LIE ALGEBRAS AND LIE GROUPS

SERGEY MOZGOVOY

Contents

1. Lie algebras 2
1.1. Definition and examples 2
1.2. Derivations and homomorphisms 6
1.3. Vector fields on manifolds 8
2. Lie groups 10
2.1. Linear groups 10
2.2. Exponential map 12
2.3. Lie algebra of a Lie group 15
2.4. Properties of $L(G)$ 17
2.5. Lie group homomorphisms 19
3. Nilpotent and solvable Lie algebras 22
3.1. Ideals 22
3.2. Representations 24
3.3. Nilpotent Lie algebras 26
3.4. Solvable Lie algebras 28
4. Semisimple and reductive Lie algebras 31
4.1. Semisimple and reductive Lie algebras 31
4.2. Reductivity criterion 33
4.3. Cartan's criterion 35
4.4. Irreducible and completely reducible representations 38
4.5. Casimir operator 40
4.6. Complete reducibility (Weyl's theorem) 41
4.7. Representations of $\mathfrak{sl}_2$ 43
5. Classification of semisimple Lie algebras 47
5.1. Jordan-Chevalley decomposition 47
5.2. Jordan decomposition in semisimple Lie algebras 49
5.3. Root space decomposition 52
5.4. Abstract root systems 57
5.5. Pairs of roots 59
5.6. Bases of root systems 61
5.7. Dynkin diagrams and classification 63
6. Highest weight representations 66
6.1. Universal enveloping algebra 66
6.2. Weights 68
6.3. Construction of highest weight representations 70

Date: October 9, 2025.

1. Lie algebras

1.1. **Definition and examples.** In what follows, we will consider vector spaces over a fixed field K. We are mostly interested in the cases $K = \mathbb{R}$ and $K = \mathbb{C}$. Let A be an associative algebra over K. We define the commutator

$$[-,-]\colon A \times A \to A, \qquad [x,y] = xy - yx.$$

This map is bilinear and satisfies

(1) $[x,x] = 0$ for all $x \in A$ (anti-commutativity).
(2) $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$ for all $x,y,z \in A$ (Jacobi identity).

Exercise 1.1. Verify the Jacobi identity.

Sometimes we have an operation on a vector space as above that is not induced by an associative algebra structure. This gives rise to the notion of a Lie algebra.

Definition 1.2. A *Lie algebra* over a field K is a K-vector space L equipped with a bilinear map

$$[-,-]\colon L \times L \to L$$

called a Lie bracket, such that

(1) $[x,x] = 0$ for all $x \in L$ (anti-commutativity).
(2) $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$ for all $x,y,z \in L$ (Jacobi identity).

Remark 1.3. Anti-commutativity implies $[x,y] = -[y,x]$ as

$$0 = [x+y, x+y] = [x,x] + [x,y] + [y,x] + [y,y] = [x,y] + [y,x].$$

Conversely, if $[x,y] = -[y,x]$ for all $x,y \in L$, then $[x,x] = -[x,x]$ and therefore $2[x,x] = 0$. This implies $[x,x] = 0$ if $\operatorname{char} K \neq 2$.

Definition 1.4. Let L be a Lie algebra. A *Lie subalgebra* of L is a subspace $M \subseteq L$ such that $[x,y] \in M$ for all $x,y \in M$. Note that M, equipped with the induced Lie bracket, is itself a Lie algebra.

Example 1.5. Let A be an associative K-algebra. Then the vector space A, equipped with the commutator operation

$$[x,y] = xy - yx, \qquad x,y \in A,$$

is a Lie algebra, denoted by $L(A)$ (note that A and $L(A)$ coincide as vector spaces, but have different binary operations).

Example 1.6. In particular, if V is a vector space, then the vector space of endomorphisms

$$\operatorname{End}(V) = \{f\colon V \to V \text{ a linear map}\}$$

is an associative algebra (with the product given by the composition). The corresponding Lie algebra is called the *general linear Lie algebra* and is denoted by $\mathfrak{gl}(V)$. Note that the algebra $\operatorname{End}(K^n)$ is isomorphic to the algebra $M_n(K)$ of $n \times n$ matrices over K. The corresponding Lie algebra is also called the general linear Lie algebra and is denoted by $\mathfrak{gl}_n(K)$ or $\mathfrak{gl}(n,K)$.

Example 1.7. Let us define a subspace

$$\mathfrak{sl}_n(K) = \mathfrak{sl}(n,K) = \{A \in \mathfrak{gl}_n(K) \mid \operatorname{tr} A = 0\} \subseteq \mathfrak{gl}_n(K).$$

For any $A, B \in \mathfrak{gl}_n(K)$, we have

$$\operatorname{tr}[A,B] = \operatorname{tr}(AB - BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$$

and therefore $[A, B] \in \mathfrak{sl}_n(K)$. This implies that $\mathfrak{sl}_n(K) \subseteq \mathfrak{gl}_n(K)$ is a Lie subalgebra. It is called the *special linear Lie algebra*.

In particular, let us consider the algebra

$$\mathfrak{sl}_2(K) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathfrak{gl}_2(K) \,|\, a + d = 0\}.$$

It is a 3-dimensional algebra with a basis

$$x = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \qquad y = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right), \qquad h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right).$$

The multiplication rules are

$$[x, y] = h, \qquad [h, x] = 2x, \qquad [h, y] = -2y.$$

Note that the matrix product of $x, y \in M_2(K)$ is

$$xy = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) \notin \mathfrak{sl}_2(K).$$

This means that $\mathfrak{sl}_2(K) \subseteq M_2(K)$ is not closed under the matrix multiplication, although it is closed under the commutator. This is an example of a Lie algebra that is not induced by an associative structure.

Example 1.8. Let us define the algebra of upper-triangular matrices

$$\mathfrak{b}_n(K) = \{(a_{ij}) \in \mathfrak{gl}_n(K) \,|\, a_{ij} = 0 \text{ for } i > j\},$$

the algebra of strictly upper-triangular matrices

$$\mathfrak{n}_n(K) = \{(a_{ij}) \in \mathfrak{gl}_n(K) \,|\, a_{ij} = 0 \text{ for } i \geq j\},$$

the algebra of diagonal matrices

$$\mathfrak{t}_n(K) = \{(a_{ij}) \in \mathfrak{gl}_n(K) \,|\, a_{ij} = 0 \text{ for } i \neq j\}.$$

All of them are Lie subalgebras of $\mathfrak{gl}_n(K)$.

Example 1.9. Let V be a K-vector space of dimension n and $\sigma\colon V \times V \to K$ be a bilinear form. We define

$$\mathfrak{o}(V, \sigma) = \{A \in \mathfrak{gl}(V) \,|\, \sigma(Av, w) = -\sigma(v, Aw)\ \forall v, w \in V\}.$$

It is a Lie subalgebra of $\mathfrak{gl}(V)$, called the *orthogonal Lie algebra*. Indeed, if $A, B \in \mathfrak{o}(V, \sigma)$, then

$$\begin{aligned}\sigma([A, B]v, w) = \sigma(ABv, w) - \sigma(BAv, w) &= -\sigma(Bv, Aw) + \sigma(Av, Bw) \\ &= \sigma(v, BAw) - \sigma(v, ABw) = -\sigma(v, [A, B]w)\end{aligned}$$

and therefore $[A, B] \in \mathfrak{o}(V, \sigma)$.

In particular, for $V = K^n$, there exists a (unique) matrix $S = (s_{ij})$ such that

$$\sigma(v, w) = v^{\mathsf{T}} S w \qquad \forall v, w \in V,$$

given by $s_{ij} = \sigma(e_i, e_j)$. The condition $\sigma(Av, w) = -\sigma(v, Aw)$ can be written in the form $(Av)^{\mathsf{T}} S w = v^{\mathsf{T}} A^{\mathsf{T}} S w = -v^{\mathsf{T}} S A w$. As it is satisfied for all $v, w \in V$, we conclude that $A^{\mathsf{T}} S = -SA$. Therefore

$$\mathfrak{o}(V, \sigma) = \mathfrak{o}(n, S) = \left\{A \in \mathfrak{gl}_n(K) \,\middle|\, A^{\mathsf{T}} S = -SA\right\}.$$

If σ is non-degenerate (meaning that S is invertible), then the above condition can be written as $S^{-1} A^{\mathsf{T}} S = -A$. Therefore

$$-\operatorname{tr}(A) = \operatorname{tr}(S^{-1} A^{\mathsf{T}} S) = \operatorname{tr}(S S^{-1} A^{\mathsf{T}}) = \operatorname{tr}(A^{\mathsf{T}}) = \operatorname{tr}(A),$$

hence $\operatorname{tr}(A) = 0$ and $A \in \mathfrak{sl}_n(K)$. For this reason the Lie algebra $\mathfrak{o}(V, \sigma)$ is also denoted by $\mathfrak{so}(V, \sigma) = \mathfrak{so}(n, S)$ and is also called the *special orthogonal Lie algebra*.

(1) If σ is symmetric non-degenerate and $K = \mathbb{C}$ (or any algebraically closed field with $\operatorname{char} K \neq 2$), then one can choose a basis of V such that σ is given by the identity matrix $S = I_n$. The algebra

$$\mathfrak{so}(n) = \mathfrak{so}(n, I_n) = \left\{ A \in \mathfrak{gl}_n(K) \,\middle|\, A^{\mathsf{T}} = -A \right\}$$

consists of skew-symmetric matrices and is called the *(special) orthogonal Lie algebra*.

(2) If σ is symmetric non-degenerate and $K = \mathbb{R}$ then one can choose a basis of V such that σ is given by the diagonal matrix $S = I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$ for some $p + q = n$. The algebra

$$\mathfrak{so}(p, q) = \mathfrak{so}(n, I_{p,q}) = \left\{ A \in \mathfrak{gl}_n(K) \,\middle|\, A^{\mathsf{T}} I_{p,q} = -I_{p,q} A \right\}$$

is called the *indefinite special orthogonal Lie algebra* of signature (p, q). The Lie algebra $\mathfrak{so}(1, 3)$ is called the *Lorentz Lie algebra*.

(3) Let σ be a skew-symmetric non-degenerate form. One can show that $n = 2\ell$ is even and there exists a basis of V such that σ is given by the matrix $S = \left(\begin{smallmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{smallmatrix}\right)$. The algebra

$$\mathfrak{sp}(2\ell) = \mathfrak{so}(n, S) = \left\{ A \in \mathfrak{gl}_n(K) \,\middle|\, A^{\mathsf{T}} S = -SA \right\}$$

is called the *symplectic Lie algebra*.

Example 1.10. For a matrix $A \in \mathfrak{gl}_n(\mathbb{C})$, the matrix $A^* = \bar{A}^{\mathsf{T}}$, with components $a^*_{ij} = \bar{a}_{ji}$, is called *conjugate transpose*. The set

$$\mathfrak{u}(n) = \{ A \in \mathfrak{gl}_n(\mathbb{C}) \,|\, A^* = -A \},$$

is a vector space over $\mathbb{R}$ (but not over $\mathbb{C}$). It is a Lie subalgebra of $\mathfrak{gl}(n, \mathbb{C})$ (over $\mathbb{R}$), called the *unitary Lie algebra*. For example

$$\mathfrak{u}(1) = \{ x + iy \in \mathbb{C} \,|\, x - iy = -(x + iy) \} = \{ x + iy \in \mathbb{C} \,|\, x = 0 \} = i\mathbb{R}.$$

We define the *special unitary Lie algebra* to be

$$\mathfrak{su}(n) = \mathfrak{u}(n) \cap \mathfrak{sl}(n, \mathbb{C}).$$

More generally, consider a matrix $S \in M_n(\mathbb{C})$ and a sesquilinear form on $V = \mathbb{C}^n$ given by $\sigma(v, w) = \bar{v}^{\mathsf{T}} S w$ (linear in the second argument and conjugate-linear in the first argument). Define

$$\mathfrak{u}(n, S) = \{ A \in \mathfrak{gl}_n(\mathbb{C}) \,|\, \sigma(Av, w) = -\sigma(v, Aw) \ \forall v, w \} = \{ A \in \mathfrak{gl}_n(\mathbb{C}) \,|\, A^* S = -SA \}$$

which is a Lie algebra over $\mathbb{R}$ (but not over $\mathbb{C}$). Note that if S is invertible, then $S^{-1} A^* S = -A$, hence $-\operatorname{tr}(A) = \operatorname{tr}(S^{-1} A^* S) = \operatorname{tr}(A^*) = \overline{\operatorname{tr}(A)}$. This implies that $\operatorname{tr}(A) \in i\mathbb{R}$.

Example 1.11 (Heisenberg algebra)**.** Let $\mathbb{C}[t, t^{-1}]$ be the algebra of Laurent polynomials and let $\mathcal{H} = \mathbb{C}[t, t^{-1}] \oplus \mathbb{C}K$, where $\mathbb{C}K$ is a one dimensional space generated by an element K. Define the bracket on $\mathcal{H}$

$$[f, g] = \operatorname{Res}(f'g) \cdot K = \operatorname{coeff}_{t^{-1}}(f'g) \cdot K \qquad \forall f, g \in \mathbb{C}[t, t^{-1}]$$

and $[f, K] = 0$ for all f (K is called a *central* element in this case).

Note that $(t^n)' = nt^{n-1}$ has zero coefficient in degree -1 for all $n \in \mathbb{Z}$. Therefore $\operatorname{Res}(f') = 0$ for all $f \in \mathbb{C}[t, t^{-1}]$. We conclude that $\operatorname{Res}(f'g) + \operatorname{Res}(fg') = \operatorname{Res}((fg)') = 0$, hence

$$\operatorname{Res}(f'g) = -\operatorname{Res}(fg'),$$

$$[f, g] = \operatorname{Res}(f'g)K = -\operatorname{Res}(g'f)K = -[g, f],$$

meaning that the above bracket is anti-symmetric. The vector space $\mathcal{H}$ equipped with the above bracket is a Lie algebra called the *Heisenberg algebra*. We have

$$[t^m, t^n] = \operatorname{Res}(mt^{m+n-1})K = m\delta_{m,-n}K.$$

Remark 1.12. More generally, let V be a vector space and $\omega\colon V \times V \to \mathbb{C}$ be a skew-symmetric form, meaning that

$$\omega(x, y) = -\omega(y, x), \qquad x, y \in V.$$

We define the *Heisenberg algebra* $\mathcal{H} = V \oplus \mathbb{C}K$ with the Lie bracket

$$[x, y] = \omega(x, y)K, \qquad x, y \in V,$$

and central K. The initial example is obtained by considering $V = \mathbb{C}[t, t^{-1}]$ and the skew-symmetric bilinear form ω on V given by

$$\omega(f, g) = \operatorname{Res}(f'g) = -\operatorname{Res}(fg') = -\omega(g, f), \qquad f, g \in \mathbb{C}[t, t^{-1}].$$

Exercise 1.13. Consider the vector space $V = \mathbb{C}x \oplus \mathbb{C}y$ with the skew-symmetric form ω given by $\omega(x, y) = 1 = -\omega(y, x)$. Then the Heisenberg algebra $\mathcal{H} = V \oplus \mathbb{C}K$ is 3-dimensional, with $[x, y] = K$ and central K. Show that

(1) $\mathcal{H}$ is **not** isomorphic to the Lie algebra $\mathfrak{b}_2(\mathbb{C})$ of upper-triangular 2×2 matrices or to the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$.
(2) $\mathcal{H}$ is isomorphic to the Lie algebra $\mathfrak{n}_3(\mathbb{C})$ of strictly upper-triangular 3×3 matrices.

1.2. **Derivations and homomorphisms.**

Definition 1.14. The algebras in this definition are not necessarily associative or Lie algebras.

(1) A *K-algebra* (or an algebra over a field K) is a K-vector space A equipped with a bilinear map
$$A \times A \to A, \qquad (a, b) \mapsto a \cdot b.$$
(2) A K-linear map $f\colon A \to B$ between two K-algebras is called an *algebra homomorphism* if $f(a \cdot b) = f(a) \cdot f(b)$ for all $a, b \in A$. It is called an *algebra isomorphism* if f is bijective.
(3) Let A be a K-algebra. A linear map $\delta\colon A \to A$ is called a *derivation* if it satisfies the *Leibniz rule*
$$\delta(a \cdot b) = \delta(a) \cdot b + a \cdot \delta(b) \qquad \forall a, b \in A.$$
The set of all derivations of A is a vector space, denoted by $\operatorname{Der} A$.

Remark 1.15. Note that

(1) An algebra homomorphism between Lie algebras is also called a *Lie (algebra) homomorphism*.
(2) In the case of unital associative algebras, a homomorphism $f\colon A \to B$ should also map 1_A to 1_B.

Lemma 1.16. *Let A be a K-algebra. For any $\delta, \delta' \in \operatorname{Der} A$, the commutator*
$$[\delta, \delta'] = \delta \circ \delta' - \delta' \circ \delta$$
is again a derivation. In particular, $\operatorname{Der} A$ *is a Lie subalgebra of* $\mathfrak{gl}(A)$.

Proof. Let $\delta, \delta' \in \operatorname{Der} A$. It is left as an exercise to show that, for any $x, y \in A$, we have
$$[\delta, \delta'](xy) = [\delta, \delta'](x) \cdot y + x \cdot [\delta, \delta'](y).$$
This implies that $[\delta, \delta'] \in \operatorname{Der} A$. One can also show that if $\delta, \delta' \in \operatorname{Der} A$ and $c \in K$ then
$$\delta + \delta' \in \operatorname{Der} A, \qquad c\delta \in \operatorname{Der} A.$$
This implies that $\operatorname{Der} A \subseteq \mathfrak{gl}(A)$ is a subspace closed under the bracket operation. □

Example 1.17 (Witt algebra)**.** Consider the algebra of polynomials $\mathbb{C}[t]$ (or the algebra of Laurent polynomials $\mathbb{C}[t, t^{-1}]$) and its Lie algebra of derivations $\operatorname{Der} \mathbb{C}[t]$, called the *Witt algebra*. There is an obvious derivation $\partial = \frac{d}{dt}$. Moreover, for any $f \in \mathbb{C}[t]$, the map
$$f(t)\partial\colon \mathbb{C}[t] \to \mathbb{C}[t], \qquad g(t) \mapsto f(t)g'(t),$$
is also a derivation. Conversely, given a derivation δ, let $f = \delta(t) \in \mathbb{C}[t]$. Then
$$\delta(t^n) = \delta(\underbrace{t \cdot t \cdot \ldots \cdot t}_{n \text{ times}}) = \sum_{i=1}^{n} t^{n-1}\delta(t) = nt^{n-1}f(t) = f(t)\partial(t^n)$$
and therefore $\delta = f(t)\partial$. We proved that $\operatorname{Der} \mathbb{C}[t] = \mathbb{C}[t]\partial$. One can show that the bracket is given by
$$[f\partial, g\partial] = (fg' - gf')\partial.$$
Note that $\operatorname{Der} \mathbb{C}[t]$ is not closed under the usual composition. Indeed, the map $\partial^2 = \partial \circ \partial$ is not a derivation:
$$\partial^2(t \cdot t) = (t^2)'' = 2 \neq \partial^2(t) \cdot t + t \cdot \partial^2(t) = 0.$$

Lemma 1.18. *Let L be a Lie algebra. Then, for any $x \in L$, the map*

$$\operatorname{ad}(x) = \operatorname{ad}_x \colon L \to L, \qquad y \mapsto [x, y]$$

is a derivation, called an inner derivation. The map

$$\operatorname{ad} \colon L \to \operatorname{Der} L \subseteq \mathfrak{gl}(L), \qquad x \mapsto \operatorname{ad}(x)$$

is a homomorphism of Lie algebras, called the adjoint representation of L.

Proof. The Leibniz rule for ad_x can be written as

$$\operatorname{ad}_x([y, z]) = [\operatorname{ad}_x(y), z] + [y, \operatorname{ad}_x(z)] \qquad \forall y, z \in L$$

or

$$[x, [y, z]] = [[x, y], z] + [y, [x, z]].$$

Using anti-commutativity we can rewrite this in the form

$$[x, [y, z]] = -[z, [x, y]] - [y, [z, x]]$$

which is equivalent to the Jacobi identity. This implies that $\operatorname{ad}_x \in \operatorname{Der} L$.

For any derivation $\delta \in \operatorname{Der} L$, we have

$$[\delta, \operatorname{ad}_y](z) = \delta(\operatorname{ad}_y(z)) - \operatorname{ad}_y(\delta(z)) = \delta([y, z]) - [y, \delta(z)] = [\delta(y), z] = \operatorname{ad}_{\delta(y)}(z)$$

and therefore

$$[\delta, \operatorname{ad}_y] = \operatorname{ad}_{\delta(y)}\,.$$

In particular, for $\delta = \operatorname{ad}_x$, we obtain

$$[\operatorname{ad}_x, \operatorname{ad}_y] = \operatorname{ad}_{[x,y]}$$

and this means that $\operatorname{ad} \colon L \to \operatorname{Der} L$ is a Lie algebra homomorphism. □

1.3. **Vector fields on manifolds.**

1.3.1. *Manifolds.*

Definition 1.19. Let M be a topological space and $d \geq 1$. A *chart* (or coordinate system) on M is a pair (U, ϕ), where $U \subseteq M$ is open and $\phi\colon U \to \mathbb{R}^d$ is a map such that $\phi(U) \subseteq \mathbb{R}^d$ is open and $\phi\colon U \to \phi(U)$ is a homeomorphism. An *atlas* (or differentiable structure) on M is a collection of charts $(U_i, \phi_i)_{i \in I}$ such that

(1) $M = \bigcup_{i \in I} U_i$.

(2) The composition $\phi_i(U_i \cap U_j) \xrightarrow{\phi_i^{-1}} U_i \cap U_j \xrightarrow{\phi_j} \phi_j(U_i \cap U_j)$ is C^∞-differentiable (smooth).

A Hausdorff topological space M equipped with an atlas as above (up to some equivalence) is called a (smooth differentiable) *manifold* of dimension d.

1.3.2. *Tangent vectors for embedded manifolds.* Assume that M is a manifold of dimension d embedded into $\mathbb{R}^n$. Given a point $p \in M$, we say that a vector $v \in \mathbb{R}^n$ is *tangent* to M at p if there exists a (smooth) curve $\gamma\colon (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = p$ and

$$\dot{\gamma}(0) = \lim_{t \to 0} \frac{\gamma(t) - \gamma(0)}{t} = v.$$

All such vectors form a d-dimensional subspace $T_pM \subseteq \mathbb{R}^n$, called the *tangent space* of M at the point p.

1.3.3. *Intrinsic definition of tangent vectors.* We can also give an intrinsic definition of tangent vectors at $p \in M$, independent of an embedding $M \subseteq \mathbb{R}^n$. We consider an equivalence relation on the set of (smooth) curves $\gamma\colon (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = p$, where

$$\gamma \sim \nu \qquad \Longleftrightarrow \qquad \lim_{t \to 0} \frac{\gamma(t) - \nu(t)}{t} = 0$$

and the limit is taken in some coordinate system (U, ϕ) on M around p, meaning that we substitute γ and ν by $\underline{\gamma} = \phi\gamma\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^d$ and $\underline{\nu} = \phi\nu\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^d$. The equivalence class of γ is denoted by $\dot{\gamma}(0)$ and is called a *tangent vector* at $p \in M$. We define the *tangent space* T_pM to be the set of all such equivalence classes. We can approximate $\underline{\gamma}(t) = \phi(p) + t\dot{\underline{\gamma}}(0) + o(t)$, for $|t| \ll 1$ and $\dot{\underline{\gamma}}(0) \in \mathbb{R}^d$. Then the equivalence class $\dot{\gamma}(0) \in T_pM$ is uniquely determined by the vector $\dot{\underline{\gamma}}(0) \in \mathbb{R}^d$ and we can identify T_pM with $T_{\phi(p)}\mathbb{R}^d \simeq \mathbb{R}^d$.

Remark 1.20. Tangent vectors can be written in local coordinates as follows. Assume that we have a coordinate system (U, ϕ) on M around p as before. It is customary to denote by

$$\left.\frac{\partial}{\partial x_1}\right|_p, \ldots, \left.\frac{\partial}{\partial x_d}\right|_p$$

the basis of T_pM corresponding to the standard basis of $\mathbb{R}^d$ under the identification $T_pM \simeq \mathbb{R}^d$ constructed above. A curve $\gamma\colon (-\varepsilon, \varepsilon) \to M$ with local coordinates

$$\underline{\gamma} = \phi\gamma = (\gamma_1, \ldots, \gamma_d)\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^d$$

corresponds to the tangent vector

$$\dot{\underline{\gamma}}(0) = (\gamma_i'(0))_i \in \mathbb{R}^d, \qquad \dot{\gamma}(0) = \sum_{i=1}^d \gamma_i'(0) \left.\frac{\partial}{\partial x_i}\right|_p \in T_pM.$$

1.3.4. *Tangent vectors as directional derivatives.* We can also interpret tangent vectors as directional derivatives. Given $v \in T_pM$, we define the directional derivative of a function $f \in C^\infty(M)$ as (we locally identify M with $\mathbb{R}^d$ using a chart)

$$D_v(f) = \lim_{t\to 0} \frac{f(p+tv) - f(p)}{t} \in \mathbb{R}.$$

If v corresponds to a curve $\gamma\colon (-\varepsilon, \varepsilon) \to M$ with $\gamma(0) = p$, then D_v is equal to

$$D_\gamma\colon C^\infty(M) \to \mathbb{R}, \qquad f \mapsto (f\gamma)'(0) = \sum_{i=1}^d \frac{\partial f}{\partial x_i}(p)\gamma_i'(0) = \sum_{i=1}^d \gamma_i'(0)\left.\frac{\partial}{\partial x_i}\right|_p (f).$$

This explains our notation $\left.\frac{\partial}{\partial x_i}\right|_p$ for the basis vectors of T_pM. The map D_γ is $\mathbb{R}$-linear and satisfies an analogue of the Leibniz rule

$$D_\gamma(fg) = D_\gamma(f)\cdot g(p) + f(p)\cdot D_\gamma(g). \tag{1}$$

There is a bijection between T_pM and such derivations. In what follows, we will identify tangent vectors $\dot\gamma(0) \in T_pM$ and derivations D_γ.

1.3.5. *Tangent space and vector fields.* The union $TM = \bigcup_{p\in M} T_pM$ has a structure of a smooth manifold of dimension $2d$. It is called the *tangent bundle* of M. There is smooth map $\pi\colon TM \to M$, sending T_pM to p and called the natural projection. A smooth section

$$X\colon M \to TM, \qquad p \mapsto X_p \in T_pM$$

of π is called a *vector field* on M. Given a vector field X, we define a map

$$L_X\colon C^\infty(M) \to C^\infty(M), \qquad (L_Xf)(p) = X_p(f), \qquad f \in C^\infty(M),\ p \in M,$$

called a *Lie derivative*. It is linear and satisfies the Leibniz rule because of (1). Therefore $L_X \in \operatorname{Der} C^\infty(M)$. One can show that this induces a bijection between the set of all vector fields on M and the set of derivations $\operatorname{Der} C^\infty(M)$. For any vector fields X, Y the commutator $[L_X, L_Y]$ is again a Lie derivative, hence it has a form $L_{[X,Y]}$ for some vector field $[X, Y]$. This bracket equips the space of vector fields with a structure of a Lie algebra, which was historically one of the first examples of Lie algebras.

Remark 1.21. Let us describe the bracket $[X, Y]$ in local coordinates. Given a chart (U, ϕ) around $p \in M$, we can write

$$X = \sum_{i=1}^d X_i\partial_i, \qquad Y = \sum_{i=1}^d Y_i\partial_i,$$

where $\partial_i = \frac{\partial}{\partial x_i}$ and $X_i, Y_i \in C^\infty(U)$. Then

$$\begin{aligned}[L_X, L_Y]f = \sum_{i,j}(X_i\partial_i(Y_j\partial_j(f)) - Y_i\partial_i(X_j\partial_j(f))) &= \sum_{i,j}(X_i\partial_i(Y_j)\partial_j(f) - Y_i\partial_i(X_j)\partial_j(f)) \\ &= \sum_{j=1}^d\left(\sum_{i=1}^d (X_i\partial_i(Y_j) - Y_i\partial_i(X_j))\right)\partial_j(f).\end{aligned}$$

Therefore

$$[X, Y] = \sum_{j=1}^d\left(\sum_{i=1}^d (X_i\partial_i(Y_j) - Y_i\partial_i(X_j))\right)\partial_j.$$

In particular, in dimension 1 we can write $X = f\partial$, $Y = g\partial$, where $\partial = \frac{\partial}{\partial x}$. Then $[X, Y] = (fg' - gf')\partial$.

