by lemma 2.1.5 find a basis $\{v_1, \cdots, v_n\}$ of V under which the matrix of x_s is $\operatorname{diag}(\lambda_1, \cdots, \lambda_n)$ and $x_s e_i = \lambda_i e_i$. Let $\{E_{kl}\}$ be the standard basis of $\mathfrak{gl}(V)$. Then

$$\begin{aligned} (\operatorname{ad} x_s)(E_{kl}) &= \left[\sum_{i=1}^n \lambda_i E_{ii}, E_{kl}\right] = \sum_{i=1}^n \lambda_i [E_{ii}, E_{kl}] \\ &\overset{(1.2)}{=\!=} \lambda_k E_{kl} - \lambda_l E_{kl} = (\lambda_k - \lambda_l) E_{kl}. \end{aligned}$$

Thus, the matrix of $\operatorname{ad} x_s$ is diagonal. ♠

Corollary 2.1.11. *Let $x \in \mathfrak{gl}(V)$ ($\dim V < \infty$) with Jordan decomposition $x = x_s + x_n$. Then $\operatorname{ad} x = \operatorname{ad} x_s + \operatorname{ad} x_n$ is the Jordan decomposition of $\operatorname{ad} x$ (in $\mathfrak{gl}(\mathfrak{gl}(V))$).*

Proof. Due to the remark above, $\operatorname{ad} x_n$ is nilpotent and $\operatorname{ad} x_s$ is semisimple. x_n, x_s commute due to proposition 2.1.7 (a), i.e., $[x_n, x_s] = 0$. Thus, $[\operatorname{ad} x_n, \operatorname{ad} x_s] = \operatorname{ad}[x_n, x_s] = 0 \implies \operatorname{ad} x_n, \operatorname{ad} x_s$ commute. Thus, the uniqueness of proposition 2.1.7 (a) concludes. ∎

2.1.3 Cartan's Criterion

Remark 2.1.12. We collect some facts.

1. L is solvable $\overset{\text{Cor.2.1.4}}{\iff}$ $[L, L]$ is nilpotent $\overset{\text{Engel}}{\iff}$ $\operatorname{ad} y$ nilp, $\forall y \in [L, L]$.
2. $\operatorname{tr}(xy) = \operatorname{tr}(yx)$.
   $\operatorname{tr}(\mathrm{AB}) = (ab)_{ii} = a_{ij}b_{ji} = b_{ji}a_{ij} = (ba)_{jj} = \operatorname{tr}(\mathrm{BA})$
3. Similar matrices have same trace: $\operatorname{tr}(A) = \operatorname{tr}(SBS^{-1}) = \operatorname{tr}((S^{-1})(SB)) = \operatorname{tr}(B)$.
4. $\operatorname{tr}([x, y]z) = \operatorname{tr}(x[y, z])$.
   To show this, we write $[x, y]z = xyz - yxz$ and $x[y, z] = xyz - xzy$ and use $\operatorname{tr}(y(xz)) = \operatorname{tr}((xz)y)$ due to the second fact.

♠

Proposition 2.1.13. *If L is solvable, then $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$, $\forall x \in L$, $\forall y \in [L, L]$.*

Proof. We first observe a fact:
If $x \in \mathfrak{b}$ and $y \in \mathfrak{u}$, then $xy \in \mathfrak{u}$: $\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix}\begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}$ Thus, $\operatorname{tr}(xy) = 0$.

Now L solvable.
$\Longrightarrow$ the homomorphic image $\operatorname{ad}(L) \subseteq \mathfrak{gl}(L)$ is solvable & $[L, L]$ is nilpotent.
$\xRightarrow[\text{Cor.2.1.2}]{\text{rmk.1.3.15 \& Cor.1.3.21}}$ there is a basis of L s.t. the matrices of elements of $\operatorname{ad}(L)$ are upper-triangular in this basis & every ele. of $[L, L]$ is nilp, i.e., every ele. of $\operatorname{ad}([L, L])$ is nilp. and thus matrices of elements of $\operatorname{ad}([L, L])$ are strictly upper-triangular.
$\xRightarrow{\text{above fact}}$ $\operatorname{tr}(\operatorname{ad} x, \operatorname{ad} y) = 0$. ■

Example 2.1.14. Recall the solvable Lie algebra

$$\mathfrak{sl}_2(\mathbb{C}) = \{A \in \mathfrak{gl}_2(\mathbb{C}) \mid \operatorname{tr}(A) = 0\} = \mathbb{C}\left\{ \overbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}^{x}, \overbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}^{h}, \overbrace{\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}}^{y} \right\}.$$

where $[h, x] = 2x$, $[x, y] = h$, $[h, y] = -2y$ and

$$\operatorname{ad} x = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \operatorname{ad} y = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \operatorname{ad} h = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Also $[\mathfrak{sl}_2(\mathbb{C}), \mathfrak{sl}_2(\mathbb{C})] = \mathfrak{sl}_2(\mathbb{C})$. We can easily verify above proposition that $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$, $\forall x \in L$, $\forall y \in [L, L]$. ♣

We let $F = \mathbb{C}$ for simplicity now.

Proposition 2.1.15. *Let L be a subalgebra of $\mathfrak{gl}(V)$, $\dim V = n < \infty$. If $\operatorname{tr}(xy) = 0$ for every $x, y \in L$, then L is solvable.*

Proof. Note that every $x \in [L, L]$ is nilpotent $\xRightarrow{\text{lem.1.3.17}}$ every $\operatorname{ad} x$ is nilpotent $\xRightarrow{\text{rmk.2.1.12}}$ L solvable.

Thus, we show every $x \in [L, L]$ is nilpotent. Let $x = x_s + x_n$ be the Jordan decomposition of x. We need $x_s = 0$. We fix a basis of V in which x_s is diagonal, i.e., $x_s = \operatorname{diag}(\lambda_1, \cdots, \lambda_m)$, and x_n is strictly upper triangular, i.e., a matrix with 1 or 0 just above the diagonal and all other entries zero. Since x_n is strictly upper triangular, we see $\operatorname{tr}(dx_n) = \operatorname{tr}(x_n d) = 0$ for any diagonal matrix d.

We want to show $\sum_{i=1}^m \lambda_i\bar{\lambda}_i = \sum_{i=1}^m |\lambda_i|^2 = 0$ because this implies $\lambda_i = 0\ \forall i \implies x_s = 0$.

Let $\overline{x}_s = \operatorname{diag}(\bar{\lambda}_1, \cdots, \bar{\lambda}_m)$. We note that $\overline{x}_s \in \mathfrak{gl}(V)$ is not necessarily in L. We compute

$$\operatorname{tr}(\overline{x}_s x) = \operatorname{tr}(\overline{x}_s(x_s + x_n)) = \operatorname{tr}(\overline{x}_s x_s) + \underbrace{\operatorname{tr}(\overline{x}_s x_n)}_{=0} = \sum_{i=1}^m \lambda_i\bar{\lambda}_i.$$

Now, as $x \in [L, L]$, we may express x as a linear combination of commutators $[y, z]$ with $y, z \in L$, so we need to show that $\operatorname{tr}(\overline{x}_s[y, z]) \overset{\text{rmk.2.1.12(4)}}{=\!=\!=} \operatorname{tr}([\overline{x}_s, y]z) = 0$. This will hold by our hypothesis, provided we can show that $[\overline{x}_s, y] \in L$. In other words, we must show that $\operatorname{ad}\overline{x}_s$ maps L into L.

By Cor. 2.1.11, the Jordan decomposition of $\operatorname{ad} x$ is $\operatorname{ad} x_s + \operatorname{ad} x_n$. Therefore, by part (b) of [4] Lemma 16.8, there is a polynomial $p(t) \in \mathbb{C}[t]$ such that $p(\operatorname{ad} x) = \overline{\operatorname{ad} x_s} = \operatorname{ad}\overline{x}_s$. Now ad x maps L into itself, so $p(\operatorname{ad} x)$ does also. ■

Theorem 2.1.16 (Cartan's First Criterion)**.** *L is a Lie algebra such that $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$, $\forall x \in [L, L]$, $\forall y \in L$. $\iff$ L is solvable.*

Proof. The $\impliedby$ is by proposition 2.1.13.

$\implies$: Consider $\operatorname{ad}([L, L])$. By assumption, $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0, \forall x, y \in [L, L]$. Proposition 2.1.15 implies $\operatorname{ad}([L, L])$ is solvable. Since $\operatorname{ad}([L, L]) \cong [L, L]/Z([L, L])$ and $Z([L, L])$ is abelian, Proposition 1.3.3 (2) shows that $[L, L]$ is solvable. Since $[L, L] = L^{(1)}$, we have L solvable. ■

Remark 2.1.17. For the more general theory, see [1] section 4.3, where Cartan's First Criterion is named for a result generalizing proposition 2.1.15 and above theorem is listed as a corollary. ♠

2.2 Killing Form

Let F be an algebraically closed field with characteristic zero.

2.2.1 Semisimplicity

Let L be a Lie algebra. Define **Killing form** on L by

$$\begin{aligned} \kappa : L \times L &\to \mathbb{C} \\ (x, y) &\mapsto \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) \end{aligned}$$

- It is a symmetric bilinear form on L.
- It is also associative: $\kappa([x, y], z) = \kappa(x, [y, z])$ by ad as a homomorphism and Rmk.2.1.12(4).
- Cartan's First Criterion translates to: $\kappa(x, y) = 0$, $\forall x \in L$, $\forall y \in [L, L] \iff L$ solvable.

Remark 2.2.1. A simple fact from linear algebra: If W is a subspace of a (finite dimensional) vector space V, and ϕ an endomorphism of V mapping V into W, then $\operatorname{tr}\phi = \operatorname{tr}\left(\phi|_W\right)$. (To see this, extend a basis of W to a basis of V and look at the resulting matrix of ϕ.) ♠

Lemma 2.2.2. *If $I \subseteq L$ is an ideal and κ_I is the Killing form on I, then*

$$\kappa_I = \kappa|_{I \times I}.$$

Proof. If $x, y \in I$, then $(\operatorname{ad} x)(\operatorname{ad} y)$ is an endomorphism of L, mapping L into I, so remark above says its trace $\kappa(x, y)$ coincides with the trace $\kappa_I(x, y)$ of $(\operatorname{ad} x)(\operatorname{ad} y)|_I = (\operatorname{ad}_I x)(\operatorname{ad}_I y)$. ■

In general, a symmetric bilinear form $\beta(x, y)$ is called **nondegenerate** if its **radical** $\operatorname{rad}(\beta)$ is 0, where $\operatorname{rad}(\beta) := \{x \in L \mid \beta(x, y) = 0 \text{ for all } y \in L\}$. Alternatively, $\operatorname{rad}(\beta) = \ker(\widetilde{\beta})$ where $\widetilde{\beta} : L \to L^*$; $\widetilde{\beta}(x)(y) = \beta(x, y)$. Because the Killing form is associative, its radical is more than just a subspace: $\operatorname{rad}(k)$ is an ideal of L.

Remark 2.2.3. From linear algebra (theory of bilinear form), a practical way to test nondegeneracy is as follows: Fix a basis $x_1, \ldots, x_n$ of L. Then κ is nondegenerate if and only if the $n \times n$ matrix whose i, j entry is $\kappa(x_i, x_j)$ has nonzero determinant. ♠

Example 2.2.4. We compute the Killing form of $\mathfrak{sl}(2, F)$, using the standard basis (x, h, y). The matrices of $\operatorname{ad} x$, $\operatorname{ad} h$, $\operatorname{ad} y$ are shown in Example 2.1.14. Therefore κ has matrix

$$\begin{bmatrix} \kappa(x,x) & \kappa(x,h) & \kappa(x,y) \\ \kappa(h,x) & \kappa(h,h) & \kappa(h,y) \\ \kappa(y,x) & \kappa(y,h) & \kappa(y,y) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{bmatrix}$$

with determinant $\det = -128$, and κ is nondegenerate. (This is still true so long as char $F \neq 2$.) ♣

Theorem 2.2.5 (Cartan's Second Criterion)**.** *Let L be a Lie algebra. Then L is semisimple, i.e., $\operatorname{Rad}(L) = 0$, if and only if its Killing form κ is nondegenerate.*

Proof. Suppose first that Rad $L = 0$. Let S be the radical of κ. By definition, $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$ for all $x \in S, y \in L$ (in particular, for $y \in [S, S]$). According to Cartan's 1st Criterion, S is solvable. But we remarked before that S is an ideal of L and $\operatorname{Rad}(L)$ is the maximal solvable ideal, so $S \subset \operatorname{Rad} L = 0 \implies S = 0$, and κ is nondegenerate.

Conversely, let $S = 0$. To prove that L is semisimple, it will suffice to prove that every abelian ideal I of L is included in S (see Proposition 1.3.8). Suppose $x \in I, y \in L$. Then consider $(\operatorname{ad} x \operatorname{ad} y \operatorname{ad} x \operatorname{ad} y)z$ for $z \in L$: $[y, [x, [y, z]]] \in I$, so $[x, [y, [x, [y, z]]]] \in [I, I] = 0$ and $\operatorname{ad} x \operatorname{ad} y \operatorname{ad} x \operatorname{ad} y = (\operatorname{ad} x \operatorname{ad} y)^2 = 0$. This means that $\operatorname{ad} x \operatorname{ad} y$ is nilpotent, hence its matrix can be written as upper-triangular and $0 = \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = \kappa(x, y)$, so $I \subset S = 0$. ■

The proof shows that we always have $S \subset \operatorname{Rad} L$; however, the reverse inclusion need not hold.

