M4P46: LIE ALGEBRAS

LECTURES BY PROF. MARTIN LIEBECK; NOTES BY ALEKSANDER HORAWA

These are notes from the course M4P46: Lie Algebras taught by Prof. Martin Liebeck, in Winter 2016 at Imperial College London. They were LaTeX'd by Aleksander Horawa.

This version is from May 10, 2016. Please check if a new version is available at my website https://sites.google.com/site/aleksanderhorawa/. If you find a mistake, please let me know at aleksander.horawa13@ic.ac.uk.

References:

(main) K. Erdmann and M. Wildon, *Introduction to Lie algebras*
- N. Jacobson, *Lie algebras*
- J. Humphreys, *Introduction to Lie algebras and representation theory*

Website: http://wwwf.imperial.ac.uk/~mwl/m4p46/

Contents

Introduction

Recall that a field $(F, +, \times)$ satisfies

$(F, +)$ abelian group,
$(F^*, \times)$ abelian group,
$a(b+c) = ab + ac$.

For example, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$ are fields.

Finite fields: $\mathbb{F}_p = \mathbb{Z}_p$ (integers modulo p where p is a prime) and, more generally, for $q = p^a$ there exists a unique field $\mathbb{F}_q$ of order q.

The *characteristic* of a field F is the smallest $n \in \mathbb{N}$ such that

$$\underbrace{1 + 1 + \cdots + 1}_{n \text{ times}} = 0$$

and is 0 if no such n exists.

For example, $\text{char}(\mathbb{R}) = 0$, $\text{char}(\mathbb{F}_q) = p$.

Definition. An *algebra* over F is a vector space over F, with a multiplication ($vw \in V$ for all $v, q \in V$) satisfying bilinearity rules:

(1) $v(w_1 + w_2) = vw_1 + vw_2$ and $(v_1 + v_2)w = v_1 w + v_2 w$
(2) $\lambda(vw) = (\lambda v)w = v(\lambda w)$

for all $v_i, w_i, v, w \in V$, $\lambda \in F$.

Examples.

(1) $V = M_n(F)$, $n \times n$ matrices over F with matrix multiplication.
(2) $V = \mathbb{R}^3$, with the vector product $vw = v \times w$.
(3) V any vector space, product $vw = 0$ for all $v, w \in V$.

Definition. An algebra V is *associative* if

$$(xy)z = x(yz)$$

for all $x, y, z \in V$.

For example, $M_n(F)$ is associative, but $(\mathbb{R}^3, \times)$ is not.

Definition. A *Lie algebra* is an algebra V, with product $[xy]$ (sometimes written $[x, y]$ for clarity), satisfying:

(1) $[xx] = 0$ for all $x \in V$ (*skew-symmetric*),
(2) $[[xy]z] + [[yz]x] + [[zx]y] = 0$ for all $x, y, z \in V$ (*Jacobi identity*).

Note. For $x, y \in V$, (1) implies

$$\begin{aligned} 0 &= [x+y, x+y] && \text{by (1)} \\ &= [xx] + [xy] + [yx] + [yy] && \text{by bilinearity} \\ &= [xy] + [yx] && \text{by (1)} \end{aligned}$$

and therefore $[yx] = -[xy]$ (skew-symmetry). However, the converse implication does not hold in general.

We call $[vw]$ a *Lie product* or a *Lie bracket*.

Examples.

(1) Any V with the zero product $[vw] = 0$ is an *abelian* Lie algebra.

(2) We have that $(\mathbb{R}^3, \times)$ is a Lie algebra; indeed, recall that
   (1) $v \times v = 0$ for all $v \in \mathbb{R}^3$,
   (2) since $(u \times v) \times w = (u \cdot w)v - (v \cdot w)u$, we have that

$$(u\times v)\times w+(v\times w)\times u+(w\times u)\times v = ((u\cdot w)v-(v\cdot w)u)+((v\cdot u)w-(w\cdot u)v)+((w\cdot v)u-(u\cdot v)w) = 0.$$

(3) Let V be an associative algebra with the product vw. We define a bracket on V by

$$[vw] = vw - wv$$

for all $v, w \in V$. Then:
   (1) $[vv] = vv - vv = 0$,
   (2) $[[uv]w] + [[vw]u] + [[wu]v] = ((uv - vu)w - w(uv - vu)) + ((vw - wv)u - u(vw - wv)) + ((wu - uw)v - v(wu - uw)) = 0$ as V is associative.

Hence $[vw]$ is a Lie product on V.

(4) Apply example (3) to the associative algebra $V = M_n(F)$, i.e. define $[AB] = AB-BA$ for all $A, B \in V$. This makes V a Lie algebra, called the *general linear* Lie algebra, denoted $\mathfrak{gl}(n, F)$. **Note.** $\dim \mathfrak{gl}(n, F) = n^2$,

Definition. Let V be an algebra with product vw. We say $W \subseteq V$ is a *subalgebra* if W is a subspace and is closed under multiplication, i.e. $x, y \in W$ implies $xy \in W$.