2. Lie groups

Definition 2.1. A *Lie group* G is a group equipped with a smooth manifold structure such that the following maps are smooth (C^∞-differentiable)

$$G \times G \to G, \quad (x, y) \mapsto xy, \qquad G \to G, \quad x \mapsto x^{-1}.$$

Remark 2.2. Alternatively, one can require just that the following map is smooth

$$G \times G \to G, \qquad (x, y) \mapsto xy^{-1}.$$

Example 2.3.

(1) $(\mathbb{R}, +)$ is a Lie group, where the group operation is given by addition. More generally $(\mathbb{R}^n, +)$ and $(\mathbb{C}^n, +) \simeq (\mathbb{R}^{2n}, +)$ are Lie groups.
(2) Consider the group of invertible $n \times n$ matrices over $\mathbb{R}$

$$\mathrm{GL}_n(\mathbb{R}) = \mathrm{GL}(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) \,|\, \det A \neq 0\},$$

where the group operation is given by the product of matrices. It is an open subset of $M_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$, hence it is a manifold and a Lie group, called the *general linear group*.
(3) Similarly, the group

$$\mathrm{GL}_n(\mathbb{C}) = \mathrm{GL}(n, \mathbb{C}) = \{A \in M_n(\mathbb{C}) \,|\, \det A \neq 0\}$$

is an open subset of $M_n(\mathbb{C}) \simeq \mathbb{C}^{n^2} \simeq \mathbb{R}^{2n^2}$, hence is a manifold and a Lie group.
(4) The group

$$\mathrm{U}(1) = \{z \in \mathbb{C} \,|\, z\bar{z} = 1\} = \{z \in \mathbb{C} \,|\, |z| = 1\} \subseteq \mathbb{C}^* = \mathrm{GL}_1(\mathbb{C})$$

is a Lie group. Note that it can be identified with the unit circle S^1.

2.1. Linear groups.

An important family of Lie groups is the family of linear groups.

Definition 2.4. A *linear group* is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ (meaning a subgroup that is a closed subset with respect to the usual topology on $\mathrm{GL}_n(\mathbb{R}) \subseteq M_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$).

We will see later that every linear group is a Lie group and in particular it has a structure of a smooth manifold.

Example 2.5.

(1) The group $\mathrm{GL}_n(\mathbb{R})$ is a linear group, called the *general linear group*.
(2) The group

$$\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) \,|\, \det A = 1\}$$

is called the *special linear group*.
(3) The group

$$\mathrm{O}(n) = \left\{A \in \mathrm{GL}_n(\mathbb{R}) \,\middle|\, A^\mathsf{T} A = 1\right\}$$

is called the *orthogonal group*. More generally, consider a non-degenerate matrix $S \in M_n(\mathbb{R})$ and the corresponding bilinear form σ on $V = \mathbb{R}^n$

$$\sigma(v, w) = v^\mathsf{T} S w, \qquad v, w \in V.$$

A linear map $A\colon V \to V$ (or the corresponding matrix $A \in M_n(\mathbb{R})$) is called *σ-orthogonal* if

$$\sigma(Av, Aw) = \sigma(v, w) \qquad \forall v, w \in V$$

We define the orthogonal group

$$\begin{aligned}\mathrm{O}(V,\sigma) = \mathrm{O}(n,S) &= \{A \in M_n(\mathbb{R}) \,|\, \sigma(Av,Aw) = \sigma(v,w)\} \\ &= \big\{A \in M_n(\mathbb{R}) \,\big|\, A^{\mathsf{T}} S A = S\big\}.\end{aligned}$$

Note that $\det S = \det(A^{\mathsf{T}} S A) = \det(A)^2 \det S$, hence $\det(A) = \pm 1$ and A is invertible. Therefore $\mathrm{O}(n,S) \subseteq \mathrm{GL}_n(\mathbb{R})$ is a linear group (one needs to check that it is indeed a subgroup and a closed subset). Given a non-degenerate symmetric bilinear form σ on a vector space V over $\mathbb{R}$, one can find a basis of V such that the corresponding matrix is $S = I_{p,q} = \left(\begin{smallmatrix} I_p & 0 \\ 0 & -I_q \end{smallmatrix}\right)$ for some $p+q=n$. We define

$$\mathrm{O}(p,q) = \mathrm{O}(n, I_{p,q}) = \big\{A \in \mathrm{GL}_n(\mathbb{R}) \,\big|\, A^{\mathsf{T}} I_{p,q} A = I_{p,q}\big\}.$$

The group $\mathrm{O}(1,3)$ is called the Lorentz group.

(4) Define the *special orthogonal group*

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}_n(\mathbb{R}) = \big\{A \in \mathrm{GL}_n(\mathbb{R}) \,\big|\, A^{\mathsf{T}} A = 1, \det A = 1\big\}.$$

More generally, define

$$\mathrm{SO}(n,S) = \mathrm{O}(n,S) \cap \mathrm{SL}_n(\mathbb{R}), \qquad \mathrm{SO}(p,q) = \mathrm{O}(p,q) \cap \mathrm{SL}_{p+q}(\mathbb{R}).$$

(5) Let $\mathrm{GL}_n(\mathbb{C}) \subseteq M_n(\mathbb{C})$ be the group of invertible complex matrices. Every matrix $A \in \mathrm{GL}_n(\mathbb{C})$ induces a bijective $\mathbb{C}$-linear map $A\colon \mathbb{C}^n \to \mathbb{C}^n$, hence a bijective $\mathbb{R}$-linear map $R(A)\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$. In this way we obtain an injective group homomorphism

$$R\colon \mathrm{GL}_n(\mathbb{C}) \to \mathrm{GL}_{2n}(\mathbb{R}).$$

We can describe the image of this map as follows. Consider $V = \mathbb{C}^n = \mathbb{R}^{2n}$ and the linear map $J\colon V \to V$ given by multiplication by i. For $n=1$, we have $x+iy \mapsto -y+ix$, hence $J = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. An $\mathbb{R}$-linear map $A\colon V \to V$ is $\mathbb{C}$-linear if and only if $Ai = iA$, or equivalently if $AJ = JA$. We obtain

$$\mathrm{GL}_n(\mathbb{C}) \simeq \operatorname{Im} R = \{A \in \mathrm{GL}_{2n}(\mathbb{R}) \,|\, AJ = JA\} \subseteq \mathrm{GL}_{2n}(\mathbb{R}).$$

Therefore $\mathrm{GL}_n(\mathbb{C})$ is a closed subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$, hence a linear group. Note that $\mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{GL}_n(\mathbb{C})$ is a closed subgroup, hence every closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ is automatically a closed subgroup of $\mathrm{GL}_n(\mathbb{C})$. For this reason we could define linear groups also as closed subgroups of $\mathrm{GL}_n(\mathbb{C})$.

(6) Define the *unitary group*

$$\mathrm{U}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \,|\, A^* A = 1\}, \qquad A^* = \overline{A}^{\mathsf{T}},$$

and the *special unitary group*

$$\mathrm{SU}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \,|\, A^* A = 1, \det A = 1\}.$$

For example

$$\mathrm{U}(1) = \{z \in \mathbb{C}^* \,|\, \overline{z} z = 1\} = \{z \in \mathbb{C}^* \,||z| = 1\} \simeq S^1$$

and $\mathrm{SU}(1) = \{1\}$.

(7) The additive group $(\mathbb{R},+)$ can be interpreted as a linear group using an injective group homomorphism

$$(\mathbb{R},+) \to M_2(\mathbb{R}), \qquad x \mapsto \left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right).$$

Alternatively, we can consider a group isomorphism $\exp\colon (\mathbb{R},+) \to (\mathbb{R}_{>0},*)$ and note that $(\mathbb{R}_{>0},*)$ is a closed subgroup of $\mathrm{GL}_1(\mathbb{R})$.

2.2. **Exponential map.** The goal of this section is to study properties of the exponential map

$$\exp\colon \mathfrak{gl}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R}), \qquad A \mapsto \exp(A) = \sum_{n\geq 0} \frac{A^n}{n!} = I + A + \frac{A^2}{2} + \dots.$$

To prove that the above series is convergent, we need to recall first operator norms. Let us equip $\mathbb{R}^n$ with some norm $\|-\|$. For example,

$$\|x\|_2 = \left(\sum_i x_i^2\right)^{1/2}, \qquad \|x\|_1 = \sum_i |x_i|, \qquad \|x\|_\infty = \max_i |x_i|.$$

Any two norms on $\mathbb{R}^n$ are equivalent, meaning that there exist $C_1, C_2 > 0$ such that

$$C_1 \|x\| \leq \|x\|' \leq C_2 \|x\| \qquad \forall x \in \mathbb{R}^n.$$

Any matrix $A \in M_n(\mathbb{R})$ can be interpreted as a linear map (also called a linear operator) $A\colon \mathbb{R}^n \to \mathbb{R}^n$. This map is automatically bounded as

$$\left\|A\Big(\sum x_i e_i\Big)\right\| \leq \sum |x_i| \cdot \|Ae_i\| \leq \max_i \|Ae_i\| \cdot \|x\|_1 = C\|x\|_1, \qquad x = \sum x_i e_i,$$

and every norm is equivalent to the norm $\|-\|_1$. We define the operator norm (which depends on the fixed norm $\|-\|$ on $\mathbb{R}^n$) by

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}.$$

It is a norm on the vector space $M_n(\mathbb{R})$:

(1) $\|A\| = 0 \iff A = 0$.
(2) $\|\lambda A\| = |\lambda| \cdot \|A\|$, $\lambda \in \mathbb{R}$.
(3) $\|A + B\| \leq \|A\| + \|B\|$.

It also satisfies the property

(4) $\|AB\| \leq \|A\| \cdot \|B\|$.

Remark 2.6. Alternatively, we can define a particular norm on $M_n(\mathbb{R})$

$$\|A\| = \left(\sum_{i,j} a_{ij}^2\right)^{1/2}$$

and prove directly that $\|AB\| \leq \|A\| \cdot \|B\|$ for all $A, B \in M_n(\mathbb{R})$.

For any matrix $A \in M_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R})$, we define

$$e^A = \exp(A) = \sum_{n\geq 0} \frac{A^n}{n!} = I + A + \frac{A^2}{2} + \dots$$

This series is absolutely convergent (hence convergent) as

$$\sum_{n\geq 0} \frac{\|A^n\|}{n!} \leq \sum_{n\geq 0} \frac{\|A\|^n}{n!} = e^{\|A\|} < \infty.$$

Lemma 2.7. *We have*

(1) If $AB = BA$, then $e^A e^B = e^{A+B} = e^B e^A$.
(2) $e^A e^{-A} = I$. In particular, $e^A \in \mathrm{GL}_n(\mathbb{R})$.

Proof. **(1)** Using $AB = BA$, we obtain

$$e^{A+B} = \sum_n \frac{(A+B)^n}{n!} = \sum_n \frac{\sum_{i+j=n} \frac{n!}{i!j!} A^i B^j}{n!} = \sum_i \frac{A^i}{i!} \sum_j \frac{B^j}{j!} = e^A e^B.$$

(2) We have $e^A e^{-A} = e^{A+(-A)} = e^0 = I$. This implies that e^A is invertible. □

Lemma 2.8. *Assume that $A \in \mathfrak{gl}_n(\mathbb{R})$ has norm < 1. Then*

(1) $I + A$ is invertible and $(I+A)^{-1} = \sum_{n\geq 0}(-A)^n$ is (absolutely) convergent.
(2) The series $\log(I+A) = \sum_{n\geq 1} \frac{(-1)^{n-1}A^n}{n}$ is (absolutely) convergent.
(3) $e^{\log(I+A)} = I + A$.
(4) If $\|A\| < \log 2$, then $\|e^A - I\| < 1$ and $\log e^A = A$.

Proof. **(4)** For $B = e^A - I$, we have

$$\|B\| = \left\| \sum_{n\geq 1} \frac{A^n}{n} \right\| \leq \sum_{n\geq 1} \frac{\|A\|^n}{n} = e^{\|A\|} - 1 < 2 - 1 = 1.$$

This implies that the series $\log e^A = \log(I+B)$ is well-defined. We obtain the required equality by analyzing the summands of the series. □

Theorem 2.9. *The map $\exp\colon \mathfrak{gl}_n \to \mathrm{GL}_n$ is a local homeomorphism at $0 \in \mathfrak{gl}_n$. This means that there exists an open set $0 \in U \subseteq \mathfrak{gl}_n$ such that $V = \exp(U) \subseteq \mathrm{GL}_n$ is open and $\exp\colon U \to V$ is a homeomorphism.*

Proof. Let us choose $0 < \varepsilon < \log 2$ and define open subsets

$$U = \{A \in \mathfrak{gl}_n \mid \|A\| < \varepsilon\} \subseteq \mathfrak{gl}_n, \qquad W = \{C \in \mathrm{GL}_n \mid \|C - I\| < 1\} \subseteq \mathrm{GL}_n\,.$$

By the previous lemma, we have continuous maps

$$\exp\colon \mathfrak{gl}_n \to \mathrm{GL}_n, \qquad \log\colon W \to \mathfrak{gl}_n$$

such that $\exp(U) \subseteq W$, $\log e^A = A$ for $A \in U$, and $e^{\log C} = C$ for $C \in W$.

Let $V = \exp(U) \subseteq W$. If $C \in V$, then $C = e^A$, for some $A \in U$, and $\log C = \log e^A = A \in U$. Therefore the map log sends V to U and we conclude that the maps $\exp\colon U \to V$ and $\log\colon V \to U$ are inverse to each other. They are both continuous maps, hence $\exp\colon U \to V$ is a homeomorphism. To see that V is open, we note that $V = \exp(U) = \log^{-1}(U)$: if $C = e^A$, $A \in U$, then $\log C = \log e^A = A$, hence $C \in \log^{-1}(U)$; if $C \in \log^{-1}(U) \subseteq W$, then $\log(C) \in U$ and $C = e^{\log C} \in \exp(U)$. As log is continuous, we conclude that $\log^{-1}(U)$ is open. □

We have $e^{A+B} = e^A e^B$ if A and B commute, but not in general. Nevertheless, we have the following result.

Theorem 2.10. *We have $e^{A+B} = \lim_{n\to\infty}(e^{A/n}e^{B/n})^n$ for arbitrary $A, B \in M_n(\mathbb{R})$.*

Proof. We have $e^{A/n}e^{B/n} = I + (A+B)/n + O(1/n^2)$. If $n \gg 0$, then $e^{A/n}e^{B/n}$ is close to I and we can take the logarithm and obtain $\log(e^{A/n}e^{B/n}) = (A+B)/n + O(1/n^2)$. Multiplying by n and taking the exponential, we obtain

$$\exp\big(n\log(e^{A/n}e^{B/n})\big) = (e^{A/n}e^{B/n})^n = \exp(A + B + O(1/n)) \to \exp(A+B), \qquad n \to \infty.$$ □

Lemma 2.11. *For any $A \in \mathfrak{gl}_n(\mathbb{R})$ and $g \in \mathrm{GL}_n(\mathbb{R})$, we have*

$$e^{gAg^{-1}} = ge^Ag^{-1}.$$

Proof. We have $e^{gAg^{-1}} = \sum_{n\geq 0} \frac{(gAg^{-1})^n}{n!} = \sum_{n\geq 0} \frac{gA^ng^{-1}}{n!} = ge^Ag^{-1}$. □

Lemma 2.12. *For any $A \in \mathfrak{gl}_n(\mathbb{R})$, we have* $\det e^A = e^{\operatorname{tr} A}$.

Proof. Any matrix $A \in M_n(\mathbb{R})$ can be made upper triangular by conjugation in $M_n(\mathbb{C})$ (we can consider the Jordan normal form). Let $B = gAg^{-1}$ be upper triangular and $\lambda_i \in \mathbb{C}$ be its eigenvalues. Then e^B is an upper triangular matrix with eigenvalues e^{λ_i} and the determinant

$$\det e^B = \prod e^{\lambda_i} = e^{\sum \lambda_i} = e^{\operatorname{tr} B}.$$

We have

$$\det e^B = \det e^{gAg^{-1}} = \det(ge^A g^{-1}) = \det e^A, \qquad \operatorname{tr} B = \operatorname{tr}(gAg^{-1}) = \operatorname{tr} A.$$

Therefore $\det e^A = e^{\operatorname{tr} A}$. □

2.3. **Lie algebra of a Lie group.**

Definition 2.13. Given a linear group $G \subseteq \mathrm{GL}_n(\mathbb{R})$, we define

$$L(G) = \big\{ A \in \mathfrak{gl}_n(\mathbb{R}) \,\big|\, e^{tA} \in G \ \forall t \in \mathbb{R} \big\} \subseteq \mathfrak{gl}_n(\mathbb{R}).$$

We will see later that it is a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{R})$.

Example 2.14.

(1) Let $G = \mathrm{GL}_n(\mathbb{R})$. Then $e^{tA} \in G$ for all $A \in \mathfrak{gl}_n(\mathbb{R})$ and $t \in \mathbb{R}$, hence $A \in L(G)$. We conclude that $L(G) = \mathfrak{gl}_n(\mathbb{R})$.

(2) Let $G = \mathrm{SL}_n(\mathbb{R})$. If $A \in L(G)$, then $e^{tA} \in \mathrm{SL}_n(\mathbb{R})$, hence $\det e^{tA} = e^{t \operatorname{tr} A} = 1$ for all t. This implies $\operatorname{tr} A = 0$, hence $A \in \mathfrak{sl}_n(\mathbb{R})$. Conversely, if $A \in \mathfrak{sl}_n(\mathbb{R})$, then $\operatorname{tr} A = 0$, hence $\det e^{tA} = e^{t \operatorname{tr} A} = e^0 = 1$ and $e^{tA} \in G$ for all t. We conclude that $A \in L(G)$. This implies that $L(G) = \mathfrak{sl}_n(\mathbb{R})$.

Example 2.15. Let $G = \mathrm{O}(n) = \big\{ A \in \mathrm{GL}_n(\mathbb{R}) \,\big|\, A^{\mathsf{T}} A = 1 \big\}$. If $A \in L(G)$, then $I = (e^{sA})^{\mathsf{T}} e^{sA}$ for all s. Taking the derivative, we obtain

$$0 = \frac{d}{ds} (e^{sA})^{\mathsf{T}} e^{sA}|_{s=0} = \frac{d}{ds} (I + sA)^{\mathsf{T}} (I + sA)|_{s=0} = A^{\mathsf{T}} + A.$$

Conversely, if $A^{\mathsf{T}} + A = 0$ then A and A^{T} commute with each other and

$$(e^{sA})^{\mathsf{T}} e^{sA} = e^{sA^{\mathsf{T}}} e^{sA} = e^{s(A^{\mathsf{T}} + A)} = I.$$

This implies $e^{sA} \in G$, hence $A \in L(G)$. We conclude that

$$L(G) = \big\{ A \in \mathfrak{gl}_n(\mathbb{R}) \,\big|\, A^{\mathsf{T}} + A = 0 \big\} = \mathfrak{o}(n),$$

the orthogonal Lie algebra defined earlier.

We proved earlier that $\mathfrak{o}(n) = \mathfrak{o}(n) \cap \mathfrak{sl}_n(\mathbb{R}) = \mathfrak{so}(n)$: if $A \in \mathfrak{o}(n)$, then $A^{\mathsf{T}} + A = 0$, hence $2 \operatorname{tr} A = 0$ and $A \in \mathfrak{sl}_n(\mathbb{R})$. If $G = \mathrm{SO}(n) = O(n) \cap \mathrm{SL}_n(\mathbb{R})$, then $L(G) = \mathfrak{o}(n) \cap \mathfrak{sl}_n(\mathbb{R}) = \mathfrak{so}(n)$. We conclude that $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ have the same Lie algebra. Note that these groups are different, for example

$$\mathrm{O}(1) = \big\{ x \in \mathbb{R}^* \,\big|\, x^2 = 1 \big\} = \{\pm 1\}, \qquad \mathrm{SO}(1) = \{1\}.$$

Example 2.16. Similarly, consider a (non-degenerate) bilinear form $\sigma \colon V \times V \to \mathbb{R}$ and the corresponding orthogonal group

$$G = \mathrm{O}(V, \sigma) = \{ A \in \mathrm{GL}(V) \,|\, \sigma(Av, Aw) = \sigma(v, w) \ \forall v, w \in V \} \subseteq \mathrm{GL}(V).$$

If $A \in L(G)$, then $e^{tA} \in G$, hence

$$\sigma(e^{tA} v, e^{tA} w) = \sigma(v, w) \qquad \forall t \in \mathbb{R}.$$

Taking the derivative at $t = 0$ we obtain

$$\frac{d}{dt} \sigma((1 + tA)v, (1 + tA)w)|_{t=0} = \sigma(Av, w) + \sigma(v, Aw) = 0.$$

This implies that $A \in \mathfrak{o}(V, \sigma)$, the orthogonal Lie algebra defined earlier.

Conversely, let us show that every $A \in \mathfrak{o}(V, \sigma)$ satisfies $e^{tA} \in G$ for all $t \in \mathbb{R}$. Let S be the matrix of the bilinear form σ, meaning that $\sigma(v, w) = v^{\mathsf{T}} S w$, for $v, w \in V = \mathbb{R}^n$. Then we have $A^{\mathsf{T}} S + SA = 0$, or equivalently $SAS^{-1} = -A^{\mathsf{T}}$. We have to show that $(e^{tA})^{\mathsf{T}} S e^{tA} = S$ for all $t \in \mathbb{R}$, or equivalently $S e^{tA} S^{-1} = e^{-tA^{\mathsf{T}}}$. But

$$S e^{tA} S^{-1} = e^{tSAS^{-1}} = e^{-tA^{\mathsf{T}}}.$$

We conclude that

$$L(G) = \mathfrak{o}(V, \sigma) = \big\{ A \in \mathfrak{gl}_n(\mathbb{R}) \,\big|\, A^{\mathsf{T}} S + SA = 0 \big\}. \tag{2}$$

Example 2.17. Consider the (real) Lie group (the *unitary group*)

$$\mathrm{U}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \,|\, AA^* = 1\}, \qquad A^* = \overline{A}^{\mathsf{T}}.$$

Its Lie algebra

$$\mathfrak{u}(n) = \{A \in \mathfrak{gl}_n(\mathbb{C}) \,|\, A + A^* = 0\}$$

is a Lie algebra over $\mathbb{R}$ (but not over $\mathbb{C}$) (the *unitary Lie algebra*). For example, if $n = 1$, then

$$\mathrm{U}(1) = \{z \in \mathbb{C}^* \,|\, |z| = 1\}, \qquad \mathfrak{u}(1) = \{z \in \mathbb{C} \,|\, z + \overline{z} = 0\} = i\mathbb{R} \simeq \mathbb{R}.$$

The exponential map $\exp\colon \mathfrak{u}(1) \to \mathrm{U}(1)$ is given by

$$\mathbb{R} \to \mathrm{U}(1), \qquad \phi \mapsto e^{i\phi}.$$

Example 2.18. Consider the (real) Lie group (the *special unitary group*)

$$\mathrm{SU}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \,|\, AA^* = 1, \det A = 1\}.$$

Its Lie algebra

$$\mathfrak{su}(n) = \{A \in \mathfrak{gl}_n(\mathbb{C}) \,|\, A + A^* = 0, \operatorname{tr} A = 0\}$$

is a Lie algebra over $\mathbb{R}$ (but not over $\mathbb{C}$) (the *special unitary Lie algebra*).

2.4. **Properties of $L(G)$.** Our goal is to show that for any linear group $G \subseteq \mathrm{GL}_n(\mathbb{R})$, the subset $L(G) \subseteq \mathfrak{gl}_n(\mathbb{R})$ is a Lie subalgebra.

Lemma 2.19. *Let $(X_k)_k$ be a sequence of (nonzero) matrices such that $X_k \to 0$, $X_k/\|X_k\| \to X$ and $e^{X_k} \in G$. Then $X \in L(G)$.*

Proof. For any $t \in \mathbb{R}$, we will show that $e^{tX} \in G$. We have $\|X_k\| \to 0$, hence we can choose integers m_k such that $m_k \|X_k\| \to t$. Then $m_k X_k = m_k \|X_k\| \frac{X_k}{\|X_k\|} \to tX$, hence $G \ni (e^{X_k})^{m_k} = e^{m_k X_k} \to e^{tX}$ and $e^{tX} \in G$ as G is closed. □

Lemma 2.20. *Let $A\colon (0,\varepsilon) \to G$ be a continuous map such that $A(t) = I + tX + o(t)$ for some $X \in M_n(\mathbb{R})$. Then $X \in L(G)$.*

Proof. Define

$$X_k = \log A(1/k) = \log(I + X/k + o(1/k)) = X/k + o(1/k)$$

(we can do this for sufficiently large k). Then $e^{X_k} = A(1/k) \in G$. We have $kX_k \to X$, hence $X_k \to 0$ and $X_k/\|X_k\| \to X/\|X\|$ (assuming $X \neq 0$). By the previous result $X/\|X\| \in L(G)$, hence $X \in L(G)$. □

Remark 2.21 (Alternative proof)**.** It is enough to show that $e^{tX} \in G$ for $t \in (0,\varepsilon)$. Then $e^{ntX} = (e^{tX})^n \in G$, hence $e^{sX} \in G$ for $s \in (0, n\varepsilon)$ and any $n \geq 1$. This implies that $e^{sX} \in G$ for $s \geq 0$, hence also $e^{-sX} = (e^{sX})^{-1} \in G$. This implies $X \in L(G)$.

We know that $\lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n = e^x$. Similarly, one can show that if $x_n \to x$, then $\left(1 + \frac{x_n}{n}\right)^n \to e^x$. Indeed, $\forall \varepsilon > 0\ \exists N > 0\colon\ |x_n - x| < \varepsilon\ \forall n \geq N$. Then $\left(1 + \frac{x_n}{n}\right)^n \leq \left(1 + \frac{x+\varepsilon}{n}\right)^n \leq e^{x+\varepsilon} + \varepsilon$ for $n \gg 0$. Similarly, $\left(1 + \frac{x_n}{n}\right)^n \geq e^{x-\varepsilon} - \varepsilon$ for $n \gg 0$ and we conclude that $\left(1 + \frac{x_n}{n}\right)^n \to e^x$. In the same way one can show for the matrices that if $X_n \to X$, then $\left(I + \frac{X_n}{n}\right)^n \to e^X$. We can write $A(t/n) = I + \frac{t}{n}(X + X_n)$ for some matrices $X_n \to 0$. Therefore

$$\lim_{n\to\infty} A(t/n)^n = \lim_{n\to\infty}\left(I + \frac{t(X + X_n)}{n}\right)^n = e^{tX}.$$

We have $A(t/n) \in G$, hence $A(t/n)^n \in G$ and the limit $e^{tX} = \lim_{n\to\infty} A(t/n)^n \in G$.

Theorem 2.22. *The subset $L(G) \subseteq \mathfrak{gl}_n(\mathbb{R})$ is a Lie subalgebra.*

Proof. **(1)** Let $A \in L(G)$ and $s \in \mathbb{R}$. For any $t \in \mathbb{R}$, we have $e^{t(sA)} = e^{(ts)A} \in G$, hence $sA \in L(G)$.
(2) Let $A, B \in L(G)$. Then $e^{tA} = I + tA + o(t) \in G$, $e^{tB} = I + tB + o(t) \in G$ and

$$G \ni e^{tA}e^{tB} = (I + tA + o(t))(I + tB + o(t)) = I + t(A + B) + o(t)$$

By the previous result $A + B \in L(G)$. This proves that $L(G) \subseteq \mathfrak{gl}_n(\mathbb{R})$ is a vector space.
(3) Using the previous notation we have

$$\begin{aligned} G \ni e^{tA}e^{tB}e^{-tA}e^{-tB} &= \left(I + tA + \frac{t^2A^2}{2}\right)\left(I + tB + \frac{t^2B^2}{2}\right)\left(I - tA + \frac{t^2A^2}{2}\right)\left(I - tB + \frac{t^2B^2}{2}\right) + o(t^2) \\ &= I + t^2(AB - BA) + o(t^2) = I + s[A,B] + o(s), \end{aligned}$$

where $s = t^2$. By the previous result $[A, B] \in L(G)$. □

Theorem 2.23. *The map* $\exp\colon L(G) \to G$ *is a local homeomorphism at* $0 \in L(G)$.