2.2.2 Simple Ideals

Definition 2.2.6. *A Lie algebra L is said to be the* **direct sum** *of ideals $I_1, \ldots, I_t$ provided $L = I_1 \oplus \cdots \oplus I_t$ as direct sum of subspaces. That is $I_i \cap I_j = 0$ if $i \neq j$. This condition forces $[I_i, I_j] \subset I_i \cap I_j = 0$ if $i \neq j$ (so the algebra L can be viewed as gotten from the Lie algebras I_i by defining Lie bracket componentwise).*

Theorem 2.2.7. *L semisimple $\implies$ there exists simple ideals L_i's s.t. $L = L_1 \oplus \ldots \oplus L_t$. Every simple ideal of L coincides with one of the L_i. Moreover, the Killing form of L_i is the restriction of κ to $L_i \times L_i$.*

Proof. As a first step, let I be an arbitrary ideal of L. Then $I^\perp = \{x \in L \mid \kappa(x, y) = 0 \text{ for all } y \in I\}$ is also an ideal, by the associativity of κ. Cartan's Criterion, applied to the Lie algebra I, shows that the ideal $I \cap I^\perp$ of L is solvable (hence 0). Therefore, since $\dim I + \dim I^\perp = \dim L$, we must have $L = I \oplus I^\perp$.

Now proceed by induction on $\dim L$ to obtain the desired decomposition into direct sum of simple ideals. If L has no nonzero proper ideal, then L is simple already and we're done. Otherwise let L_1 be a minimal nonzero ideal; by the preceding paragraph, $L = L_1 \oplus L_1^\perp$. In particular, any ideal of L_1 is also an ideal of L, so L_1 is semisimple (hence simple, by minimality). For the same reason, $L_1^\perp$ is semisimple; by induction, it splits into a direct sum of simple ideals, which are also ideals of L. The decomposition of L follows.

Next we have to prove that these simple ideals are unique. If I is any simple ideal of L, then $[I, L]$ is also an ideal of I, nonzero because $Z(L) = 0$; this forces $[I, L] = I$. On the other hand, $[I, L] = [I, L_1] \oplus \ldots \oplus [I, L_t]$, so all but one summand must be 0.Say $[I, L_i] = I$. Then $I \subset L_i$, and $I = L_i$ (because L_i is simple).

The last assertion of the theorem follows from Lemma 2.2.2. ■

Corollary 2.2.8. *If L is semisimple, then $L = [L, L]$, and all ideals and homomorphic images of L are semisimple. Moreover, each ideal of L is a sum of certain simple ideals of L.*

2.2.3 Abstract Jordan Decomposition

We recall that $\operatorname{ad} : L \to \operatorname{Der}(L) \subseteq \mathfrak{gl}(L)$, where $\operatorname{Der}(L) = \{\delta \in \mathfrak{gl}(L) | \delta([a, b]) = [a, \delta(b)] + [\delta(a), b], \forall a, b \in L\}$. Furthermore, $\operatorname{ad}(L)$ is an ideal of $\operatorname{Der}(L)$. This is because

$$(*) \qquad \forall \delta \in \operatorname{Der}(L),\ x \in L,\ [\delta, \operatorname{ad} x] = \operatorname{ad}(\delta x).$$

proof of (*). We compute

$$\begin{aligned}
[\delta, \operatorname{ad} x](y) &= (\delta \operatorname{ad}(x))(y) - (\operatorname{ad} x \delta)(y) \\
&= \delta([x, y]) - [x, \delta(y)] \\
&= [\delta(x), y] - \operatorname{ad}(\delta(x))(y) \\
&\implies [\delta, \operatorname{ad} x] = \operatorname{ad}(\delta(x))
\end{aligned}$$

■

Theorem 2.2.9. *If L is semisimple, then ad L = Der L (i.e., every derivation of L is inner).*

Proof. Since L is semisimple, the abelian ideal $Z(L)$ is 0. Therefore, $L \to \operatorname{ad} L$ is an isomorphism of Lie algebras. In particular, $M = \operatorname{ad} L$ itself has nondegenerate Killing form (Cartan's 2nd Criterion). If $D = \operatorname{Der} L$, we just remarked that $[D, M] \subset M$. This implies (by Lemma 2.2.2) that κ_M is the restriction to $M \times M$ of the Killing form κ_D of D. In particular, if $I = M^\perp$ is the subspace $\{\delta \in D | \kappa_D(\delta, \tau) = 0, \forall \tau \in M\}$ of D orthogonal to M under κ_D, then the nondegeneracy of $\kappa_M \implies I \cap M = 0 \implies [M, I] \subseteq M \cap I = 0$. If $\delta \in I$, then (*) gives $\operatorname{ad}(\delta(x)) = 0$ for all $x \in L$, so in turn $\delta(x) = 0,\ \forall x \in L$. Because ad is $1-1$, we have $\delta = 0$. Conclusion: $I = 0$, Der $L = M = \operatorname{ad} L$. ■

Proposition 2.2.10. *Let A be a finite-dimensional F-algebra. Prove that Lie algebra of derivations $\operatorname{Der}(A) \subseteq \mathfrak{gl}(A)$ contains the semisimple and nilpotent parts of its elements. That is, given a decomposition*

$$\operatorname{Der}(A) \ni \delta = \underbrace{\nu}_{\textit{nilp.}} + \underbrace{\sigma}_{\textit{s.s.}},$$

we have $\nu, \sigma \in \operatorname{Der}(A)$.

Proof. See [1] Lemma B of 4.2. ■

We then have a decomposition for element in semisimple Lie algebra (note that our original Jordan decomposition is for linear Lie algebra $L \subseteq \mathfrak{gl}(V)$.)

Proposition 2.2.11 (Abstract Jordan Decomposition)**.** *Let L be a semisimple Lie algebra. For each $x \in L$:*

$$\begin{aligned}
&\exists! n, s \in L \\
&\quad x = n + s \\
&\quad \operatorname{ad} n \textit{ nilp.} \\
&\quad \operatorname{ad} s \textit{ s.s.} \\
&\quad [n, s] = 0
\end{aligned}$$

Proof. As remarked in the proof of above proposition, ad is an isomorphism in this case.

$$\begin{aligned} \operatorname{Der}(L) = \operatorname{ad} L &\longleftrightarrow L \\ \operatorname{ad} x &\longleftarrow x \end{aligned}$$

Decompose $\operatorname{ad} x$ as $\nu + \sigma$. $\nu, \sigma \in \operatorname{Der}(L) = \operatorname{ad} L \implies \exists! n, s$ s.t. $\operatorname{ad} n = \nu,\ \operatorname{ad} s = \sigma$ due to isomorphic ad. Thus, $\operatorname{ad} x = \operatorname{ad} n + \operatorname{ad} s = \operatorname{ad}(n+s) \implies x = n + s$. Besides, $0 = [\operatorname{ad} n, \operatorname{ad} s] = \operatorname{ad}[n, s] \implies [n, s] = 0$ again due to isomorphic ad. ■

We can write $s = x_s$, $n = x_n$, and (by abuse of language) call these the **semisimple** and **nilpotent** parts of x. The alert one will object at this point that the notation x_s, x_n is ambiguous in case L happens to be a linear Lie algebra. It will be shown that the abstract decomposition of x just obtained does in fact agree with the usual Jordan decomposition in all such cases. For the moment we shall be content to point out that this is true in the special case $L = \mathfrak{sl}(V)$ $(\dim V < \infty)$: Write $x = x_s + x_n$ in End V (usual Jordan decomposition), $x \in L$. Since x_n is a nilpotent endomorphism, its trace is 0 and therefore $x_n \in L$. This forces x_s also to have trace 0, so $x_s \in L$. Moreover, $\operatorname{ad}_{\mathfrak{gl}(V)} x_s$ is semisimple (rmk.2.1.10), so $\operatorname{ad}_L x_s$ is a fortiori semisimple; similarly $\operatorname{ad}_L x_n$ is nilpotent (lem.1.3.17), and $[\operatorname{ad}_L x_s, \operatorname{ad}_L x_n] = \operatorname{ad}_L [x_s x_n] = 0$. By the uniqueness of the abstract Jordan decomposition in $L, x = x_s + x_n$ must be it.

2.3 Complete Reducibility of Representations

2.3.1 Modules

A (finite-dimensional) representation of Lie algebra L is a Lie algebra homomorphism

$$\varphi : L \to \mathfrak{gl}(V)$$

where V is a finite dimensional F-vector space.

Definition 2.3.1. *A vector space V with an operation $L \times V \to V$; $(x, v) \mapsto x.v$ is an L-module if*

- $(ax + by).v = ax.v + by.v$
- $x.(av + bw) = ax.v + bx.w$
- $[x, y].v = x.y.v - y.x.v = x.(y.v) - y.(x.v)$

We have the identification

$$\begin{aligned} \left\{ \begin{array}{c} \text{representations} \\ \varphi : L \to \mathfrak{gl}(V) \end{array} \right\} &\longleftrightarrow \{L-\text{modules}\} \\ \varphi &\longrightarrow x.v := \varphi(x)(v) \\ \left(x \mapsto \left(\begin{array}{c} V \to V \\ v \mapsto x.v \end{array} \right) \right) &\longleftarrow x.v \end{aligned}$$

Definition 2.3.2.

- *A* ***homomorphism of*** L***-modules*** *is a linear map $\phi : V \to W$ such that $\phi(x.v) = x.\phi(v)$. The kernel of such a homomorphism is then an L-submodule of V (a vector subspace with operation closed in it). When ϕ is an isomorphism of vector spaces, we call it an* ***isomorphism of*** L***-modules****; in this case, the two modules are said to afford equivalent representations of L. The standard isomorphism theorems all go through without difficulty.*

- *An L-module V is called **irreducible** if it has precisely two L-submodules (itself and 0). We do not regard a zero dimensional vector space as an irreducible L-module. Every one dimensional vector space on which L acts is an irreducible L-module (because one-dimensional vector space does not have any nonzero proper subspace.)*
- *V is called **completely reducible** if V is a direct sum of irreducible L-submodules, or equivalently (Exercise 2.3.3), if each L-submodule W of V has a complement W' (an L-submodule such that $V = W \oplus W'$). When W, W' are arbitrary L-modules, we can of course make their direct sum an L-module in the obvious way, by defining $x.(w, w') = (x.w, x.w')$. These notions are all standard and also make sense when $\dim V = \infty$.*
- *The terminology "irreducible" and "completely reducible" applies equally well to representations of L. Namely, a representation $\phi : L \to \mathfrak{gl}(V)$ is irreducible and completely reducible if the corresponding L-module is irreducible and completely reducible respectively.*

Exercise 2.3.3. *Let V be an L-module. Prove that V is a direct sum of irreducible submodules if and only if each L-submodule of V possesses a complement.*

Let L be a Lie algebra. It is an L-module corresponding to the adjoint representation $\mathrm{ad} : L \to \mathfrak{gl}(L)$. An L-submodule is just an ideal, so it follows that the L-module arising in this way with L a simple algebra is irreducible, and if L is semisimple, the module is completely reducible.

Suppose that S and T are irreducible Lie modules and that $\theta : S \to T$ is a non-zero module homomorphism. Then $\mathrm{Im}\,\theta$ is a non-zero submodule of T, so $\mathrm{im}\,\theta = T$. Similarly, $\ker\theta$ is a proper submodule of S, so $\ker\theta = 0$. It follows that θ is an isomorphism from S to T, so there are no non-zero homomorphisms between non-isomorphic irreducible modules.

Now we consider the homomorphism from an irreducible module to itself.

Lemma 2.3.4 (Schur's Lemma)**.** *Let L be a Lie algebra and let S be a finite-dimensional irreducible L-module over V. Then A map $\theta : S \to S$ is an L-module homomorphism if and only if θ is a scalar multiple of the identity transformation; that is, $\theta = \lambda 1_S$ for some $\lambda \in F$*

Remark 2.3.5. $\theta : S \to S$ is an L-module homomorphism by definition means $\forall x \in L$, $\forall v \in V$, $\theta(x.v) = x.(\theta(v))$. If we let $\phi : L \to \mathfrak{gl}(V)$ be the corresponding representation of S, then this condition means $\forall x \in L$, $\forall v \in V$, $\theta(\phi(x)(v)) = \phi(x)(\theta(v))$, or $\forall x \in L$, $[\theta, \phi(x)] = 0$. Thus, an L-module homomorphism $\theta : S \to S$ is precisely an endomorphism $\theta \in \mathfrak{gl}(V)$ that commutes with every $\phi(x)$ in $\mathrm{Im}(\phi)$. ♠

Proof. The "if" direction should be clear. For the "only if" direction, suppose that $\theta : S \to S$ is an L-module homomorphism. Then $\theta \in \mathfrak{gl}(V)$. As a matrix with entries in algebraically closed field, it has at least one eigenvector ξ with an eigenvalue λ. Now $\theta - \lambda 1_S$ is also an L-module homomorphism. The kernel of this map contains ξ and is thus a nonzero submodule of S. As S is irreducible, $S = \ker(\theta - \lambda 1_S)$; that is, $\theta = \lambda 1_S$. ■

Constructions of L-modules:

Let V and W be f.d. L-modules.