(5) Define

$$\mathfrak{sl}(n, F) = \{A \in \mathfrak{gl}(n, F) : \mathrm{Tr}(A) = 0\}$$

where $\mathrm{Tr}(A) = \sum_i a_{ii}$ for $A = (a_{ij})$. We claim that $\mathfrak{sl}(n, F)$ is a subalgebra of $\mathfrak{gl}(n, F)$. It is clearly a subspace. We need to show that for $A, B \in \mathfrak{sl}(n, F)$, we have that $[AB] \in \mathfrak{sl}(n, F)$, i.e. $\mathrm{Tr}(A) = \mathrm{Tr}(B) = 0$ implies that $\mathrm{Tr}(AB - BA) = 0$. To see this, note that

$$\mathrm{Tr}(AB) = \sum_i \sum_j a_{ij} b_{ji},$$
$$\mathrm{Tr}(BA) = \sum_j \sum_i b_{ji} a_{ij},$$

so $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. We call $\mathfrak{sl}(n, F)$ the *special linear* Lie algebra.
Note. $\dim \mathfrak{sl}(n, F) = n^2 - 1$.

(6) Suppose that $\mathrm{char}(F) \neq 2$ (i.e. $1 \neq -1$). Define

$$\mathfrak{o}(n, F) = \{A \in \mathfrak{gl}(n, F) : A^T = -A\}.$$

We claim that this is a Lie subalgebra of $\mathfrak{gl}(n, F)$, called the *orthogonal Lie algebra.* It is clear that $\mathfrak{o}(n, F)$ is a subspace. For $A, B \in \mathfrak{o}(n, F)$, we have that

$$[AB]^T = (AB - BA)^T = B^T A^T - A^T B^T = (-B)(-A) - (-A)(-B) = BA - AB = -[AB],$$

and so $[AB] = -[AB]$.
Note. $\dim \mathfrak{o}(n, F) = \frac{1}{2}(n^2 - n)$. (Exercise.)

(7) Define
$$\mathfrak{t}(n,F) = \{A \in \mathfrak{gl}(n,F) : A \text{ is upper triangular}\}.$$
It is easy to check this is a Lie subalgebra of $\mathfrak{gl}(n,F)$.

(8) Define
$$\mathfrak{u}(n,F) = \{A \in \mathfrak{t}(n,F)) : a_{ii} = 0 \text{ for all } i\}.$$
It is easy to check this is a Lie subalgebra of $\mathfrak{t}(n,F)$.

(9) Let V be a vector space of dimension n over F, and let
$$L = \mathrm{End}(V),$$
the vector space of al linear transformations $V \to V$.

Then L is an associative algebra with product being composition of functions. The corresponding Lie algebra with
$$[fg] = fg - gf \qquad (f, g \in L)$$
is called $\mathfrak{gl}(V)$.

Note. The Lie algebras $\mathfrak{gl}(V)$ and $\mathfrak{gl}(n,F)$ are isomorphic (the definition of an isomorphism is below).

Similarly, we can define Lie subalgebras of $\mathfrak{gl}(v)$: $\mathfrak{sl}(V)$, $\mathfrak{o}(V)$, $\mathfrak{t}(V)$, $\mathfrak{u}(V)$.

Why study Lie algebras? We will see more later, but Lie algebras have fundamental connections with the following topics:

- Lie groups and differential geometry
- Finite simple groups
- Algebraic groups
- Root systems
- Reflection groups

1. Basic theory

A. **Homomorphisms and Ideals.**

Definition. Let L be a Lie algebra. We say $I \subseteq L$ is an *ideal* of L if

(1) I is a subspace of L
(2) $[IL] \subseteq I$, where $[IL] = \mathrm{Span}\{[il] : i \in I, l \in L\}$ (i.e. $[il] \in I$ for all $i \in I$, $l \in L$).

Note. $[li] = -[il]$, so also $[IL] \subseteq I$.

Examples.

(1) 0, L are ideals.
(2) $\mathfrak{sl}(n)$ is an ideal of $\mathfrak{gl}(n)$ as $[AB] = AB - BA \in \mathfrak{sl}(n)$ for all $A \in \mathfrak{sl}(n)$.[1]

Note. If I is an ideal then I is a subalgebra of L. Not conversely, for example $\mathfrak{t}(2)$ is a subalgebra of $\mathfrak{gl}(2)$ which is not an ideal.

[1]We write $\mathfrak{gl}(n)$, $\mathfrak{sl}(n)$, ... for $\mathfrak{gl}(n,F)$, $\mathfrak{sl}(n,F)$, ..., if we do not wish to specify the underlying field F.

Definition. Let L be a Lie algebra. The *centre* of L is

$$Z(L) = \{x \in L : [xl] = 0 \text{ for all } l \in L\}.$$

For example, $Z(\mathfrak{gl}(n)) = \{A : [AB] = 0 \text{ for all } B \in \mathfrak{gl}(n)\} = \{A : AB = BA \text{ for all } B \in \mathfrak{gl}(n)\}$. In fact, $Z(\mathfrak{gl}(n)) = \{\lambda I_n : \lambda \in F\}$ (which is an exercise on Sheet 1).

Note. $Z(L)$ is an ideal of L.

Definition. Let L, M be Lie algebras over F. We say $\varphi\colon L \to M$ is a *(Lie) homomorphism* if

(1) φ is linear,
(2) $\varphi([xy]) = [\varphi(x), \varphi(y)]$ for all $x, y \in L$.