Proof. Let $L = L(G)$. We know from theorem 2.9 that for $0 < \varepsilon < 1$ the sets

$$U_\varepsilon = \{A \in \mathfrak{gl}_n \,|\, \|A\| < \varepsilon\} \subseteq \mathfrak{gl}_n, \qquad V_\varepsilon = \exp(U_\varepsilon) \subseteq \mathrm{GL}_n$$

are open and there is a homeomorphism $\exp\colon U_\varepsilon \to V_\varepsilon$ with the inverse map log. This map restricts to $\exp\colon U_\varepsilon \cap L \to V_\varepsilon \cap G$ and to prove the theorem, we need to show that $\log(V_\varepsilon \cap G) \subseteq L$ for some $0 < \varepsilon \ll 1$. If this is not the case, then we can find a sequence $g_n \in V_{1/n} \cap G$ such that $\log g_n \notin L$. Consider a vector space decomposition $\mathfrak{gl}_n = L \oplus M$. Then the map $L \oplus M \to \mathrm{GL}_n$, $(A, B) \to e^A e^B$ is a local homeomorphism at 0. Therefore we can represent $g_n = e^{A_n} e^{B_n}$. If $B_n = 0$, then $\log g_n = A_n \in L$, a contradiction. Therefore $B_n \neq 0$ and we can choose a subsequence such that $B_n / \|B_n\|$ converges to some $B \in M$ having norm 1. We have $e^{B_n} = e^{-A_n} g_n \in G$, hence by Lemma 2.19 we have $B \in L$. This implies $B \in L \cap M = 0$, hence $B = 0$, a contradiction. □

Remark 2.24. Let us choose an open set $0 \in U \subseteq L = L(G)$ such that $V = \exp(U) \subseteq G$ is open and $\exp\colon U \to V$ is a homeomorphism. Then $I \in V$ and $\log\colon V \to U \subseteq L$ is a homeomorphism. Note that $L \subseteq \mathfrak{gl}_n(\mathbb{R})$ is a vector space, hence $L \simeq \mathbb{R}^d$ for some d. This means that $(V, \log)$ is a chart on G around I. We can also construct a chart around any $g \in G$. Consider $gV = \{gx \,|\, x \in V\}$ and a homeomorphism $L_g\colon V \to gV$, $x \mapsto gx$ (left multiplication). Then we have a map $\log \circ L_g^{-1}\colon gV \to V \to L$ and a chart $(gV, \log \circ L_g^{-1})$ around g. Collection of all such charts defines a manifold structure on G (although one has to prove that the gluing maps are smooth). In this way we equip G with a Lie group structure. The map $\exp\colon L \to G$ is smooth and is a local diffeomorphism at $0 \in L$. Note that the map $\exp\colon U \to V$ induces an isomorphism of tangent spaces $T_0 L = T_0 U \to T_I V = T_I G$. As L is a vector space, we can identify $T_0 L$ with L. We conclude that $T_I G \simeq L$. For an arbitrary Lie group G, one defines $L(G)$ to be the tangent space $T_I G$ and equips it with a Lie algebra structure.

2.5. **Lie group homomorphisms.**

Definition 2.25. Given Lie groups G and H, we define a *Lie group homomorphism* between them to be a smooth group homomorphism $f\colon G \to H$. We define a 1-*parameter subgroup* of G to be a Lie group homomorphism $\gamma\colon \mathbb{R} \to G$.

Our goal is to show that a Lie group homomorphism $f\colon G \to H$ induces a Lie algebra homomorphism $\bar{f}\colon L(G) \to L(H)$.

Remark 2.26. Note that a smooth map $f\colon G \to H$ induces a linear map between tangent spaces $df\colon T_IG \to T_IH$ (given by the Jacobian matrix of f at $I \in G$). Using our interpretation $L(G) = T_IG$, we automatically obtain a linear map $\bar{f} = df\colon L(G) \to L(H)$. However, we will use a different approach for linear groups.

Theorem 2.27. *Let G be a linear group and $L = L(G)$ be its Lie algebra. Then there is a 1-1 correspondence between elements of L and 1-parameter subgroups of G. More precisely,*

(1) For any $A \in L$, there exists a unique 1-parameter subgroup $\gamma_A\colon \mathbb{R} \to G$ such that

$$\dot{\gamma}_A(0) = \frac{d}{dt}\gamma_A(t)|_{t=0} = A.$$

(2) Any 1-parameter subgroup $\gamma\colon \mathbb{R} \to G$ is of the form $\gamma(t) = e^{tA}$ for some $A \in L$.

Proof. **(1)** To prove existence, we consider the group homomorphism $\gamma_A\colon \mathbb{R} \to G$, $t \mapsto e^{tA}$. It satisfies

$$\dot{\gamma}_A(0) = \frac{d}{dt}e^{tA}|_{t=0} = (e^{tA} \cdot A)|_{t=0} = A.$$

To prove uniqueness of the 1-parameter subgroup $\gamma\colon \mathbb{R} \to G$ satisfying $\dot{\gamma}(0) = A$, we note that $\gamma(s+t) = \gamma(s)\gamma(t)$, hence

$$\dot{\gamma}(s) = \frac{d}{dt}\gamma(s+t)|_{t=0} = \gamma(s)\frac{d}{dt}\gamma(t)|_{t=0} = \gamma(s)A.$$

We obtain a differential equation with the initial condition $\gamma(0) = I$. It has a unique solution.
(2) Given a 1-parameter group $\gamma\colon \mathbb{R} \to G$, define $A = \dot{\gamma}(0) \in \mathfrak{gl}_n$. Then $\gamma(t) = I + tA + o(t)$ and $\gamma(t) \in G$ for all t. By Lemma 2.20 this implies $A \in L$. The 1-parameter group $\gamma_A\colon t \mapsto e^{tA}$ also satisfies $\dot{\gamma}_A(0) = A$, hence we conclude that $\gamma = \gamma_A$ by uniqueness. □

Remark 2.28. Given $A, B \in \mathfrak{gl}_n(\mathbb{R})$, we can identify the bracket $[A, B]$ with

$$\frac{d}{dt}e^{tA}Be^{-tA}\Big|_{t=0} = \frac{d}{dt}(I+tA)B(I-tA)\Big|_{t=0} = AB - BA = [A, B].$$

This formula will be used in the following theorem.

Theorem 2.29. *For a Lie group homomorphism $f\colon G \to H$, there exists a unique linear map $\bar{f} = df\colon L(G) \to L(H)$ (the differential of f) such that the following diagram commutes*

$$\begin{array}{ccc} L(G) & \xrightarrow{\bar{f}} & L(H) \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{f} & H \end{array}$$

meaning that $e^{\bar{f}(A)} = f(e^A)$ for all $A \in L(G)$. The map $\bar{f}$ satisfies

(1) $\bar{f}\colon L(G) \to L(H)$ is a Lie algebra homomorphism.
(2) $\bar{f}(A) = \frac{d}{dt}f(e^{tA})|_{t=0}$.
(3) $\bar{f}(gAg^{-1}) = f(g)\bar{f}(A)f(g)^{-1}$ for $A \in L(G)$ and $g \in G$.

Proof. **(1)** Let us prove uniqueness of a linear map $\bar{f}\colon L(G) \to L(H)$ such that the above diagram commutes. We have $e^{t\bar{f}(A)} = f(e^{tA})$. Taking the derivative at $t = 0$, we obtain

$$\bar{f}(A) = \frac{d}{dt} e^{t\bar{f}(A)}\Big|_{t=0} = \frac{d}{dt} f(e^{tA})\Big|_{t=0}.$$

This proves uniqueness of $\bar{f}$ as well as (2).
(2) For every $A \in L(G)$, let us show that $B = \bar{f}(A) \coloneqq \frac{d}{dt} f(e^{tA})|_{t=0}$ is contained in $L(H)$. The maps $t \mapsto e^{tB}$ and $t \mapsto f(e^{tA})$ are two 1-parameter groups having the same derivative at $t = 0$. Therefore they are equal and $e^{tB} = f(e^{tA}) \in H$. This implies that $I + tB + o(t) \in H$, hence $B \in L(H)$. We obtain the map $\bar{f}\colon L(G) \to L(H)$ such that $e^{\bar{f}(A)} = f(e^A)$, hence the above diagram is commutative.
(3) By the chain rule, we have

$$\bar{f}(A) = \frac{d}{dt} f(e^{tA})|_{t=0} = J_f \cdot \frac{d}{dt} e^{tA}|_{t=0} = J_f \cdot A,$$

where $J_f = J_f(I)$ is the Jacobian matrix of f at $I = e^{0 \cdot A} \in G$. Therefore $\bar{f}$ is linear.
(4) To prove (3), note that $e^{tgAg^{-1}} = ge^{tA}g^{-1} \in G$, hence $gAg^{-1} \in L(G)$. We have

$$e^{t\bar{f}(gAg^{-1})} = f(e^{tgAg^{-1}}) = f(ge^{tA}g^{-1}) = f(g)f(e^{tA})f(g^{-1}) = f(g)e^{t\bar{f}(A)}f(g^{-1})$$

Taking the derivative at $t = 0$, we get $\bar{f}(gAg^{-1}) = f(g)\bar{f}(A)f(g)^{-1}$. This proves (3).
(5) For $A, B \in L(G)$, we have $[A, B] = \frac{d}{dt}(e^{tA}Be^{-tA})|_{t=0}$. Therefore, using linearity of $\bar{f}$ and (3),

$$\begin{aligned}\bar{f}([A, B]) = \frac{d}{dt}\bar{f}(e^{tA}Be^{-tA})|_{t=0} &= \frac{d}{dt}(f(e^{tA})\bar{f}(B)f(e^{-tA}))|_{t=0} \\ &= \frac{d}{dt}(e^{t\bar{f}(A)}\bar{f}(B)e^{-t\bar{f}(A)})|_{t=0} = [\bar{f}(A), \bar{f}(B)].\end{aligned}$$

This implies that $\bar{f}\colon L(G) \to L(H)$ is a Lie algebra homomorphism. □

Let $G \subseteq \mathrm{GL}_n(\mathbb{R})$ be a linear group and $L = L(G) \subseteq \mathfrak{gl}_n(\mathbb{R})$ be its Lie algebra.

Lemma 2.30. *For $g \in G$ and $A \in L$, we have $\mathrm{Ad}_g(A) = gAg^{-1} \in L$ and $\mathrm{Ad}_g \in \mathrm{GL}(L)$. The map*

$$\mathrm{Ad}\colon G \to \mathrm{GL}(L), \qquad g \mapsto \mathrm{Ad}_g,$$

is a Lie group homomorphism, called th adjoint representation of the group G.

Proof. We have $e^{tgAg^{-1}} = ge^{tA}g^{-1} \in G$ for all $t \in \mathbb{R}$, hence $gAg^{-1} \in L$. The map $\mathrm{Ad}_g\colon L \to L$, $A \mapsto gAg^{-1}$, is linear and invertible, hence $\mathrm{Ad}_g \in \mathrm{GL}(L)$. For $g, h \in G$, we have $\mathrm{Ad}_{gh}(A) = ghAh^{-1}g^{-1} = \mathrm{Ad}_g \mathrm{Ad}_h(A)$. Therefore $\mathrm{Ad}_{gh} = \mathrm{Ad}_g \mathrm{Ad}_h$ and $\mathrm{Ad}\colon G \to \mathrm{GL}(L)$ is a group homomorphism. It is also smooth. □

Lemma 2.31. *The differential of $\mathrm{Ad}\colon G \to \mathrm{GL}(L)$ is equal to $\mathrm{ad}\colon L \to \mathfrak{gl}(L)$. For all $A \in L$, we have $\mathrm{Ad}_{e^A} = e^{\mathrm{ad}_A}$ in $\mathrm{GL}(L)$ (Campbell identity). This means that*

$$e^A Be^{-A} = e^{\mathrm{ad}_A}(B) = \sum_{n \geq 0} \frac{1}{n!} \mathrm{ad}_A^n(B) = B + [A, B] + \frac{1}{2}[A, [A, B]] + \ldots \qquad \forall B \in L.$$

Proof. For $f = \mathrm{Ad}\colon G \to \mathrm{GL}(L)$, the map $\bar{f}\colon L \to \mathfrak{gl}(L)$ is given by $\bar{f}(A) = \frac{d}{dt}\mathrm{Ad}(e^{tA})|_{t=0}$, hence

$$\bar{f}(A)(B) = \frac{d}{dt} e^{tA}Be^{-tA}|_{t=0} = [A, B] = \mathrm{ad}_A(B).$$

Therefore $\bar{f} = \mathrm{ad}$. By Theorem 2.29, we have $f(e^A) = e^{\bar{f}(A)}$ for $A \in L$, hence $\mathrm{Ad}_{e^A} = e^{\mathrm{ad}_A}$. □

Remark 2.32. One can write $e^A e^B = e^Z$ with $Z = A + B + \frac{1}{2}[A, B] + \ldots$ (Baker-Campbell-Hausdorff formula). We will show this in the case where $[A, B]$ commutes with A and B.

Lemma 2.33. *If $[A, B]$ commutes with A and B, then $e^A e^B = e^{A+B+\frac{1}{2}[A,B]}$.*

Proof. Using $C = [A, B]$, we have $e^{tA} B e^{-tA} = e^{t \operatorname{ad}_A}(B) = B + tC$. Then $f(t) = e^{tA} e^{tB}$ satisfies $f(0) = I$ and $f'(t) = A e^{tA} e^{tB} + e^{tA} B e^{tB} = (A + e^{tA} B e^{-tA}) f(t) = (A + B + tC) f(t)$. This differential equation has a unique solution $f(t) = e^{t(A+B)+t^2 C/2}$, hence $e^{tA} e^{tB} = e^{t(A+B)+t^2 C/2}$. □

Theorem 2.34. *Let G be a Lie group. There is a bijection between connected Lie subgroups $H \subseteq G$ and Lie subalgebras $L \subseteq L(G)$. It is given by sending H to its Lie algebra*

$$L = \{x \in L(G) \mid \exp(tx) \in H \ \forall t \in \mathbb{R}\}.$$

Theorem 2.35. *There is a bijection between isomorphism classes of connected, simply connected Lie groups and (finite-dimensional) Lie algebras.*

Example 2.36. Let us consider Lie groups $G = (\mathbb{R}, +)$ and $H = U(1) = \{z \in \mathbb{C}^\times \mid |z| = 1\}$ and the group homomorphism $f \colon \mathbb{R} \to U(1)$, $t \mapsto e^{2\pi i t}$. We have $L(G) = \mathbb{R}$ and $L(H) = \mathbb{R}$ (with the zero Lie bracket). The group homomorphism $f \colon G \to H$ induces the isomorphism of Lie algebras $\bar{f} \colon L(G) \to L(H)$. Note that $G = \mathbb{R}$ is simply connected, while $H = U(1)$ is not.

Theorem 2.37. *Let G be a connected linear group and $L = L(G)$ be its Lie algebra. Then $\exp(L) \subseteq G$ generates G, meaning that every element of G can be written in the form*

$$e^{A_1} \dots e^{A_n}, \qquad A_i \in L.$$

Proof. As G is connected, for any $g \in G$, we can find a path $X \colon [0, 1] \to G$ such that $X(0) = I$ and $X(1) = g$. This map is continuous, hence uniformly continuous, meaning that $\forall \varepsilon > 0$ $\exists \delta > 0$ such that $|s - t| < \delta \implies \|X(s) - X(t)\| < \varepsilon$. There exists $\varepsilon > 0$, such that $\exp(L) \subseteq G$ contains all $g' \in G$ with $\|g' - I\| < \varepsilon$. We have

$$\left\|X(s)X(t)^{-1} - I\right\| \le \|X(s) - X(t)\| \cdot \left\|X(t)^{-1}\right\| \le \|X(s) - X(t)\| \cdot C,$$

where $C = \sup_{t \in [0,1]} \|X(t)^{-1}\|$. We can find $\delta > 0$ such that $|s - t| < \delta$ implies $\|X(s) - X(t)\| < \varepsilon / C$, hence $\|X(s)X(t)^{-1} - I\| < \varepsilon$. Let us divide the interval $[0, 1]$ into $0 = t_0 < \dots < t_n = 1$ such that $t_{i+1} - t_i < \delta$, hence $\|X(t_{i+1})X(t_i)^{-1} - I\| < \varepsilon$. Then we can write $X(t_{i+1})X(t_i)^{-1} = e^{A_{i+1}}$ for some $A_{i+1} \in L$. This implies $g = X(t_n) = e^{A_n} X(t_{n-1}) = \dots = e^{A_n} \dots e^{A_1}$. □

Remark 2.38. Even if G is connected, the map $\exp \colon L(G) \to G$ is not surjective in general. For example, consider $g = \left(\begin{smallmatrix} -1 & 1 \\ 0 & -1 \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{C})$ and assume that $g = e^A$ for some $A \in \mathfrak{sl}_2(\mathbb{C})$. If A has distinct eigenvalues, then it is diagonalizable, hence e^A is diagonalizable, a contradiction. If A has a multiple eigenvalue λ, then $e^{\operatorname{tr} A} = \det(e^A) = 1$, hence $2\lambda = \operatorname{tr} A = 0$. But then $g = e^A$ has a multiple eigenvalue $e^\lambda = 1$, a contradiction.

Remark 2.39. The above results imply that many questions about Lie groups can be reduced to questions about the corresponding Lie algebras. In particular, this applies to representations. Let G be a Lie group and $L = L(G)$ be its Lie algebra. A *Lie group representation* is a Lie group homomorphism $\rho \colon G \to \mathrm{GL}(V) = \mathrm{GL}_n(\mathbb{R})$ for $V = \mathbb{R}^n$ and a *Lie algebra representation* is a Lie algebra homomorphism $\phi \colon L \to \mathfrak{gl}(V) = \mathfrak{gl}_n(\mathbb{R})$. Applying previous results, we see that a Lie group representation $\rho \colon G \to \mathrm{GL}(V)$ induces a Lie algebra representation $\bar{\rho} = d\rho \colon L \to \mathfrak{gl}(V)$ such that $\bar{\rho}(A) = \frac{d}{dt} \rho(e^{tA})|_{t=0}$ and $\rho(e^A) = e^{\bar{\rho}(A)}$. This implies that if we know $\bar{\rho} \colon L \to \mathfrak{gl}(V)$, then we can determine $\rho(e^A)$ for all $A \in L$. Therefore we can also determine the value of ρ on the group generated by $\exp(L) \subseteq G$. If G is connected, then this is the whole group G, hence $\rho \colon G \to \mathrm{GL}(V)$ is uniquely determined by the Lie algebra representation $\bar{\rho} \colon L \to \mathfrak{gl}(V)$. Note, however, that there can be more Lie algebra representations than Lie group representations.

3. Nilpotent and solvable Lie algebras

In this section we fix a field K of characteristic zero and we assume that all our Lie algebras and vector spaces are finite-dimensional.

3.1. **Ideals.** Given a Lie algebra L and subspaces $A, B \subseteq L$, we define $[A, B] \subseteq L$ to be the subspace generated by elements $[x, y]$ for all $x \in A$, $y \in B$. It consists of all finite linear combinations of elements $[x, y]$. Note that $[A, B] = [B, A]$ by anti-symmetry. For $x \in L$ and a subspace $B \subseteq L$, we define $[x, B] = \{[x, y] \,|\, y \in B\}$ which is a subspace of L.

Definition 3.1. Let L be a Lie algebra.

(1) A subspace $I \subseteq L$ is called an *ideal* of L if $[x, y] \in I$ for all $x \in L$, $y \in I$. Equivalently, $[L, I] \subseteq I$.
(2) The *center* of L is $Z(L) = \{x \in L \,|\, [x, y] = 0 \ \forall y \in L\} = \{x \in L \,|\, [x, L] = 0\}$.
(3) The *derived algebra* of L is $L' = [L, L]$.
(4) L is called *abelian* (or commutative) if $[x, y] = 0$ for all $x, y \in L$. Equivalently, $[L, L] = 0$.

Remark 3.2. Note that

(1) Every ideal $I \subseteq L$ is also a Lie subalgebra.
(2) For a subspace $I \subseteq L$, we have $[L, I] \subseteq I$ if and only if $[I, L] \subseteq I$.
(3) The center $Z(L)$ is an ideal of L.
(4) If $A, B \subseteq L$ are ideals, then $[A, B]$ is an ideal (exercise). In particular, the derived algebra $L' = [L, L]$ is an ideal of L.

Example 3.3. If $f\colon L \to M$ is a Lie algebra homomorphism, then $\operatorname{Ker} f$ is an ideal of f. Indeed, given $x \in L$ and $y \in \operatorname{Ker} f$, we have

$$f([x, y]) = [f(x), f(y)] = [f(x), 0] = 0$$

and therefore $[x, y] \in \operatorname{Ker} f$.

Example 3.4. If L is abelian then any subspace $I \subseteq L$ is an ideal of L.

Example 3.5. Consider the algebra $L = \mathfrak{sl}_2(K)$. We know that it has a basis x, y, h with the bracket

$$[h, x] = 2x, \qquad [h, y] = -2y, \qquad [x, y] = h.$$

This implies that $L' = [L, L] = L$. Let us show that the only ideals of L are zero and L. Let $I \subseteq L$ be a non-zero ideal and $ax + by + ch \in I$ be a non-zero element.

(1) If $a \neq 0$, we can apply ad_y twice and obtain $[y, [y, ax]] = -2ay \in I$, hence $y \in I$.
(2) If $a = 0$ and $b \neq 0$, we can apply ad_h and obtain $[h, by] = -2by \in I$, hence $y \in I$.
(3) If $a = b = 0$ and $c \neq 0$, we can apply ad_y and obtain $[y, ch] = 2cy \in I$, hence $y \in I$.

Therefore $y \in I$, hence $h = [x, y] \in I$ and $2x = [h, x] \in I$. We obtain $I = L$.

Example 3.6. Let $L = \mathfrak{gl}_n(K)$. Every element of the form λI_n (for $\lambda \in K$) is central as $[\lambda I_n, x] = \lambda [I_n, x] = 0$. One can show that $Z(L) = KI_n$. One can also show that the only ideals of $L = \mathfrak{gl}_n(K)$ are 0, $Z(L)$ and L.

Given an ideal $I \subseteq L$, consider the quotient space L/I. Recall that L/I consists of equivalence classes with respect to the equivalence on L given by

$$x \sim y \iff y - x \in L.$$

Every equivalence class has the form $x + I = \{x + a \,|\, a \in I\}$ for some $x \in L$. We equip L/I with the Lie algebra structure having the Lie bracket

$$[x + I, y + I] = [x, y] + I, \qquad x, y \in L.$$

This Lie algebra is called the *quotient algebra*. The map

$$\pi \colon L \to L/I, \qquad x \mapsto x + I$$

is a Lie algebra homomorphism.

Remark 3.7. One needs to show that the above bracket is well-defined. If $x \sim x'$ and $y \sim y'$, then $a = x' - x \in I$, $b = y' - y \in I$ and

$$[x', y'] + I = [x + a, y + b] + I = [x, y] + ([x, b] + [a, y] + [a, b]) + I = [x, y] + I.$$

Example 3.8. Let $L = \mathfrak{b}_2(K) = \{\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \,|\, a, b, d \in K\}$ be the Lie algebra of upper-triangular matrices and $I = \{\left(\begin{smallmatrix} 0 & b \\ 0 & 0 \end{smallmatrix}\right) \,|\, b \in K\}$. Then I is an ideal of L and L/I is isomorphic to the abelian Lie algebra K^2 (of diagonal matrices).

3.2. **Representations.**

Definition 3.9. Let L be a Lie algebra.

(1) An *L-representation* (or representation of L) is a pair $(V, \rho) = (V, \rho_V)$, where V is a vector space and $\rho\colon L \to \mathfrak{gl}(V)$ is a Lie algebra homomorphism. We call V the *space of the representation*.
(2) An *L-module* (or module over L) is a vector space V equipped with a bilinear map

$$L \times V \to V, \qquad (x, v) \mapsto xv$$

such that

$$[x, y]v = x(yv) - y(xv), \qquad \forall x, y \in L,\ v \in V.$$

Remark 3.10. There is a $1-1$ correspondence between L-representations and L-modules. Given a representation (V, ρ), we equip V with an L-module structure by

$$(x, v) \mapsto xv := \rho(x)v, \qquad x \in L,\ v \in V.$$

This map satisfies the axiom of an L-module

$$[x, y]v = \rho([x, y])v = [\rho(x), \rho(y)]v = (\rho(x)\rho(y) - \rho(y)\rho(x))v = x(yv) - y(xv).$$

We will usually identify L-representations and L-modules and we will use notation $\rho(x)v$ and xv interchangeably. We will say that L acts on V.

Example 3.11.

(1) Any vector space V (in particular K) can be equipped with the *trivial action* of L by setting $xv = 0$ for all $x \in L$ and $v \in V$.
(2) If V is a vector space, then the Lie algebra $L = \mathfrak{gl}(V)$ acts on V in an obvious way. The representation V is called the *standard representation* of $\mathfrak{gl}(V)$.
(3) We have seen that the map

$$\mathrm{ad}\colon L \to \mathfrak{gl}(L), \qquad x \mapsto \mathrm{ad}(x) = \mathrm{ad}_x$$

is a Lie algebra homomorphism, hence it defines a representation of L on the vector space L. We call it the *adjoint representation* of L. Note that

$$\mathrm{Ker}(\mathrm{ad}) = \{x \in L \,|\, \mathrm{ad}(x) = 0\} = \{x \in L \,|\, [x, L] = 0\} = Z(L)$$

is the center of L.

Example 3.12. Let V be an L-representation. A subspace $U \subseteq V$ is called a *subrepresentation* of V if $xu \in U$ for all $x \in L$, $u \in U$. In this case the quotient space V/U can be equipped with an L-representation structure

$$x(v + U) = xv + U, \qquad x \in L,\ v \in V.$$

Example 3.13. Given vector spaces V, W, we define their tensor product $V \otimes W = V \otimes_K W$ as follows. It is the vector space generated by elements of the form $v \otimes w$ for $v \in V$, $w \in W$ subject to the relations

(1) $(v + v') \otimes w = v \otimes w + v' \otimes w$ for $v, v' \in V$ and $w \in W$.
(2) $v \otimes (w + w') = v \otimes w + v \otimes w'$ for $v \in V$ and $w, w' \in W$.
(3) $a(v \otimes w) = (av) \otimes w = v \otimes (aw)$ for $a \in K$, $v \in V$ and $w \in W$.

If V has a basis $(e_1, \dots, e_m)$ and W has a basis $(f_1, \dots, f_n)$, then $V \otimes W$ has a basis

$$(e_i \otimes f_j \mid 1 \le i \le m,\, 1 \le j \le n).$$

If V, W are L-representations, then $V \otimes W$ can be equipped with an L-representation structure by the rule

$$x(v \otimes w) = (xv) \otimes w + v \otimes (xw), \qquad x \in L,\, v \in V,\, w \in W.$$

Example 3.14. Given vector spaces V, W, we have the vector space $\mathrm{Hom}(V, W) = \mathrm{Hom}_K(V, W)$ of all linear maps $f\colon V \to W$. If V, W are L-representations, then $\mathrm{Hom}(V, W)$ can be equipped with an L-representation structure by the rule

$$(xf)(v) = x(f(v)) - f(xv), \qquad x \in L,\, f \in \mathrm{Hom}(V, W),\, v \in V.$$

One can verify that this action indeed satisfies the axiom of an L-representation. Note that the axiom would not be satisfied without the above minus sign.

In particular, the dual space $V^* = \mathrm{Hom}(V, K)$ (where K is equipped with the trivial representation structure) has an L-representation structure given by

$$(xf)(v) = -f(xv), \qquad x \in L,\, f \in V^*,\, v \in V.$$

Definition 3.15. Let L be a Lie algebra.

(1) A linear map $f\colon V \to W$ between two L-representations is called a *homomorphism* of L-representations if

$$f(xv) = x(f(v)) \qquad \forall x \in L,\, v \in V.$$

It is called an isomorphism of representations if it is an isomorphism of vector spaces.

(2) Given an L-representation V, an element $v \in V$ is called *L-invariant* if $xv = 0$ for all $x \in L$. The vector space of all L-invariant elements in V is denoted by V^L.

Remark 3.16. To give a motivation for the last definition, let G be a Lie group and $L = L(G)$ be its Lie algebra. One defines a representation of G to be a pair (V, ρ), where V is a vector space and $\rho\colon G \to \mathrm{GL}(V)$ is a Lie group homomorphism. The map ρ induces a Lie algebra homomorphism $\bar{\rho}\colon L \to \mathfrak{gl}(V)$, hence $(V, \bar{\rho})$ is an L-representation. A vector $v \in V$ is called G-invariant if $gv = v$ for all $g \in G$. Then for all $A \in L$ and $t \in \mathbb{R}$, we have $e^{tA} \in G$, hence $e^{tA}v = v$. Taking the derivative at $t = 0$, we obtain

$$\tfrac{d}{dt} e^{tA} v|_{t=0} = e^{tA} Av|_{t=0} = Av = 0.$$

This implies that $Av = 0$ for all $A \in L$, hence $v \in V$ is L-invariant.

Remark 3.17. Let V, W be L-representations and $\mathrm{Hom}(V, W)$ be equipped with an L-representation structure as before. Then $f \in \mathrm{Hom}(V, W)$ is L-invariant $\iff$ $(xf)(v) = 0$ for all $x \in L, v \in V$ $\iff$ $f(xv) = x(f(v))$ for all $x \in L, v \in V$ $\iff$ $f\colon V \to W$ is a homomorphism of L-representations.

3.3. **Nilpotent Lie algebras.**

Definition 3.18. Let L be a finite-dimensional Lie algebra.