- The **dual vector space**

$$V^* = \{f : V \to F \mid f \text{ is linear } \}$$

is an L-module if we define for $f \in V^*, v \in V, x \in L:\ (x.f)(v) = -f(x.v)$. The first two axioms are immediate. The third is true as

$$\begin{aligned}([x,y].f)(v) &= -f([x,y].v)\\ &= -f(x.y.v - y.x.v)\\ &= -f(x.y.v) + f(y.x.v)\\ &= (x.f)(y.v) - (y.f)(x.v)\\ &= -(y.x.f)(v) + (x.y.f)(v)\\ &= ((x.y - y.x).f)(v)\end{aligned}$$

- If V, W are L-modules, let $V \otimes W$ be the **tensor product** over F of the underlying vector spaces. Recall that if V, W have respective bases $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_n)$, then $V \otimes W$ has a basis consisting of the mn vectors $v_i \otimes w_j$. One recalls how to give a module structure to the tensor product of two modules for a group G: on the generators $v \otimes w$, require $g.(v \otimes w) = g.v \otimes g.w$. For Lie algebras the correct definition is gotten by "differentiating" this one: $x.(v \otimes w) = x.v \otimes w + v \otimes x.w$. As before, the crucial axiom to verify is the third axiom:

$$\begin{aligned}[x,y].(v\otimes w) &= [x,y].v \otimes w + v \otimes [x,y].w\\ &= (x.y.v - y.x.v)\otimes w + v \otimes (x.y.w - y.x.w)\\ &= (x.y.v\otimes w + v\otimes x.y.w) - (y.x.v \otimes w + v \otimes y.x.w)\end{aligned}$$

Expanding $(x.y - y.x).(v \otimes w)$ yields the same result.

Remark 2.3.6. There is a canonical isomorphism $\theta : V^* \otimes V \to \mathrm{End}(V)$ by $\theta(f,v)(w) = (f(w))(v)$. We can use this som. to nuke $\mathrm{End}(V)$ into an L-module:

$$x.\theta(f,v) := \theta(x.f, v) + \theta(f, x.v)$$

Note that $x.\theta(f,v)$ acts on $w \in V$ by

$$\begin{aligned}\left[\overbrace{x}^{\in L} \cdot \overbrace{\theta(f,v)}^{\in \mathrm{End}(V)}\right](w) &= \left[\theta(\overbrace{x.f}^{V^*\ L-\text{mod}}, v)\right](w) + \left[\theta(f, \overbrace{x.v}^{V\ L-\text{mod}})\right](w)\\ &= (x.f)(w)v + f(w)(x.v)\\ &= -f(x.w)v + x.(f(w)v)\\ &= -\theta(f,v)(x.w) + x.(\theta(f,v)(w))\end{aligned}$$

Thus, the action of L on $\varphi \in \mathrm{End}(V)$ is given by

$$(x.\varphi)(w) = x.\varphi(w) - \varphi(x.w)$$

♠

Remark 2.3.7. More generally, if V and W are two L-modules, then L acts naturally on $\varphi \in \mathrm{Hom}(V,W) \cong V^* \otimes W$ of linear maps by the same rule

$$(x.\varphi)(w) = x.\varphi(w) - \varphi(x.w)$$

♠

Exercise 2.3.8. *Check that above action satisfies the three axioms of L-module action.*

Exercise 2.3.9. *If $\varphi : L \to \mathfrak{gl}(V)$ is a representation and $x = s + n$ is the Jordan decomposition then $\varphi(x) = \varphi(s) + \varphi(n)$ is the Jordan decomposition of $\varphi(x)$.*

2.3.2 Casimir Element of a Representation

Field F, algebraically closed, $\mathrm{char}(F) = 0$.

A representation $\varphi : L \to \mathfrak{gl}(V)$ is **faithful** if $\ker \varphi = 0$

Given a faithful representation φ of semisimple L, we define

$$\beta_\varphi : L \times L \to F$$
$$\beta_\varphi(x, y) = \mathrm{tr}(\varphi(x)\varphi(y)).$$

Remark 2.3.10.

(1) β_φ is a symmetric bilinear form on L.

(2) β_φ is associative: $\beta_\varphi([x, y], z) = \beta_\varphi(x, [y, z])$ due to Remark 2.1.12 (4).

(3) Thus its radical $S_\varphi = \{x \in L \mid \beta_\varphi(x, y) = 0,\ \forall y \in L\}$ is an ideal of L.

(4) $\beta_{\mathrm{ad}} = k$ is the Killing form.

(5) β_φ is nondegenerate, i.e., $S_\varphi = 0$:
Since φ is faithful, we see $\varphi : S_\varphi \to \varphi(S_\varphi)$ is bijective, so $\forall x', y' \in \varphi(S_\varphi)$, we can find $x, y \in S_\varphi$ s.t. $x' = \varphi(x), y' = \varphi(y)$ and $\mathrm{tr}(\varphi(x)\varphi(y)) = \beta_\varphi(x, y) = 0$. Proposition 2.1.15 then says that $\varphi(S_\varphi) \cong S_\varphi$ is solvable. Thus, S_φ, as an ideal of semisimple L, is 0.

♠

Basic setting:

Now let L be semisimple, β any nondegenerate symmetric associative bilinear form on L. If $(x_1, \ldots, x_n)$ is a basis of L, there is a uniquely determined basis of L $(y_1, \ldots, y_n)$ relative to β, satisfying $\beta(x_i, y_j) = \delta_{ij}$. ($\beta$ nondegenerate so we have isomorphism $\widetilde{\beta} : L \to L^*$; $y \mapsto \beta(\,\cdot\,, y)$. Then for basis x_i of L, there is a unique dual basis θ_j of L^* s.t. $\theta_j(x_i) = \delta_{ij}$. Now take $y_j = \widetilde{\beta}(\theta_j)$, i.e., $\theta_j = \beta(\,\cdot\,, y_j)$. The y_j's serve the unique basis of L determined by x_i's wrt β s.t. $\beta(x_i, y_j) = \delta_{ij}$.)

Lemma 2.3.11. *If $x \in L$, we can write $[x, x_i] = \sum_j a_{ij} x_j$ and $[x, y_i] = \sum_j b_{ij} y_j$. Using the associativity of β, we compute:*

$$a_{ik} = \sum_j a_{ij} \beta(x_j, y_k) = \beta([x, x_i], y_k) = \beta(-[x_i, x], y_k) = \beta(x_i, -[x, y_k]) = -\sum_j b_{kj} \beta(x_i, y_j) = -b_{ki}.$$

Definition 2.3.12. *Let (x_i) and (y_i) be bases of L wrt β as above. The **Casimir operator** of a representation $\varphi : L \to \mathfrak{gl}(V)$ is the linear map*

$$c_\varphi = c_\varphi(\beta) := \sum_{i=1}^n \varphi(x_i) \varphi(y_i) \in \mathfrak{gl}(V)$$

Lemma 2.3.13.

(1) c_φ *commutes with* $\varphi(x)$ $\quad \forall x \in L$

(2) $\mathrm{tr}(c_\varphi) = \dim L$.

Proof. (1):

$$\begin{aligned}
[\varphi(x), c_{\varphi}] &= \sum_{i} [\varphi(x), \varphi(x_i)\varphi(y_i)] \\
&= \sum_{i} [\varphi(x), \varphi(x_i)]\varphi(y_i) + \sum_{i} \varphi(x_i)[\varphi(x), \varphi(y_i)] \\
&= \sum_{ij} a_{ij}\varphi(x_j)\varphi(y_i) + \sum_{i} \varphi(x_i) b_{ij}\varphi(y_j) \\
&= \sum_{ij} (a_{ij} + b_{ji})\varphi(x_j)\varphi(y_i) \\
&= 0.
\end{aligned}$$

where for the second step we used

$$[x, yz] = xyz - yzx = xyz - yxz + yxz - yzx = [x, y]z + y[x, z].$$

(2):

$$\begin{aligned}
\operatorname{tr}(c_{\varphi}) &= \sum_{i} \operatorname{tr}(\varphi(x_i)\varphi(y_i)) \\
&= \sum_{i=1}^{n} \beta(x_i, y_i) = n = \dim L.
\end{aligned}$$

■

Remark 2.3.14. In case φ is also irreducible, Schur's Lemma implies that c_{φ} is a scalar (equal to $\dim L/\dim V$, in view of (2)); in this case we see that c_{φ} is independent of the basis of L which we chose. ♠

Example 2.3.15. $L = \mathfrak{sl}(2, F), V = F^2, \varphi$ the identity map $L \to \mathfrak{gl}(V)$. Let (x, h, y) be the standard basis of L. It is quickly seen that the dual basis relative to the trace form is $(y, h/2, x)$, so $c_{\varphi} = xy + (1/2)h^2 + yx = \begin{pmatrix} 3/2 & 0 \\ 0 & 3/2 \end{pmatrix}$. Notice that $3/2 = \dim L/\dim V$. ♣

2.3.3 Weyl's Theorem

Lemma 2.3.16. *Let $\varphi : L \to \mathfrak{gl}(V)$ be a representation of a semisimple Lie algebra L. Then $\varphi(L) \subseteq \mathfrak{sl}(V)$. In particular, semisimple L acts trivially on any one-dimensional L-module.*

Proof. Semisimplicity implies $L = [L, L]$. Then $\varphi(L) = \varphi([L, L]) \subseteq [\mathfrak{gl}(V), \mathfrak{gl}(V)] = \mathfrak{sl}(V)$. When V is one-dimensional, $\mathfrak{sl}(V) \cong \mathfrak{sl}_1(F) = \{0\}$. ■

Theorem 2.3.17 (Weyl)**.** *Let $\phi : L \to \mathfrak{gl}(V)$ be a (finite dimensional) representation of a semisimple Lie algebra. Then ϕ is completely reducible.*

Proof. special case: V has an L-submodule W of codimension one.

By the lemma, L acts trivially on V/W, so we may denote this module F without misleading the reader: $0 \to W \to V \to F \to 0$ is therefore exact. Using induction on $\dim W$, we can reduce to the case where W is an irreducible L-module, as follows. Let W' be a proper nonzero submodule of W. This yields an exact sequence: $0 \to W/W' \to V/W' \to F \to 0$. By induction, this sequence "splits", i.e., there exists a one dimensional L-submodule of V/W' (say $\widetilde{W}/W'$) complementary to W/W'. So we get another exact sequence: $0 \to W' \to \widetilde{W} \to F \to 0$. This is like the original situation, except that $\dim W' < \dim W$, so induction provides a (one dimensional) submodule X complementary to W' in $\widetilde{W} : \widetilde{W} = W' \oplus X$. But $V/W' = W/W' \oplus \widetilde{W}/W'$. It follows that $V = W \oplus X$, since the dimensions add up to $\dim V$ and since $W \cap X = 0$.

further case: Now we may assume that W is irreducible. (We may also assume without loss of generality that L acts faithfully on V.)

Let $c = c_\phi$ be the Casimir element of ϕ. Since c commutes with $\phi(L)$, c is actually an L-module endomorphism of V; in particular, $c(W) \subset W$ and $\operatorname{Ker} c$ is an L-submodule of V. Because L acts trivially on V/W (i.e., $\phi(L)$ sends V into W), c must do likewise (as a linear combination of products of elements $\phi(x)$). So c has trace 0 on V/W. On the other hand, c acts as a scalar on the irreducible L-submodule W (Schur's Lemma); this scalar cannot be 0, because that would force $\operatorname{Tr}_V(c) = 0$, contrary to lemma 2.3.13. It follows that Ker c is a one dimensional L-submodule of V which intersects W trivially. This is the desired complement to W.

fgeneral case: Let W be a nonzero submodule of $V : 0 \to W \to V \to V/W \to 0$. Let $\operatorname{Hom}(V, W)$ be the space of linear maps $V \to W$, viewed as L-module (6.1). Let $\mathscr{V}$ be the subspace of Hom (V, W) consisting of those maps whose restriction to W is a scalar multiplication. $\mathscr{V}$ is actually an L-submodule: Say $f|_W = a.1_W$; then for $x \in L, w \in W$, $(x.f)(w) = x.f(w) - f(x.w) = a(x.w) - a(x.w) = 0$, so $x.\ f|_w = 0$. Let $\mathscr{W}$ be the subspace of $\mathscr{V}$ consisting of those f whose restriction to W is zero. The preceding calculation shows that $\mathscr{W}$ is also an L-submodule and that L maps $\mathscr{V}$ into $\mathscr{W}$. Moreover, $\mathscr{V} \mid \mathscr{W}$ has dimension one, because each $f \in \mathscr{V}$ is determined (modulo $\mathscr{W}$) by the scalar $f|_W$. This places us precisely in the situation $0 \to \mathscr{W} \to \mathscr{V} \to F \to 0$ already treated above.

According to the first part of the proof, $\mathscr{V}$ has a one dimensional submodule complementary to $\mathscr{W}$. Let $f : V \to W$ span it, so after multiplying by a nonzero scalar we may assume that $f|_W = 1_W$. To say that L kills f is just to say that $0 = (x.f)(v) = x.f(v) - f(x.v)$, i.e., that f is an L-homomorphism. Therefore Ker f is an L-submodule of V. Since f maps V into W and acts as 1_W on W, we conclude that $V = W \oplus \operatorname{Ker} f$, as desired. ■

2.3.4 Preservation of Jordan Decomposition

We promised to show the following result to resolve the ambiguity of usual and abstract Jordan decompositions of semisimple linear Lie algebra.