Moreover, φ is an *isomorphism* if it is a bijective homomorphism.

Example. If $V = F^n$ then $\mathfrak{gl}(V) \cong \mathfrak{gl}(n, F)$. Fix a basis B of V. Then for $T \in \mathfrak{gl}(v)$

$$\phi\colon T \mapsto [T]_B$$

is a (Lie) isomorphism. (Exercise.)

The adjoint homomorphism. Let L be a Lie algebra. For $x \in L$, define

$$\operatorname{ad} x\colon L \to L$$

by

$$(\operatorname{ad} x)(y) = [xy] \qquad \text{for all } y \in L.$$

Then $\operatorname{ad} x \in \mathfrak{gl}(L)$.

Proposition 1.1. *The map*

$$\operatorname{ad}\colon L \to \mathfrak{gl}(L)$$

(sending x to $\operatorname{ad} x$) is a Lie homomorphism.

Proof. Clearly, ad is linear, as

$$\operatorname{ad}(x + y) = \operatorname{ad} x + \operatorname{ad} y,\ \operatorname{ad}(\lambda x) = \lambda \operatorname{ad} x.$$

So we need to check that

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

To check this, let $l \in L$. Then

$$\begin{aligned}
[\operatorname{ad} x, \operatorname{ad} y](l) &= [x[yl]] - [y[xl]] \\
&= [[ly]x] + [[xl]y] \\
&= -[[yx]l] && \text{by the Jacobi identity} \\
&= \operatorname{ad}[xy]\ (l)
\end{aligned}$$

□

Proposition 1.2. *If $\varphi\colon L \to M$ is a Lie homomorphism, then $\operatorname{Ker}\varphi$ is an ideal of L and $\operatorname{Im}\varphi$ is a Lie subalgebra of M.*

Proof. We have that $\operatorname{Ker}\varphi$ is a subspace, and for $x \in \operatorname{Ker}\varphi$, $l \in L$, we have that

$$\varphi[xl] = [\varphi(x)\varphi(l)] = 0,$$

and hence $[xl] \in \operatorname{Ker}\varphi$. Thus $\operatorname{Ker}\varphi$ is an ideal of L.

Moreover, $\operatorname{Im}\varphi$ is a subspace, and for $x, y \in L$

$$[\varphi(x)\varphi(y)] = \varphi[xy] \in \operatorname{Im}\varphi,$$

hence $\operatorname{Im}\varphi$ is a Lie subalgebra. □

Note. $\operatorname{Ker}\operatorname{ad} = Z(L)$.

B. **Derivations.**

Definition. Let A be an algebra over F. We say a linear map $D\colon A \to A$ is a *derivation* if

$$D(xy) = xD(y) + D(x)y$$

for all $x, y \in A$.

Let $\operatorname{Der} A = \{\text{all derivations of } A\}$.

Proposition 1.3. *The set* $\operatorname{Der} A$ *is a Lie subalgebra of* $\mathfrak{gl}(A)$.

Proof. Clearly, $\operatorname{Der} A$ is a subspace: $0 \in \operatorname{Der} A$, and if $D, E \in \operatorname{Der} A$, then $D + A \in \operatorname{Der} A$ and $\lambda D \in \operatorname{Der} A$ for $\lambda \in F$. To show it is a subalgebra, for $D, E \in \operatorname{Der} A$, check that $[DE] = DE - ED \in \operatorname{Der} A$. □

Examples.

(1) Let $A = C^\infty(\mathbb{R})$ be the vector space of infinitely differentiable functions $\mathbb{R} \to \mathbb{R}$. This is an algebra with product $(fg)(x) = f(x)g(x)$. The map $D\colon A \to A$, $D(f) = f'$, is a derivation.

(2) Let L be a Lie algebra. Let $x \in L$. We claim that $\operatorname{ad} x\colon L \to L$ (sending y to $[xy]$) is a derivation of L. Indeed, we know that $\operatorname{ad} x$ is linear, and

$$\begin{aligned}(\operatorname{ad} x)[yz] &= [x[yz]] \\ &= -[[yz]x] \\ &= [[xy]z] + [[zx]y] && \text{by the Jacobi identity} \\ &= [(\operatorname{ad} x)(y), z] + [y, (\operatorname{ad} x)(z)]\end{aligned}$$

Therefore, we obtain the following result.

Proposition 1.4. *If L is a Lie algebra, then*

$$\operatorname{ad} L \subseteq \operatorname{Der} L \subseteq \mathfrak{gl}(L)$$

(where $\operatorname{ad} L = \{\operatorname{ad} x : x \in L\}$*), are both Lie subalgebras of* $\mathfrak{gl}(L)$.

C. **Structure constants.** Let L be a Lie algebra. Assume that $x_1, \dots, x_n$ is a basis of L. The brackets $[x_i x_j]$ for all i, j determine the Lie bracket on L. We write

$$[x_i x_j] = \sum_{k=1}^{n} a_{ijk} x_k \qquad (a_{ijk} \in F).$$

Then the scalars a_{ijk} are the *structure constants* of L with respect to the basis $x_1, \dots, x_n$.

Remarks.