(1) Define the *central series* of ideals in L by

$$L^1 = L, \qquad L^n = [L, L^{n-1}], \quad n \geq 2.$$

(2) L is called *nilpotent* if $L^n = 0$ for some $n \geq 1$.

Remark 3.19. For $n \geq 1$, the following conditions are equivalent

(1) $L^n = 0$.
(2) $[x_1, [x_2, \ldots, [x_{n-1}, x_n]]] = 0 \qquad \forall x_1, \ldots, x_n \in L$.
(3) $\mathrm{ad}_{x_1} \circ \cdots \circ \mathrm{ad}_{x_{n-1}} = 0 \qquad \forall x_1, \ldots, x_{n-1} \in L$.

Recall that a linear operator $f\colon V \to V$ is called nilpotent if $f^n = 0$ for some $n \geq 1$. The above conditions imply that if L is a nilpotent Lie algebra, then $\mathrm{ad}_x\colon L \to L$ is nilpotent for all $x \in L$. We will see later that the converse is also true (Engel theorem).

Example 3.20.

(1) Any abelian Lie algebra is nilpotent.
(2) Let

$$L = \{(a_{ij}) \in \mathfrak{gl}_n(K) \,|\, a_{ij} = 0 \text{ for } i \geq j\} \subseteq \mathfrak{gl}_n(K)$$

be the Lie algebra of strictly upper-triangular matrices. Let us show that L is nilpotent. Let E_{ij} denote the elementary matrix having entry 1 on the intersection of the i-th row and the j-th column and zero otherwise. The vector space $L^1 = L$ is spanned by the matrices E_{ij} with $i < j$, hence $j - i \geq 1$. The vector space $L^2 = [L, L]$ is spanned by

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

where $i < j$, $k < l$ and we assumed that $j \leq l$. Note that if $i < j = k < l$, then $l - i \geq 2$. Continuing this process we obtain that L^m is spanned by E_{ij} with $j - i \geq m$. This implies that $L^n = 0$ and L is nilpotent.

Theorem 3.21 (Engel)**.** *Let L be a Lie algebra and (V, ρ) be a non-zero finite-dimensional representation of L such that $\rho(x) \in \mathrm{End}(V)$ is nilpotent for all $x \in L$. Then*

(1) There exists $0 \neq v \in V$ such that $\rho(x)v = 0$ for all $x \in L$ Equivalently, $V^L \neq 0$.
(2) There exists a basis of V such that all $\rho(x)$ are strictly upper-triangular in this basis.
(3) There exists a filtration by subrepresentations

$$0 = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_n = V$$

such that $LV_k \subseteq V_{k-1}$ (equivalently, L acts trivially on V_k/V_{k-1}) for $1 \leq k \leq n$.

Proof. We postpone the proof of (1) and show first that it implies (2) and (3).

1 $\implies$ 2. We choose $v_1 = v$ from (1). Then $Lv_1 = 0$ and therefore $Kv_1 \subseteq V$ is a subrepresentation. Applying induction by dimension of V to the representation $V' = V/Kv_1$, we obtain a basis $v_2', \ldots, v_n'$ of V' such that $\rho_{V'}(x)$ are strictly upper-triangular with respect to this basis. We can lift these basis vectors to vectors $v_2, \ldots, v_n$ in V. Then $v_1, v_2, \ldots, v_n$ form a basis of V such that all $\rho(x)$ are strictly-upper triangular with respect to this basis.

2 $\implies$ 3. We define $V_k = \langle v_1, \ldots, v_k \rangle$ for $1 \leq k \leq n$. As the matrix of $\rho(x)$ is strictly upper-triangular, it maps v_k to a linear combination of $v_1, \ldots, v_{k-1}$. This implies that $xv_k \in V_{k-1}$ and generally $xV_k \subseteq V_{k-1}$. $\square$

Remark 3.22. Given a nilpotent operator $A \in \operatorname{End} V$, one can find a basis of V making A a strictly upper-triangular matrix (consider the Jordan normal form of A). However, it is not true that given nilpotent operators $A, B \in \operatorname{End} V$, one can find a basis of V making A and B simultaneously strictly upper-triangular. We can not apply the previous theorem as it may happen that the Lie algebra generated by A, B contains non-nilpotent elements. For example consider the standard representation $V = K^2$ of $\mathfrak{sl}_2(K)$. The elements $x, y \in \mathfrak{sl}_2(K)$ induce nilpotent operators on V. But one can not make them simultaneously strictly upper-triangular.

Theorem 3.23 (Engel)**.** *A Lie algebra L is nilpotent if and only if $\operatorname{ad}_x \in \mathfrak{gl}(L)$ is nilpotent for all $x \in L$.*

Proof. If $L^n = 0$, then $\operatorname{ad}_x^{n-1} = 0$ for all $x \in L$. Conversely, assume that ad_x is nilpotent for all $x \in L$. Applying the above result to the adjoint representation $\operatorname{ad}\colon L \to \mathfrak{gl}(L)$, we can find a filtration by subspaces

$$0 = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_n = L$$

such that $\operatorname{ad}_x(V_k) \subseteq V_{k-1}$ for all $x \in L$ and $k \geq 1$. Then $[L, V_k] \subseteq V_{k-1}$ and therefore $[L, [L, \ldots]] = 0$ (n products). □

Lemma 3.24. *If $x \in L = \mathfrak{gl}(V)$ is a nilpotent operator on a vector space V, then $\operatorname{ad}_x\colon L \to L$ is a nilpotent operator on L.*

Proof. We have

$$\operatorname{ad}_x(y) = xy - yx = L_x(y) - R_x(y), \qquad L_x(y) = xy, \qquad R_x(y) = yx.$$

There exists $n \geq 1$ such that $x^n = 0$. Therefore $L_x^n = R_x^n = 0$. The operators L_x and R_x commute, hence $\operatorname{ad}_x^{2n} = (L_x - R_x)^{2n} = \sum_{i+j=2n} \binom{2n}{i} (-1)^j L_x^i R_x^j = 0$. □

Proof of Theorem 3.21 (1). We will prove the statement by induction on $\dim L$. Taking $\rho(L) \subseteq \operatorname{End}(V)$ instead of L, we can assume that $L \subseteq \operatorname{End}(V)$. If $\dim L = 1$ and $0 \neq x \in L$, then x is nilpotent, hence $xv = 0$ for some $0 \neq v \in V$. Therefore $Lv = 0$. Now assume that $\dim L \geq 2$ and let $I \subsetneq L$ be a maximal proper subalgebra (it is nonzero as L contains 1-dimensional subalgebras).

(1) Let us show that I is an ideal. Let $N = \{x \in L \mid [x, I] \subseteq I\}$ be the normalizer of I (it is a Lie subalgebra of L). We have $I \subseteq N \subseteq L$ and we claim that $N = L$. The Lie algebra I acts on L/I nilpotently by Lemma 3.24 and the assumption that $L \subseteq \operatorname{End}(V)$ consists of nilpotent operators. We conclude by induction (as $\dim I < \dim L$) that there exists a nonzero element $x + I \in L/I$ (with $x \notin I$) such that $[I, x + I] \subseteq I$. This implies that $[x, I] \subseteq I$ and $x \in N$. Therefore $I \subsetneq N$ and $N = L$ (by maximality of I). Hence $[L, I] \subseteq I$ and $I \subseteq L$ is an ideal.

(2) The vector space $V^I = \{v \in V \mid Iv = 0\}$ is nonzero by induction (as $\dim I < \dim L$) and is an L-subrepresentation of V:

$$zxv = xzv - [x, z]v = 0, \qquad x \in L,\ v \in V^I,\ z \in I,$$

hence $I(xv) = 0$ and $xv \in V^I$.

(3) We have $IV^I = 0$, hence the Lie algebra L/I acts on V^I. By induction (as $\dim L/I < \dim L$) there exists $0 \neq v \in V^I$ such that $(L/I)v = 0$. Then also $Lv = 0$. □

3.4. **Solvable Lie algebras.**

Definition 3.25. Let L be a finite-dimensional Lie algebra.

(1) We define the *derived series* of ideals in L by

$$L^{(0)} = L, \qquad L^{(n)} = [L^{(n-1)}, L^{(n-1)}], \quad n \geq 1.$$

(2) L is called *solvable* if $L^{(n)} = 0$ for some $n \geq 1$.

Remark 3.26. If $I, J \subseteq L$ are ideals then, using the Jacobi identity, we obtain

$$[x, [a, b]] = [[x, a], b] + [a, [x, b]] \in [I, J] + [I, J] = [I, J] \qquad \forall x \in L,\ a \in I,\ b \in J.$$

This implies that $[I, J] \subseteq L$ is also an ideal. In particular, the above spaces $L^{(n)}$ are ideals.

Remark 3.27. We have $L^{(0)} = L = L^1$ and generally $L^{(n)} \subseteq L^{n+1}$ by induction

$$L^{(n+1)} = [L^{(n)}, L^{(n)}] \subseteq [L, L^{n+1}] = L^{n+2}.$$

If L is nilpotent, then $L^{(n)} \subseteq L^{n+1} = 0$ for some $n \geq 1$, hence L is solvable. Moreover, if $L' = [L, L]$ is nilpotent, then L is solvable. Indeed, if L' is nilpotent, then it is solvable. Therefore, $L^{(n+1)} = (L')^{(n)} = 0$, for some $n \geq 1$, hence L is solvable. We will see later that the converse is true if $\operatorname{char} K = 0$.

Example 3.28. The algebra of upper-triangular matrices

$$L = \{(a_{ij}) \in \mathfrak{gl}_n(K) \,|\, a_{ij} = 0 \text{ for } i > j\}$$

is solvable. Indeed, $L' = [L, L]$ consists of strictly upper-triangular matrices and we have seen that this algebra is nilpotent. We conclude that L is solvable by the previous remark.

Lemma 3.29. *Let L be a Lie algebra.*

(1) If $I \subseteq L$ is an ideal and L is solvable then $I, L/I$ are solvable.
(2) If $I \subseteq L$ is an ideal and $I, L/I$ are solvable then L is solvable.
(3) If $I, J \subseteq L$ are solvable ideals then $I + J$ is solvable.

Proof. **(1)** Let $L^{(n)} = 0$. Then $I^{(n)} \subseteq L^{(n)} = 0$ and I is solvable. Let $\pi\colon L \to L/I$ be the projection. Then we prove by induction that $\pi(L^{(k)}) = (L/I)^{(k)}$

$$\pi(L^{(k+1)}) = \pi([L^{(k)}, L^{(k)}]) = [\pi(L^{(k)}), \pi(L^{(k)})] = [(L/I)^{(k)}, (L/I)^{(k)}] = (L/I)^{(k+1)}.$$

In particular $(L/I)^{(n)} = \pi(L^{(n)}) = 0$ and L/I is solvable.
(2) Assume that $I^{(m)} = 0$ and $(L/I)^{(n)} = 0$. Then $\pi(L^{(n)}) = (L/I)^{(n)} = 0$ in L/I, hence $L^{(n)} \subseteq I$. Therefore $L^{(m+n)} \subseteq I^{(m)} = 0$ and L is solvable.
(3) There is a natural isomorphism $(I + J)/I \simeq J/(I \cap J)$. If J is solvable then the quotient $J/(I \cap J)$ is also solvable by (1). Therefore $(I + J)/I$ is solvable and if I is solvable then $I + J$ is solvable by (2). □

Remark 3.30. Applying the above result, we obtain that the subspace

$$\operatorname{rad} L = \sum_{\substack{I \subseteq L \\ \text{solvable}}} I \subseteq L$$

is the maximal solvable ideal of L. It is called the *radical* of L.

Remark 3.31. The quotient algebra $L/[L,L]$ is abelian. This implies that if L is solvable then there is a filtration by ideals

$$L = L^{(0)} \supset L^{(1)} \supset \ldots \supset L^{(n)} = 0$$

such that each factor $L^{(k)}/L^{(k+1)}$ is abelian. Conversely, if there exists a filtration by ideals

$$L = I_0 \supset I_1 \supset \ldots \supset I_n = 0$$

such that each factor I_k/I_{k+1} is abelian, then L is solvable. Indeed, assume that we proved that I_{k+1} is solvable. As I_k/I_{k+1} is abelian, it is solvable. By the previous result we conclude that I_k is solvable. Continuing this process we obtain that $I_0 = L$ is solvable.

Theorem 3.32 (Lie)**.** *Let L be a solvable Lie algebra over an algebraically closed field K of characteristic 0. Let (V,ρ) be a non-zero representation of L. Then*

(1) There exists a common eigenvector $0 \neq v \in V$ of L, that is, an eigenvector of $\rho(x)$ for all $x \in L$. Equivalently, $Kv \subseteq V$ is a subrepresentation.
(2) There exists a basis of V such that all $\rho(x)$ are upper-triangular in this basis.
(3) There exists a chain of subrepresentations

$$0 = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_n = V$$

such that $\dim V_k/V_{k-1} = 1$ for $1 \leq k \leq n$.

Proof. **(1) $\Longrightarrow$ (2).** Let $v_1 = v \in V$ from (1). Then $Kv_1 \subseteq V$ is a subrepresentation. Applying induction to the quotient representation $V' = V/Kv_1$ we find a basis $(v_2', \ldots, v_n')$ of V' such that $\rho_{V'}(x)$ are upper-triangular in this basis. We lift this basis to some vectors $v_2, \ldots, v_n$ in V. Then $v_1, \ldots, v_n$ form a basis of V. And $\rho(x)$ are upper-triangular in this basis.
(1) $\Longrightarrow$ (3). Choose a basis as above and define $V_k = \langle v_1, \ldots, v_k \rangle$. These vector spaces satisfy the required conditions.
(1) As L is solvable, $[L,L] \neq L$. Indeed, if $L^{(1)} = [L,L] = L$, then $L^{(2)} = [L^{(1)}, L^{(1)}] = [L,L] = L$ and generally $L^{(k)} = L$, contradicting to solvability of L. The quotient $L/[L,L]$ is a nonzero abelian Lie algebra. Therefore we can choose a codimension 1 subspace $[L,L] \subseteq I \subseteq L$ which is automatically an ideal. By induction there exists $0 \neq v \in V$ and $\lambda\colon I \to K$ such that $yv = \lambda(y)v$ for all $y \in I$ (λ is automatically linear). The space

$$W = \{v \in V \mid yv = \lambda(y)v \ \forall y \in I\}$$

is nonzero and is an L-subrepresentation by the next lemma. Given $x \in L \backslash I$, the map $x\colon W \to W$ has an eigenvector $0 \neq v \in W$ (as K is algebraically closed). Then v is a common eigenvector of $Kx + I = L$. □

Lemma 3.33. *Let L be a Lie algebra over a field K of characteristic zero, $I \subseteq L$ be an ideal, V be an L-representation. For any linear map $\lambda\colon I \to K$, the space*

$$W = V_\lambda = \{v \in V \mid xv = \lambda(x)v \ \forall x \in I\}$$

is an L-subrepresentation of V.

Proof. We can assume that $W \neq 0$. Let us show first that $[L,I]W = 0$. For any $x \in L$, we have to show that $\lambda([x,I]) = 0$. Let $0 \neq v \in W$ and let $v_k = x^{k-1}v$ for $k \geq 1$. Let us show by induction that, for any $y \in I$,

$$yv_k = \lambda(y)v_k + \sum_{i=1}^{k-1} a_i v_i, \qquad a_i = a_i(y) \in K. \tag{3}$$

For $k = 1$, we have $yv = \lambda(y)v$. Further, we have

$$yv_{k+1} = yx^k v = (xy - [x,y])x^{k-1}v = x\left(\lambda(y)v_k + \sum_{i=1}^{k-1} a_i v_i\right) - \sum_{i=1}^{k} a_i' v_i = \lambda(y)v_{k+1} + \sum_{i=1}^{k} b_i v_i.$$

Let $n \geq 1$ be maximal such that $v_1, \ldots, v_n$ are linearly independent and let $U \subseteq V$ be the vector space generated by these vectors. Then $xv_k = v_{k+1} \in U$ for $1 \leq k \leq n$. Therefore $xU \subseteq U$. Equation (3) implies that every $y \in I$ induces an endomorphism of U given by an upper-triangular matrix with entries $\lambda(y)$ on the diagonal (with respect to the basis $v_1, \ldots, v_n$). This implies

$$\mathrm{Tr}_U(y) = n\lambda(y) \qquad \forall y \in I$$

and in particular

$$0 = \mathrm{Tr}_U([x,y]) = n\lambda([x,y]).$$

Therefore $\lambda([x,y]) = 0$. This proves $[L,I]W = 0$.

Finally, for any $x \in L$, $y \in I$, $v \in W$, we have $[x,y]v = 0$, hence

$$yxv = xyv - [x,y]v = xyv = \lambda(y)xv,$$

and $xv \in W$. This means that W is an L-subrepresentation. □

Corollary 3.34. *Let L be a Lie algebra over a field of characteristic zero. Then L is solvable if and only if $[L,L]$ is nilpotent.*

Proof. Assume that L is solvable. Consider the adjoint representation $\mathrm{ad}\colon L \to \mathfrak{gl}(L)$. By Lie's theorem we can find a basis of L (over an algebraic closure of K) such that all operators ad_x are upper-triangular. Then each $\mathrm{ad}_{[x,y]} = [\mathrm{ad}_x, \mathrm{ad}_y]$ is strictly upper-triangular. This implies that for all $x \in [L,L]$ the matrix of ad_x is strictly upper-triangular and therefore nilpotent. By Engel's theorem $[L,L]$ is a nilpotent Lie algebra. Conversely, if $L' = [L,L]$ is nilpotent, then it is solvable. Therefore $L^{(n+1)} = (L')^{(n)} = 0$ for some $n \geq 1$, hence L is solvable. □

4. Semisimple and reductive Lie algebras

In this section we will consider only finite-dimension Lie algebras over a field of characteristic zero.

4.1. **Semisimple and reductive Lie algebras.** Recall that $\operatorname{rad} L \subseteq L$ denotes the maximal solvable ideal of L. In particular, the center $Z(L)$ is abelian, hence solvable and $Z(L) \subseteq \operatorname{rad} L$.

Definition 4.1. Let L be a Lie algebra.

(1) L is called *simple* if it is not abelian and it does not contain ideals except 0 and L.
(2) L is called *semisimple* if it does not contain abelian ideals except zero.
(3) L is called *reductive* if $Z(L) = \operatorname{rad} L$.

Lemma 4.2. *Let L be a Lie algebra. The following conditions are equivalent.*

(1) L is semisimple.
(2) L does not contain solvable ideals except zero.
(3) $\operatorname{rad} L = 0$.

Proof. **(1) $\Longrightarrow$ (2).** Let $0 \neq I \subseteq L$ be a solvable ideal. Then there exists $n \geq 1$ such that $I^{(n)} = 0$ and $I^{(n-1)} \neq 0$. We have $[I^{(n-1)}, I^{(n-1)}] = I^{(n)} = 0$, hence $I^{(n-1)}$ is a non-trivial abelian ideal in L. A contradiction.
(2) $\Longrightarrow$ (1). If L does not contain solvable ideals (except zero) then in particular it does not contain abelian ideals (except zero).
(2) $\Longleftrightarrow$ (3) is trivial. □

Lemma 4.3. *We have: simple $\Longrightarrow$ semisimple $\Longrightarrow$ reductive.*

Proof. **(1)** If L is simple, then the only nonzero ideal of L is L and by assumption it is not abelian. Therefore L is semisimple.
(2) If L is semisimple, then $\operatorname{rad} L = 0$. Then $Z(L) \subseteq \operatorname{rad} L$ implies $Z(L) = \operatorname{rad} L = 0$. □

Example 4.4. (1) We have seen that $\mathfrak{sl}_2$ does not contain ideals except zero and $\mathfrak{sl}_2$. This implies that $\mathfrak{sl}_2$ is a simple and semisimple Lie algebra.
(2) Consider the algebra $L = L_1 \oplus L_2 = \mathfrak{sl}_2 \oplus \mathfrak{sl}_2$. The only ideals of L are $0, \mathfrak{sl}_2 \oplus 0$, $0 \oplus \mathfrak{sl}_2$ and L. This implies that L is not simple. All nonzero ideals of L are non-abelian, hence L is semisimple.
(3) Consider the algebra $L = \mathfrak{gl}_2$. We know that $Z(L) = kI$. As $Z(L) \subseteq \operatorname{rad} L$, we conclude that $\operatorname{rad} L \neq 0$, hence L is not semisimple. We have $\mathfrak{gl}_2 = kI \oplus \mathfrak{sl}_2$ (both summands are ideals), hence $\operatorname{rad} \mathfrak{gl}_2 = \operatorname{rad}(kI) \oplus \operatorname{rad}(\mathfrak{sl}_2) = kI \oplus 0$. Alternatively, kI is solvable, hence $(\operatorname{rad} \mathfrak{gl}_2)/kI \simeq \operatorname{rad}(\mathfrak{gl}_2/kI) = \operatorname{rad} \mathfrak{sl}_2 = 0$. Therefore $\operatorname{rad} \mathfrak{gl}_2 = kI$. We obtain $\operatorname{rad} \mathfrak{gl}_2 = Z(\mathfrak{gl}_2)$, hence $\mathfrak{gl}_2$ is reductive.

Exercise 4.5. Let $L_1, \ldots, L_n$ be Lie algebras and

$$L = L_1 \oplus \cdots \oplus L_n = \{x = (x_1, \ldots, x_n) \,|\, x_i \in L_i \ \forall 1 \leq i \leq n\}$$

be their direct sum. Show that it has a Lie bracket $[x, y] = \sum_{i=1}^{n} [x_i, y_i]$ for $x, y \in L$.

Exercise 4.6. Show that

(1) If $f \colon L \to M$ is a surjective Lie algebra homomorphism, then $f(\operatorname{rad} L) \subseteq \operatorname{rad} M$.
(2) If $I \subseteq L$ is a solvable ideal, then $\operatorname{rad}(L/I) \simeq (\operatorname{rad} L)/I$. In particular, $L/\operatorname{rad} L$ is semisimple.
(3) $\operatorname{rad}(L_1 \oplus L_2) = \operatorname{rad}(L_1) \oplus \operatorname{rad}(L_2)$.

Proof. **(1)** $f(\operatorname{rad} L)$ is a quotient of $\operatorname{rad} L$, hence it is solvable. Therefore $f(\operatorname{rad} L) \subseteq \operatorname{rad} M$.
(2) Let $p\colon L \to L/I$ be the projection and $J = p^{-1}(\operatorname{rad}(L/I))$. Then $J/I \simeq \operatorname{rad}(L/I)$ is solvable, hence J is solvable. Therefore $J \subseteq \operatorname{rad} L$. We have $p(\operatorname{rad} L) \subseteq \operatorname{rad}(L/I)$ by (1), hence $\operatorname{rad} L \subseteq J$. Therefore $\operatorname{rad} L = J$ and $(\operatorname{rad} L)/I = J/I \simeq \operatorname{rad}(L/I)$.
(3) Let $L = L_1 \oplus L_2$ and $p_i\colon L \to L_i$ be projections. Then $p_i(\operatorname{rad} L) \subseteq \operatorname{rad} L_i$ by (1), hence $\operatorname{rad}(L) \subseteq \operatorname{rad}(L_1) \oplus \operatorname{rad}(L_2)$. The ideal $\operatorname{rad}(L_1) \oplus \operatorname{rad}(L_2) \subseteq L$ is solvable, hence $\operatorname{rad}(L_1) \oplus \operatorname{rad}(L_2) \subseteq L$. □

Remark 4.7.

(1) If L is reductive, then to prove that it is semisimple we need to show that $Z(L) = 0$.
(2) If L is abelian, then $Z(L) = L$ and therefore $Z(L) = \operatorname{rad} L = L$, hence L is reductive.
(3) We will see later that every reductive Lie algebra is a direct sum of a semisimple Lie algebra and an abelian Lie algebra.

4.2. **Reductivity criterion.**

Definition 4.8. Let $\sigma\colon L \times L \to K$ be a bilinear form. Then

(1) σ is called symmetric if $\sigma(x,y) = \sigma(y,x)$ for all $x, y \in L$.
(2) σ is called L-invariant if $\sigma([x,y],z) = \sigma(x,[y,z])$ for all $x, y, z \in L$.
(3) σ is called non-degenerate if $\sigma(x, L) = 0$ for $x \in L$ implies $x = 0$.

Remark 4.9. One can show that a bilinear map $\sigma\colon L \times L \to K$ is L-invariant if and only if the corresponding linear map $\sigma\colon L \otimes L \to K$ is a homomorphism of L-representations, where L is equipped with the adjoint representation structure and K is equipped with the trivial representation structure.

Lemma 4.10. *Let $\sigma\colon L \times L \to K$ be a bilinear L-invariant form. Then for every ideal $I \subseteq L$, the subspace*

$$I^{\perp} = \{x \in L \,|\, \sigma(x,y) = 0 \ \forall y \in I\}$$

is also an ideal.

Lemma 4.11. *Let (V, ρ) be an L-representation. Then the bilinear form*

$$\sigma_V\colon L \times L \to K, \qquad \sigma_V(x,y) = \operatorname{tr}(\rho(x)\rho(y)) \qquad \forall x, y \in L$$

is symmetric and L-invariant.

Theorem 4.12. *Let L be a Lie algebra over a field K of characteristic zero and (V, ρ) be an L-representation such that $\sigma_V\colon L \times L \to K$ is non-degenerate. Then L is reductive.*

Proof. We claim that for any representation V, we have $\sigma_V([I, L], L) = 0$, where $I = \operatorname{rad} L$. Under our assumptions this will imply $[I, L] = 0$ (as σ_V is non-degenerate) $\implies$ $\operatorname{rad} L = I \subseteq Z(L)$ $\implies$ $\operatorname{rad} L = Z(L)$ (as $Z(L) \subseteq \operatorname{rad} L$ automatically) $\implies$ L is reductive.

If $U \subseteq V$ is a subrepresentation, then $\sigma_V = \sigma_U + \sigma_{V/U}$ (compare the traces). Therefore by induction on dimension it is enough to show that $\sigma_V([I, L], L) = 0$ for an irreducible representation V (the one that does not contain subrepresentations except 0 and V). By Lie's Theorem applied to the algebra $I = \operatorname{rad} L$ (over the algebraic closure of K), there exists $0 \neq v_0 \in V$ such that $Iv_0 \subseteq Kv_0$. Let $\lambda\colon I \to K$ be defined by $xv_0 = \lambda(x)v_0$, $x \in I$ (it is linear). Then the subspace

$$V_\lambda = \{v \in V \,|\, xv = \lambda(x)v \ \forall x \in I\}$$

is nonzero (it contains $v_0 \neq 0$) and is an L-subrepresentation of V (see the proof of Lie's theorem). As V is irreducible, we conclude that $V_\lambda = V$. Therefore, for any $x \in I$, $y \in L$, $v \in V$ we have

$$[x,y]v = xyv - yxv = \lambda(x)yv - y(\lambda(x)v) = 0$$

and therefore

$$\sigma_V([x,y],z) = \operatorname{tr}(\rho([x,y])\rho(z)) = 0 \qquad \forall x \in I,\ y, z \in L.$$

□

Remark 4.13. Note that in the above lemma the map $\rho\colon L \to \mathfrak{gl}(V)$ is injective (we say that the representation V is faithful). Indeed, if $x \in \operatorname{Ker}\rho$, then $\sigma_V(x,y) = \operatorname{tr}(\rho(x)\rho(y)) = 0$ for all $y \in L$, hence $x = 0$. Conversely, we will see in Lemma 4.32 that if L is semisimple and V is a faithful representation, then σ_V is non-degenerate.

Theorem 4.14. *Let L have a faithful irreducible representation (V, ρ) (faithful means that $\rho\colon L \to \mathfrak{gl}(V)$ is injective, irreducible means that $0, V$ are the only subrepresentations of V). Then L is reductive and $\dim Z(L) \leq 1$. If moreover $\rho(L) \subseteq \mathfrak{sl}(V)$, then L is semisimple.*

Proof. We can assume that $L \subseteq \mathfrak{gl}(V)$. Using the same argument as in the previous theorem, we find a linear map $\lambda\colon I = \operatorname{rad} L \to K$ such that

$$V_\lambda = \{v \in V \mid xv = \lambda(x)v \ \forall x \in I\} = V.$$

This implies that every $x \in I$ acts on V as $\lambda(x)\operatorname{id}_V$. Therefore $I \subseteq K\operatorname{id}_V$, hence $I \subseteq Z(L)$ and $\dim I \leq 1$. This implies that $I = \operatorname{rad} L = Z(L)$ and L is reductive. Moreover, $\dim Z(L) \leq 1$.