Theorem 2.3.18. *Let $L \subset \mathfrak{gl}(V)$ be a semisimple linear Lie algebra (V finite dimensional). Then L contains the semisimple and nilpotent parts in $\mathfrak{gl}(V)$ of all its elements. In particular, the abstract and usual Jordan decompositions in L coincide.*

Proof. The last assertion follows from the first, because each type of Jordan decomposition is unique.

Let $x \in L$ have ordinary Jordan decomposition $x = x_s + x_n$ in $\mathfrak{gl}(V)$. Then

$$\operatorname{ad} x = \operatorname{ad} x_s + \operatorname{ad} x_n$$

is the Jordan decomposition of $\operatorname{ad} x$ inside $\operatorname{End}(\mathfrak{gl}(V))$. The statement is just to show $x_n, x_s \in L$.

Since $(\operatorname{ad} x)(L) \subset L$, it follows from Cor.2.1.9 that $(\operatorname{ad} x_s)(L) \subset L$ and $(\operatorname{ad} x_n)(L) \subset L$, where $\operatorname{ad} = \operatorname{ad}_{\mathfrak{gl}(V)}$. In other words, $x_s, x_n \in$ normalizer $N_{\mathfrak{gl}(V)}(L) = N$, which is a Lie subalgebra of $\mathfrak{gl}(V)$ including L as an ideal. If we could show that $N = L$ we'd be done, but unfortunately this is false: e.g., since $L \subset \mathfrak{sl}(V)$ (Lemma 2.3.16), the scalars lie in N but not in L. Therefore we need to get x_s, x_n into a smaller subalgebra than N, which can be shown to equal L. If W is any L-submodule of V, define

$$L_W = \{y \in \mathfrak{gl}(V) \mid y(W) \subset W \text{ and } \operatorname{tr}(y|_W) = 0\}.$$

For example, $L_V = \mathfrak{sl}(V)$. Then $L \subseteq L_W$ because W is an L-submodule and for $x \in L$,

$$\begin{aligned} \operatorname{tr}(x|_W) &\xlongequal{L=[L,L]} \operatorname{tr}([y,z]|_W) \\ &= \operatorname{tr}(z|_W\, y|_W - y|_W\, z|_W) = 0. \end{aligned}$$

Set L' =intersection of N with all spaces L_W:

$$L' = N \cap \left(\bigcap_{\substack{L-\text{submod} \\ W \subseteq V}} L_w \right)$$

Clearly, L' is a subalgebra of N including L as an ideal (but notice that L' excludes the scalars). Even more is true: If $x \in L$, then x_s, x_n also lie in L_W, and therefore in L'.

It remains to prove that $L = L'$. L' being a finite dimensional L-module and L is an ideal of L', Weyl's Theorem permits us to write $L' = L \oplus M$ for some L-submodule M. But $[L, L'] \subset L$ (since $L' \subset N$), so the action of L on M is trivial. Let W be any irreducible L-submodule of V. If $y \in M$, then $[L, y] = 0$, so Schur's Lemma implies that y acts on W as a scalar. On the other hand, $\operatorname{tr}(y|_W) = 0$ because $y \in L_W$. Therefore y acts on W as zero. V can be written as a direct sum of irreducible L-submodules (by Weyl's Theorem), so in fact $y = 0$. This means $M = 0, L = L'$. ■

Corollary 2.3.19. *Let L be a semisimple Lie algebra, $\phi : L \to \mathfrak{gl}(V)$ a (finite dimensional) representation of L. If $x = s + n$ is the abstract Jordan decomposition of $x \in L$, then $\phi(x) = \phi(s) + \phi(n)$ is the usual Jordan decomposition of $\phi(x)$.*

Proof. By Cor.2.2.8, $\phi(L)$ is semisimple. Thus, it makes sense to talk about the abstract Jordan decomposition of $\phi(x)$. Then our strategy is to show $\phi(x) = \phi(s) + \phi(n)$ is the abstract Jordan decomposition of $\phi(x)$. Then Theorem 2.3.18 concludes.

We show $\operatorname{ad}_{\phi(L)} \phi(s)$ is semisimple and $\operatorname{ad}_{\phi(L)} \phi(n)$ is nilpotent.

$\operatorname{ad} s$ is semisimple, so it is diagonalizable and its eigenvectors $e_1, \cdots, e_n$ are linearly independent and form a basis for L. Then $\phi(e_1), \cdots, \phi(e_n)$ are eigenvectors for $\operatorname{ad}_{\phi(L)} \phi(s)$ and are linearly independent, spanning the algebra $\phi(L)$:

$$\left(\operatorname{ad}_{\phi(L)} \phi(s)\right)(\phi(e_i)) = [\phi(s), \phi(e_i)] = \phi([s, e_i]) = \lambda_i \phi(e_i)$$

Thus, $\operatorname{ad}_{\phi(L)} \phi(s)$ is diagonalizable, i.e., semisimple.

$\operatorname{ad} n$ is nilpotent, so $(\operatorname{ad} n)^m = 0$ for some $m > 0$. Then, for all $y \in L$,

$$\begin{aligned}(\operatorname{ad}_{\phi(L)} \phi(n))^m \phi(y) &= [\phi(n), [\phi(n), \ldots, [\phi(n), \phi(y)] \ldots]] \\ &= \phi([n, [n, \ldots, [n, y] \ldots]]) \\ &= \phi\left((\operatorname{ad} n)^m y\right) = 0\end{aligned}$$

As $\phi : L \to \phi(L)$ is surjective, i.e., every element in $\phi(L)$ is of the form $\phi(y)$, this shows that $(\operatorname{ad}_{\phi(L)} \phi(n))^m \in \mathfrak{gl}(\phi(L))$ is 0. Thus, $\operatorname{ad}_{\phi(L)} \phi(n)$ is nilpotent.

Moreover, $\operatorname{ad}_{\phi(L)} \phi(s)$ and $\operatorname{ad}_{\phi(L)} \phi(n)$ commute:

$$[\operatorname{ad}_{\phi(L)} \phi(s), \operatorname{ad}_{\phi(L)} \phi(n)] = \operatorname{ad}_{\phi(L)}[\phi(s), \phi(n)] = \operatorname{ad}_{\phi(L)} \theta([s, n]) = 0$$

By uniqueness of abstract Jordan decomposition 2.2.11, we see $\phi(s) + \phi(n)$ is the abstract Jordan decomposition of $\phi(x)$. ■

2.4 Representation of $\mathfrak{sl}(2, F)$

In this section, L denotes $\mathfrak{sl}(2, F)$ with standard basis (x, h, y).

$$x = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

Let V be an arbitrary L-module. We have the corresponding representation

$$\begin{aligned}\phi : L &\to \mathfrak{gl}(V)\\ x &\mapsto \left(\begin{array}{c}\phi_x : V \to V\\ v \mapsto x.v\end{array}\right)\end{aligned}$$

By Cor.2.3.19, h semisimple $\Rightarrow \phi_h$ is semisimple, i.e., the endomorphism $\phi_h : v \mapsto x.v$ acts diagonally on V. We can then write V as a direct sum of eigenspaces $V_\lambda := \{v \in V | x.v = \lambda v\}$ of ϕ_h. The expression V_λ still makes sense if λ is not an eigenvalue of ϕ_h (then $V_\lambda = 0$). When $V_\lambda \neq 0$ for a $\lambda \in F$, we call λ a **weight** of h in V and we call V_λ a **weight space**. We have

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

Note that the linear Lie algebras $L \subseteq \mathfrak{gl}(V)$ are naturally L-modules with obvious actions on V. In that case, a weight space is the same as an **eigenspace**.

Example 2.4.1. L denotes $\mathfrak{sl}(2, F)$ with standard basis (x, h, y). We compute the weight spaces for h. Let $v \in V_\lambda$, then

$$h.(x.v) = [h, x].v + x.h.v = 2x.v + \lambda x.v = (\lambda + 2)x.v$$

and

$$h.(y.v) = [h, y].v + y.h.v = -2y.v + \lambda y.v = (\lambda - 2)y.v$$

Thus, $x.v \in V_{\lambda+2}$ and $y.v \in V_{\lambda-2}$. ♣

Remark 2.4.2. Since there are finite number of weight spaces, we see $\exists \lambda$ s.t. $V_\lambda \neq 0$ and $V_{\lambda+2} = 0$. Then, pick $v \in V_\lambda$ and we have $x.v \in V_{\lambda+2} = 0 \implies x.v = 0$.

For such λ, any nonzero vector in V_λ will be called a **maximal vector** of weight λ. ♠

2.4.1 Module V_d

We construct a family of irreducible representations of $\mathfrak{sl}(2, F)$. Consider the vector space $F[X, Y]$ of polynomials in two variables X, Y with complex coefficients. For each integer $d \geqslant 0$, let V_d be the subspace of homogeneous polynomials in X and Y of degree d. So V_0 is the 1-dimensional vector space of constant polynomials, and for $d \geqslant 1$, the space V_d has as a basis the monomials $X^d, X^{d-1}Y, \ldots, XY^{d-1}, Y^d$. This basis shows that V_d has dimension $d + 1$ as a F-vector space.

We now make V_d into an $\mathfrak{sl}(2, F)$-module by specifying a Lie algebra homomorphism $\varphi : \mathfrak{sl}(2, F) \to \mathfrak{gl}(V_d)$. Since $\mathfrak{sl}(2, F)$ is linearly spanned by the matrices x, y, h, the map φ will be determined once we have specified $\varphi(x)$, $\varphi(y)$, $\varphi(h)$.

We let

$$\varphi(x) := X \frac{\partial}{\partial Y}$$

that is, $\varphi(x)$ is the linear map which first differentiates a polynomial with respect to Y and then multiplies it with X. This preserves the degrees of polynomials and so maps V_d into V_d. Similarly, we let

$$\varphi(y) := Y \frac{\partial}{\partial X}$$

Finally, we let

$$\varphi(h) := X \frac{\partial}{\partial X} - Y \frac{\partial}{\partial Y}$$

Notice that

$$\varphi(h)\left(X^a Y^b\right) = (a - b)X^a Y^b$$

so h acts diagonally on V_d with respect to our chosen basis. For any $v = \sum_{a+b=d} c_{ab} X^a Y^b \in V_d$, we see that $\varphi(h)(v) = \sum_{a+b=d} c_{ad}(a-b) X^a Y^b$. To let v be an eigenvector, we need $a - b = a' - b'$ for all pairs. But $a + b = a' + b' = d$ forces $a = a', b = b'$. Thus, any candidate $v \in V_d$ for an eigenvector of $\varphi(h)$ must be a multiple of a basis of V_d. The eigenspaces of $\varphi(h)$ are all one-dimensional.

Note also

$$\varphi(x)\left(X^a Y^b\right) = bX^{a+1}Y^{b-1}$$
$$\varphi(y)\left(X^a Y^b\right) = aX^{a-1}Y^{b+1}$$

Proposition 2.4.3. *With these definitions, φ is a representation of $\mathfrak{sl}(2, F)$.*

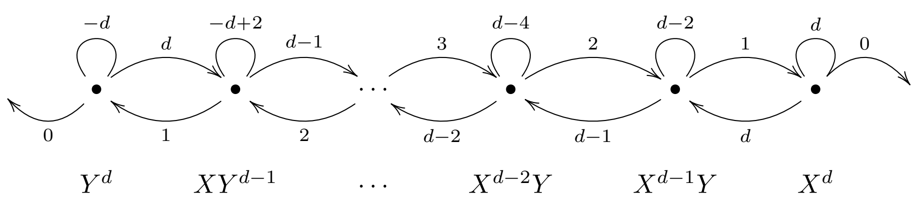
Proof. See [4] Theorem 8.1. ■

It can be useful to know the matrices of $\varphi(x), \varphi(y), \varphi(h)$ on V_d wrt. the basis $X^d, X^{d-1}Y, \ldots, Y^d$ of V_d. They are

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & d \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ d & 0 & \ldots & 0 & 0 \\ 0 & d-1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}, \begin{pmatrix} d & 0 & \ldots & 0 & 0 \\ 0 & d-2 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & -d+2 & 0 \\ 0 & 0 & \ldots & 0 & -d \end{pmatrix}$$

where the diagonal entries of the last are the numbers $d - 2k$, where $\kappa = 0, 1, \ldots, d$.

Another way to represent the action of x, y, h on V_d is to draw a diagram like



where loops represent the action of h, arrows to the right represent the action of x, and arrows to the left represent the action of y.

The diagram above show that $\mathfrak{sl}(2, F)$-submodule of V_d generated by any particular basis element $X^a Y^b$ contains all the basis elements and so is all of V_d. In general, if S is any non-zero $\mathfrak{sl}(2, F)$-submodule of V_d. Then $h.s \in S$ for all $s \in U$. We saw h acts diagonalizably on V_d, it also acts diagonalizably on S, so there is an eigenvector of h which lies in S. We have seen that all eigenspaces of h on V_d are one-dimensional, and each eigenspace is spanned by some monomial $X^a Y^b$, so the submodule S must contain some monomial, and by the observation on diagram remarked just now, S contains a basis for V_d. Hence $S = V_d$. We thus showed the following.