(1) The fact that $[xy]$ is a Lie bracket implies conditions on the a_{ijk}. For example, $[x_i x_i] = 0$ implies that $a_{iik} = 0$ for all k. (For more, see Sheet 1.)
(2) If L_1, L_2 are Lie algebras over F, with bases B_1, B_2, having the same structure constants, then $L_1 \cong L_2$. (For details, see Sheet 1.)

Examples.

(1) Consider $\mathfrak{sl}(2)$, the Lie algebra of 2×2 matrices of trace 0. It has dimension 3, and standard basis

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

It has structure constants:

$$[ef] = ef - fe = h$$
$$[eh] = eh - he = -2e$$
$$[fh] = fh - hf = 2f$$

(and $[ee] = 0$, $[fe] = -[ef]$). We can get all the other structure constants from these.
(2) Consider $(\mathbb{R}^3, \times)$, a 3-dimensional Lie algebra over $\mathbb{R}$. It has standard basis:

$$e_1 = (1,0,0), \; e_2 = (0,1,0), \; e_3 = (0,0,1)$$

and structure constants

$$e_1 \times e_2 = e_3, \; e_2 \times e_3 = e_1, \; e_3 \times e_1 = e_2.$$

(3) Consider $\mathfrak{o}(3)$, the orthogonal Lie algebra of 3×3 matrices A such that $A^T = -A$ (assuming char $F \neq 2$). It has dimension 3, and standard basis:

$$S_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \; S_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \; S_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It has structure constants:

$$[S_1 S_2] = S_1 S_2 - S_2 S_1 = S_3,$$
$$[S_2 S_3] = S_1,$$
$$[S_3 S_1] = S_2.$$

These are the same structure constants as for $(\mathbb{R}^3, \times)$. Hence

$$(\mathbb{R}^3, \times) \cong \mathfrak{o}(3, \mathbb{R}).$$

D. **Quotient Algebras.** Let be L a Lie algebra and I be an ideal of L (so I is a subspace and $[IL] \subseteq I$).

We have the quotient vector space L/I:

vectors are cosets $x + I$ (for $x \in L$)
addition: $(x + I) + (y + I) = (x + y) + I$
scalar multiplication: $\lambda(x + I) = (\lambda x) + I$

Proposition 1.5. *The vector space L/I is a Lie algebra with bracket*

$$[x + I, y + I] = [xy] + I \qquad \text{for all } x, y \in L.$$

Proof. We check that this bracket is well-defined. Suppose $x_1 + I = x_2 + I$, $y_1 + I, y_2 + I$. This implies that $x_1 - x_2, y_1 - y_2 \in I$, and hence

$$[(x_1 - x_2)y_1] + [x_2(y_1 - y_2)] \in I \text{ (since } I \text{ is an ideal),}$$

so $[x_1y_1] - [x_2y_2] \in I$, which means that $[x_1y_1] + I = [x_2y_2] + I$.

Finally, we note that this bracket is bilinear, skew-symmetric, and satisfies the Jacobi identity, since the original bracket had these properties. $\square$

Proposition 1.6 (First Isomorphism Theorem). *If $\varphi\colon L \to M$ is a Lie homomorphism, then* $\operatorname{Ker}\varphi$ *is an ideal of L,* $\operatorname{Im}\varphi$ *is a subalgebra of M, and there is a Lie isomorphism*

$$\frac{L}{\operatorname{Ker}\varphi} \cong \operatorname{Im}\varphi.$$

Proof. Write $I = \operatorname{Ker}\varphi$ and define $\alpha\colon \frac{L}{I} \to \operatorname{Im}\varphi$ by

$$\alpha(x + I) = \varphi(x) \quad \text{for all } x \in L.$$

It is easy to check that α is well-defined, bijective, and a Lie homomorphism. $\square$

Example. Let L be a Lie algebra. We know $\operatorname{ad}\colon L \to \mathfrak{gl}(L)$ is a Lie homomorphism, and $\operatorname{Ker}(\operatorname{ad}) = Z(L)$, so

$$\frac{L}{Z(L)} = \frac{L}{\operatorname{Ker}(\operatorname{ad})} = \operatorname{ad} L.$$

Definition. A Lie algebra L is *simple* if $\dim L \geq 1$ and its only ideals are 0 and L.

Example. The Lie algebra $\mathfrak{sl}(2, F)$ is simple, provided $\operatorname{char}(F) \neq 2$. (See Coursework 1.)

2. Lie algebras of small dimension

We classify Lie algebras of dimension less than or equal to 3.

Abelian Lie algebras ($[xy] = 0$ for all x, y) of same dimension over F are isomorphic. So, up to isomorphism, there exists a unique Lie algebra over any F of dimension n for all n.

Thus we study non-abelian Lie algebras. We organise the study via two ideals: the centre

$$Z(L) = \{x \in L \mid [xy] = 0 \text{ for all } y \in L\}$$

and the *derived algebra*

$$L' = \operatorname{Span}\{[xy] : x, y \in L\}.$$

Also, if L_1, L_2 are Lie algebras over F, their *direct sum* is

$$L_1 \oplus L_2 = \{(x_1, x_2) : x_i \in L_i\}$$

with bracket

$$[(x_1, x_2), (y_1, y_2)] = ([x_1 y_1], [x_2 y_2])$$

which makes $L_1 \oplus L_2$ a Lie algebra.

Dimension 1. Any Lie algebra of dimension 1 is abelian (as $[xx] = 0$).