Assume that $L \subseteq \mathfrak{sl}(V)$. If $I = K\operatorname{id}_V$, then $\operatorname{id}_V \in I \subseteq L \subseteq \mathfrak{sl}(V)$, a contradiction. Therefore $I = \operatorname{rad} L = 0$ and L is semisimple. □

Theorem 4.15. *Lie algebras $\mathfrak{gl}_n(K), \mathfrak{u}(n)$ are reductive. Lie algebras*

$$\mathfrak{sl}_n(K), \mathfrak{so}_n(K), \mathfrak{sp}_{2n}(K), \mathfrak{su}(n)$$

are semisimple.

Proof. Consider the standard representation $V = K^n$ of $\mathfrak{gl}_n = \mathfrak{gl}_n(K)$. Then σ_V is non-degenerate as

$$\sigma_V(AB) = \sum_{i,j} a_{ij}b_{ji}.$$

By the previous results we conclude that $\mathfrak{gl}_n$ is reductive, hence $\operatorname{rad}\mathfrak{gl}_n = Z(\mathfrak{gl}_n)$. One can easily show that $Z(\mathfrak{gl}_n) = KI_n$ and there is an orthogonal decomposition $\mathfrak{gl}_n = \mathfrak{sl}_n \oplus KI_n$. This implies that $\mathfrak{sl}_n \simeq \mathfrak{gl}_n/KI_n = \mathfrak{gl}_n/\operatorname{rad}(\mathfrak{gl}_n)$ is semisimple. The proof for other algebras is similar. □

4.3. **Cartan's criterion.**

Theorem 4.16 (Cartan)**.** *Let $L \subseteq \mathfrak{gl}(V)$ be a Lie algebra. Then L is solvable if and only if*

$$\operatorname{tr}(xy) = 0 \qquad \forall x \in [L, L],\ y \in L.$$

Remark 4.17. Alternatively, let (V, ρ) be a faithful representation (with injective ρ) and $\sigma_V(x, y) = \operatorname{tr}(\rho(x)\rho(y))$ be the corresponding bilinear form. Then L is solvable if and only if $\sigma_V([L, L], L) = 0$.

Proof. $\Longrightarrow$ If L is solvable then there exists a basis of V such that all elements of L are upper-triangular in this basis. Then any $x \in [L, L]$ is strictly upper triangular. Therefore, for any $y \in L$, the matrix xy is also strictly upper-triangular and $\operatorname{tr}(xy) = 0$.
$\Longleftarrow$ We will show that every $x \in [L, L] \subseteq \mathfrak{gl}(V)$ is nilpotent. This will imply that ad_x is also nilpotent, hence by Engel's theorem the algebra $[L, L]$ is nilpotent. Therefore L is solvable.

Let

$$A = [L, L] \subseteq B = L \subseteq \mathfrak{gl}(V), \qquad M = \{y \in \mathfrak{gl}(V) \mid [y, L] \subseteq [L, L]\}.$$

If $x = \sum_i [x_i, x_i'] \in [L, L] \subseteq M$ and $y \in M$, then

$$\operatorname{tr}(xy) = \sum \operatorname{tr}([x_i, x_i']y) = \sum \operatorname{tr}(x_i[x_i', y]).$$

Here $x_i' \in L$ and $y \in M$ imply $[x_i', y] \in [L, L]$ (by the definition of M). Therefore $\operatorname{tr}(x_i[x_i', y]) = 0$ by theorem assumption. We conclude that $\operatorname{tr}(xy) = 0$ for all $y \in M$. By the next lemma this implies that x is nilpotent and this is what we wanted to show. □

Lemma 4.18. *Let $A \subseteq B \subseteq \mathfrak{gl}(V)$ be subspaces and let*

$$M = \{x \in \mathfrak{gl}(V) \mid [x, B] \subseteq A\}.$$

If $x \in M$ satisfies $\operatorname{tr}(xy) = 0$ for all $y \in M$, then x is nilpotent.

Proof. Let $x = s + n$ be the Jordan decomposition (see §5.1). Then $\operatorname{ad}_x = \operatorname{ad}_s + \operatorname{ad}_n$ is the Jordan decomposition of ad_x. As $\operatorname{ad}_x(B) = [x, B] \subseteq A$, we have also $\operatorname{ad}_s(B) \subseteq A$ (ad_s is a polynomial of ad_x with the zero constant term). Let $(v_1, \dots, v_n)$ be a basis for the Jordan normal form of x. Let $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues. Define $E \subseteq K$ to be a $\mathbb{Q}$-vector space generated by the elements $\lambda_1, \dots, \lambda_n \in K$. We want to show that x is nilpotent, that is, $\lambda_1 = \dots = \lambda_n = 0$ or equivalently, $E = 0$. Let $f\colon E \to \mathbb{Q}$ be a linear map. Define $y \in \operatorname{End} V$ such that $yv_i = f(\lambda_i)v_i$. Let $E_{ij} \in \operatorname{End} V$ be defined by $E_{ij}(v_k) = \delta_{jk}v_i$ (elementary matrix). Then

$$\operatorname{ad}_y(E_{ij}) = \left[\sum_k f(\lambda_k)E_{kk}, E_{ij}\right] = f(\lambda_i - \lambda_j)E_{ij}.$$

Similarly $\operatorname{ad}_s(E_{ij}) = (\lambda_i - \lambda_j)E_{ij}$. This means that ad_y and ad_s are diagonalizable in the same basis of $\operatorname{End} V$. Let $\lambda_{ij} = \lambda_i - \lambda_j$. Applying Chinese remainder theorem, we can find a polynomial $p \in K[t]$ such that

$$p(t) \equiv f(\lambda_{ij}) \pmod{t - \lambda_{ij}} \qquad \forall i, j, \qquad p(t) \equiv 0 \pmod{t}.$$

This polynomial does not have a constant term and satisfies $p(\lambda_{ij}) = f(\lambda_{ij})$. Therefore $p(\operatorname{ad}_s) = \operatorname{ad}_y$. We obtain from $\operatorname{ad}_s(B) \subseteq A \subseteq B$ that $\operatorname{ad}_y(B) \subseteq A$. This implies that $y \in M$. By assumption

$$0 = \operatorname{tr}(xy) = \sum_i f(\lambda_i)\lambda_i.$$

Applying f to this equation, we obtain $\sum_i f(\lambda_i)^2 = 0$. As $f(\lambda_i) \in \mathbb{Q}$, we obtain $f(\lambda_i) = 0$ for all i. This implies that every linear map $f\colon E \to \mathbb{Q}$ is zero. Therefore $E = 0$ as required. □

Consider the adjoint representation $\operatorname{ad}\colon L \to \mathfrak{gl}(L)$ and the corresponding symmetric L-invariant bilinear form, called the *Killing form*

$$\kappa(x,y) = \kappa_L(x,y) = \operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_y) \qquad x,y \in L.$$

Later we will need the following result about Killing forms.

Lemma 4.19. *Let $I \subseteq L$ be an ideal of a Lie algebra. Then the Killing form κ_I of I coincides with the restriction of the Killing form κ_L to I.*

Proof. If $A \in \operatorname{End}(L)$ satisfies $A(L) \subseteq I$, then $\operatorname{tr}(A) = \operatorname{tr}(A|_I)$ (write A in a basis of L that starts with a basis of I). In particular, for any $x, y \in I$ the operators $\operatorname{ad}_x, \operatorname{ad}_y$ map L to I. Applying the first observation to $A = \operatorname{ad}_x \operatorname{ad}_y$, we obtain

$$\kappa_L(x,y) = \operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_y) = \operatorname{tr}(A) = \operatorname{tr}(A|_I) = \kappa_I(x,y).$$

□

Theorem 4.20 (Cartan)**.** *A Lie algebra L is solvable if and only if $\kappa([L,L],L) = 0$.*

Proof. $\Longrightarrow$ Assume that L is solvable. For any $x, y, z \in L$, we have

$$\operatorname{tr}([\operatorname{ad}_x, \operatorname{ad}_y] \operatorname{ad}_z) = \operatorname{tr}(\operatorname{ad}_{[x,y]} \operatorname{ad}_z) = \kappa([x,y],z) = 0.$$

The image $\bar{L} = \operatorname{ad}(L) \subseteq \mathfrak{gl}(L)$ is a solvable Lie algebra as L is solvable. Applying Theorem 4.16 to $\bar{L} \subseteq \mathfrak{gl}(L)$, we obtain $\operatorname{tr}([\operatorname{ad}_x, \operatorname{ad}_y] \operatorname{ad}_z) = 0$ for all $x, y, z \in L$. Therefore $\kappa([x,y],z) = 0$.
$\Longleftarrow$ If $\kappa([L,L],L) = 0$, then $\operatorname{tr}([\operatorname{ad}_x, \operatorname{ad}_y] \operatorname{ad}_z) = 0$ for all $x, y, z \in L$. Applying Theorem 4.16 to $\bar{L} \subseteq \mathfrak{gl}(L)$, we conclude that $\bar{L}$ is solvable. We know that $\operatorname{Ker}(\operatorname{ad}) = Z(L)$, hence $\bar{L} \simeq L/\operatorname{Ker}(\operatorname{ad}) = L/Z(L)$. As $\bar{L}$ and $Z(L)$ are solvable, we conclude that L is solvable. □

Theorem 4.21. *Let L be a Lie algebra. The following conditions are equivalent.*

1. *L is semisimple.*
2. *The Killing form κ is non-degenerate (Cartan's criterion).*
3. *L is a direct sum $I_1 \oplus \cdots \oplus I_n$ of simple ideals (simple as Lie algebras).*
4. *L is a direct sum $L_1 \oplus \cdots \oplus L_n$ of simple Lie algebras.*

Proof. **(1) $\Longrightarrow$ (2).** Let L be a semisimple Lie algebra. Define

$$I = L^\perp = \{x \in L \mid \kappa(x, L) = 0\}$$

which is an ideal in L. By Lemma 4.19 $\kappa_I = \kappa|_I = 0$. Therefore by the Cartan criterion I is solvable. As L is semisimple, we obtain that $I = 0$. This means that κ is non-degenerate.
(2) $\Longrightarrow$ (1). If κ is non-degenerate, then L is reductive by Theorem 4.12, hence $\operatorname{rad} L = Z(L)$. If $x \in Z(L)$, then $\operatorname{ad}_x = 0$, hence $\kappa(x,y) = \operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_y) = 0$ for all $y \in L$. Therefore $x = 0$ as κ is non-degenerate. We conclude that $\operatorname{rad} L = Z(L) = 0$, hence L is semisimple.

Alternatively, if L is not semisimple, then there exists an abelian ideal $0 \neq I \subseteq L$. For any $x, z \in L$ and $y \in I$, we have

$$(\operatorname{ad}_x \operatorname{ad}_y)^2 z = [x,[y,[x,[y,z]]]] \in [x,[I,I]] = 0$$

as $y \in I$ and $[x,[y,z]] \in I$. This implies that $(\operatorname{ad}_x \operatorname{ad}_y)^2 = 0$ and $\operatorname{ad}_x \operatorname{ad}_y$ is a nilpotent operator. Therefore its trace is zero and

$$\kappa(x,y) = \operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_y) = 0 \qquad \forall x \in L,\ y \in I.$$

This implies that κ is degenerate, a contradiction.

(1) $\Longrightarrow$ (3). If L is simple then we are done. Assume that $I \subseteq L$ is a non-trivial ideal. Then

$$I^\perp = \{x \in L \,|\, \kappa(x, I) = 0\}.$$

is an ideal (exercise).The Killing form is zero on the ideal $I \cap I^\perp$. Therefore this is a solvable ideal. By semisimplicity of L, we obtain $I \cap I^\perp = 0$. This implies that $L = I \oplus I^\perp$ (compare dimensions). Moreover $[I, I^\perp] \subseteq I \cap I^\perp = 0$. Ideals of $I, I^\perp$ are automatically ideals of L. This implies that there are no solvable ideals in $I, I^\perp$ and these algebras are semisimple. Then by induction we can decompose $I, I^\perp$ into a direct sum of simple ideals.
(3) $\Longrightarrow$ (4). Trivial.
(4) $\Longrightarrow$ (1). Let $L = L_1 \oplus \cdots \oplus L_n$ be a direct sum of simple Lie algebras. Let $I \subseteq L$ be an abelian ideal. Then the image of I under the projection $\pi_k \colon L \to L_k$ is an abelian ideal, but the only abelian ideal of L_k is zero. This implies that $\pi_k(I) = 0$ for all $1 \leq k \leq n$. Therefore $I = 0$. This implies that L is semisimple. □

Remark 4.22. If L is a semisimple Lie algebra with a decomposition $L = I_1 \oplus \cdots \oplus I_n$ into a sum of simple ideals as above, then every simple ideal $I \subseteq L$ is equal to one of I_k (not just isomorphic). Indeed, $[I, L] \subseteq I$ is non-zero as I is non-abelian. Therefore $I = [I, L] = [I, I_1] \oplus \cdots \oplus [I, I_n]$. Only one of the summands $[I, I_k]$ can be nonzero and then we have $I = [I, I_k] = I_k$. Note that $[I_j, I_k] \subseteq I_j \cap I_k = 0$ for $j \neq j$ and $[I_k, I_k] = I_k$. Therefore

$$L' = [L, L] = \sum_k [I_k, I_k] = \sum_k I_k = L.$$

Theorem 4.23. *Let L be a semisimple Lie algebra. Then the adjoint map* $\operatorname{ad} \colon L \to \operatorname{Der} L$ *is an isomorphism.*

Proof. We have $\operatorname{Ker}(\operatorname{ad}) = Z(L)$ which is an abelian ideal. Therefore $\operatorname{Ker}(\operatorname{ad}) = 0$ and ad is injective. Let $M = \operatorname{Der} L$ and $I = \operatorname{ad}(L) \subseteq M$. For all $\delta \in M$ and $x \in L$, we have

$$[\delta, \operatorname{ad}_x] = \operatorname{ad}_{\delta x} \in I.$$

Therefore $I \subseteq M$ is an ideal. The subspace

$$I^\perp = \{x \in M \,|\, \kappa_M(x, I) = 0\}$$

is also an ideal. As $I \simeq L$ is semisimple, we obtain that $\kappa_M|_I = \kappa_I$ is non-degenerate, hence $I \cap I^\perp = 0$. For any $\delta \in I^\perp$ and $x \in L$, we have

$$\operatorname{ad}_{\delta x} = [\delta, \operatorname{ad}_x] \in [I^\perp, I] \subseteq I \cap I^\perp = 0.$$

But ad is injective, hence $\delta x = 0$ for all $x \in L$. Therefore $\delta = 0$ and $I^\perp = 0$. This implies $I = M$, hence $\operatorname{ad}(L) = \operatorname{Der} L$ and $\operatorname{ad} \colon L \to \operatorname{Der} L$ is an isomorphism. □

4.4. **Irreducible and completely reducible representations.**

Definition 4.24. Let L be a Lie algebra and V be an L-representation. Then

(1) V is called an *irreducible representation* (or a *simple module*) if it does not contain subrepresentations except 0 and V.
(2) V is called a *completely reducible representation* (or a *semisimple* module) if it is a direct sum of irreducible representations.

Remark 4.25. Any ideal of a Lie algebra L can be considered as a subrepresentation of an adjoint representation on L. This implies that (non-abelian) L is a simple algebra if and only if the adjoint representation is irreducible. We have seen that every semisimple algebra is a direct sum of simple ideals. This implies that the adjoint representation is a direct sum of irreducible representations and is therefore completely reducible (a semisimple module). Conversely, if L is a Lie algebra such that its adjoint representation is completely reducible (without summands that are trivial representations), then L is a direct sum of simple ideals. This implies that L is a semisimple Lie algebra.

Lemma 4.26. *Let L be a Lie algebra and V be an L-module. The following conditions are equivalent*

(1) V is a sum of simple L-modules.
(2) V is a direct sum of simple L-modules (that is, semisimple).
(3) Any submodule $W \subseteq V$ has a complement (a submodule $W' \subseteq V$ such that $V = W \oplus W'$).

Proof. **(1) $\Longrightarrow$ (2).** Let $V = \sum_{i=1}^{n} S_i$ with simple S_i and with minimal possible n. By induction on n, $W = S_1 \oplus \cdots \oplus S_{n-1}$ is a direct sum. We have $W \cap S_n \subseteq S_n \implies W \cap S_n = 0$ or $W \cap S_n = S_n$. If $W \cap S_n = 0 \implies V = W \oplus S_n = S_1 \oplus \cdots \oplus S_n$. If $W \cap S_n = S_n \implies S_n \subseteq W \implies V = W = S_1 \oplus \cdots \oplus S_{n-1}$.
(2) $\Longrightarrow$ (3). Let $W \subseteq V = S_1 \oplus \cdots \oplus S_n$. Consider the projection $\pi\colon V \to S_1 \oplus \cdots \oplus S_{n-1} =: V'$ with $\operatorname{Ker}\pi = S_n$. By induction $V' = \pi(W) \oplus W'$ for some submodule $W' \subseteq V'$. We have $W \cap S_n \subseteq S_n \implies W \cap S_n = 0$ or $W \cap S_n = S_n$. If $W \cap S_n = 0 \implies W \to \pi(W)$ is an isomorphism and $V = V' \oplus S_n = W \oplus W' \oplus S_n$. If $W \cap S_n = S_n \implies S_n \subseteq W \implies V = W \oplus W'$.
(3) $\Longrightarrow$ (1). Choose a simple submodule $S_1 \subseteq V$. By assumption, it has a complement $V_1 \subseteq V$. Choose a simple submodule $S_2 \subseteq V_1$. By assumption, the module $S_1 \oplus S_2$ has a complement $V_2 \subseteq V$. Continuing this procedure, we obtain $V = S_1 \oplus \cdots \oplus S_n$. □

Example 4.27. Consider a one dimensional Lie algebra $L = Kx$ and a 2-dimensional L-module $V = Kv_1 \oplus Kv_2$ with an action of x given by the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Then the subspace $W = Kv_1 \subseteq V$ is an L-submodule. But is does not have a complement. Therefore V is not-semisimple.

On the other hand, if $L = Kx \subseteq \mathfrak{gl}(V)$ and x is a semisimple endomorphism (diagonalisable), then V is a direct sum of 1-dimensional L-submodules. Therefore V is a semisimple L-module.

Example 4.28. The standard representation K^n of $L = \mathfrak{gl}_n(K)$ is irreducible. Indeed, for any non-zero $v, w \in \mathfrak{gl}_n(K)$ there exists an endomorphism $x \in \mathfrak{gl}_n(K)$ such that $xv = w$. This implies that any submodule $W \subseteq V$ that contains v should be equal to the whole space V.

Remark 4.29. Any submodule and any quotient module of a semisimple module is again semisimple. If $\pi\colon V \to W$ is an epimorphism and $V = \sum_i S_i$ is a sum of simple modules, then $W = \sum_i \pi(S_i)$ is a sum of simple (and zero) modules and therefore W is semisimple. If $W \subseteq V$ is a submodule of a semisimple module, then there exists a complement $V = W \oplus W'$. Therefore there exists a surjective homomorphism $V \to W$ and W is semisimple.

Lemma 4.30 (Schur lemma)**.** *Let V be a simple module over a Lie algebra L. Then* $\operatorname{End}_L(V) = K \operatorname{id}_V$.

Proof. Let $f\colon V \to V$ be a homomorphism of modules. Let $\lambda \in K$ be an eigenvalue of f and let

$$V_\lambda = \{v \in V \mid f(v) = \lambda v\} \subseteq V$$

be the corresponding eigenspace. It is a submodule of V: if $x \in L$ and $v \in V_\lambda$ then

$$f(xv) = xf(v) = x\lambda v = \lambda xv$$

and therefore $xv \in V_\lambda$. As V is simple, we obtain $V_\lambda = V$. This implies $f = \lambda \operatorname{id}_V$. Therefore $\operatorname{End}_L(V) = K \operatorname{id}_V$. □

4.5. **Casimir operator.** Let $\sigma\colon L \times L \to K$ be a symmetric non-degenerate L-invariant form. Let $(x_1, \dots, x_n)$ be a basis of L and $(y_1, \dots, y_n)$ be the dual basis of L with respect to σ

$$\sigma(x_i, y_j) = \delta_{ij}, \qquad 1 \leq i, j \leq n.$$

For any L-repsentation V, define a linear operator, called the *Casimir operator*

$$C_\sigma = C_\sigma^V : V \to V, \qquad C_\sigma(v) = \sum_{i=1}^{n} x_i y_i v \qquad \forall v \in V.$$

Lemma 4.31.

(1) The Casimir operator C_σ is independent of the choice of a basis in L.
(2) $C_\sigma\colon V \to V$ is a homomorphism of L-representations.

Proof. **(1)** Let (x_i') be a different basis of L and let (y_i') be the dual basis of L with respect to σ. Let $x_k' = \sum_i a_{ki} x_i$, $y_l' = \sum_j b_{jl} y_j$. Then

$$\delta_{kl} = \sigma(x_k', y_l') = \sum_{i,j} a_{ki} b_{jl} \sigma(x_i, y_j) = \sum_{i,j} a_{ki} b_{jl} \delta_{ij} = \sum_i a_{ki} b_{il} = (AB)_{kl},$$

hence $AB = I_n$. Therefore

$$\sum_k x_k' y_k' = \sum_{i,j} \left(\sum_k a_{ki} b_{jk} \right) x_i y_j = \sum_{i,j} (BA)_{ji} x_i y_j = \sum_i x_i y_i.$$

(2) For any $x \in L$, let

$$[x_k, x] = \sum_i a_{ki} x_i, \qquad [x, y_l] = \sum_j b_{jl} y_j.$$

Then

$$a_{kl} = \sigma([x_k, x], y_l) = \sigma(x_k, [x, y_l]) = b_{kl}.$$

Therefore

$$C_\sigma(xv) - xC_\sigma(v) = \sum_k (x_k y_k x - x x_k y_k) v = \sum_k (x_k [y_k, x] + [x_k, x] y_k) v$$
$$= \left(-\sum_{k,i} b_{ik} x_k y_i + \sum_{k,i} a_{ki} x_i y_k \right) v = 0.$$

□

Lemma 4.32. *Let L be a semisimple Lie algebra and (V, ρ) be its faithful representation. Then*

(1) The bilinear form $\sigma = \sigma_V\colon L \times L \to K$, $\sigma(x, y) = \operatorname{tr}(\rho(x)\rho(y))$ is non-degenerate.
(2) The Casimir operator $C_\sigma\colon V \to V$ satisfies $\operatorname{tr} C_\sigma = \dim L$.

Proof. **(1)** Consider the ideal $I = L^\perp = \{x \in L \mid \sigma(x, L) = 0\}$. We have $\sigma([I, I], I) \subseteq \sigma(I, I) = 0$. By Theorem 4.16 this implies that I is solvable. As L is semisimple, we obtain $I = 0$ and therefore σ is non-degenerate.
(2) We have $C_\sigma = \sum_i \rho(x_i)\rho(y_i)$, hence $\operatorname{tr}(C_\sigma) = \sum_i \operatorname{tr}(\rho(x_i)\rho(y_i)) = \sum_i \sigma(x_i, y_i) = \dim L$. □

4.6. **Complete reducibility (Weyl's theorem).**

Theorem 4.33 (Weyl)**.** *Every (finite-dimensional) representation of a semisimple Lie algebra is completely reducible (i.e., semisimple).*

We will need some preparations before proving this theorem. Given an L-representation V, we can construct its subrepresentations

$$V^L = \{v \in V \mid Lv = 0\}, \qquad LV = \langle xv : x \in L, v \in V\rangle,$$

where the angular brackets denote the subspace spanned by given vectors.

Lemma 4.34. *Let L be a semisimple Lie algebra. Then*

(1) For any L-representation V, we have $V = V^L \oplus LV$.
(2) For any surjective homomorphism $V \to W$ of L-representations, the induced map $V^L \to W^L$ is surjective.

Proof. **(1)** We can substitute L by $\rho(L) \subseteq \mathfrak{gl}(V)$ which is also a semisimple Lie algebra (L is a sum of simple ideals, hence $\rho(L)$ is also a sum of simple ideals). If $L = 0$, then $V^L = V$ and $LV = 0$, hence $V = V^L \oplus LV$. Assume that $L \neq 0$. The map $\rho\colon L \to \mathfrak{gl}(V)$ is injective. By Lemma 4.32, the corresponding bilinear form $\sigma = \sigma_V$ on L is non-degenerate. Let $C = C_\sigma\colon V \to V$ be the Casimir operator. Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of C having multiplicities $m_1, \dots, m_k$. Generalized eigenspaces

$$V_i = \operatorname{Ker}(C - \lambda_i)^{m_i} \subseteq V$$

are subrepresentations of V (as $Cxv = xCv$ for all $x \in L$, $v \in V$) and $V = \bigoplus_{i=1}^k V_i$. If $k > 1$, then by induction on $\dim V$ we have $V_i = V_i^L \oplus LV_i$ and the statement of the theorem follows. If $k = 1$, then $\operatorname{tr} C = \lambda_1 \dim V = \dim L$ by Lemma 4.32. Therefore $\lambda_1 \neq 0$ and C is invertible. Injectivity of C implies that $V^L = 0$ (C acts as zero on V^L). Surjectivity of C implies $LV = V$. Therefore $V = V^L \oplus LV$.
(2) The homomorphism $V \to W$ maps V^L to W^L and maps LV to LW. We obtain from the above decomposition that if $V \to W$ is surjective, then $V^L \to W^L$ is surjective. □

Proof of the Weyl theorem. We will show that any subrepresentation $W \subseteq V$ has a complement. Consider the surjective linear map

$$\phi\colon \operatorname{Hom}(V, W) \to \operatorname{Hom}(W, W), \qquad f \mapsto f|_W.$$

The Hom-spaces are equipped with L-representation structure and ϕ is a homomorphism between them. By the previous result it induces a surjective map $\operatorname{Hom}(V, W)^L \to \operatorname{Hom}(W, W)^L$.

The space $\operatorname{Hom}(V, W)^L$ consists of linear maps $f\colon V \to W$ that are L-invariant, that is, $f(xv) = xf(v)$ for all $x \in L$ and $v \in V$. This means that $f\colon V \to W$ is a homomorphism of L-representations. By the surjectivity of the above map, we can find $f \in \operatorname{Hom}(V, W)^L$ such that $f|_W = \mathrm{id}_W$. Then $V = W \oplus \operatorname{Ker} f$: if $v \in V$, then $f(v) \in W$ and $f(v - f(v)) = f(v) - f(v) = 0$, hence $v - f(v) \in \operatorname{Ker} f$; if $v \in W \cap \operatorname{Ker} f$, then $v = f(v) = 0$, hence $W \cap \operatorname{Ker} f = 0$. This implies that W has a complement. □

Corollary 4.35. *Let L be a Lie algebra. The following conditions are equivalent*

(1) L is reductive.
(2) Adjoint representation of L is a direct sum of irreducible representations.
(3) L is a direct sum of $Z(L)$ and a semisimple Lie algebra.

Proof. **(1) $\Longrightarrow$ (2).** Consider the adjoint representation $\mathrm{ad}\colon L \to \mathfrak{gl}(L)$. Its kernel is $\operatorname{Ker}(\mathrm{ad}) = Z(L) = \operatorname{rad}(L)$. We can consider L as a representation of $M = L/\operatorname{Ker}(\mathrm{ad}) = L/\operatorname{rad}(L)$ which

is semisimple. By the Weyl Theorem L can be decomposed as a direct sum of irreducible M-representations (which are also irreducible L-representations).

(2) $\Longrightarrow$ (3). Let $L = I_1 \oplus \cdots \oplus I_n$ be a decomposition into a direct sum of irreducible representations (which are automatically ideals). We have $[L, I_k] \subseteq I_k$, hence $[I_j, I_k] \subseteq I_j \cap I_k = 0$ for $j \neq k$. If $J \subseteq I_k$ is an ideal of the algebra I_k, then $[L, J] = [I_k, J] \subseteq J$ and $J \subseteq I_k$ is an L-subrepresentation, hence $J = 0$ or $J = I_k$. This implies that I_k is either a simple or an abelian algebra. If I_k is abelian, then $[L, I_k] = [I_k, I_k] = 0$ and $I_k \subseteq Z(L)$. This implies that we can decompose $L = I \oplus J$, where $I \subseteq Z(L)$ and J is a semisimple Lie algebra. The algebra $L/I \simeq J$ is semisimple, hence I is the maximal solvable ideal. This implies that $I = Z(L) = \operatorname{rad} L$, hence $L = Z(L) \oplus J$.

(3) $\Longrightarrow$ (1). Our assumption implies that $L/Z(L)$ is a semisimple Lie algebra. Therefore $Z(L)$ is the maximal solvable ideal and $Z(L) = \operatorname{rad} L$. $\square$

4.7. **Representations of $\mathfrak{sl}_2$.** We consider the algebra $\mathfrak{sl}_2 = \mathfrak{sl}_2(K)$ over an algebraically closed field K of characteristic zero. We know that $\mathfrak{sl}_2$ is simple, hence semisimple. By the Weyl's Theorem every representation of $\mathfrak{sl}_2$ is a direct sum of simple representations. Therefore we need to classify just simple representations of $\mathfrak{sl}_2$.