Proposition 2.4.4. *V_d is an irreducible $\mathfrak{sl}(2, F)$-module.*

Exercise 2.4.5. *([4] ex.8.2) Find explicit isomorphisms between*

(i) the trivial representation of $\mathfrak{sl}(2, \mathbb{C})$ and V_0;

(ii) the natural representation of $\mathfrak{sl}(2, \mathbb{C})$ and V_1;

(iii) the adjoint representation of $\mathfrak{sl}(2, \mathbb{C})$ and V_2.

Exercise 2.4.6. *([4] ex.8.3) Show that the subalgebra of $\mathfrak{sl}(3, \mathbb{C})$ consisting of matrices of the form*

$$\begin{pmatrix} \star & \star & 0 \\ \star & \star & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$. *We may therefore regard* $\mathfrak{sl}(3,\mathbb{C})$ *as a module for* $\mathfrak{sl}(2,\mathbb{C})$, *with the action given by* $x \cdot y = [x,y]$ *for* $x \in \mathfrak{sl}(2,\mathbb{C})$ *and* $y \in \mathfrak{sl}(3,\mathbb{C})$. *Show that as an* $\mathfrak{sl}(2,\mathbb{C})$*-module*

$$\mathfrak{sl}(3,\mathbb{C}) \cong V_2 \oplus V_1 \oplus V_1 \oplus V_0.$$

2.4.2 Classification of Irreducible $\mathfrak{sl}(2,F)$-Modules

Let $L = \mathfrak{sl}(2,F)$. We recall some representations of L.

- We have seen in Lemma 2.3.16 that a 1-dim representation of semisimple Lie algebra is trivial. Thus, a representation $\phi : L \to \mathfrak{gl}(V)$ with $\dim V = 1$ has image $\mathrm{Im}(\phi) = 0$.
- We have also seen in example 2.3.15 the 2-dim representation of L by the inclusion map $\phi : L \to L \subseteq \mathfrak{gl}(F^2) \cong \mathfrak{gl}(2,F)$.
- Since $\dim L = \dim \mathfrak{sl}(2,F) = 3$, we see $\mathrm{ad} : L \to \mathfrak{gl}(L)$ is a 3-dim representation of L.
- For general dimension, we have seen V_d is an irreducible representation of L with dimension $d+1$.

In fact, it can be shown that every irreducible $\mathfrak{sl}(2,F)$-module is isomorphic to V_d with some d (see [4] Theorem 8.5).

Let V be any irreducible L-module. We choose a maximal vector $v_0 \in V_\lambda$ (recall that the λ is such that $V_\lambda \neq 0$ and $V_{\lambda+2} = 0$). We set $v_{-1} = 0$, $v_i = \left(\frac{1}{i!}\right) y^i_. v_0$ for $i \geqslant 0$. Then one can compute that

Lemma 2.4.7.

(a) $h_. v_i = (\lambda - 2i) v_i$

(b) $y_. v_i = (i+1) v_{i+1}$,

(c) $x_. v_i = (\lambda - i + 1) v_{i-1}$ *for* $i \geqslant 0$.

Proof. See [1] Lemma 7.2. ■

(a) shows that all nonzero v_i are linearly independent. But $\dim V < \infty$. Let m be the smallest integer for which $v_m \neq 0, v_{m+1} = 0$; evidently $v_{m+i} = 0$ for all $i > 0$. Taken together, formulas (a)-(c) show that the subspace of V with basis $(v_0, v_1, \ldots, v_m)$ is an L-submodule, different from 0. Because V is irreducible, this subspace must be all of V. Moreover, relative to the ordered basis $(v_0, v_1, \ldots, v_m)$, the matrices of the endomorphisms ϕ_x, ϕ_y, ϕ_h representing x, y, h can be written down explicitly; notice that h yields a diagonal matrix, while x and y yield (respectively) upper and lower triangular nilpotent matrices (see what we did for V_d for example.)

A closer look at formula (c) reveals a striking fact: for $i = m+1$, the left side is 0, whereas the right side is $(\lambda - m)v_m$. Since $v_m \neq 0$, we conclude that $\lambda = m$. In other words, the weight of a maximal vector is a nonnegative integer (one less than $\dim V$). We call it the **highest weight** of V. Moreover, each weight μ occurs with multiplicity one (i.e., $\dim V_\mu = 1$ if $V_\mu \neq 0$), by formula (a); in particular, since V determines λ uniquely ($\lambda = \dim V - 1$), the maximal vector v_0 is the only possible one in V (apart from nonzero scalar multiples). To summarize :

Theorem 2.4.8. *Let* V *be an irreducible module for* $L = \mathfrak{sl}(2,F)$.

(a) Relative to h, V *is the direct sum of weight spaces* $V_\mu, \mu = m, m-2, \ldots, -(m-2), -m$, *where* $m+1 = \dim V$ *and* $\dim V_\mu = 1$ *for each* μ.

(b) V *has (up to nonzero scalar multiples) a unique maximal vector, whose weight (called the highest weight of* V*) is* m.

(c) *The action of L on V is given explicitly by the above formulas, if the basis is chosen in the prescribed fashion. In particular, there exists at most one irreducible L-module (up to isomorphism) of each possible dimension $m+1$, $m \geqslant 0$*

Corollary 2.4.9. *Let V be any (finite dimensional) L-module, $L = \mathfrak{sl}(2,\ \mathrm{F})$. Then the eigenvalues of h on V are all integers, and each occurs along with its negative (an equal number of times). Moreover, in any decomposition of V into direct sum of irreducible submodules, the number of summands is precisely* $\dim V_0 + \dim V_1$

Proof. If $V = 0$, there is nothing to prove. Otherwise use Weyl's Theorem to write V as direct sum of irreducible submodules. The latter are described by the theorem, so the first assertion of the corollary is obvious. For the second, just observe that each irreducible L-module has a unique occurrence of either the weight 0 or else the weight 1 (but not both). ■

2.5 Root Space Decomposition

L: f.d s.s Lie alg over F with $F = \bar{F}$ and $\operatorname{char}(F) = 0$. We will study the structure of such L in this section. Further abusing the notation for s, n as in the abstract Jordan decomposition, in a semisimple Lie algebra L, we will say an ad-nilpotent element nilpotent and an ad-s.s. element semisimple.

2.5.1 Maximal Toral Subalgebras and Roots

Definition 2.5.1. *A Lie subalgebra $H \subseteq L$ is **toral** if all its elements are semisimple.*

Lemma 2.5.2. *Toral subalgebras exist. There exists some element $x \in L$ such that x is semisimple. Then $F\{x\}$ toral.*

Lemma 2.5.3. *A toral subalgebra of L is abelian.*

Proof. Let $H \subseteq L$ be a toral subalg. Let $x \in H$, we must show $[x, y] = 0,\ \forall y \in H$; i.e., $\operatorname{ad}_H x = 0$. Since $\operatorname{ad}_H x$ is semistmple, it suffices to show all its e.values (eigenvalues) are all 0: Let $y \in H$ be an e.vector (eigenvector) of $\operatorname{ad}_H x$, so $[x, y] = (\operatorname{ad}_H x)(y) = \alpha y$ for some $\alpha \in F$. We have $[y, x] = -\alpha y \Rightarrow (\operatorname{ad}_H y)(x) = -\alpha y \implies (\operatorname{ad}_H y)^2 (x) = 0$. Since $ad_H y$ is semsimple, so there exists a basis $\{x_1, \ldots, x_n\}$ consisting of e-vec of $\operatorname{ad}_H y$. Let $x = \sum_i c_i x_i$ with $c_i \in F$.

$$\begin{aligned}
&(\operatorname{ad}_H y)(x) = \sum_i c_i \lambda_i x_i, \lambda_i \neq 0 \\
&\Rightarrow 0 = (\operatorname{ad}_H y)^2 (x) = \sum_i c_i \lambda_i^2 x_i \\
&\Rightarrow c_i \lambda_i^2 = 0,\ \forall i \\
&\Rightarrow c_i = 0,\ \forall i \text{ s.t } \lambda_i \neq 0
\end{aligned}$$

Thus, $-\alpha y = (\operatorname{ad}_H y)(x) = \sum_i c_i \lambda_i x_i = 0 \Rightarrow \alpha = 0$. ■

Let H be a **maximum toral subalgebra**, i.e., a toral subalgebra not properly contained in any other. Then, the matrices $\{\operatorname{ad}_H h : h \in H\}$ are simutinuasly diagonalizable. Thus, L decomposes:

$$L = \bigoplus_{\alpha \in H*} L_\alpha, \text{ where } L_\alpha = \{x \in L | [h, x] = \alpha(h) x, \forall h \in H\}.$$

where $H*$ is the dual space of H. Notes:

(1) $L_0 = C_L(H)$.

(2) $H \subseteq L_0$ by the lemma.

(3) If $0 \neq \alpha \in H^*$ s.t. $L_\alpha \neq 0$, then α is a **root** of L relative to H. Let $\Phi \subseteq H^*$ be the **set of roots**.

Definition 2.5.4. *The* ***root space decomposition*** *of L is*

$$L = C_L(H) \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

Proposition 2.5.5. *For all $\alpha, \beta \in H^*$,*

(1) $[L_\alpha, L_\beta] \subset L_{\alpha+\beta}$.

(2) If $x \in L_\alpha, \alpha \neq 0$, then ad x is nilpotent.

(3) If $\alpha, \beta \in H^*$, and $\alpha + \beta \neq 0$, then L_α is orthogonal to L_β, relative to the Killing form κ of L.

Proof. (1) follows from the Jacobi identity: $x \in L_\alpha, y \in L_\beta$, $h \in H$ imply that ad $h([x, y]) = [[h, x], y] + [x, [h, y]] = \alpha(h)[x, y] + \beta(h)[x, y] = (\alpha + \beta)(h)[x, y]$.

(2): $\exists k > 0$ s.t. $L_{\alpha+\beta} = 0,\ \forall \beta \in \Phi$ and

$$\beta = 0 \Rightarrow (\operatorname{ad} x)^k = 0$$

(3): Find $h \in H$ for which $(\alpha + \beta)(h) \neq 0$. Then if $x \in L_\alpha, y \in L_\beta$, associativity of the form allows us to write $\kappa([h, x], y) = -\kappa([x, h], y) = -\kappa(x, [h, y])$, or $\alpha(h)\kappa(x, y) = -\beta(h)\kappa(x, y)$, or $(\alpha + \beta)(h)\kappa(x, y) = 0$. This forces $\kappa(x, y) = 0$. ■

Corollary 2.5.6. *The restriction of the Killing form to $L_0 = C_L(H)$ is nondegenerate.*

Proof. We know from Theorem 2.2.5 that κ is nondegenerate. Let $z \in L_0$ and suppose $\kappa(z, L_0) = 0$. Proposition (3) $\Rightarrow \kappa(z, L_\alpha) = 0,\ \forall \alpha \in \Phi$, so $k(z, L) = 0 \Rightarrow z = 0$. ■

Definition 2.5.7. *A* ***Cartan subalgebra (CSA)*** *of a Lie algebra L is a nilpotent subalgebra H of L that equals to the normalizer of it in L, i.e., $N_L(H) = H$.*

Remark 2.5.8. If L is semisimple, then maximal toral subalgebra H is a CSA of L. (HW). Furthermore, CSA $\Rightarrow$ maximal toral. ♠

2.5.2 Centralizer of H

We shall need a fact from linear algebra, whose proof is trivial:

Lemma 2.5.9. *If x, n are commuting endomorphisms of a finite dimensional vector space, with n nilpotent, then xn is nilpotent; in particular, $\operatorname{tr}(xn) = 0$.*

Remark 2.5.10. If H is a toral subalgebra with $H = C_L(H)$, then H is maximal: suppose $H = C_L(H) \subsetneq H'$ with H' another toral subalgebra, then Lemma 2.5.3 implies that H' is abelian, so every element in H' commutes with every element in $H \subsetneq H'$. In particular, $H' \subseteq C_L(H)$. Thus, H has to be maximal. ♠

Proposition 2.5.11. *Let H be a maximal toral subalgebra of L. Then $H = C_L(H)$.*

Proof. Let $C = C_L(H)$.

(1) We show that given abstract decomposition of $x \in C$ by $x = s + n$, we have $s, n \in C$:

Then the Jordan decomposition of $\operatorname{ad} x$ is $\operatorname{ad} s + \operatorname{ad} n$. $x \in C \Rightarrow [x, H] = 0 \Rightarrow (\operatorname{ad} x)(H) = 0 \Rightarrow (\operatorname{ad} n)(H) = 0,\ (\operatorname{ad} s)(H) = 0 \Rightarrow n, s \in C$

(2) If $x \in C$ semisimple, then $x \in H$.

x s.s. $\Rightarrow H + F\{x\}$ is abelian and thus toral as sum of commuting semisimple elements is again semisimple by remark 2.1.6. By maximality of H, we see $x \in H$.

(3) The restriction of κ to H is nondegenerate.