Dimension 2. There is a unique abelian and a unique non-abelian Lie algebra of dimension 2.

Proposition 2.1. *Up to isomorphism, there exists a unique non-abelian Lie algebra of dimension 2 over any F: it has a basis x, y with $[xy] = x$ (which determines all the structure constants).*

Proof. If u, v is a basis of L, then $[uv] \neq 0$ (as L non-abelian), so $L' = \mathrm{Span}([uv])$ is a 1-dimensional Lie algebra. Let $x = [uv]$ so $L' = \mathrm{Span}(x)$. Extend to a basis x, y of L. Since L' is an ideal:

$$[xy] = \lambda x \quad \text{for some } \lambda \in F^*.$$

Replace y by $\lambda^{-1} y$ to get $[xy] = x$.

Finally, we check that the axioms (skew-symmetry, Jacobi) hold for this Lie bracket. (See Sheet 1, Question 2 (ii)) ☐

Dimension 3. Let L be non-abelian of dimension 3. Then $Z(L)$ has dimension 0, 1 or 2, and L' has dimension 1, 2 or 3.

Organise the study by cases for $\dim L'$ and containment between L' and $Z(L)$.

(A) $\dim L' = 1$. We consider two cases
  (a) $L' \subseteq Z(L)$,
  (b) $L' \not\subseteq Z(L)$.

Proposition 2.2. *Up to isomorphism, there exists a unique 3-dimensional Lie algebra L over F such that* $\dim L' = 1$ *and $L' \subseteq Z(L)$, namely*

$$\mathfrak{u}(3, F) = \left\{ \begin{pmatrix} 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix} \right\} \subseteq \mathfrak{gl}(3, F).$$

Proof. Let $L' = \mathrm{Span}(z)$ where $z = [fg]$. We claim that f, g, z is a basis of L. Since L is 3-dimensional, we only have to check that they are linearly independent. Suppose that

$$\alpha f + \beta g + \gamma z = 0.$$

By assumption, $z \in Z(L)$. Bracketing the equation with f, we obtain $\beta = 0$, and bracketing the equation with g, we obtain $\alpha = 0$. Thus also $\gamma = 0$. Therefore, L has a basis f, g, z with structure constants

$$[fg] = z, \ [fz] = [gz] = 0.$$

Finally, $\mathfrak{u}(3,F)$ is such a Lie algebra, taking

$$f = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ g = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},\ z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(One can verify this by multiplying the matrices.) □

Definition. We call $\mathfrak{u}(3,F)$ the *Heisenberg* Lie algbera over F.

Proposition 2.3. *Up to isomorphism, there exists a unique 3-dimensional Lie algebra L over F such that $\dim L' = 1$ and $L' \nsubseteq Z(L)$, namely $L_1 \oplus L_2$, where $\dim L_1 = 1$ and $\dim L_2 = 2$, L_2 non-abelian (as in Proposition 2.1).*

Proof. Let $L' = \mathrm{Span}(x)$. As $x \notin Z(L)$, there exists $y \in L$ such that

$$[xy] = \alpha x,\ \ \alpha \neq 0.$$

Replace y by $\alpha^{-1} y$ to get $[xy] = y$.

Extend x, y to a basis x, y, w of L. Let

$$[xw] = ax,\ \ [yw] = bx.$$

We claim that there exists $z \in Z(L)$ such that x, y, z is a basis of L. To show this, observe that

$$[x, \alpha x + \beta y + \gamma z] = (\beta + \gamma a)x,$$
$$[y, \alpha x + \beta y + \gamma z] = (-\alpha + \gamma b)x.$$

Take $\gamma = 1$, $\alpha = b$, $\beta = -a$. So if

$$z = bx - ay + w$$

then $z \in Z(L)$ and x, y, z is a basis of L. Finally

$$L \cong \mathrm{Span}(x,y) \oplus \mathrm{Span}(z) \cong L_2 \oplus L_1,$$

as required. □

Note. $L \cong \mathfrak{t}(2,F) = \left\{ \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \right\} \subseteq \mathfrak{gl}(2,F)$. (Exercise)

(B) $\dim L' = 2$.

Lemma 2.4. *Suppose $\dim L = 3$, $\dim L' = 2$. Then*
(1) *L' is abelian,*
(2) *for $x \in L \setminus L'$ the map $\mathrm{ad}\, x \colon L' \to L'$ is an isomorphism.*

Proof. For (1), let $L' = \mathrm{Span}(y,z)$ and extend to a basis x, y, z of L. Let $[yz] = \alpha y + \beta z$. The matrix

$$(\mathrm{ad}\, y)_{x,y,z} = \begin{pmatrix} 0 & 0 & 0 \\ * & 0 & \alpha \\ * & 0 & \beta \end{pmatrix}.$$

Now, $\mathrm{Tr}(\mathrm{ad}\, y) = 0$ (Sheet 1, Question 4). So $\beta = 0$. Similarly $\alpha = 0$ by considering $\mathrm{ad}\, z$. Therefore, $[yz] = 0$ and L' is abelian.

For (2), note that as $[yz] = 0$, $L' = \mathrm{Span}([xy],[xz])$. Hence $\mathrm{ad}\, x \colon L' \to L'$ is surjective, thus an isomorphism. □

Assume now $F = \mathbb{C}$. Two cases:

(a) There exists $x \in L \setminus L'$ such that $\operatorname{ad} x \colon L' \to L'$ is diagonalisable.