Theorem 4.36. *Let x, y, h be the standard generators of $\mathfrak{sl}_2$ with*

$$[x, y] = h, \qquad [h, x] = 2x, \qquad [h, y] = -2y. \tag{4}$$

For any $n \geq 0$, there exists exactly one simple representation $V(n)$ of $\mathfrak{sl}_2$ of dimension $n+1$ (up to an isomorphism). Representation $V = V(n)$ is a direct sum of 1-dimensional h-eigenspaces

$$V = V_{-n} \oplus V_{-n+2} \oplus \cdots \oplus V_{n-2} \oplus V_n,$$

where V_k is the subspace of h-eigenvectors with eigenvalue k, and x, y induce isomorphisms

$$x \colon V_k \to V_{k+2}, \qquad y \colon V_{k+2} \to V_k$$

whenever $V_k, V_{k+2} \neq 0$.

Proof. Let V be a simple representation. For any $\lambda \in K$ define the eigenspace

$$V_\lambda = \{v \in V \mid hv = \lambda v\}.$$

Then

$$xV_\lambda \subseteq V_{\lambda+2}, \qquad yV_\lambda \subseteq V_{\lambda-2}.$$

Indeed, if $v \in V_\lambda$, then

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

Therefore $xv \in V_{\lambda+2}$. Similarly $yv \in V_{\lambda-2}$. There exists $\lambda \in K$ such that $V_\lambda \neq 0$ and $V_{\lambda+2} = 0$. Let $0 \neq v \in V_\lambda$. Then $xv \in V_{\lambda+2} = 0$ and we obtain

$$hv = \lambda v, \qquad xv = 0.$$

We will prove by induction that

$$hy^i v = (\lambda - 2i)y^i v, \qquad xy^i v = i(\lambda - i + 1)y^{i-1} v. \tag{5}$$

The statement for $i = 0$ was just proved. As $v \in V_\lambda$, we obtain $yv \in V_{\lambda-2}$ and generally $y^i v \in V_{\lambda-2i}$. Therefore $hy^i v = (\lambda - 2i)y^i v$. Assume that the second formula is proved for $i \geq 0$. Then

$$\begin{aligned} xy^{i+1}v = xy(y^i v) = ([x, y] + yx)y^i v = (h + yx)y^i v = (\lambda - 2i)y^i v + i(\lambda - i + 1)y^i v \\ = (\lambda(i+1) - i^2 - i)y^i v = (i+1)(\lambda - i)y^i v. \end{aligned}$$

This proves the induction step. Let $n \geq 0$ be the maximal integer such that $y^n v \neq 0$. Then the vector space

$$\langle v, yv, \ldots, y^n v \rangle \subseteq V$$

is a submodule of V (it is closed under the action of x, y, h). As V is simple, the above subspace should be the whole space V. The above vectors are linearly independent as $y^i v \in V_{\lambda-2i}$ are contained in different eigenspaces. This implies that $\dim V = n + 1$. As $y^{n+1} v = 0$, we obtain

$$0 = xy^{n+1} v = (n+1)(\lambda - n)y^n v,$$

hence $\lambda - n = 0$ and $\lambda = n$. The above formulas (5) uniquely determine representation V.

To show existence, we use the above formulas as a motivation for our construction. Let V be a vector space with a basis $(v_0, \dots, v_n)$ (v_i corresponds to $y^i v \in V_{n-2i}$ in the above discussion). Define the linear operators x, y, h on V by the rule

$$xv_i = i(n-i+1)v_{i-1}, \qquad yv_i = v_{i+1}, \qquad hv_i = (n-2i)v_i, \qquad 0 \le i \le n, \tag{6}$$

where $v_{-1} = v_{n+1} = 0$. In order to show that this rule defines a representation of $\mathfrak{sl}_2$, we need to verify the commutator relations (4). Let us check just the first one

$$[x,y]v_i = (xy-yx)v_i = xv_{i+1} - yxv_i = (i+1)(n-i)v_i - i(n-i+1)v_i = (n-2i)v_i = hv_i.$$

This implies that V is indeed a representation of $\mathfrak{sl}_2$. We have to check V is a simple representation. Let $W \subseteq V$ be a non-zero subrepresentation and $0 \ne v \in W$. Assume that $v = \sum_i a_i v_i$ and $0 \le i \le n$ is minimal with the property $a_i \ne 0$. Then $y^{n-i}v = a_i v_n \in W$, hence $v_n \in W$. Applying the operator x, we obtain that $v_{n-1}, \dots, v_1, v_0$ are also contained in W. Therefore $W = V$ and V is a simple module. □

Remark 4.37. Note that

(1) The 1-dimensional representation $V(0)$ is the trivial representation K.
(2) The 2-dimensional representation $V(1)$ is the standard representation K^2 of $\mathfrak{sl}_2(K)$.
(3) The 3-dimensional representation $V(2)$ is the adjoint representation of $\mathfrak{sl}_2$.

Corollary 4.38. *Let V be a finite-dimensional representation of $\mathfrak{sl}_2$. Then*

(1) $V = \bigoplus_{k\in\mathbb{Z}} V_k$, where $V_k = \{v \in V \mid hv = kv\}$ is the h-eigenspace (also called a weight space) with eigenvalue k (also called a weight).
(2) x maps V_k to V_{k+2} and y maps V_k to V_{k-2}.
(3) $x^k \colon V_{-k} \to V_k$ and $y^k \colon V_k \to V_{-k}$ are isomorphisms, for any $k \ge 1$.

Proof. **(1)** Every finite-dimensional representation V of $\mathfrak{sl}_2$ is a direct sum of simple representations $V(n)$ classified in the previous theorem. By this theorem, if $V = V(n)$, then $V = \bigoplus_{k\in\mathbb{Z}} V_k$, the direct sum of h-eigenspaces. This implies the first statement of the corollary for simple modules, hence also for direct sums of simple modules.
(2) We have seen that it is true for simple modules. Therefore it is also true for arbitrary modules.
(3) It is enough to prove the statement for a simple module $V = V(n)$. We know that

$$V = V_{-n} \oplus V_{-n+2} \oplus \dots \oplus V_{n-2} \oplus V_n.$$

If k and n have different parities, then $V_k = V_{-k} = 0$ and the statement is automatically true. If $k > n$, then again $V_k = V_{-k} = 0$ and the statement is true. Assume that k and n have the same parity and $k \le n$. By the previous theorem, there is a sequence of isomorphisms

$$V_{-n} \to^x V_{-n+2} \to^x \dots \to^x V_{n-2} \to^x V_n.$$

In particular, $x^k \colon V_{-k} \to V_k$ is an isomorphism. The proof for y^k is the same. □

Remark 4.39. Let V be a representation of $\mathfrak{sl}_2$. Then we can decompose V as a direct sum of simple modules $V = \bigoplus_{n\ge 0} V(n)^{m_n}$. We can determine multiplicities m_n as follows. Let $V = \bigoplus_{k\in\mathbb{Z}} V_k$ be the weight space decomposition. A weight $k \ge 0$ appears (with multiplicity one) only in representations

$$V(k),\ V(k+2),\ \dots$$

This implies

$$\dim V_k = m_k + m_{k+2} + \dots$$

Therefore

$$m_k = \dim V_k - \dim V_{k+2}.$$

Example 4.40. Consider representation $V = V(1) \otimes V(2)$. The weights of $V(1)$ ar $(-1, 1)$ and the weights of $V(2)$ are $-2, 0, 2$. Therefore the weights of the tensor product V (they are sums of pairs of weights) are $(-3, -1 * 2, 1 * 2, 3)$, where $k * m$ means that the weight k has multiplicity m. Comparing the weights we conclude that $V \simeq V(3) \oplus V(1)$.

4.7.1. *Action of $\mathfrak{sl}_2$ on the ring of polynomials.* Define an $\mathfrak{sl}_2$ representation on $K[u, v]$ by the map $\rho\colon \mathfrak{sl}_2 \to \operatorname{Der} K[u, v] \subseteq \mathfrak{gl}(K[u, v])$

$$\rho(x) = u\partial_v, \qquad \rho(y) = v\partial_u, \qquad \rho(h) = u\partial_u - v\partial_v.$$

To verify that ρ is indeed a representation, we note that

$$\begin{gathered}[\rho(x), \rho(y)] = [u\partial_v, v\partial_u] = u(\partial_v v)\partial_u - v(\partial_u u)\partial_v = u\partial_u - v\partial_v = \rho(h),\\ [\rho(h), \rho(x)] = [u\partial_u, u\partial_v] - [v\partial_v, u\partial_v] = u\partial_v - (-u\partial_v) = 2\rho(x),\\ [\rho(h), \rho(x)] = [u\partial_u, v\partial_u] - [v\partial_v, v\partial_u] = -v\partial_u - v\partial_u = -2\rho(y)\end{gathered}$$

Representation $K[u, v]$ is infinite dimensional and can be decomposed as a direct sum of finite-dimensional spaces

$$K[u, v]_n = \left\langle u^0 v^n, u^1 v^{n-1}, \dots, u^n v^0 \right\rangle$$

which are spaces of homogeneous polynomials of total degree n. It is clear that operators x, y, h don't change the total degree, hence $K[u, v]_n$ are subrepresentations.

Remark 4.41. Consider the action of $\mathfrak{sl}_2$ on $K[u, v]_1 = Ku \oplus Kv$. We have

$$\rho(x)u = u\partial_v u = 0, \qquad \rho(x)v = u\partial_v v = u,$$

hence $\rho(x) = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Similarly, $\rho(y) = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, $\rho(h) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and we obtain the standard representation of $\mathfrak{sl}_2$ on K^2.

Let us show that there is an isomorphism between representations $K[u, v]_n$ and $V(n)$ constructed in the previous theorem. We consider an isomorphism of vector spaces

$$\phi\colon V(n) \to K[u, v]_n, \qquad v_i \mapsto \frac{1}{(n-i)!} u^{n-i} v^i.$$

Note that

$$h\phi(v_i) = \frac{1}{(n-i)!}(u\partial_u - v\partial_v)(u^{n-i}v^i) = \frac{n-i-i}{(n-i)!} u^{n-i}v^i = (n-2i)\phi(v_i).$$

$$x\phi(v_i) = \frac{1}{(n-i)!} u\partial_v(u^{n-i}v^i) = \frac{i(n-i+1)}{(n-i+1)} u^{n-i+1}v^{i-1} = i(n-i+1)\phi(v_{i-1}).$$

$$y\phi(v_i) = \frac{1}{(n-i)!} v\partial_u(u^{n-i}v^i) = \frac{1}{(n-i-1)!} u^{n-i-1}v^{i+1} = \phi(v_{i+1}).$$

This implies that ϕ is an isomorphism of $\mathfrak{sl}_2$-representations.

Remark 4.42. Alternatively, we can exponentiate the action of $\mathrm{SL}_2(\mathbb{R})$ on $\mathbb{R}[u, v]$ to the action of $\mathfrak{sl}_2(\mathbb{R})$. The Lie group $\mathrm{SL}_2(\mathbb{R})$ acts on $\mathbb{R}^2$ in the obvious way and induces an action on $C^\infty(\mathbb{R}^2)$ (as well as on $\mathbb{R}[u, v]$)

$$(gf)(p) = f(g^T p), \qquad g \in \mathrm{SL}_2(\mathbb{R}),\ f \in C^\infty(\mathbb{R}^2),\ p \in \mathbb{R}^2.$$

This action induces the action of $\mathfrak{sl}_2(\mathbb{R})$ on $C^\infty(\mathbb{R})$ determined by the rule

$$(Af)(p) = \frac{d}{dt} f(\exp(tA)^T p)|_{t=0} = (\partial_u f, \partial_v f) \circ A^T \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)$$
$$= (u,v) A \left(\begin{smallmatrix} \partial_u \\ \partial_v \end{smallmatrix}\right) f, \qquad A \in \mathfrak{sl}_2(\mathbb{R}),\ f \in C^\infty(\mathbb{R}^2),\ p = (u,v) \in \mathbb{R}^2,$$

where $\partial_u = \frac{\partial}{\partial u}$, $\partial_v = \frac{\partial}{\partial v}$ are considered as differential operators on $C^\infty(\mathbb{R}^2)$. The rule

$$A \mapsto (u,v) A \left(\begin{smallmatrix} \partial_u \\ \partial_v \end{smallmatrix}\right)$$

defines a representation of $\mathfrak{sl}_2(\mathbb{R})$ on $C^\infty(\mathbb{R})$ or $\mathbb{R}[u,v]$. In particular, x, y, h act as operators

$$x = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \mapsto (u,v) \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} \partial_u \\ \partial_v \end{smallmatrix}\right) = u\partial_v, \qquad y = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) \mapsto (u,v) \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) \left(\begin{smallmatrix} \partial_u \\ \partial_v \end{smallmatrix}\right) = v\partial_u,$$
$$h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \mapsto (u,v) \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \left(\begin{smallmatrix} \partial_u \\ \partial_v \end{smallmatrix}\right) = u\partial_u - v\partial_v.$$

5. Classification of semisimple Lie algebras

5.1. **Jordan-Chevalley decomposition.** The Jordan normal form of a linear operator $x \in \mathfrak{gl}(V)$ is a sum of Jordan blocks

$$J_{n,\lambda} = \begin{pmatrix} \lambda & 1 & 0 & \dots\dots & 0 \\ 0 & \lambda & 1 & \dots\dots & 0 \\ 0 & 0 & \lambda & \dots\dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

Each Jordan block $J_{n,\lambda}$ can be written as a sum of a diagonal matrix λI_n and a nilpotent matrix which commute. This implies that x also can be written as a sum of a diagonal and a nilpotent matrix which commute.

Definition 5.1. Let V be a vector space over a field K. An endomorphism $x \in \mathfrak{gl}(V)$ is called *semisimple* if it is diagonalizable.

Remark 5.2. Given $x \in \mathfrak{gl}(V)$, we can consider a one-dimensional Lie algebra $L = Kx \subseteq \mathfrak{gl}(V)$ and consider V as a representation of L. Then x is semisimple if and only if V is a direct sum of 1-dimensional (therefore simple) modules over L. This is the reason for the name semisimple.

Lemma 5.3. *Let V be a vector space and $x \in \operatorname{End} V$ be an endomorphism.*

(1) x is semisimple if and only if the minimal polynomial of x (i.e., a monic polynomial p of minimal degree such that $p(x) = 0$) does not have multiple roots.
(2) If $W \subseteq V$ is an x-invariant subspace, then the resctriction $x|_W$ is a semisimple endomorphism of W.

Proof. 1. The minimal polynomial of the Jordan block $J_{n,\lambda}$ is $(t-\lambda)^n$. If x has distinct eigenvalues $\lambda_1, \dots, \lambda_k$ and its Jordan normal form has corresponding maximal Jordan blocks of sizes $n_1, \dots, n_k$, then the minimal polynomial of x is $\prod_{i=1}^k (t-\lambda_i)^{n_i}$. In particular, if x is semisimple, then the minimal polynomial of x is $\prod_{i=1}^k (t-\lambda_i)$ and has no multiple roots. Conversely, if x is not semisimple, then $n_i \geq 2$ for some $1 \leq i \leq k$ and the minimal polynomial of x has root λ_i of multiplicity ≥ 2.
2. If $p(t)$ is the minimal polynomial of x and $q(t)$ is the minimal polynomial of $x|_W$, then $p(x|_W) = 0$ and therefore $q \mid p$. As x is semisimple, p does not have multiple roots. This implies that q also does not have multiple roots and $x|_W$ is semisimple. For an alternative approach see Remark 4.29, where semisimple modules are considered. □

Theorem 5.4 (Jordan-Chevalley decomposition)**.** *Let V be a vector space and $x \in \operatorname{End} V$. Then*

(1) There exists a unique decomposition $x = s + n$, where s is semisimple, n is nilpotent and $sn = ns$.
(2) There exist polynomials $p, q \in K[t]$ without constant term, such that $s = p(x)$ and $n = q(x)$.

Proof. **(1)** Let us prove uniqueness of the decomposition first. Let $(\lambda_1, \dots, \lambda_k)$ be distinct eigenvalues of x with multiplicities $(m_1, \dots, m_k)$. The characteristic polynomial of x is $\det(tI - x) = \prod (t-\lambda_i)^{m_i} \in K[t]$. The vector spaces

$$V_i = \operatorname{Ker}(x - \lambda_i)^{m_i}$$

are x-stable and $V = \bigoplus_{i=1}^k V_i$. As s commutes with n, it also commutes with $x = s + n$ and with $(x - \lambda_i)^{m_i}$. Therefore $sV_i \subseteq V_i$. As s, x commute and $x - s$ is nilpotent, the eigenvalues of x and s on V_i coincide (we can decompose V_i into the eigenspaces of s, put x to Jordan normal form on each summand and then compare the eigenvalues). But x has the only eigenvalue λ_i on V_i. As s is semisimple, it acts on V_i by multiplication with λ_i. This proves uniqueness of s, hence of n.

Taking the above description of s as a definition and putting $n = x - s$, we obtain the required decomposition.

(2) Applying Chinese remainder theorem, we can find a polynomial $p \in K[t]$ such that

$$p(t) \equiv \lambda_i \pmod{(t - \lambda_i)^{m_i}}, \qquad p(t) \equiv 0 \pmod{t} \qquad 1 \leq i \leq k.$$

The operator $(x - \lambda_i)^{m_i}$ is zero on V_i. Therefore $p(x)$ acts on V_i by multiplication with λ_i and $p(x) = s$. Setting $q(t) = t - p(t)$, we obtain $q(x) = x - p(x) = n$. □

Lemma 5.5. *Let $x \in \mathfrak{gl}(V)$ have the Jordan decomposition $x = s + n$. If $W \subseteq V$ is x-invariant (that is, $xW \subseteq W$), then W is also s- and n-invariant.*

Proof. The statement follows from the presentation $s = p(x)$, $n = q(x)$, where p, q are polynomials without constant term. □

Lemma 5.6. *Let $x \in L = \mathfrak{gl}(V)$ have the Jordan decomposition $x = s + n$. Then $\mathrm{ad}_x \in \mathfrak{gl}(L)$ has the Jordan decomposition $\mathrm{ad}_x = \mathrm{ad}_s + \mathrm{ad}_n$.*

Proof. As $[s, n] = 0$, we have

$$[\mathrm{ad}_s, \mathrm{ad}_n] = \mathrm{ad}_{[s,n]} = 0,$$

hence ad_s and ad_n commute. We have to show that ad_s is semisimple and ad_n is nilpotent. We have seen before (see the proof of Engel's theorem 3.21), that if n is nilpotent, then ad_n is nilpotent. Let us choose a basis $(v_1, \ldots, v_n)$ of V such that s is diagonal with respect to it and has eigenvalues $\lambda_1, \ldots, \lambda_n$. Define elementary matrices E_{ij} by the rule $E_{ij} v_k = \delta_{jk} v_i$. Then $s = \sum_i \lambda_i E_{ii}$ and

$$\mathrm{ad}_s(E_{ij}) = \sum_k \lambda_k [E_{kk}, E_{ij}] = (\lambda_i - \lambda_j) E_{ij}.$$

This implies that $\mathrm{ad}_s \in \mathfrak{gl}(L)$ is diagonal in the basis consisting of elementary matrices. □

5.2. **Jordan decomposition in semisimple Lie algebras.**

Remark 5.7. When we classified representations of $\mathfrak{sl}_2$, we noticed an interesting phenomenon. The action of h on every simple $\mathfrak{sl}_2$-representation (hence all representations) was diagonal. It turns out that this is a general rule for semisimple Lie algebras. Namely, if $L \subseteq \mathfrak{gl}(V)$ is a semisimple algebra and $h \in L$ is a semisimple element (our example is $h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in \mathfrak{sl}_2$) then, for any L-representation, the action of h on it is semisimple.

Lemma 5.8. *Let $L \subseteq \mathfrak{gl}(V)$ be a semisimple subalgebra. If $x \in L$ and $x = s + n$ is the Jordan decomposition in $\mathfrak{gl}(V)$, then $s, n \in L$.*

Proof. Let

$$N = \{y \in \mathfrak{gl}(V) \mid [y, L] \subseteq L\}$$

be the normalizer of L. Then N is a Lie algebra and $L \subseteq N$ is an ideal. For any L-submodule $W \subseteq V$, define the Lie algebra

$$L_W = \{y \in \mathfrak{gl}(V) \mid yW \subseteq W,\ \mathrm{tr}(y|_W) = 0\}.$$

As $L = [L, L]$ (for semisimple L), we obtain $\mathrm{tr}(y|_W) = 0\ \forall y \in L$. Therefore $L \subseteq L_W$. Let M be the intersection of N and all L_W (for all L-submodules $W \subseteq V$). Then M is a Lie algebra and $L \subseteq M$ is an ideal.

We claim that $s, n \in M$. There is the Jordan decomposition $\operatorname{ad} x = \operatorname{ad} s + \operatorname{ad} n$ in $\mathfrak{gl}(\mathfrak{gl}(V))$. As $\operatorname{ad} x(L) = [x, L] \subseteq L$, we obtain that $\operatorname{ad} s(L) \subseteq L$ and $\operatorname{ad} n(L) \subseteq L$ and therefore $s, n \in N$. For every L-submodule $W \subseteq V$, we have $xW \subseteq W$ and $\mathrm{tr}(x|_W) = 0$ as $x \in L = [L, L]$. Therefore the same is true for s and n, hence $s, n \in L_W$. This implies $s, n \in M$.

We claim that $M = L$ (hence $s, n \in L$). We can write $M = L \oplus L'$ for some L-submodule $L' \subseteq M$. Then $[L, L'] \subseteq L \cap L' = 0$, therefore every $y \in L'$ acts on a simple L-submodule $W \subseteq V$ as an L-endomorphism, hence as a scalar (by Schur's lemma). As $\mathrm{tr}(y|_W) = 0$, we obtain $y|_W = 0$. But V is a direct sum of simple L-submodules, hence $y = 0$. This implies $L' = 0$ and $M = L$. □

Definition 5.9. Let L be a semisimple Lie algebra. Then

(1) $x \in L$ is called *semisimple* if $\operatorname{ad}_x \in \mathfrak{gl}(L)$ is semisimple.
(2) $x \in L$ is called *nilpotent* if $\operatorname{ad}_x \in \mathfrak{gl}(L)$ is nilpotent.

Theorem 5.10. *Let L be a semisimple Lie algebra.*

(1) For any $x \in L$, there exists a unique decomposition $x = s + n$ in L (called the abstract Jordan decomposition) such that $[s, n] = 0$, $\operatorname{ad} s$ is semisimple and $\operatorname{ad} n$ is nilpotent.
(2) If $f\colon L \to M$ is a homomorphism of semisimple Lie algebras and $x = s + n$ is an abstract Jordan decomposition in L, then $f(x) = f(s) + f(n)$ is an abstract Jordan decomposition in M.
(3) If $\rho\colon L \to \mathfrak{gl}(V)$ is a representation and $x = s + n$ is an abstract Jordan decomposition in L, then $\rho(x) = \rho(s) + \rho(n)$ is a usual Jordan decomposition in $\mathfrak{gl}(V)$.

Proof. **(1)** The adjoint map $\operatorname{ad}\colon L \to \mathfrak{gl}(L)$ is injective as its kernel is equal to $Z(L) = 0$ (as L is semisimple). Applying the previous lemma to $L \simeq \operatorname{ad}(L) \subseteq \mathfrak{gl}(L)$, we obtain a decomposition $x = s + n$ in L such that $\operatorname{ad} x = \operatorname{ad} s + \operatorname{ad} n$ is the Jordan decomposition.
(2) Consider M as an L-module via f and let $V = L \oplus M$ be the direct sum of L-modules. The representation map $\rho\colon L \to \mathfrak{gl}(V)$ is injective as already $\operatorname{ad}\colon L \to \mathfrak{gl}(L)$ is injective. Applying the previous lemma to $L \simeq \rho(L) \subseteq \mathfrak{gl}(V)$, we obtain a decomposition $x = s + n$ in L such that $[s, n] = 0$ and

(1) $\rho(s) \in \mathfrak{gl}(V)$ is semisimple, hence $\operatorname{ad} s$, $\operatorname{ad} f(s)$ are semisimple.
(2) $\rho(n) \in \mathfrak{gl}(V)$ is nilpotent, hence $\operatorname{ad} n$, $\operatorname{ad} f(n)$ are nilpotent.

This implies that $\operatorname{ad} x = \operatorname{ad} s + \operatorname{ad} n$ is the Jordan decomposition in $\mathfrak{gl}(L)$ (hence $x = s + n$ is the abstract Jordan decomposition in L) and $\operatorname{ad} f(x) = \operatorname{ad} f(s) + \operatorname{ad} f(n)$ is the Jordan decomposition in $\mathfrak{gl}(M)$ (hence $f(x) = f(s) + f(n)$ is the abstract Jordan decomposition in M).
(3) As L is semisimple, we have $L = [L, L]$. This implies that $\rho(L) \subseteq \mathfrak{sl}(V)$. We proved that the Lie algebra homomorphism $\rho\colon L \to \mathfrak{sl}(V)$ preserves abstract Jordan decompositions. Therefore we can assume that $L = \mathfrak{sl}(V) \subseteq \mathfrak{gl}(V) = M$. Let $x \in L$ and $x = s + n$ be the usual Jordan decomposition in $M = \mathfrak{gl}(V)$. We will show that it is also the abstract Jordan decomposition of x in L. By the previous lemma $s, n \in L$. By Lemma 5.6, we have a Jordan decomposition $\operatorname{ad} x = \operatorname{ad} s + \operatorname{ad} n$ in $\mathfrak{gl}(M)$. It is also a Jordan decomposition in $\mathfrak{gl}(L)$, hence $x = s + n$ is an abstract Jordan decomposition in L as required. □

Remark 5.11. Let L be a semisimple Lie algebra and $\rho\colon L \to \mathfrak{gl}(V)$ be a representation. The previous theorem implies that if $x \in L$ is semisimple (nilpotent), then $\rho(x)$ is semisimple (nilpotent).

5.2.1. *Weight spaces.* Let $x \in \mathfrak{gl}(V)$ be a semisimple endomorphism. For any $\lambda \in K$, define the eigenspace

$$V_\lambda = \{v \in V \mid xv = \lambda v\}.$$

Then $V = \bigoplus_\lambda V_\lambda$. We will say that x acts on V_λ with weight λ. On the other hand, if $x, y \in \mathfrak{gl}(V)$ are semisimple and commute, then y preserves each summand V_λ:

$$x(yv) = yxv = \lambda(yv) \qquad \forall v \in V_\lambda.$$

This implies that we can consider y as a semisimple endomorphism of V_λ and decompose V_λ according to the eigenvalues of y. In this way V decomposes into a direct sum $V = \bigoplus_{\lambda,\mu} V_{\lambda,\mu}$ of simultaneous eigenspaces

$$V_{\lambda,\mu} = \{v \in V \mid xv = \lambda v,\ yv = \mu v\}.$$

Generally, we obtain the following result.

Lemma 5.12. *Let $H \subseteq \mathfrak{gl}(V)$ be an abelian subalgebra consisting of semisimple elements. Then V has a weight space decomposition (or simultaneous eigenspace decomposition)*

$$V = \bigoplus_{\lambda \in H^*} V_\lambda,$$

where the weight spaces V_λ are defined by

$$V_\lambda = \{v \in V \mid hv = \lambda(h)v\ \forall h \in H\}, \qquad \lambda \in H^*.$$

Proof. Let $H = Kx \oplus J$, where $0 \neq x \in H$. By induction we can decompose V into a direct sum of simultaneous eigenspaces of J. Then x acts on each of these eigespaces as a semisimple endomorphism and we can decompose them further into a direct sum of x-eigenspaces. These smaller summands will be simultaneous eigespaces of H. □

Remark 5.13. Let L be a semisimple Lie algebra, and $H \subseteq L$ be an abelian subalgebra consisting of semisimple elements. For any representation $\rho\colon L \to \mathfrak{gl}(V)$, the image $\rho(H)$ consists of commuting semisimple endomorphisms by Theorem 5.10. Therefore we also obtain a *weight space decomposition*

$$V = \bigoplus_{\lambda \in H^*} V_\lambda, \qquad V_\lambda = \{v \in V \mid \rho(h)v = \lambda(h)v\ \forall h \in H\}.$$

The spaces V_λ are called *weight spaces* of the representation V and the elements $\lambda \in H^*$ with $V_\lambda \neq 0$ are called *weights* of V.

5.3. **Root space decomposition.**

Definition 5.14. A subalgebra $H \subseteq L$ of a semisimple Lie algebra L is called a *Cartan subalgebra* if

(1) H is abelian and consists of semisimple elements.
(2) H is maximal among all subalgebras satisfying (1).