That is, if $\kappa(h, H) = 0$ for some $h \in H$ then we must show that $h = 0$. If $x \in C$ is nilpotent, then the fact that $[x, H] = 0$ and the fact that ad x is nilpotent together imply (by the above lemma) that $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$ for all $y \in H$, or

$$(*) : \quad \kappa(x, H) = 0$$

We then claim that $\kappa(h, H) = 0 \implies \kappa(h, C) = 0$. Indeed, for $x = s + n \in C$, we have $\kappa(h, s + n) = \kappa(h, s) + \kappa(h, n)$. The second part is zero because (2) $\Rightarrow n \in C$ and $(*)$; the first part being zero because of (1) and the given condition $\kappa(h, H) = 0$.

Now, Corollary 2.5.6 implies $h = 0$.

(4) C is nilpotent.

If $x \in C$ is semisimple, then $x \in H$ by (2), and $\operatorname{ad}_C x (= 0)$ is certainly nilpotent, so semisimple elements are nilpotent in C. On the other hand, if $x \in C$ is nilpotent, then $\operatorname{ad}_C x$ is nilpotent. Now let $x \in C$ be arbitrary, $x = x_s + x_n$. Since both x_s, x_n lie in C by (1), $\operatorname{ad}_C x$ is the sum of commuting nilpotents and is therefore itself nilpotent by remark 2.1.6. By Engel's Theorem, C is nilpotent.

(5) $H \cap [C, C] = 0$.

Since κ is associative and $[H, C] = 0, \kappa(H, [C, C]) = 0$. Now use (3).

(6) C is abelian.

Otherwise $[C, C] \neq 0$. C being nilpotent, by (4), $Z(C) \cap [C, C] \neq 0$ (Lemma 1.3.22). Let $z \neq 0$ lie in this intersection. By (2) and (5), z cannot be semisimple. Its nilpotent part n is therefore nonzero and lies in C, by (1), hence also lies in $Z(C)$ by Corollary 2.1.9. But then our lemma implies that $\kappa(n, C) = 0$, contrary to Corollary 2.5.6.

(7) $C = H$.

Otherwise C contains a nonzero nilpotent element, x, by (1), (2). According to the lemma and (6), $\kappa(x, y) = \operatorname{tr}(\operatorname{ad} x \text{ ad } y) = 0$ for all $y \in C$, contradicting Corollary 2.5.6.

■

Corollary 2.5.12. *The restriction of κ to H is nondegenerate.*

The corollary allows us to identify H with H^* canonically:

$$\begin{aligned} H &\to H^* \\ h &\mapsto \kappa(h, \cdot) \end{aligned}$$

This is an isomorphism by Riesz representation theorem (see [2] Theorem 11.5), so for each $\phi \in H^*$ corresponds the (unique) element $t_\phi \in H$ satisfying $\phi(h) = \kappa(t_\phi, h)$ for all $h \in H$. In particular, Φ corresponds to the subset $\{t_\alpha; \alpha \in \Phi\}$ of H.

2.5.3 Orthogonality Properties

In this subsection we shall obtain more precise information about the root space decomposition, using the Killing form. We already saw in Proposition 2.5.5 that $\kappa(L_\alpha, L_\beta) = 0$ if $\alpha, \beta \in H^*, \alpha + \beta \neq 0$; in particular, $\kappa(H, L_\alpha) = 0$ for all $\alpha \in \Phi$, so that (Proposition 2.5.11) the restriction of κ to H is nondegenerate.

Proposition 2.5.13.

(a) Φ *spans* H^*.

(b) *If* $\alpha \in \Phi$, *then* $-\alpha \in \Phi$.

(c) *Let* $\alpha \in \Phi, x \in L_\alpha, y \in L_{-\alpha}$. *Then* $[x, y] = \kappa(x, y)t_\alpha$.

(d) *If* $\alpha \in \Phi$, *then* $[L_\alpha, L_{-\alpha}]$ *is one dimensional, with basis* t_α.

(e) $\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \neq 0$, *for* $\alpha \in \Phi$.

(f) *If* $\alpha \in \Phi$ *and* x_α *is any nonzero element of* L_α, *then there exists* $y_\alpha \in L_{-\alpha}$ *such that* $x_\alpha, y_\alpha, h_\alpha = [x_\alpha, y_\alpha]$ *span a three dimensional simple subalgebra of* L *isomorphic to* $\mathfrak{sl}(2, F)$ *via* $x_\alpha \mapsto \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), y_\alpha \mapsto \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), h_\alpha \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.

(g) $h_\alpha = \dfrac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}; \quad h_\alpha = -h_{-\alpha}$.

Proof. (a) If Φ fails to span H^*, then there exists nonzero $h \in H$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$ (this is by duality: let $V = \{t_\alpha; \alpha \in \Phi\}$ be the corresponding subspace of H, then we can write $H = V \oplus V^\perp$ wrt. the symmetric bilinear form κ. Then that h is in $V^\perp$. [4] Lemma 10.11 explains this more clearly.) But this means that $[h, L_\alpha] = 0$ for all $\alpha \in \Phi$. Besides, H abelian, so $[h, L_0]$ is also 0. Thus, $[h, L] = 0$, or $h \in Z(L) = 0$ (by for example prop.1.3.8), which is absurd.

(b) Let $\alpha \in \Phi$. If $-\alpha \notin \Phi$ (i.e., $L_{-\alpha} = 0$), then for all $\beta \in H^*$, including $-\alpha$, we have $\kappa(L_\alpha, L_\beta) = 0$ by Proposition 2.5.5. Therefore $\kappa(L_\alpha, L) = 0$, contradicting the nondegeneracy of κ.

(c) Let $\alpha \in \Phi, x \in L_\alpha, y \in L_{-\alpha}$. Let $h \in H$ be arbitrary. The associativity of κ implies:

$$\kappa(h, [x, y]) = \kappa([h, x], y) = \alpha(h)\kappa(x, y) = \kappa(t_\alpha, h)\kappa(x, y) = \kappa(\kappa(x, y)t_\alpha, h) = \kappa(h, \kappa(x, y)t_\alpha). \tag{2.1}$$

This says that H is orthogonal to $[x, y] - \kappa(x, y)t_\alpha$, forcing $[x, y] = \kappa(x, y)t_\alpha$.

(d) Part (c) shows that t_α spans $[L_\alpha, L_{-\alpha}]$, provided $[L_\alpha, L_{-\alpha}] \neq 0$. Let $0 \neq x \in L_\alpha$. If $\kappa(x, L_{-\alpha}) = 0$, then $\kappa(x, L) = 0$ (cf. proof of (b)), which is absurd since κ is nondegenerate. Therefore we can find $0 \neq y \in L_{-\alpha}$ for which $\kappa(x, y) \neq 0$. By (c), $[x, y] \neq 0$.

(e) Suppose $\alpha(t_\alpha) = 0$, so that $[t_\alpha, x] = 0 = [t_\alpha, y]$ for all $x \in L_\alpha, y \in L_{-\alpha}$. As in (d), we can find such x, y satisfying $\kappa(x, y) \neq 0$. Modifying one or the other by a scalar, we may as well assume that $\kappa(x, y) = 1$. Then $[x, y] = t_\alpha$, by (c). It follows that the subspace S of L spanned by x, y, t_α is a three dimensional solvable algebra (for reasons similar to $\mathfrak{u}_3$ in example 1.3.2), $S \cong \operatorname{ad}_L S \subset \mathfrak{gl}(L)$. In particular, $\operatorname{ad}_L s$ is nilpotent for all $s \in [S, S]$ (Corollary 2.1.4), so $\operatorname{ad}_L t_\alpha$ is both semisimple and nilpotent, i.e., $\operatorname{ad}_L t_\alpha = 0$. This says that $t_\alpha \in Z(L) = 0$ (by for example prop.1.3.8), contrary to choice of t_α.

(f) Given $0 \neq x_\alpha \in L_\alpha$, find $y_\alpha \in L_{-\alpha}$ such that $\kappa(x_\alpha, y_\alpha) = \frac{2}{\kappa(t_\alpha, t_\alpha)}$. This is possible in view of (e) and the fact that $\kappa(x_\alpha, L_{-\alpha}) \neq 0$. Set $h_\alpha = 2t_\alpha / \kappa(t_\alpha, t_\alpha)$. Then $[x_\alpha, y_\alpha] = h_\alpha$, by (c). Moreover, $[h_\alpha, x_\alpha] = \frac{2}{\alpha(t_\alpha)}[t_\alpha, x_\alpha] = \frac{2\alpha(t_\alpha)}{\alpha(t_\alpha)}x_\alpha = 2x_\alpha$, and similarly, $[h_\alpha, y_\alpha] = -2y_\alpha$. So $x_\alpha, y_\alpha, h_\alpha$ span a three dimensional subalgebra of L with the same multiplication table as $\mathfrak{sl}(2, F)$.

(g) h_α is defined in (f). Recall that t_α is defined by $\kappa(t_\alpha, h) = \alpha(h)(h \in H)$. This shows that $t_\alpha = -t_{-\alpha}$ and thus $h_\alpha = -h_{-\alpha}$.

■

Definition 2.5.14. *For pair of roots* $\alpha, -\alpha$, *we denote the three dimensional simple subalgebra of* L *spanned by* $x_\alpha, y_\alpha, h_\alpha = [x_\alpha, y_\alpha] = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$ *as in (f) by*

$$S_\alpha = F\{x_\alpha, y_\alpha, h_\alpha\}.$$

Remark 2.5.15. Several small facts will be used.

(1) By above proposition and facts 2.5.5, we see h_α spans $[L_\alpha, L_{-\alpha}] \subseteq L_{\alpha+(-\alpha)} = L_0 = H$.

(2) Note that $\alpha(h_\alpha) = 2$: just observe $[h_\alpha, x_\alpha] = 2x_\alpha$ from the multiplication table of $S_\alpha \cong \mathfrak{sl}(2, F)$. Then $h_\alpha \in H$, $x_\alpha \in L_\alpha$ give $[h_\alpha, x_\alpha] = \alpha(h_\alpha)x_\alpha$.

♠

2.5.4 Integerality Properties

We are now in a position to apply the representation theory of $\mathfrak{sl}(2, F)$. Let $\alpha \in \Phi$. We may regard L as an S_α-module via restriction of the adjoint representation. That is, for $a \in S_\alpha$ and $y \in L$ we define the action as

$$a \cdot y = (\operatorname{ad} a)y = [a, y]$$

Note that the S_α-submodules of L are precisely the vector subspaces M of L such that $[s, m] \in M$ for all $s \in S_\alpha$ and $m \in M$. Of course, it is enough to check this when s is one of the standard basis elements $x_\alpha, y_\alpha, h_\alpha$. We shall also need the following lemma.

Lemma 2.5.16. *If M is an S_α-submodule of L, then the eigenvalues of h_α acting on M are integers.*

Proof. By Weyl's Theorem, M as a module by semisimple Lie algebra S_α may be decomposed into a direct sum of irreducible S_α-modules; for irreducible $\mathfrak{sl}(2, F)$-modules, the result follows from corollary 2.4.9. ■

Example 2.5.17.
(1) (cf. Exercise 2.4.6.) It is an exercise to show that the set $H = \left\{\begin{pmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix}\right\}$ consisting of all diagonal matrices in $\mathfrak{sl}(n, F)$ is the maximal toral subalgebra of $\mathfrak{sl}(n, F)$ (see Math547 HW3). Then we write

$$L = \mathfrak{sl}_3 = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

Note that

$$\begin{aligned} [E_{ii}, E_{kl}] &= \delta_{ik} E_{kl} - \delta_{il} E_{kl} \\ &= (\delta_{ik} - \delta_{il}) E_{kl} \\ &= (\varepsilon_k - \varepsilon_l)(E_{ii}) E_{kl} \end{aligned}$$

where $\varepsilon_i : H \to F$ s.t. $\varepsilon_i(\operatorname{diag}(h_1, h_2, h_3)) = h_i$. Then

$$L_{\varepsilon_k - \varepsilon_l} = \{x \in \mathfrak{sl}_3 \mid [h, x] = (\varepsilon_k - \varepsilon_l)(h)x, \ \forall h \in H\} = F\{E_{kl}\}.$$

$$L = \mathfrak{sl}_3 = \left\{\begin{pmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix}\right\} \oplus \bigoplus_{(k,l), k \neq l} L_{\varepsilon_k - \varepsilon_l}.$$

and

$$\Phi = \{\varepsilon_1 - \varepsilon_2, \varepsilon_1 - \varepsilon_3, \varepsilon_2 - \varepsilon_3, \varepsilon_3 - \varepsilon_2, \varepsilon_3 - \varepsilon_1, \varepsilon_2 - \varepsilon_1\} \subseteq H^*$$

(2) Let $U = H + S_\alpha$. Let $K = \ker \alpha \subseteq H$. By the rank-nullity formula, $\dim K = \dim H - 1$. (We know that $\dim \operatorname{Im} \alpha = 1$ as $\alpha(h_\alpha) = 2 \neq 0$ due to fact 2.5.15 (2).) Note that S_α acts trivially on $K = \ker(\alpha)$: for $k \in \ker(\alpha)$, we have

$$\begin{aligned} &\operatorname{ad}(x_\alpha)(k) = [x_\alpha, k] = -[k, x_\alpha] \xlongequal{x_\alpha \in L_\alpha} -\alpha(k)s \xlongequal{k \in K} 0 \\ &\operatorname{ad}(y_\alpha)(k) = [y_\alpha, k] = -[k, y_\alpha] \xlongequal{y_\alpha \in L_{-\alpha}} \alpha(k)s \xlongequal{k \in K} 0 \\ &\operatorname{ad}(h_\alpha)(k) = 0 \text{ bc. } h_\alpha \in H, \ k \in K \subseteq H, \text{ and } H \text{ is abelian.} \end{aligned}$$

Thus every element of S_α acts trivially on K. It follows that $U = K \oplus S_\alpha$ is a decomposition of U into S_α-modules. By Exercise 2.4.5(iii), the adjoint representation of S_α is isomorphic to V_2, so U is isomorphic to the direct sum of $\dim H - 1$ copies of the trivial representation, V_0, and one copy of the adjoint representation, V_2.