(b) Not (a).

Proposition 2.5. *Suppose* $\dim L = 3$, $\dim L' = 2$ *and (a) holds. Then there exists* $\mu \in F^*$ *and a basis* x, y, z *of* L *such that*

$$[xy] = y,\ [xz] = \mu z,\ [yz] = 0.$$

These structure constants define a Lie algebra L_μ*, and* $L_\mu \cong L_\nu$ *if and only if* $\mu = \nu^{\pm 1}$*.*

Proof. By (a), there exist basis y, z of L' such that

$$[xy] = \lambda y,\ [xz] = \mu z$$

and $\lambda, \mu \neq 0$ by Lemma 2.4 (2). Rescale x to take $\lambda = 1$. Also $[yz] = 0$ by Lemma 2.4 (1). Finally, we need to check the last assertion. (Sheet 1, Questions 2 and 6). ☐

Proposition 2.6. *Suppose* $F = \mathbb{C}$, $\dim L = 3$, $\dim L' = 2$*, and (b) holds. There is a unique (up to isomorphism) Lie algebra* L *with basis* x*,* y*,* z *and*

$$[xy] = y,\ [xz] = y + z,\ [yz] = 0.$$

Proof. Let $x \in L \setminus L'$. As $F = \mathbb{C}$, $\operatorname{ad} x \colon L' \to L'$ has an eigenvector $y \neq 0$, and we can rescale x to take $[xy] = y$. Let $L' = \operatorname{Span}(y, z)$ and

$$[xz] = \lambda y + \mu z.$$

By (b), $\lambda \neq 0$. Scale z to take $\lambda = 1$. So

$$(\operatorname{ad} x)_{y,z} = \begin{pmatrix} 1 & 1 \\ 0 & \mu \end{pmatrix}.$$

By (b), $\mu = 1$. Hence

$$[xy] = y,\ [xz] = y + z,\ [yz] = 0.$$

Finally, we check these define a Lie algebra. (Sheet 1, Question 2(ii)) ☐

(C) $\dim L' = 3$. Assume $F = \mathbb{C}$.

Proposition 2.7. *Assume* $F = \mathbb{C}$, $\dim L = 3$ *and* $L' = L$*. Then* $L \cong \mathfrak{sl}(2, \mathbb{C})$*.*

Proof. (1) For $0 \neq x \in L$, $\operatorname{ad} x \colon L \to L$ has rank 2.
If x, y, z is a basis, then $L' = \operatorname{Span}([xy], [xz], [yz])$, so $[xy]$ and $[xz]$ are linearly independent.

(2) There exists $h \in L$ such that $\operatorname{ad} h$ has a nonzero eigenvalue.
Let $0 \neq x \in L$ and assume $\operatorname{ad} x$ has only eigenvalue 0. As $\operatorname{ad} x$ has rank 2, by (1), the Jordan Canonical Form of $\operatorname{ad} x$ is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

So there exists a basis x, y, z such that $[xy] = x$, $[xz] = y$. So $\operatorname{ad} y$ has eigenvalue -1.

By (2), there exist $h, x \in L$ such that $[hx] = \alpha x$ for $\alpha \neq 0$. As $\mathrm{Tr}(\mathrm{ad}\, h) = 0$ (Sheet 1, Question 4), the eigenvalues of $\mathrm{ad}\, h$ are 0, α, $-\alpha$. So there exists a basis h, x, y with

$$[hx] = \alpha x,\ [hy] = -\alpha y.$$

To find $[xy]$, note that

$$\begin{aligned} [h[xy]] &= [[hx]y] + [x[hy]] \quad \text{by the Jacobi identity} \\ &= \alpha[xy] - \alpha[xy] \\ &= 0 \end{aligned}$$

Therefore, $[xy] \in \mathrm{Ker}(\mathrm{ad}\, x) = \mathrm{Span}(h)$, so $[xy] = \lambda h$ for $\lambda \neq 0$ (as $L' = L$). Rescale h ($h \mapsto \frac{2}{\alpha} h$) to take $\alpha = 2$. Then rescale x ($x \mapsto \frac{1}{\lambda} x$) to take $\lambda = 1$. So now

$$[hx] = 2x,\ [hy] = -2y, [xy] = h.$$

These are the structure constants for the basis e, f, h of $\mathfrak{sl}(2, \mathbb{C})$. So $L \cong \mathfrak{sl}(2, \mathbb{C})$, as required. $\square$

Note. Proposition 2.7 is not true over other fields F.

3. Soluble and nilpotent Lie algebras

Let L be a Lie algebra, I, J be ideals. Define:

$$I + J = \{i + j : i \in I, j \in J\},$$
$$[IJ] = \mathrm{Span}\{[ij] : i \in I, j \in J\} \subseteq I \cap J.$$

Proposition 3.1. *The sets $I \cap J$, $I + J$ and $[IJ]$ are ideals.*

Proof. The fact that $I \cap J$, $I + J$ are ideals is left as an exercise. Note that $[IJ]$ is a subspace, and for $l \in L$

$$\begin{aligned} [[ij]l] &= [[il]j] - [[jl]i] && \text{by the Jacobi idenitity} \\ &\in [IJ] && \text{since } [il] \in I \text{ and } [jl] \in J \end{aligned},$$

so $[IJ]$ is an ideal. $\square$

Proposition 3.2.