Remark 5.15. If the centralizer $C_L(H) = \{x \in L \,|\, [x, H] = 0\}$ is equal to H, then the second condition above is automatically satisfied. We will see later that conversely, if H is a Cartan subalgebra, then $C_L(H) = H$.

Definition 5.16. Let L be a semisimple Lie algebra and $H \subseteq L$ be a Cartan subalgebra. We have a decomposition corresponding to $\operatorname{ad}\colon H \to \mathfrak{gl}(L)$

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

The finite set

$$R = \{\alpha \in H^* \backslash \{0\} \,|\, L_\alpha \neq 0\} \subseteq H^*$$

is called the *root system* of L relative to H, the elements of R are called *roots* and the spaces L_α are called *root spaces*. Decomposition

$$L = L_0 \oplus \bigoplus_{\alpha \in R} L_\alpha$$

is called the *root space decomposition* (or Cartan decomposition).

Example 5.17. Let $L = \mathfrak{sl}_n$ and let $H \subseteq L$ consist of all diagonal matrices. Then H is abelian and all of its elements are semisimple. It is not difficult to see that $C_L(H) = H$, hence H is a Cartan subalgebra. Let $\varepsilon_i \in H^*$, for $1 \leq i \leq n$, be defined by

$$\varepsilon_i(h) = a_i \qquad h = \sum_{i=1}^{n} a_i E_{ii} \in H.$$

These vectors in H^* are not linearly independent as $\dim H = n - 1$. Given an element $h = \sum_i a_i E_{ii} \in H$, we have

$$\operatorname{ad}_h(E_{ij}) = [h, E_{ij}] = (a_i - a_j)E_{ij} = (\varepsilon_i - \varepsilon_j)(h) \cdot E_{ij}.$$

This means that there are root spaces $L_{\alpha_{ij}} = KE_{ij}$ for the roots

$$\alpha_{ij} = \varepsilon_i - \varepsilon_j \in H^* \qquad \forall i \neq j.$$

The case $i = j$ corresponds to the zero weight $0 \in H^*$. Therefore L_0 consists of all diagonal matrices in L, hence $L_0 = H$. We obtain the following root space decomposition

$$L = H \oplus \bigoplus_{i \neq j} KE_{ij}.$$

Note that $L_0 = H$ and all root spaces $L_{\alpha_{ij}} = KE_{ij}$ are one-dimensional. We will see later that this is true in general.

Lemma 5.18. *We have*

(1) $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$.
(2) If κ *is the Killing form on* L*, then* $\kappa(L_\alpha, L_\beta) = 0$ *whenever* $\alpha + \beta \neq 0$.
(3) The Killing form κ *induces a non-degenerate pairing* $L_\alpha \times L_{-\alpha} \to K$*. In particular, the restriction of* κ *to* L_0 *is non-degenerate.*

Proof. **(1)** If $x \in L_\alpha$, $y \in L_\beta$ and $h \in H$, then

$$[h,[x,y]] = [[h,x],y] + [x,[h,y]] = \alpha(h)[x,y] + \beta(h)[x,y] = (\alpha+\beta)(h)\cdot[x,y].$$

This implies that $[x,y] \in L_{\alpha+\beta}$.
(2) If, additionally, $z \in L_\gamma$, then

$$\operatorname{ad}_x \operatorname{ad}_y(z) = [x,[y,z]] \in L_{\alpha+\beta+\gamma}.$$

This implies that $\operatorname{ad}_x \operatorname{ad}_y$ maps L_γ to $L_{\alpha+\beta+\gamma}$. Therefore, if $\alpha+\beta \neq 0$, then $\operatorname{ad}_x \operatorname{ad}_y$ is nilpotent. In particular, it has trace zero and

$$\kappa(x,y) = \operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_y) = 0.$$

(3) If $x \in L_\alpha$ and $\kappa(x, L_{-\alpha}) = 0$, then $\kappa(x,L) = 0$ by (2). As κ is non-degenerate on L, we obtain $x = 0$. Similarly, if $x \in L_{-\alpha}$ and $\kappa(L_\alpha, x) = 0$, then $x = 0$. This implies that the restriction of κ to $L_\alpha \times L_{-\alpha}$ is non-degenerate. □

Remark 5.19. For any $\alpha \in R$, the root space L_α consists of nilpotent elements. Indeed, for every $x \in L_\alpha$, we proved that ad_x maps L_β to $L_{\alpha+\beta}$. As L is finite dimensional, we can find some $n \geq 1$ such that $L_{n\alpha} = 0$ and $L_{n\alpha+\beta} = 0$ for all $\beta \in R$, hence $\operatorname{ad}_x^n = 0$.

Lemma 5.20. *Let $C_L(H) = \{x \in L \mid [x,H] = 0\}$ be the centralizer of H in L. Then*

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

Proof. The first equation follows from the definitions, namely,

$$L_0 = \{x \in L \mid \operatorname{ad}_h(x) = 0 \ \forall h \in H\} = \{x \in L \mid [H,x] = 0\}.$$

For the second equation, we note first that $H \subseteq L_0$ as H is abelian. Let $x \in L_0$ and $x = s + n$ be its Jordan decomposition in L. Then $\operatorname{ad} x = \operatorname{ad} s + \operatorname{ad} n$ is the Jordan decomposition in $\mathfrak{gl}(L)$. As

$$\operatorname{ad} x(H) = [x,H] = 0,$$

we conclude that also $[s,H] = \operatorname{ad} s(H) = 0$. This implies that $H + Ks$ is abelian and consists of semisimple elements. By the maximality of H, we obtain $s \in H$. Note also that $s, n \in L_0$. To show that $L_0 = H$, we need to show that the only nilpotent element of L_0 is zero as this would imply $n = 0$ and therefore $x = s \in H$.

We claim that L_0 is nilpotent. Given $x \in L_0$ with a Jordan decomposition $x = s + n$, we proved that $s \in H$, hence $\operatorname{ad} s|_{L_0} = 0$. This implies that $\operatorname{ad} x|_{L_0} = \operatorname{ad} n|_{L_0}$ is nilpotent on L_0. By Engel theorem L_0 is nilpotent. Therefore L_0 is solvable and by Lie theorem we can find a basis of L such that $\operatorname{ad} x$ is upper-triangular for all $x \in L_0$. If $x \in L_0$ is nilpotent, then $\operatorname{ad} x$ is nilpotent and is strictly upper-triangular. Therefore for any $y \in L_0$, $\operatorname{ad}_x \operatorname{ad}_y$ is strictly upper-triangular and has zero trace. Therefore

$$\kappa(x,y) = \operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_y) = 0 \qquad \forall y \in L_0.$$

By the previous lemma, κ is non-degenerate on L_0, hence $x = 0$. This implies that the only nilpotent element of L_0 is trivial, as required. □

Denote the standard pairing between H and H^* by

$$\langle h, \alpha\rangle = \langle \alpha, h\rangle \coloneqq \alpha(h), \qquad h \in H,\ \alpha \in H^*.$$

From now on we will denote the Killing form on L by $(-,-)$. Its restriction to $H = L_0$ is non-degenerate, hence induces an isomorphism

$$H \simeq H^*, \qquad h \mapsto (h,-).$$

We will denote the induced symmetric bilinear form on H^* also by $(-,-)$. For any $\alpha \in H^*$, let h'_α denote the corresponding element in H (under the above isomorphism), that is $\alpha = (h'_\alpha, -)$ and

$$\langle \alpha, h \rangle = \alpha(h) = (h'_\alpha, h) \qquad h \in H.$$

In particular,

$$(\alpha, \beta) = (h'_\alpha, h'_\beta) = \langle \alpha, h'_\beta \rangle \qquad \forall \alpha, \beta \in H^*.$$

Lemma 5.21. *Let $\alpha \in R$. Then*

(1) If $x \in L_\alpha$ and $y \in L_{-\alpha}$, then

$$[x, y] = (x, y) h'_\alpha.$$

(2) $(\alpha, \alpha) \neq 0$.

(3) If $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ are such that $(x_\alpha, y_\alpha) = \frac{2}{(\alpha,\alpha)}$ (they are unique up to a scalar) and $h_\alpha = \frac{2h'_\alpha}{(\alpha,\alpha)}$, then $\langle \alpha, h_\alpha \rangle = 2$ and the elements $x_\alpha, y_\alpha, h_\alpha$ satisfy the relations of $\mathfrak{sl}_2$. We will denote this algebra by $(\mathfrak{sl}_2)_\alpha$.

Proof. **(1)** For any $h \in H$, we have

$$([x, y], h) = (x, [y, h]) = -(x, [h, y]) = \langle \alpha, h \rangle (x, y) = (h'_\alpha, h)(x, y) = ((x, y) h'_\alpha, h).$$

Using the fact that $[x, y] \in L_0 = H$, $(x, y) h'_\alpha \in H$ and the Killing form is non-degenerate on H, we conclude that $[x, y] = (x, y) h'_\alpha$.

(2) Assume that $(\alpha, \alpha) = \langle \alpha, h'_\alpha \rangle = 0$. Let $0 \neq x \in L_\alpha$. As the Killing form is non-degenerate on $L_\alpha \times L_{-\alpha}$, there exists $y \in L_{-\alpha}$ such that $(x, y) \neq 0$. Then

$$h = [x, y] = (x, y) h'_\alpha \neq 0.$$

Consider the vector space M generated by x, y, h. We have

$$[h, x] = \langle \alpha, h \rangle x = \langle \alpha, (x, y) h'_\alpha \rangle x = 0$$

as $\langle \alpha, h'_\alpha \rangle = 0$. Similarly $[h, y] = 0$. This implies that M is a solvable Lie algebra. By Lie theorem, there exists a basis of L such that M acts on L by upper-triangular matrices. This implies that $\mathrm{ad}_h = [\mathrm{ad}_x, \mathrm{ad}_y]$ acts by a strictly upper-triangular matrix, hence is nilpotent. But $h \in H$ is also semisimple, hence $h = 0$. A contradiction.

(3) We have

$$\langle \alpha, h_\alpha \rangle = \frac{2}{(\alpha, \alpha)} \langle \alpha, h'_\alpha \rangle = \frac{2}{(\alpha, \alpha)} (\alpha, \alpha) = 2.$$

Moreover,

$$[x_\alpha, y_\alpha] = (x_\alpha, y_\alpha) h'_\alpha = \frac{2h'_\alpha}{(\alpha, \alpha)} = h_\alpha,$$

$$[h_\alpha, x_\alpha] = \langle \alpha, h_\alpha \rangle x_\alpha = 2x_\alpha.$$

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

Theorem 5.22. *We have*

(1) R spans H^*.

(2) The vector space

$$V = K h_\alpha \oplus_{k \in \mathbb{Z} \setminus \{0\}} L_{k\alpha}$$

is an irreducible representation of $\mathfrak{sl}_2 = \langle x_\alpha, y_\alpha, h_\alpha \rangle$.

(3) For any root $\alpha \in R$, $\dim L_\alpha = 1$.

(4) If $\alpha, \beta \in R$, then $\langle \beta, h_\alpha \rangle = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}$.

(5) For any root $\alpha \in R$, the only multiples of α which are also roots are $\pm\alpha$.

(6) *For any $\alpha \in R$, define a reflection*

$$s_\alpha \colon H^* \to H^*, \qquad s_\alpha(\lambda) = \lambda - \langle \lambda, h_\alpha \rangle \alpha = \lambda - \frac{2(\alpha,\lambda)}{(\alpha,\alpha)}\alpha,$$

which leaves the hyperplane $\alpha^\perp = \{\lambda \in H^* \mid (\alpha,\lambda) = 0\}$ invariant and maps α to $-\alpha$. Then $s_\alpha(R) \subseteq R$.

(7) *For any roots α and $\beta \neq \pm\alpha$, the vector space*

$$V = \bigoplus_{k\in\mathbb{Z}} L_{\beta+k\alpha}$$

is an irreducible representation of $\mathfrak{sl}_2 = \langle x_\alpha, y_\alpha, h_\alpha \rangle$.

(8) *If $\alpha, \beta, \alpha+\beta \in R$ then $[L_\alpha, L_\beta] = L_{\alpha+\beta}$.*

Proof. (1). Assume that R does not span H^*. Then there exists $0 \neq h \in H$ such that $\langle R, h \rangle = 0$. Then, for any $\alpha \in R$ and $z \in L_\alpha$,

$$[h, z] = \langle \alpha, h \rangle z = 0,$$

hence $[h, L_\alpha] = 0$. This implies that $[h, L] = 0$ and $h \in Z(L)$. As L is semisimple, its center is zero, hence $h = 0$. A contradiction.

(2). The vector space V is a representation of $\mathfrak{sl}_2$. Indeed, h_α preserves each $L_{k\alpha}$ and $[h_\alpha, h_\alpha] = 0$. The element $x_\alpha \in L_\alpha$ maps $L_{k\alpha}$ to $L_{(k+1)\alpha}$ for $k \neq 0, -1$ and maps h_α to L_α. On the other hand, for any $z \in L_{-\alpha}$, we have $[x_\alpha, z] = (x_\alpha, z)h'_\alpha \in Kh_\alpha$. Therefore $[x_\alpha, L_{-\alpha}] \subseteq Kh_\alpha$. This implies $[x_\alpha, V] \subseteq V$. Similarly $[y_\alpha, V] \subseteq V$. For any $z \in L_{k\alpha}$, we have

$$[h_\alpha, z] = \langle k\alpha, h_\alpha \rangle z = 2kz,$$

hence z has weight $2k$. On the other hand, the elements of Kh_α have weight zero. Representation V of $\mathfrak{sl}_2$ is irreducible as $V_0 = Kh_\alpha$ has dimension one (here we use classification of irreducible representations of $\mathfrak{sl}_2$).

(3). From the previous statement we obtain that the weight space $V_2 = L_\alpha$ has dimension one.

(4). Let us consider L as a representation of $\mathfrak{sl}_2 = \langle x_\alpha, y_\alpha, h_\alpha \rangle$. We know that all weights of this representation are integers. In particular, for $z \in L_\beta$,

$$[h_\alpha, z] = \langle \beta, h_\alpha \rangle z,$$

hence z has weight $\langle \beta, h_\alpha \rangle$ which is an integer.

(5). Assume that $\alpha, \beta \in R$ and $\beta = c\alpha$ for some $c \in K \backslash \{\pm 1\}$. Then by ((4))

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = 2c \in \mathbb{Z}, \qquad \frac{2(\alpha,\beta)}{(\beta,\beta)} = \frac{2}{c} \in \mathbb{Z}.$$

Therefore $c = \pm\frac{1}{2}, \pm 2$. If $c = \pm\frac{1}{2}$, we can exchange α and β and assume $c = \pm 2$. If $c = -2$, we can substitute β by $-\beta$ and assume $\beta = 2\alpha$. Consider the irreducible representation

$$V = Kh_\alpha \oplus_{k\in\mathbb{Z}\backslash\{0\}} L_{k\alpha}$$

from ((2)). By the property of irreducible $\mathfrak{sl}_2$ representations, the map $x_\alpha \colon V_2 \to V_4$ is an isomorphism if $V_4 \neq 0$. But $V_2 = L_\alpha$ is one-dimensional, generated by x_α, and $[x_\alpha, x_\alpha] = 0$. Therefore $x_\alpha \colon V_2 \to V_4$ is zero and $V_4 = 0$. We obtain from $L_\beta = L_{2\alpha} = V_4 = 0$, that β is not a root.

(6). Let $\beta \in R$ and $n = \langle \beta, h_\alpha \rangle$. We have to show that $s_\alpha(\beta) = \beta - n\alpha$ is a root. As before, we consider $V = L$ as a representation of $\mathfrak{sl}_2$. We have seen that elements of L_β have weight

$\langle \beta, h_\alpha \rangle = n$. Assume that $n \geq 0$. Then $y_\alpha^n \colon V_n \to V_{-n}$ is an isomorphism. In particular, for any $0 \neq z \in L_\beta \subseteq V_n$, the element $y_\alpha^n(z)$ is non-zero. But

$$y_\alpha^n(z) = \mathrm{ad}_{y_\alpha}^n(z) \in L_{\beta - n\alpha},$$

hence $\beta - n\alpha \in R$ as required. The proof for $n < 0$ is similar (consider $m = -n$ and an isomorphism $x_\alpha^m \colon V_{-m} \to V_m$).

(7). The vector space V is a representation of $\mathfrak{sl}_2$. Indeed, h_α preserves each $L_{\beta+k\alpha}$, $x_\alpha \in L_\alpha$ maps $L_{\beta+k\alpha}$ to $L_{\beta+(k+1)\alpha}$ and $y_\alpha \in L_{-\alpha}$ maps $L_{\beta+k\alpha}$ to $L_{\beta+(k-1)\alpha}$. For any $z \in L_{\beta+k\alpha}$, we have

$$[h_\alpha, z] = \langle \beta + k\alpha, h_\alpha \rangle z = (\langle \beta, h_\alpha \rangle + 2k)z,$$

hence elements of $L_{\beta+k\alpha}$ have weight $\langle \beta, h_\alpha \rangle + 2k$. All of these weight spaces are one-dimensional, hence V is an irreducible representation.

(8). We know that $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$ and the latter space is one-dimensional. We obtain from ((7)) that the map $x_\alpha \colon L_\beta \to L_{\alpha+\beta}$ between nonzero weight spaces is an isomorphism. Therefore $[x_\alpha, L_\beta] \neq 0$ and $[L_\alpha, L_\beta] = L_{\alpha+\beta}$. □

Lemma 5.23. *We have*

1. $(h, h') = \sum_{\alpha \in R} \langle \alpha, h \rangle \langle \alpha, h' \rangle$, *for any* $h, h' \in H$.
2. $(\beta, \gamma) = \sum_{\alpha \in R} (\alpha, \beta)(\alpha, \gamma)$, *for any* $\beta, \gamma \in H^*$.
3. $(\beta, \gamma) \in \mathbb{Q}$ *for any* $\beta, \gamma \in R$.
4. *Let* $E \subseteq H^*$ *be the vector space over* $\mathbb{Q}$ *generated by* R. *Then the form* $(-,-)$ *is positive definite on* E, *that is,* $(\gamma, \gamma) > 0$ *for any* $0 \neq \gamma \in E$.

Proof. **(1)** Consider the action of $\mathrm{ad}_h \, \mathrm{ad}\, h'$ on every root space L_α and take the trace.

(2) We have

$$(\beta, \gamma) = (h'_\beta, h'_\gamma) = \sum_{\alpha \in R} \langle \alpha, h'_\beta \rangle \langle \alpha, h'_\gamma \rangle = \sum_{\alpha \in R} (\alpha, \beta)(\alpha, \gamma).$$

(3) We have $(\beta, \beta) = \sum_{\alpha \in R} (\alpha, \beta)^2$, hence $\frac{1}{(\beta,\beta)} = \sum_{\alpha \in R} \frac{(\alpha,\beta)^2}{(\beta,\beta)^2} \in \mathbb{Q}$ as $\frac{2(\alpha,\beta)}{(\beta,\beta)} \in \mathbb{Z}$. We obtain from $\frac{2(\beta,\gamma)}{(\beta,\beta)} \in \mathbb{Z}$ that $(\beta, \gamma) \in \mathbb{Q}$.

(4) If $0 \neq \gamma \in E \subseteq H^*$ then

$$(\gamma, \gamma) = \sum_{\alpha \in R} (\alpha, \gamma)^2.$$

We can write $\gamma = \sum_{\beta \in R} c_\beta \beta$ for some $c_\beta \in \mathbb{Q}$. Then $(\alpha, \gamma) = \sum_{\beta \in R} c_\beta (\alpha, \beta) \in \mathbb{Q}$ and $(\alpha, \gamma)^2 \geq 0$, for all $\alpha \in R$. There exists $\alpha \in R$ such that $(\alpha, \gamma) \neq 0$ as R generates H^*. This implies that $\sum_{\alpha \in R} (\alpha, \gamma)^2 > 0$. □

5.4. **Abstract root systems.** Assume that $K = \mathbb{C}$ and let $E \subseteq H^*$ be a real vector space generated by R. Let us formalize the properties of $R \subseteq E$.

Definition 5.24. Let E be a real vector space with a positive definite inner product (we call E a Euclidean vector space). An *abstract root system* is a finite subset $R \subseteq E\backslash\{0\}$ such that

(1) R generates E.
(2) For any $\alpha, \beta \in R$, the number $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ is an integer.
(3) For any $\alpha \in R$, define a *reflection*

$$s_\alpha \colon E \to E, \qquad \lambda \mapsto \lambda - \frac{2(\alpha,\lambda)}{(\alpha,\alpha)}\alpha.$$

Then $s_\alpha(R) \subseteq R$.
(4) For any root $\alpha \in R$, the only multiples of α which are also roots are $\pm\alpha$.

Remark 5.25. Note that $s_\alpha(\alpha) = -\alpha$ and $s_\alpha(\lambda) = \lambda$ if λ is orthogonal to α (i.e., $(\alpha,\lambda) = 0$). Therefore s_α should be interpreted as a reflection with respect to the hyperplane

$$H_\alpha = \alpha^\perp = \{\lambda \in E \,|\, (\alpha,\lambda) = 0\}.$$

Note that $s_\alpha^2 = \mathrm{id}$.

For any root $\alpha \in E$, define the corresponding coroot $\alpha^\vee \in E^*$ (it was denoted by h_α previously) by

$$\langle \lambda, \alpha^\vee \rangle = \frac{2(\alpha,\lambda)}{(\alpha,\alpha)}.$$

Then

$$\langle \alpha, \alpha^\vee \rangle = 2, \qquad s_\alpha(\lambda) = \lambda - \langle \lambda, \alpha^\vee \rangle \alpha.$$

Definition 5.26. The *Weyl group* W of the root system $R \subseteq E$ is the subgroup of $\mathrm{GL}(E)$ generated by reflections s_α for $\alpha \in R$.

Lemma 5.27. *The Weyl group is a finite subgroup of $O(E)$ which preserves $R \subseteq E$.*

Proof. Every s_α is orthogonal as

$$\begin{aligned}(s_\alpha\lambda, s_\alpha\mu) &= \left(\lambda - \frac{2(\alpha,\lambda)}{(\alpha,\alpha)}\alpha, \mu - \frac{2(\alpha,\mu)}{(\alpha,\alpha)}\alpha\right) \\ &= (\lambda,\mu) - \frac{4(\alpha,\lambda)(\alpha,\mu)}{(\alpha,\alpha)} + \frac{4(\alpha,\lambda)(\alpha,\mu)}{(\alpha,\alpha)^2}(\alpha,\alpha) = (\lambda,\mu).\end{aligned}$$

This implies that $W \subseteq O(E)$. We know that $s_\alpha(R) = R$. This implies that W preserves R. If some $w \in W$ acts trivially on R, then w acts trivially on E as R generates E. Therefore W is a subgroup of the group $\mathrm{Aut}\, R$ of all permutations of R. As R is a finite set, $\mathrm{Aut}\, R$ and W are also finite. □

Example 5.28. Let $R \subseteq E$ be the root system of $\mathfrak{sl}_n$. We identify E with $\{x \in \mathbb{R}^n \,|\, \sum x_i = 0\}$ and let

$$R = \{\alpha_{ij} = e_i - e_j \,|\, i \neq j\}.$$

For any root $\alpha = \alpha_{ij}$, the corresponding reflection s_α acts as

$$s_\alpha(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n),$$

that is, permutes coordinates i and j. We can identify s_α with a transposition (ij) considered as an element of the symmetric group S_n. Then the Weyl group W, generated by all such transpositions, coincides with the symmetric group S_n.

Lemma 5.29. *For any $w \in W$ and $\alpha \in R$, we have $s_{w\alpha} = ws_\alpha w^{-1}$.*

Proof. If $\lambda \in H_\alpha$, then $(w\alpha, w\lambda) = (\alpha, \lambda) = 0$ hence $w\lambda \in H_{w\alpha}$. This implies $wH_\alpha = H_{w\alpha}$. The element $ws_\alpha w^{-1}$ acts as identity on $wH_\alpha = H_{w\alpha}$ and maps $w\alpha$ to $ws_\alpha(\alpha) = -w\alpha$. The same is true about $s_{w\alpha}$. Therefore they are identical. □

5.5. **Pairs of roots.**

Theorem 5.30. *Let $\alpha, \beta \in R$ be two roots such that $\beta \neq \pm\alpha$ and $\|\beta\| \leq \|\alpha\|$. Let θ be the angle between α and β. Then we have only the following possibilities*

	$\langle\beta, \alpha^\vee\rangle$	$\langle\alpha, \beta^\vee\rangle$	θ
1	0	0	$\pi/2$
2	1	1	$\pi/3$
3	-1	-1	$2\pi/3$
4	1	2	$\pi/4$
5	-1	-2	$3\pi/4$
6	1	3	$\pi/6$
7	-1	-3	$5\pi/6$

Proof. We have $(\alpha, \beta) = \|\alpha\| \, \|\beta\| \cos\theta$. Let $m = \langle\beta, \alpha^\vee\rangle = \frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ and $n = \langle\alpha, \beta^\vee\rangle = \frac{2(\alpha,\beta)}{(\beta,\beta)}$ and . Then $m, n \in \mathbb{Z}$ and

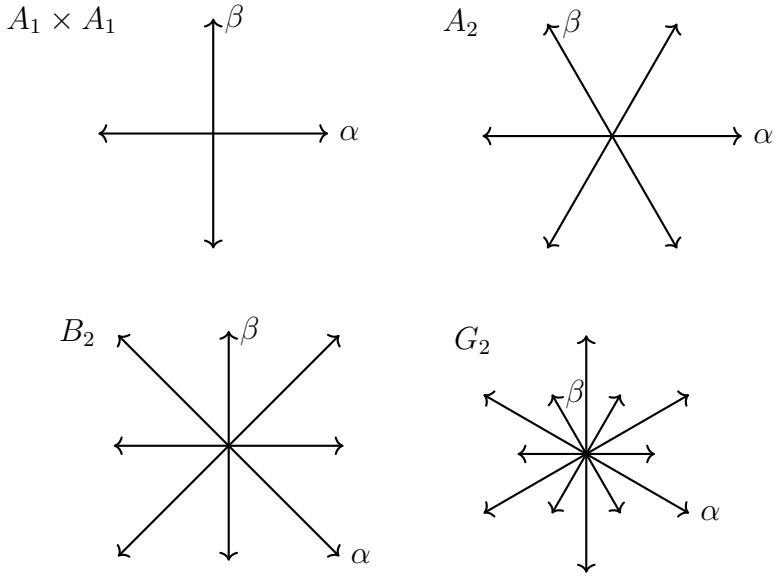
$$mn = 4\frac{(\alpha, \beta)^2}{\|\alpha\|^2 \cdot \|\beta\|^2} = 4\cos^2\theta.$$

As $\beta \neq \pm\alpha$, we have $\cos\theta \neq \pm 1$. Therefore $mn \in \{0, 1, 2, 3\}$. We have

$$\frac{m}{n} = \frac{(\beta, \beta)}{(\alpha, \alpha)} \leq 1,$$

hence $|m| \leq |n|$. This implies that we can only have $m \in \{0, \pm 1\}$. If $m = 0$ then $n = 0$. Otherwise $m = \pm 1$ and $n = \pm 1, \pm 2, \pm 3$. The angle θ is determined from $4\cos^2\theta = mn$. □

The root system that appears in (1) is called $A_1 \times A_1$, the root system in (2) and (3) is called A_2, the root system in (4) and (5) is called B_2 and the root system in (6) and (7) is called G_2.



Corollary 5.31. *Let $\alpha, \beta \in R$. Then*

(1) If $(\alpha, \beta) > 0$, then $\alpha - \beta \in R$.
(2) If $(\alpha, \beta) < 0$, then $\alpha + \beta \in R$.

Proof. It is enough to prove (1). Let $m = \langle\alpha, \beta^\vee\rangle > 0$, $n = \langle\beta, \alpha^\vee\rangle > 0$. Then $m = 1$ or $n = 1$. If $m = 1$, then

$$s_\beta(\alpha) = \alpha - \langle\alpha, \beta^\vee\rangle\beta = \alpha - \beta \in R.$$

If $n = 1$, then

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha^\vee \rangle \alpha = \beta - \alpha \in R,$$

hence $\alpha - \beta \ni R$. $\square$

5.6. **Bases of root systems.**

Definition 5.32. A subset $\Pi \subseteq R$ is called a *base* of R if

(1) Π is a basis of E.
(2) If $\beta = \sum_{\alpha\in\Pi} n_\alpha \alpha$ is a root, then all n_α are simultaneously in $\mathbb{Z}_{\geq 0}$ or simultaneously in $\mathbb{Z}_{\leq 0}$.

The set

$$R_+ = \left\{ \beta = \sum_{\alpha\in\Pi} n_\alpha \alpha \in R \,\middle|\, n_\alpha \geq 0 \ \forall \alpha \in \Pi \right\}$$

is called the *positive root system* (or the set of positive roots) and the set $R_- = -R_+$ is called the *negative root system* (or the set of negative roots). We have a decomposition $R = R_+ \sqcup R_-$, called a *polarization* of the roots system R corresponding to Π.