(3) If $\beta \in \Phi$ or $\beta = 0$, let

$$M := \bigoplus_c L_{\beta + c\alpha}$$

where the sum is over all $c \in F$ such that $\beta + c\alpha \in \Phi$. It follows from proposition 2.5.5 (1) that M is an S_α-submodule of L. This module is said to be the **α-root string through** β. ♣

Recall every irreducible S_α-module is

$$V(d) \cong F\left\{y^d\right\} \oplus F\left\{y^{d-1}x\right\} \oplus \cdots \oplus F\left\{x^d\right\}$$

and Theorem 2.4.8 writes that as

$$V(d) = V(d)_{-d} \oplus V(d)_{-d+2} \oplus \cdots V(d)_d$$

where each $V(d)_i$ is an h-eigenspace, and $\dim V(d)_i = 1$

Note:
$V(d)_0 \neq 0 \iff d$ is even;
$V(d)_1 \neq 0 \iff d$ is odd.

Let L be a semisimple Lie algebra and $L = H \oplus \bigoplus_{\alpha\in\Phi} L_\alpha$ be the root space decomposition of L. For a root $\alpha \in \Phi$, we realized L as an S_α-module.

Proposition 2.5.18.

$$S_\alpha = F\{h_\alpha\} \oplus L_\alpha \oplus L_{-\alpha}$$

L_α and $L_{-\alpha}$ are 1-dimensional and the only multiples of a root α which are roots are $\pm\alpha$.

Proof. Consider the following subspace of L,

$$M = H \oplus \bigoplus_{c\alpha\in\Phi,\ c\in F^*} L_{c\alpha}.$$

It is an S_α-submodule by Proposition 2.5.5 (i).

(i) The weights of M: eigenvalues of $\mathrm{ad}\,(h_\alpha)$. They are 0 and $2c = c\alpha\,(h_\alpha)$.
For the latter, we first recall $h_\alpha \in H$ from fact 2.5.15 (1). Then for $x \in L_{c\alpha}$, $c\alpha \in \Phi$, $c \in F^*$, we have $\mathrm{ad}(h_\alpha)x = [h_\alpha, x] = c\alpha(h_\alpha)x \implies c\alpha(h_\alpha)$ is an eigenvalue for the action of h_α on M. To see $c\alpha\,(h_\alpha) = 2c$, use fact 2.5.15 (2).

(ii) These weights are all integers by lemma 2.5.16. Thus, all c occurring here must be integral multiples of $1/2$.

(iii) We have shown this in Example 2.5.17 (2) that S_α acts trivially on $K = \ker(\alpha)$ and that $\dim K = \dim H - 1$. From fact 2.5.15 (1) we have $F\{h_\alpha\} = L_0 \subseteq H$, which is one-dimensional. $F\{h_\alpha\} \cap K = 0$ bc. $\alpha(h_\alpha) = 2 \neq 0$. Thus,

$$M_0 = \{x \in M | h_\alpha \cdot x = 0\} \overset{(i)}{=\!=} H = K \oplus F\{h_\alpha\}.$$

Also, relative to h_α, $K = \ker(\alpha)$ is the direct sum of $\dim H - 1$ copies of K_0 by Theorem 2.4.8 (a).

(v) S_α is itself an irreducible S_α-submodule of M.

Taken together, K and S_α exhaust the occurrences of the weight 0 for h_α and by Weyl's theorem we write

$$M = K \oplus S_\alpha \oplus W$$

where W is a complementary S_α-submodule. Since $H = K \oplus F\{h_\alpha\} \subseteq K \oplus F\{h_\alpha\} \oplus F\{x_\alpha\} \oplus F\{y_\alpha\} = K \oplus S_\alpha$, we see W only takes elements in $\bigoplus_{c\alpha\in\Phi, c\in F^*} L_{c\alpha}$. W as an S_α-submodule can be decomposed into irreducible submodules, which can be further written as weight spaces of h_α wrt. those irreducible submodules. However, these weights $2c \in \mathbb{Z}$ cannot be even, because by Theorem 2.4.8 (a) there will be weight spaces corresponding to zero weight, forcing some c to be zero, which is absurd as $c \in F^*$.
$\Longrightarrow$ the only even weights occurring in M are $0, \pm 2$; in particular, $2c \neq 4 \implies c \neq 2$, so $2\alpha \notin \Phi$.
$\Longrightarrow$ $c\alpha = \frac{1}{2}\alpha \notin \Phi$ (otherwise $2\beta \in \Phi$ for $\beta = \frac{1}{2}\alpha$), so $2c = 2\frac{1}{2} = 1$ is not a weight of h_α in M.
$\Longrightarrow$ W is zero and

$$M = K \oplus S_\alpha = K \oplus F\{h_\alpha\} \oplus L_\alpha \oplus L_{-\alpha}.$$

L_α and $L_{-\alpha}$ are 1-dimensional and the only multiples of a root α which are roots are $\pm\alpha$. ■

Proposition 2.5.19. *Suppose that $\alpha, \beta \in \Phi$ and $\beta \neq \pm\alpha$.*

(i) $\beta(h_\alpha) \in \mathbb{Z}$. *This is called* ***Cartan integer****.*

(ii) There are integers $r, q \geqslant 0$ such that if $k \in \mathbb{Z}$, then $\beta + k\alpha \in \Phi$ if and only if $-r \leqslant k \leqslant q$. Moreover, $r - q = \beta(h_\alpha)$.

(iii) If $\alpha + \beta \in \Phi$, then $[x_\alpha, x_\beta]$ is a non-zero scalar multiple of $x_{\alpha+\beta}$.

(iv) $\beta - \beta(h_\alpha)\alpha \in \Phi$.

Proof. Let $M := \bigoplus_k L_{\beta+k\alpha}$ be the root string of α through β. To prove (i), we note that $\beta(h_\alpha)$ is the eigenvalue of h_α acting on L_β, and so it lies in $\mathbb{Z}$.

We know from the previous proposition that $\dim L_{\beta+k\alpha} = 1$ whenever $\beta + k\alpha$ is a root, so the eigenspaces of ad h_α on M are all 1-dimensional and, since $(\beta + k\alpha)h_\alpha = \beta(h_\alpha) + 2k$, the eigenvalues of ad h_α on M are either all even or all odd. It now follows from corollary 2.4.9 that M is an irreducible S_α-module. Suppose that $M \cong V_d$. On V_d, the element h_α acts diagonally with eigenvalues

$$\{d, d-2, \ldots, -d\}$$

whereas on M, h_α acts diagonally with eigenvalues

$$\{\beta(h_\alpha) + 2k : \beta + k\alpha \in \Phi\}.$$

Equating these sets shows that if we define r and q by $d = \beta(h_\alpha) + 2q$ and $-d = \beta(h_\alpha) - 2r$, then (ii) will hold.

Suppose $v \in L_\beta$, so v belongs to the h_α-eigenspace where h_α acts as $\beta(h_\alpha)$. If $(\operatorname{ad} x_\alpha)x_\beta = 0$, then x_β is a highest-weight vector in the irreducible representation $M \cong V_d$, with highest weight $\beta(h_\alpha)$. If $\alpha + \beta$ is a root, then h_α acts on the associated root space as $(\alpha+\beta)h_\alpha = \beta(h_\alpha) + 2$. Therefore x_β is not in the highest weight space of the irreducible representation M, and so $(\operatorname{ad} x_\alpha) x_\beta \neq 0$. This proves (iii).

Finally, (iv) follows from part (ii) as

$$\beta - \beta(h_\alpha)\alpha = \beta - (r-q)\alpha$$

and $-r \leqslant -r + q \leqslant q$. ■

2.5.5 Rationality Properties

Proposition 2.5.20. *Let $\alpha, \beta \in \Phi$*

(i) $t_\alpha = \dfrac{h_\alpha}{\kappa(x_\alpha, y_\alpha)}$; $2t_\alpha\kappa(t_\alpha, t_\alpha)$;

(ii) $\kappa(h_\alpha, h_\beta) \in \mathbb{Z}$;

(iii) $\kappa(t_\alpha, t_\alpha) = \dfrac{4}{\kappa(h_\alpha, h_\alpha)} \in \mathbb{Q}$;

(iv) $\kappa(t_\alpha, t_\beta) \in \mathbb{Q}$.

Proof. (i) is from proposition 2.5.13 (c) and (g).

(ii): Since $h_\alpha, h_\beta \in H$, for any $x \in L_\gamma \subset H \oplus \bigoplus_{\gamma\in\Phi} L_\gamma$,

$$\operatorname{ad}(h_\alpha)\operatorname{ad}(h_\beta)x = [h_\alpha, [h_\beta, x]] = \gamma(h_\alpha)\gamma(h_\beta)x.$$

Thus,

$$\kappa(h_\alpha, h_\beta) = \operatorname{tr}(\operatorname{ad}(h_\alpha)\operatorname{ad}(h_\beta)) = \sum_{\gamma\in\Phi} \gamma(h_\alpha)\gamma(h_\beta) \in \mathbb{Z} \text{ by Prop.2.5.19 (i)}$$

(iii) follows immediately from (i) and (ii).

(iv) From (i) we have $t_\alpha = \frac{1}{2}h_\alpha\kappa(t_\alpha, t_\alpha)$, so

$$\kappa(t_\alpha, t_\beta) = \kappa\left(\frac{1}{2}h_\alpha\kappa(t_\alpha, t_\alpha), \frac{1}{2}h_\beta\kappa(t_\beta, t_\beta)\right)$$

which is clearly $\in \mathbb{Q}$ using (ii) and (iii). ■

We can translate the Killing form on H to obtain a non-degenerate symmetric bilinear form on H^*, denoted $(-,-)$. This form may be defined by

$$(\theta, \varphi) = \kappa(t_\theta, t_\varphi),$$

where t_θ and t_φ are the elements of H corresponding to θ and φ under the isomorphism $H \equiv H^*$ induced by κ. In particular, if α and β are roots, then

$$(\alpha, \beta) = \kappa(t_\alpha, t_\beta) \in \mathbb{Q}.$$

Exercise 2.5.21. *Show that*

$$\beta(h_\alpha) = \frac{2(\beta, \alpha)}{(\alpha, \alpha)}.$$

Solution.

$$\beta(h_\alpha) = \kappa(t_\beta, h_\alpha) = \kappa\left(t_\beta, \frac{2t_\alpha}{(t_\alpha, t_\alpha)}\right) = \frac{2(\beta, \alpha)}{(\alpha, \alpha)}.$$

♦

We know from Proposition 2.5.13 (a) that the roots of L span H^*, so H^* has a vector space basis consisting of roots, say $\{\alpha_1, \alpha_2, \ldots, \alpha_\ell\}$. We can now prove that something stronger is true as follows.

Lemma 2.5.22. *If β is a root, then β is a linear combination of the α_i with coefficients in $\mathbb{Q}$.*

Proof. Certainly we may write $\beta = \sum_{i=1}^{\ell} c_i\alpha_i$ with coefficients $c_i \in F$. For each j with $1 \leqslant j \leqslant \ell$, we have

$$(\beta, \alpha_j) = \sum_{i=1}^{\ell} (\alpha_i, \alpha_j) c_i.$$

We can write these equations in matrix form as

$$\begin{pmatrix} (\beta, \alpha_1) \\ \vdots \\ (\beta, \alpha_\ell) \end{pmatrix} = \begin{pmatrix} (\alpha_1, \alpha_1) & \ldots & (\alpha_\ell, \alpha_1) \\ \vdots & \ddots & \vdots \\ (\alpha_1, \alpha_\ell) & \ldots & (\alpha_\ell, \alpha_\ell) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_\ell \end{pmatrix}.$$

The matrix is the matrix of the non-degenerate bilinear form $(-,-)$ with respect to the chosen basis of roots, and so it is invertible (see [2] Theorem 11.3). Moreover, we have seen that its entries are rational numbers, so it has an inverse with entries in $\mathbb{Q}$. Since also $(\beta, \alpha_j) \in \mathbb{Q}$, the coefficients c_i are rational. ■

By this lemma, the $\mathbb{R}$-subspace of H^* spanned by the roots $\alpha_1, \ldots, \alpha_\ell$ contains all the roots of Φ and so does not depend on our particular choice of basis. Let E denote this subspace.