(1) *There exists a Lie isomorphism*

$$\frac{I + J}{J} \cong \frac{I}{I \cap J} \quad \textit{(second isomorphism theorem).}$$

(2) *Ideals of quotient $\frac{L}{I}$ are of the form $\frac{K}{I}$ where $I \subseteq K \subseteq L$ and K is an ideal.*

Proof. For (1), define a map $\varphi\colon I \to \frac{I+J}{J}$ by $\varphi(i) = i + J$ for $i \in I$. This is a surjective Lie homomorphism with kernel

$$\mathrm{Ker}\, \varphi = \{i \in I : i + J = J\} = I \cap J.$$

Hence (1) follows from the first isomorphism theorem 1.6.

For (2), let M be an ideal of $\frac{L}{I}$ and define

$$K = \{x \in L : x + I \in M\}.$$

To finish the proof, check that $I \subseteq K$, K is an ideal of L, and $M = \frac{K}{I}$. $\square$

Solubility. Recall $L' = \mathrm{Span}\{[xy] : x, y \in L\}$ is an ideal.

Proposition 3.3. *Let L be a Lie algebra. For an ideal I, the quotient L/I is abelian if and only if $L' \subseteq I$. In particular, L/L' is abelian.*

Proof. We have that L/I is abelian if and only if $[x+I, y+I] = I$ for all $x, y \in L$, i.e. $[xy] + I = I$ for all $x, y \in L$. But this is equivalent to $[xy] \in I$ for all $x, y \in I$, which means $L' \subseteq I$. □

Define a series of ideals of L:

$$\begin{aligned} L^{(1)} &= L' = [LL], \\ L^{(2)} &= \left[L^{(1)}, L^{(1)}\right], \\ &\vdots \\ L^{(i)} &= \left[L^{(i-1)}, L^{(i-1)}\right]. \end{aligned}$$

We have $L \supseteq L^{(1)} \supseteq L^{(2)} \supseteq \cdots$, the *derived series of* L, and $L^{(i)}/L^{(i+1)}$ is abelian by Proposition 3.3.

Definition. A Lie algebra L is *soluble* if $L^{(m)} = 0$ for some m.

Examples.

(1) If $\dim L = 2$, then by Proposition 2.1 we have $\dim L' \leq 1$, so $L^{(2)} = 0$ and L is soluble.
(2) If $\dim L = 3$ and $\dim L' \leq 2$, then by (1) we have $L^{(3)} = 0$, and hence L is soluble. If $\dim L = 3$ and $\dim L' = 3$, then $L^{(1)} = L$, so $L^{(i)} = L$ for all i, and L is not soluble.
(3) We have that $\mathfrak{t}(n, F)$ and $\mathfrak{u}(n, F)$ are soluble (Sheet 1).

Proposition 3.4. *Let L be a Lie algebra.*

(1) *If L is soluble, then all subalgebras of L and quotient algebras L/I are soluble.*
(2) *If I is an ideal, and both I and L/I are soluble, then L is soluble.*
(3) *If I, J are soluble ideals of L, then $I + J$ is soluble.*

Proof. For (1), suppose $L^{(m)} = 0$. For a subalgebra M, $M^{(m)} = 0$, so M is soluble. Also, for an ideal I,

$$\left(\frac{L}{I}\right)^{(i)} = \frac{L^{(i)} + I}{I},$$

and hence $\left(\frac{L}{I}\right)^{(m)} = 0$, so L/I is soluble.

For (2), assume $I^{(m)} = 0$ and $\left(\frac{L}{I}\right)^{(n)} = 0$. Then

$$\left(\frac{L}{I}\right)^{n} = \frac{L^{(n)} + I}{I} = I$$

so $L^{(n)} \subseteq I$. Hence

$$0 = (L^{(n)})^m = L^{(n+m)},$$

so L is soluble.

Finally, to show (3), we assume I and J are soluble. Then by the second isomorphism theorem (Proposition 3.2 (1)), there exists a Lie isomorphism

$$\frac{I+J}{I} \to \frac{J}{I\cap J}.$$

Now, $\frac{I}{I\cap J}$ is soluble by (1), and I is soluble. Therefore, $I+J$ is soluble by (2). $\square$

Corollary 3.5. *Let L be a finite-dimensional Lie algebra. Then L has a unique soluble ideal which contains every soluble ideal of L.*

Proof. Let R be a soluble ideal of L of maximal dimension. If I is any soluble ideal, then $I+R$ is a soluble ideal by Proposition 3.4 (3). But by the choice of R, $\dim(I+R) \leq \dim R$. Hence $I+R=R$, so $I \subseteq R$. $\square$

Definition. The unique maximal soluble ideal of L (finite-dimensional Lie algebra) is called the *radical* of L, written $\mathrm{Rad}(L)$.

Examples.