One can construct bases of R as follows. Let $H_\alpha = \alpha^\perp = \{\lambda \in E \,|\, (\alpha, \lambda) = 0\}$, for any root α, and let

$$\gamma \in E \setminus \bigcup_{\alpha\in R} H_\alpha,$$

called a regular vector. Define

$$R_+(\gamma) = \{\alpha \in R \,|\, (\alpha, \gamma) > 0\}, \qquad R_-(\gamma) = \{\alpha \in R \,|\, (\alpha, \gamma) < 0\}.$$

Then $R_-(\gamma) = -R_+(\gamma)$ and $R = R_+(\gamma) \sqcup R_-(\gamma)$. The sets $R_+(\gamma)$ and $R_-(\gamma)$ are called the sets of positive and negative roots respectively.

Definition 5.33. We say that an element $\alpha \in R_+(\gamma)$ is a *simple root* if it can not be written as a sum of two positive roots. Denote the set of all simple roots by $\Pi(\gamma)$.

Lemma 5.34. *Any positive root can be written as a sum of simple roots.*

Proof. If $\alpha \in R_+(\gamma)$ is not simple, then $\alpha = \alpha' + \alpha''$ for some $\alpha', \alpha'' \in R_+(\gamma)$. This implies that $(\alpha', \gamma) < (\alpha, \gamma)$ and $(\alpha'', \gamma) < (\alpha, \gamma)$. Applying induction on (α, γ) (there is a finite set of values for $\alpha \in R_+(\gamma)$), we obtain representations of α' and α'' as sums of simple roots. Then we can represent α as a sum of simple roots. □

Theorem 5.35. *For any regular $\gamma \in E$, the set of simple roots $\Pi(\gamma)$ is a base of R. Any base of the root system R has this form.*

Proof. **(1)** We have seen that any positive root is a linear combination of $\Pi = \Pi(\gamma)$. On the other hand R generates E. This implies that Π generates E.
(2) We claim that $(\alpha, \beta) \leq 0$ for (distinct) $\alpha, \beta \in \Pi$. If $(\alpha, \beta) > 0$, then we have seen that $\beta' = \alpha - \beta \in R$. If $\beta \in R_+(\gamma)$, then $\alpha = \beta + \beta'$ is not simple. If $\beta' \in R_-(\gamma)$, then $\beta = \alpha + (-\beta')$ is not simple. This is a contradiction.
(3) Let us show that elements of Π are linearly independent. Assume that $\sum_{\alpha\in\Pi} n_\alpha \alpha = 0$ for some $n_\alpha \in \mathbb{R}$. Let $S = \{\alpha \,|\, n_\alpha > 0\}$ and $T = \{\alpha \,|\, n_\alpha < 0\}$. Then $\lambda := \sum_{\alpha\in S} n_\alpha \alpha = \sum_{\beta\in T} t_\beta \beta$, where $t_\beta = -n_\beta$. Therefore

$$(\lambda, \lambda) = \sum_{\alpha\in S, \beta\in T} n_\alpha t_\beta (\alpha, \beta) \leq 0$$

as $n_\alpha > 0$, $t_\beta > 0$ and $(\alpha, \beta) \leq 0$. This implies that $\lambda = 0$. Therefore

$$0 = (\gamma, \lambda) = \sum_{\alpha\in S} n_\alpha (\gamma, \alpha).$$

But $n_\alpha > 0$ and $(\gamma, \alpha) > 0$ as $\alpha \in R_+$. We conclude that $S = \varnothing$ and similarly $T = \varnothing$. This implies that elements of Π are linearly independent.

(4) We have seen that every $\beta \in R_+(\gamma)$ is a sum of simple roots, hence $\beta = \sum_{\alpha\in\Pi} n_\alpha \alpha$ with $n_\alpha \in \mathbb{Z}_{\geq 0}$. Similarly, every $\beta \in R_-(\gamma)$ can be written in the form $\beta = \sum_{\alpha\in\Pi} n_\alpha \alpha$ with $n_\alpha \in \mathbb{Z}_{\leq 0}$. This implies that $\Pi = \Pi(\gamma)$ is a base of R.

(5) Assume that $\Pi \subseteq R$ is a base of R and let $\gamma \in E$ be such that $(\gamma, \alpha) > 0$ for all $\alpha \in \Pi$. We claim that $\Pi = \Pi(\gamma)$. We have $\Pi \subseteq R_+(\gamma)$, hence $R_+ \subseteq R_+(\gamma)$. Similarly $R_- \subseteq R_-(\gamma)$. This implies $R_+ = R_+(\gamma)$. Any $\alpha \in \Pi$ is a simple root. Indeed, if $\alpha = \alpha' + \alpha''$, for $\alpha' = \sum_{\beta\in\Pi} n'_\beta \beta \in R_+$ and $\alpha'' = \sum_{\beta\in\Pi} n''_\beta \beta \in R_+$, then $n'_\beta + n''_\beta = 0$ for $\beta \neq 0$ and $n'_\alpha + n''_\alpha = 1$. This implies $\alpha' = 0$ or $\alpha'' = 0$, a contradiction. Therefore $\Pi \subseteq \Pi(\gamma)$. We conclude that $\Pi = \Pi(\gamma)$ as they are both bases of E and have the same number of elements. □

Example 5.36. Let $R \subseteq E$ be the root system of $\mathfrak{sl}_n$. We identify E with $\{x \in \mathbb{R}^n \mid \sum x_i = 0\}$ (corresponds to the traceless diagonal matrices) and define

$$R = \{\alpha_{ij} = e_i - e_j \mid i \neq j\}.$$

Choose $\gamma = (n, n-1, \ldots, 1) \in \mathbb{R}^n$. Then

$$(\gamma, \alpha_{ij}) = (\gamma, e_i) - (\gamma, e_j) = (n+1-i) - (n+1-j) = j - i > 0$$

if and only if $i < j$. Therefore

$$R_+ = \{e_i - e_j \mid i < j\}.$$

For any $i < j$, we can write

$$e_i - e_j = (e_i - e_{i+1}) + \cdots + (e_{j-1} - e_j).$$

Therefore the simple roots are

$$\alpha_i = e_i - e_{i+1}, \qquad 1 \leq i \leq n-1$$

and $\Pi = \{\alpha_1, \ldots, \alpha_{n-1}\}$ is a base of R. Note that Π is a basis of E.

5.7. **Dynkin diagrams and classification.** Let $r = \dim E$ (called the rank of the root system $R \subseteq E$) and let $\Pi = \{\alpha_1, \dots, \alpha_r\}$ be a base of R. For any simple root $\alpha_i \in \Pi$, let $s_i = s_{\alpha_i} \in W$ be the corresponding reflection, called a *simple reflection*.

Theorem 5.37. *We have*

(1) Simple reflections generate the Weyl group W.
(2) $W\Pi = R$.

Corollary 5.38. *R can be recovered from Π.*

Proof. We can recover W as the group generated by simple reflections. Then we recover R as $R = W\Pi$. □

Definition 5.39. A *Cartan matrix* A of the base $\Pi \subseteq R$ is the $r \times r$ matrix with entries

$$a_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}.$$

Lemma 5.40. *We have*

(1) $a_{ii} = 2$.
(2) $a_{ij} \in \mathbb{Z}_{\leq 0}$ for $i \neq j$.
(3) $a_{ij} = 0$ if and only if $a_{ji} = 0$.
(4) $a_{ij}a_{ji} \in \{0, 1, 2, 3\}$.

Proof. 1. Is obvious.
2. Was proved earlier.
3. $a_{ij} = 0$ if and only if $(\alpha_i, \alpha_j) = (\alpha_j, \alpha_i) = 0$.
4. We proved more generally that $\langle \alpha, \beta^\vee \rangle \langle \beta, \alpha^\vee \rangle \in \{0, 1, 2, 3\}$ for any roots α, β with $\beta \neq \pm\alpha$. □

Example 5.41. Consider the root system R of $\mathfrak{sl}_n$ with simple roots

$$\alpha_i = e_i - e_{i+1}, \qquad 1 \leq i \leq n-1.$$

Then

$$(\alpha_i, \alpha_i) = (e_i - e_{i+1}, e_i - e_{i+1}) = 2,$$
$$(\alpha_i, \alpha_{i+1}) = (e_i - e_{i+1}, e_{i+1} - e_{i+2}) = -1$$

and $(\alpha_i, \alpha_j) = 0$ for $|i - j| > 1$. This implies $a_{ii} = 2$,

$$a_{i,i+1} = a_{i+1,i} = \frac{2(\alpha_i, \alpha_{i+1})}{(\alpha_i, \alpha_i)} = -1$$

and $a_{ij} = 0$ for $|i - j| > 1$.

Remark 5.42. The integers a_{ij} and a_{ji} are both ≤ 0 for $i \neq j$. The property $a_{ij}a_{ji} \in \{0, 1, 2, 3\}$ implies that in order to determine a_{ij} and a_{ji} it is enough to know $a_{ij}a_{ji}$ and if

$$\frac{a_{ij}}{a_{ji}} = \frac{2(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i)}{2(\alpha_i, \alpha_j)/(\alpha_j, \alpha_j)} = \frac{\|\alpha_j\|^2}{\|\alpha_i\|^2} \leq 1$$

or ≥ 1. The first inequality means that the root α_j is shorter than the root α_i.

Definition 5.43. Define the *Dynkin diagram* of (R, Π) as follows. It has vertices $1, \dots, r$ and the number of edges between i and j, for $i \neq j$, is equal to $a_{ij}a_{ji}$. Moreover, if α_i and α_j have different length, we draw an arrow in the direction of the shorter root. This diagram uniquely determines the Cartan matrix.

Remark 5.44. The Weyl group W has generators $s_1, \dots, s_r$ subject to relations

(1) $s_i^2 = 1$,
(2) $(s_i s_j)^{m_{ij}} = 1$, where m_{ij} are defined as follows

$a_{ij}a_{ji}$	0	1	2	3
m_{ij}	2	3	4	6

Groups of this form (with arbitrary m_{ij}) are called *Coxeter groups*. They are not finite in general. Note, in particular, that W is determined by the Cartan matrix.

Example 5.45. If $R \subseteq E$ is the root system of $\mathfrak{sl}_n$ and the simple roots are

$$\alpha_i = e_i - e_{i+1}, \qquad 1 \leq i \leq n-1,$$

then the simple reflections correspond to transpositions $s_i = (i, i+1)$. One can show that they generate $W = S_n$ and subject to the relations

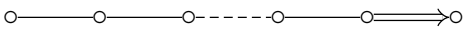
$$s_i^2 = 1, \qquad (s_i s_{i+1})^3 = 1, \qquad s_i s_j = s_j s_i \qquad |i-j| > 1.$$

Definition 5.46. A root system $R \subseteq E$ is called *reducible* if $R = R_1 \sqcup R_2$ for non-empty subset R_1, R_2 such that $(R_1, R_2) = 0$. Otherwise R is called *irreducible*.

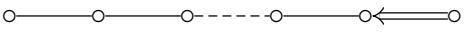
Theorem 5.47. *The Dynkin diagrams of irreducible root systems are*

(1) A_n:

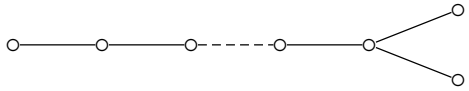
(2) B_n:



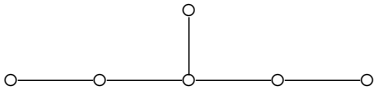
(3) C_n:



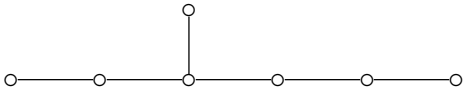
(4) D_n:



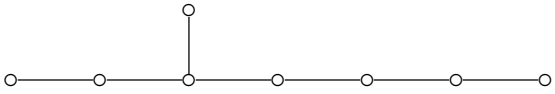
(5) E_6:



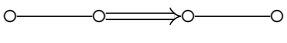
(6) E_7:



(7) E_8:



(8) F_4:



(9) G_2:



Theorem 5.48 (Serre relations)**.** *Let L be a semisimple Lie algebra, $H \subseteq L$ be a Cartan subalgebra, $R \subseteq H^*$ be the corresponding roots system, $\Pi = \{\alpha_1, \ldots, \alpha_r\} \subseteq R$ be a base of R, and $R_+ \subseteq R$ be the corresponding positive root system. For any simple root $\alpha_i \in \Pi$, let $x_i \in L_{\alpha_i}$, $y_i \in L_{-\alpha_i}$ and $h_i = h_{\alpha_i} \in H$ be such that x_i, y_i, h_i form a standard basis of $\mathfrak{sl}_2$. Then the elements x_i, y_i, h_i for $i = 1, \ldots, r$ generate the Lie algebra L and subject to the relations*

(1) $[h_i, h_j] = 0$,
(2) $[h_i, x_j] = a_{ij} x_j$ *and* $[h_i, y_j] = -a_{ij} y_j$,
(3) $[x_i, y_j] = \delta_{ij} h_i$,

(4) $(\operatorname{ad} x_i)^{1-a_{ij}} x_j = 0$ *and* $(\operatorname{ad} y_i)^{1-a_{ij}} y_j = 0$,

where $a_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle$. *Conversely, given a root system* R *with a base* $\Pi \subseteq R$ *as above, the Lie algebra* $L(R)$ *with generators* x_i, y_i, h_i *as above and the above relations is a finite-dimensional semisimple Lie algebra with a root system* R. *Irreducible root systems correspond to simple Lie algebras.*

6. Highest weight representations

6.1. **Universal enveloping algebra.** Given a vector space V over a field K, let

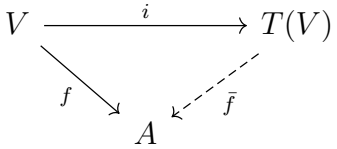
$$T^0V = K, \quad T^1V = V, \quad T^2V = V \otimes V, \quad T^nV = V \otimes \ldots \otimes V.$$

Define an associative product on $T(V) = \bigoplus_{i \geq 0} T^iV$ by

$$(v_1 \otimes \ldots v_m) \cdot (w_1 \otimes \ldots \otimes w_n) = v_1 \otimes \ldots v_m \otimes w_1 \otimes \ldots \otimes w_n.$$

Then $T(V)$ is an associative graded algebra with a unit $1 \in T^0V$, called the *tensor algebra* of V.

Lemma 6.1 (Universal property of $T(V)$)**.** *For any associative algebra A and a linear map $f\colon V \to A$, there exists a unique algebra homomorphism $\bar{f}\colon T(V) \to A$ making the following diagram commute*



where $i\colon V = T^1V \hookrightarrow T(V)$ is an embedding. The map $\bar{f}\colon T(V) \to A$ is defined on T^nV by the rule

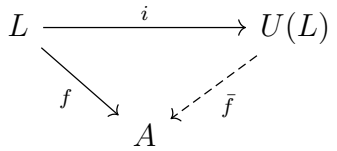
$$\bar{f}(v_1 \otimes \ldots \otimes v_n) = f(v_1)f(v_2)\ldots f(v_n).$$

Now let us consider a Lie algebra L and the corresponding tensor algebra $T(L)$. Using an embedding $i\colon L = T^1L \hookrightarrow T(L)$, we identify L with a subspace of $T(L)$. Let $I \subseteq T(L)$ be an ideal generated by elements

$$x \otimes y - y \otimes x - [x, y] \qquad \forall x, y \in L$$

and let $U(L) = T(L)/I$ be the corresponding (associative) quotient algebra, called the *universal enveloping algebra* of L.

Theorem 6.2. *For any associative algebra A and for any Lie algebra homomorphism $f\colon L \to A$ (meaning that f is linear and $f([x, y]) = f(x)f(y) - f(y)f(x)$), there exists a unique algebra homomorphism $\bar{f}\colon U(L) \to A$ (of associative algebras) such that the following diagram commutes*



Proof. We can uniquely extend the linear map $f\colon L \to A$ to an algebra homomorphism $f'\colon T(L) \to A$. For any $x, y \in L$, we have

$$f'(x \otimes y - y \otimes x - [x, y]) = f(x) \otimes f(y) - f(y) \otimes f(x) - f([x, y]) = 0$$

as $f\colon L \to A$ is a homomorphism of Lie algebras. This implies that $x \otimes y - y \otimes x - [x, y] \in \operatorname{Ker} f'$ and therefore $I \subseteq \operatorname{Ker} f'$. We conclude that $f'\colon T(L) \to A$ can be factored through $\bar{f}\colon T(L)/I \to A$. Uniqueness of $\bar{f}$ follows from the uniqueness of f'. □

Remark 6.3. Let $\rho\colon L \to \operatorname{End}(V)$ be a representation of a Lie algebra L. By the above result this map extends to an associative algebra homomorphism $\bar{\rho}\colon U(L) \to \operatorname{End}(V)$. This means that we can consider V as a module over the associative algebra $U(L)$, where multiplication is given by

$$av \coloneqq \bar{\rho}(a)v, \qquad a \in U(L),\ v \in V.$$

Conversely, if V is a module over the associative algebra $U(L)$ then it is automatically a module over the Lie algebra L, by using the embedding $i\colon L \hookrightarrow U(L)$.

Remark 6.4. Let V be a vector space equipped with an abelian Lie algebra structure. The ideal $I \subseteq T(V)$ from the above construction is generated by the elements

$$x \otimes y - y \otimes x, \qquad x, y \in V.$$

The corresponding algebra $U(V) = T(V)/I$ is called a symmetric algebra of V and is denoted by $S(V)$. If V has a basis $(x_1, \dots, x_n)$, then $S(V)$ is isomorphic to the algebra of polynomials $K[x_1, \dots, x_n]$.

The following important result tells us that one can similarly describe a basis of the universal enveloping algebra $U(L)$.

Theorem 6.5 (Poincaré–Birkhoff–Witt Theorem)**.** *Let $(x_1, \dots, x_n)$ be a basis of a Lie algebra L. Then the universal enveloping algebra $U(L)$ has a basis*

$$i(x_1)^{k_1} \dots i(x_n)^{k_n}, \qquad k_1, \dots, k_n \in \mathbb{N},$$

where $i\colon L \hookrightarrow T(L) \to U(L)$ is a canonical embedding.

Let $\pi\colon T(L) \to U(L)$ be the canonical projection. Define a filtration $0 = U_{-1} \subseteq U_0 \subseteq U_1 \subseteq \dots$ of $U(L)$ by

$$U_m = \pi(T^0 L \oplus \dots \oplus T^m L).$$

Then $U_m U_n \subseteq U_{m+n}$ and there is an induced associative algebra structure on

$$\operatorname{Gr} U(L) = \bigoplus_{i \geq 0} U_i / U_{i-1}.$$

If $x, y \in L \subseteq U_1$ then $xy - yx = [x, y]_L \in U_1$. More generally, one can show that for any $x \in U_m$ and $y \in U_n$, we have $xy - yx \in U_{m+n-1}$, hence this element induces zero in U_{m+n}/U_{m+n-1}. This implies that the algebra $\operatorname{Gr} U(L)$ is commutative. The natural embedding $L = U_1/U_0 \hookrightarrow \operatorname{Gr} U(L)$ can be considered as a homomorphism of abelian Lie algebras. It can be extended to an (associative) algebra homomorphism $S(L) \to \operatorname{Gr} U(L)$ by the universal property of $S(L)$. Then the PBW theorem can be reformulated as follows.

Theorem 6.6 (Poincaré–Birkhoff–Witt Theorem)**.** *The natural algebra homomorphism $S(L) \to \operatorname{Gr} U(L)$ is an isomorphism.*

Corollary 6.7. *Let L be a Lie algebra and $L_1, L_2 \subseteq L$ be subalgebras such that $L = L_1 \oplus L_2$ as a vector space. Then $U(L) \simeq U(L_1) \otimes U(L_2)$ as a vector space.*

Let L be a semisimple Lie algebra, $H \subseteq L$ be a Cartan subalgebra and

$$L = H \oplus \bigoplus_{\alpha \in R} L_\alpha$$

be its root space decomposition. Let $\Pi = \{\alpha_1, \dots, \alpha_r\} \subseteq R$ be a base of R, and $R_+ \subseteq R$ be the corresponding positive root system. Then we can write

$$L = L_- \oplus H \oplus L_+, \qquad L_- = \bigoplus_{\alpha \in R_-} L_\alpha, \qquad L_+ = \bigoplus_{\alpha \in R_+} L_\alpha.$$

This implies

$$U(L) \simeq U(L_-) \otimes U(H) \otimes U(L_+).$$

Note that $U(H) = S(H)$ is isomorphic to $K[h_1, \dots, h_r]$, where $h_i = \alpha_i^\vee$ are simple coroots. The algebra $U(L_+)$ is generated by the elements $x_1, \dots, x_r$ and the algebra $U(L_-)$ is generated by the elements $y_1, \dots, y_r$.

6.2. **Weights.** Let L be a semisimple Lie algebra with a choice of a Cartan algebra, root system and a base. Representations in this section are not necessarily finite-dimensional, unless otherwise stated.

For any (f.d.) L-representation $\rho\colon L \to \mathfrak{gl}(V)$, the image of H consists of commuting semisimple endomorphisms and we have a *weight space decomposition*

$$V = \bigoplus_{\lambda \in H^*} V_\lambda, \qquad V_\lambda = \{v \in V \mid hv = \lambda(h)v \ \forall h \in H\}.$$

The space V_λ are called *weight spaces* and the set

$$P(V) = \{\lambda \in H^* \mid V_\lambda \neq 0\}$$

is called the *set of weights* of V.

Lemma 6.8. *We have $L_\alpha V_\lambda \subseteq V_{\lambda+\alpha}$, for any $\alpha \in R$ and $\lambda \in P(V)$.*

Proof. Let $x \in L_\lambda$ and $v \in V_\lambda$. Then, for any $h \in H$, we have

$$h(xv) = ([h,x] + xh)v = \alpha(h)xv + x(\lambda(h)v) = (\alpha(h) + \lambda(h))xv = \langle \lambda + \alpha, h\rangle xv.$$

Therefore $xv \in V_{\lambda+\alpha}$. □

Definition 6.9. Let V be an L-representation. Then

(1) A vector $0 \neq v \in V_\lambda$ (for some $\lambda \in H^*$) is called a *highest weight vector* (or primitive vector) if $x_\alpha v = 0$ for all $\alpha \in R_+$.
(2) Representation V is called a *highest weight representation* if it is generated by its highest weight vector $v \in V_\lambda$. In this case λ is called the *highest weight* of V.

Lemma 6.10. *Every f.d. irreducible L-representation is a highest weight representation.*

Proof. The set of weights $P(V)$ is finite, hence there exists $\lambda \in P(V)$ such that $\lambda + \alpha \notin P(V)$ for all $\alpha \in R_+$. Let $0 \neq v \in V_\lambda$. Then $x_\alpha v \in V_{\lambda+\alpha} = 0$, hence $x_\alpha v = 0$, for all $\alpha \in R_+$. The vector v generates V as V is irreducible. □

Theorem 6.11. *Let V be a highest weight representation with highest weight λ. Then*

(1) V has a weight decomposition and weights of V are of the form $\lambda - \sum k_i\alpha_i$ with $k_i \in \mathbb{Z}_{\geq 0}$.
(2) $\dim V_\lambda = 1$ and $\dim V_\mu < \infty$ for $\mu \in H^*$.
(3) If V is f.d. then V is irreducible.

Proof. 1. Let $v \in V_\lambda$ be a highest weight vector generating V. We have $L_+v = 0$ and $Hv \subseteq V_\lambda$. Therefore $U(L_+)v = Kv$, $U(H)v = Kv$ and

$$V = U(L)v = U(L_-)v.$$

The space $U(L_-)v$ is generated by

$$y_{\beta_1} \dots y_{\beta_k} v, \qquad \beta_1, \dots, \beta_k \in R_+.$$

The above element has weight

$$\lambda - \sum \beta_i = \lambda_i - \sum_{i=1}^{r} m_i\alpha_i$$

for some $m_i \in \mathbb{Z}$. This proves (1).
2. If $v' = y_{\beta_1} \dots y_{\beta_k} v \in V_\lambda$, then $\lambda = \lambda - \sum \beta_i$, hence $\beta_1 = \dots = \beta_k = 0$. This implies that $v' = v$ and V_λ is generated by v. If V_μ is generated by elements of the form $y_{\beta_1} \dots y_{\beta_k} v$, then they satisfy $\mu = \lambda - \sum \beta_i$. There is a finite number of tuples (β_i) in R_+ satisfying this condition.

3. If V is not irreducible, then $V = V' \oplus V''$ for non-trivial V', V''. Then $V_\mu = V'_\mu \oplus V''_\mu$ for any weight μ. In particular, $V_\lambda = V'_\lambda \oplus V''_\lambda$. As $\dim V_\lambda = 1$, we have $V'_\lambda = V_\lambda$ or $V''_\lambda = V_\lambda$. Then $V' = V$ or $V'' = V$ as $v \in V_\lambda$ generates V. □

Theorem 6.12. *Let V be an irreducible f.d. representation. Then*

(1) V contains only one highest weight vector (up to a scalar).
(2) Two irreducible (f.d.) representations have the same highest weight if and only if they are isomorphic.

Proof. 1. Assume that $v \in V_\lambda$ and $v' \in V_{\lambda'}$ are two highest weight vectors. Then by the previous theorem

$$\lambda - \lambda' = \sum m_i \alpha_i, \qquad \lambda' - \lambda = \sum m'_i \alpha_i$$

for some $m_i, m'_i \in \mathbb{Z}_{\geq 0}$. This implies $m_i = m'_i = 0$ and $\lambda = \lambda'$.
2. Let V_1 and V_2 be irreducible representations having the same highest weight λ and having highest weight vectors v_1, v_2. Then the representation $W \subseteq V_1 \oplus V_2$ generated by the vector $v = (v_1, v_2) \in V_1 \oplus V_2$ is a highest weight representation. By the previous theorem W is irreducible. The composition $W \hookrightarrow V_1 \oplus V_2 \to V_1$ is non-trivial, hence is an isomorphism (as W and V_1 are irreducible). Similarly, $W \simeq V_2$ and we conclude that $V_1 \simeq V_2$. □

6.3. **Construction of highest weight representations.** Let V be a highest weight representation of L with the highest weight λ and a highest weight vector $v \in V_\lambda$. Then we have

$$hv = \lambda(h)v \quad \forall h \in H, \qquad xv = 0 \quad \forall x \in L_+.$$

Let us construct a universal module with these properties. Consider the left ideal $J(\lambda) \subseteq U(L)$ generated by L_+ and elements $h - \lambda(h)$ for all $h \in H$. Define the module

$$M(\lambda) = U(L)/J(\lambda),$$

called the *Verma module*. If V is a highest weight module with a highest weight vector $v \in V_\lambda$ as above, then there exists a surjective homomorphism of $U(L)$-modules (or L-modules)

$$M(\lambda) \to V, \qquad 1 \mapsto v.$$

One can construct $M(\lambda)$ also as follows. Let Kv_λ be a one-dimensional module over $H \oplus L_+$, where L_+ acts trivially and H acts as

$$hv_\lambda = \lambda(h)v_\lambda \qquad \forall h \in H.$$

Then Kv_λ is also a module over $U(H \oplus L_+)$ and we define

$$M(\lambda) = U(L) \otimes_{U(H \oplus L_+)} Kv_\lambda.$$

It is isomorphic to $U(L_-)$ as a vector space.

Theorem 6.13. *We have*

(1) $M(\lambda) = U(L_-)v_\lambda$.
(2) Weights of $M(\lambda)$ *are of the form* $\lambda - \sum m_i\alpha_i$ *for* $m_i \in \mathbb{Z}_{\geq 0}$.
(3) Each submodule of $M(\lambda)$ *has a weight space decomposition.*

Lemma 6.14. *For every* $\lambda \in H^*$ *there exists an irreducible module* $V(\lambda)$ *with highest weight* λ.

Proof. If $W \subseteq M(\lambda)$ is a submodule, then $W = \bigoplus W_\mu$ has a weight space decomposition. If $W \subseteq M(\lambda)$ is proper, then $W_\lambda = 0$ as otherwise $W_\lambda = M(\lambda)_\lambda$ generates $M(\lambda)$ and $W = M(\lambda)$. Taking the sum over all proper submodules, we obtain the maximal submodule $W \subseteq M(\lambda)$ (it is proper, as it satisfies $W_\lambda = 0$). Then the quotient $V(\lambda) = M(\lambda)/W$ is irreducible. □

Theorem 6.15. *For* $\lambda \in H^*$, *the irreducible module* $V(\lambda)$ *is finite-dimensional if and only if* $\langle \alpha_i^\vee, \lambda \rangle \in \mathbb{Z}_{\geq 0}$ *for all* $\alpha_i \in \Pi$ *(such* λ *is called a dominant integral weight).*