Proposition 2.5.23. *The form $(-,-)$ is a real-valued inner product on E, so E is a Euclidean space.*

Proof. Since $(-,-)$ is a symmetric bilinear form, we only need to check that the restriction of $(-,-)$ to E is positive definite. Let $\theta \in E$ correspond to $t_\theta \in H$. Using the root space decomposition and the fact that $(\operatorname{ad} t_\theta) e_\beta = \beta(t_\theta) e_\beta$, we get

$$(\theta, \theta) = \kappa(t_\theta, t_\theta) = \sum_{\beta \in \Phi} \beta(t_\theta)^2 = \sum_{\beta \in \Phi} \kappa(t_\beta, t_\theta)^2 = \sum_{\beta \in \Phi} (\beta, \theta)^2.$$

As (β, θ) is real, the right-hand side is real and non-negative. Moreover, if $(\theta, \theta) = 0$, then $\beta(t_\theta) = 0$ for all roots β, so by (proof of) proposition 2.5.13 (a), $\theta = 0$. ■

We summarize the results from proposition 2.5.13 (a), (b), proposition 2.5.18, and proposition 2.5.19 (i), (iv) in view of exercise 2.5.21.

Theorem 2.5.24. *L, H, Φ, E as above. Then:*

(a) Φ spans E, and 0 does not belong to Φ.

(b) If $\alpha \in \Phi$ then $-\alpha \in \Phi$, but no other scalar multiple of α is a root.

(c) If $\alpha, \beta \in \Phi$, then $\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha \in \Phi$.

(d) If $\alpha, \beta \in \Phi$, then $\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.

In the language of Chapter III, the theorem asserts that Φ is a **root system** in the real euclidean space E. We have therefore set up a correspondence $(L, H) \mapsto (\Phi, \mathrm{E})$. Pairs (Φ, E) will be completely classified in Chapter III. Later (Chapters IV and V) it will be seen that the correspondence here is actually $1-1$, and that the apparent dependence of Φ on the choice of H is not essential.

Chapter 3

Root Systems

In this chapter, we are concerned with a fixed Euclidean space E, i.e., a finite dimensional vector space over $\mathbb{R}$ endowed with a positive definite symmetric bilinear form (α, β).

3.1 Definitions

Notation:

- Any $0 \neq \alpha \in E$ defines a **hyperplance**

$$H_\alpha := \{\beta \in E \mid (\beta, \alpha) = 0\} = \mathbb{R}\{x\}^\perp$$

- $\sigma_\alpha \in \mathrm{End}(E)$ is called a **reflection** across H_α if σ_α fixes H_α pointwise, and $\sigma_\alpha(\alpha) = -\alpha$. One can verify the explicit formula,

$$\sigma_\alpha(\beta) = \beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha$$

- Since the number $\frac{2(\beta,\alpha)}{(\alpha,\alpha)}$ appears frequently, we denote it by $\langle \beta, \alpha \rangle$. Notice that $\langle \beta, \alpha \rangle$ is linearly only in the first variable.

Definition 3.1.1. *A subset R of a real inner-product space E is a **root system** if it satisfies the following axioms.*

(R1) R is finite, it spans E, and it does not contain 0.

(R2) If $\alpha \in R$, then the only scalar multiples of α in R are $\pm\alpha$.

(R3) If $\alpha \in R$, then the reflection σ_α permutes the elements of R.

(R4) If $\alpha, \beta \in R$, then $\langle \beta, \alpha \rangle \in \mathbb{Z}$.

*The elements of R are called **roots**.*

Example 3.1.2. (1) The roots Φ for semisimple Lie algebra L over algebraically closed char-0 F form a root system for the real span $E = \mathbb{R}\Phi$ of Φ.

(2) [4] Exercise 11.1: Consider $\mathbb{R}^{\ell+1}$, with the Euclidean inner product. Let ε_i be the vector in $\mathbb{R}^{\ell+1}$ with i-th entry 1 and all other entries zero. Define

$$R := \{\pm(\varepsilon_i - \varepsilon_j) : 1 \leqslant i < j \leqslant \ell + 1\}$$

and let $E = \operatorname{Span} R = \{\sum \alpha_i \varepsilon_i : \sum \alpha_i = 0\}$. Show that R is a root system in E. ♣

Notations: Let R be a root system for inner-product space E.
(1) $\dim(E)$ is the **rank** of R.
(2) $\langle \beta, \alpha \rangle = \dfrac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z}$.
(3) We also let **projection of** β **along** α be denoted by $\operatorname{proj}_\alpha \beta = \dfrac{(\beta, \alpha)}{(\alpha, \alpha)} \alpha = \dfrac{1}{2} \langle \beta, \alpha \rangle \alpha$.
(4) $\forall v \in E, \quad (v, v) = \|v\|^2$.
(5) **Weyl group** of R is the group of invertible linear transformations of E generated by the reflections σ_α for $\alpha \in R$, i.e., $W(R) := \langle \sigma_\alpha \mid \alpha \in R \rangle$.
(6) The root system R is called **decomposable** if there is a proper decomposition $R = R_1 \cup R_2$ such that $\forall \alpha_1 \in R_1, \forall \alpha_2 \in R_2 : (\alpha_1, \alpha_2) = 0$. Otherwise it is called **indecomposable** or **irreducible**.

Lemma 3.1.3 (Finiteness Lemma)**.** *Suppose that R is a root system in the real inner-product space E. Let $\alpha, \beta \in R$ with $\beta \neq \pm\alpha$. Then*

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{0, 1, 2, 3\}.$$

Proof. Thanks to (R4), the product in question is an integer: We must establish the bounds. For any non-zero $v, w \in E$, the angle θ between v and w is such that $(v, w)^2 = (v, v)(w, w) \cos^2 \theta$. This gives

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4 \cos^2 \theta \leqslant 4$$

Suppose we have $\cos^2 \theta = 1$. Then θ is an integer multiple of π and so α and β are linearly dependent, contrary to our assumption. ■

We now use this lemma to show that there are only a few possibilities for the integers $\langle \alpha, \beta \rangle$. Take two roots α, β in a root system R with $\alpha \neq \pm\beta$. We may choose the labelling so that $(\beta, \beta) \geqslant (\alpha, \alpha)$ and hence

$$|\langle \beta, \alpha \rangle| = \frac{2|(\beta, \alpha)|}{(\alpha, \alpha)} \geqslant \frac{2|(\alpha, \beta)|}{(\beta, \beta)} = |\langle \alpha, \beta \rangle|.$$

By the Finiteness Lemma, the possibilities are:

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	θ	$\dfrac{\|\beta\|^2}{\|\alpha\|^2} = \dfrac{(\beta, \beta)}{(\alpha, \alpha)} = \dfrac{\langle \beta, \alpha \rangle}{\langle \alpha, \beta \rangle}$
0	0	$\pi/2$	undetermined
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

Table 3.1: Angles between α and β

Given roots α and β, we would like to know when their sum and difference lie in R. Our table gives some information about this question.

Proposition 3.1.4. *Let $\alpha, \beta \neq \pm\alpha \in R$. Assume $(\beta, \beta) \geqslant (\alpha, \alpha)$.*

(a) $(\alpha, \beta) > 0 \iff \theta(\alpha, \beta)$ *obtuse* $\iff \alpha + \beta \in R$

(b) $(\alpha, \beta) < 0 \iff \theta(\alpha, \beta)$ *acute* $\iff \alpha - \beta \in R.$

Proof. $(\alpha, \beta) > 0 \iff \langle \alpha, \beta \rangle = \dfrac{2(\alpha, \beta)}{(\beta, \beta)} > 0.$

The table shows that if θ is acute, then $\langle\alpha, \beta\rangle = 1 > 0$, and if θ is obtuse, then $\langle\alpha, \beta\rangle = -1 < 0$.

By (R3), we know that $\sigma_\beta(\alpha) = \alpha - \langle\alpha, \beta\rangle\beta$ lies in R, which is either 1 (when θ is obstuse) or -1 (when θ is acute). ■

3.2 Examples

Call $\ell = \dim E$ the rank of the root system R. When $\ell \leqslant 2$, we can describe R by simply drawing a picture.

We shall immitate the pictures from here and here, the latter using Fulton-Harris style (see Fig.3.3 for example).

3.2.1 Root Systems of Rank 1

If we choose any non-zero vector $\alpha \in \mathbb{R}$, then $R = \{\alpha, -\alpha\}$ is a root system. Since any other non-zero vector is a multiple of α, property (R2) forbids us to add more vectors to our root system. Therefore in rank 1 there is only one possible root system - it is called A_1.

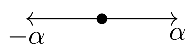


Figure 3.1: The root system A_1.

3.2.2 Root Systems of Rank 2

In rank 2 there is more freedom, because we can use any angle θ given in Table 3.1.

When the angle between the two roots is $\theta = \dfrac{\pi}{2}$, the system is called $A_1 \times A_1$, because it is a direct sum of two rank 1 root systems A_1.

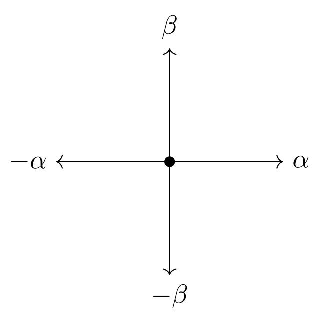


Figure 3.2: The root system $A_1 \times A_1$.

When $\theta = \dfrac{\pi}{3}, \dfrac{2\pi}{3}$, we place α on the positive x-axis and β by the $\frac{2\pi}{3}$ rotation of α (note that they have same length by the last column of the table). Then Proposition 3.1.4 says that there is also a root $\alpha + \beta$, which is drawn by parallelogram rule. Now, all of their negatives live in R too, completing the remaining three roots in the drawing. This root system is called A_2.

When $\theta = \dfrac{\pi}{4}, \dfrac{3\pi}{4}$, the root system consists of 8 vectors. They correspond to the vertices and to the midpoints of the edges of a regular square. The ratio of lengths of these roots is $\sqrt{2}$. This root system is called B_2.

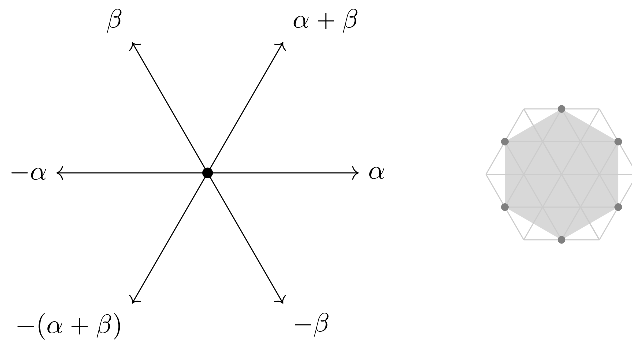


Figure 3.3: The root system A_2.

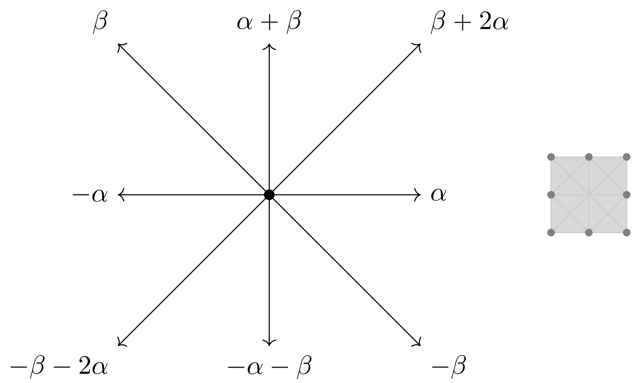


Figure 3.4: The root system B_2.

When $\theta = \dfrac{\pi}{6}, \dfrac{5\pi}{6}$, the root system consists of 12 vectors. They correspond to the vertices of two regular hexagons that have different sizes and are rotated away from each other by an angle $\pi/6$. The ratio of lengths of these vectors is $\sqrt{3}$. This is called G_2.

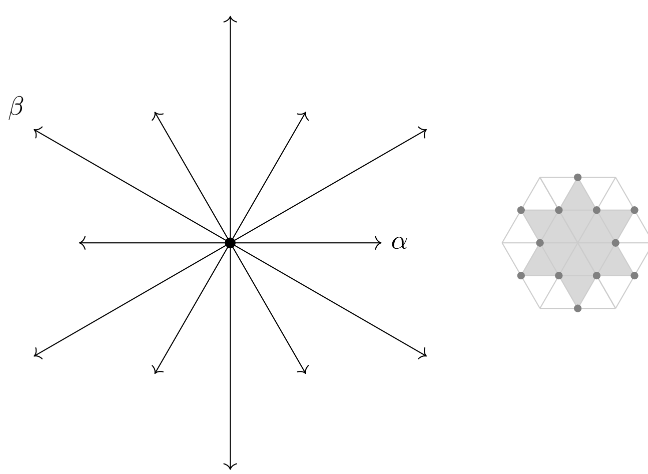


Figure 3.5: The root system G_2.

It is not hard to see, that there are no other root systems of rank 2, because in two dimensions the angle θ

determines the root system completely, i.e., once the angle is chosen, the ratio of lengths of two consecutive roots is determined (except for the case $\theta = \pi/2$), hence the root system itself.

Bibliography

[1] Humphreys, James E. *Introduction to Lie Algebras and Representation Theory*, Vol. 9. Springer Science & Business Media, 2012.

[2] Steven Roman, *Advanced Linear Algebra*, Vol. 3, New York: Springer, 2005.

[3] Roe Goodman and Nolan R. Wallach, *Symmetry, Representations, and Invariants*, Vol. 255, Dordrecht: Springer, 2009.

[4] Karin Erdmann and Mark J. Wildon, *Introduction to Lie Algebras*, Vol. 122, London: Springer, 2006.