(1) If L is soluble, then $\mathrm{Rad}(L) = L$.
(2) Take $L = \mathfrak{sl}(2,\mathbb{C})$. Here, $L = L'$ (by considering the structure constants $[ef] = h$, $[eh] = -2e$, $[fh] = 2f$, we note that any basis vector is contained in L'), so L is not soluble. Also, L is simple (see Coursework 1), so $\mathrm{Rad}(L)$ is 0 or L. Therefore, $\mathrm{Rad}(L) = 0$.
(3) Take $L = \mathfrak{gl}(2,\mathbb{C})$. Here $\mathrm{Rad}(L) = Z(L) = \{\lambda I : \lambda \in \mathbb{C}\}$ (exercise).

Definition. If $\mathrm{Rad}(L) = 0$, we call L a *semisimple* Lie algebra.

Fact. *If L is simple, then L is semisimple.*

Note that we assume that simple Lie algebras have dimension greater than 1. Otherwise, the abelian Lie algebra of dimension 1 would be a counterexample to this statement.

Proof. If L is simple, then $L' = 0$ or L. Hence $L' = L$ (as abelian Lie algebras of dimension greater than 1 are not simple). Thus L is not soluble. Also, $\mathrm{Rad}(L) = 0$ or L, and is not L, because L is not soluble. Therefore, $\mathrm{Rad}(L) = 0$, and L is semisimple. $\square$

Proposition 3.6. *Let L be a finite-dimensional Lie algebra. Then $\frac{L}{\mathrm{Rad}(L)}$ is semisimple.*

Proof. Let $R = \mathrm{Rad}(L)$, and let I be a soluble ideal of $\frac{L}{R}$. Then $I = \frac{J}{R}$, where J is an ideal of L containing I by Proposition 3.2. Now, I is soluble, R is soluble, so J is soluble by Proposition 3.4 (2). Hence $J \subseteq \mathrm{Rad}(L) = R$, and so $I = 0$. Therefore, $\mathrm{Rad}\left(\frac{L}{R}\right) = 0$. $\square$

Nilpotence. We define another series of ideals of a Lie algebra L:

$$\begin{aligned} L^1 &= L' = [LL], \\ L^2 &= [L, L^1], \\ L^3 &= [L, L^2], \\ &\vdots \\ L^i &= [L, L^{i-1}]. \end{aligned}$$

We have $L \supseteq L^1 \supseteq L^2 \supseteq \cdots$, the *lower central series series* of L.

Note. That the series is *central* means $\frac{L^i}{L^{i+1}} \subseteq Z\left(\frac{L}{L^{i+1}}\right)$. Indeed:

$$\left[\frac{L^i}{L^{i+1}}, \frac{L}{L^{i+1}}\right] = \frac{[L^i L]}{L^{i+1}} = 0.$$

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

Clearly, if L is nilpotent then L is soluble.

Examples.

(1) Suppose $\dim L = 2$, L is non-abelian. By Proposition 2.1, L has basis x, y with $[xy] = x$. So $L^1 = \mathrm{Span}(x)$, $L^2 = \mathrm{Span}(x)$, …, $L^i = \mathrm{Span}(x)$ for all i. Hence L is **not** nilpotent.
(2) The Lie algebra $\mathfrak{u}(n, F)$ is nilpotent, but $\mathfrak{t}(n, F)$ is not nilpotent (for $n \geq 2$). (Sheet 1).

The future. By Proposition 3.6, if L is a finite-dimensional Lie algebra, $\frac{L}{\mathrm{Rad}(L)}$ is semisimple. So we need to understand:

- soluble Lie algebras,
- semisimple Lie algebras.

Over $\mathbb{C}$:

(A) Structure of soluble Lie algebras is given by Lie's Theorem: every soluble Lie algebra is isomorphic to a subalgebra of $\mathfrak{t}(n, \mathbb{C})$ for some n. (Chapter 5)
(B) A theorem to come says that every semisimple Lie algebra is isomorphic to $L_1 \oplus \cdots \oplus L_r$, a direct sum of **simple** Lie algebras L_i.
We will classify the simple Lie algebras over $\mathbb{C}$: they are

classical types:	$\mathfrak{sl}_n, \mathfrak{o}_n, \mathfrak{sp}_{2n}$ (Sheet 1)
5 exceptional types:	$\mathfrak{g}_2$, $\mathfrak{f}_4$, $\mathfrak{e}_6$, $\mathfrak{e}_7$, $\mathfrak{e}_8$

4. Engel's Theorem

The theorem is an important result about Lie subalgebras of $\mathfrak{gl}(V)$.

Definition. A linear map $T\colon V \to V$ is *nilpotent* if $T^r = 0$ for some r. An $n \times n$ matrix A is *nilpotent* if $A^r = 0$ for some r.

Note. A nilpotent linear transformation $T\colon V \to V$ has only eigenvalue 0, so its characteristic polynomial is x^n, where $n = \dim V$. Hence $T^n = 0$ by Cayley–Hamilton Theorem. Also, the Jordan Canonical Form of T is of the form

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix},$$

strictly upper-triangular. Also, such a matrix is nilpotent (as its characteristic polynomial is x^n), and lies in $\mathfrak{u}(n, F)$.

Any Lie subalgebra of $\mathfrak{u}(n,F)$ consists of nilpotent matrices. Engel's theorem gives the converse.

Theorem 4.1 (Engel's Theorem)**.** *Let V be n-dimensional over F. Suppose L is a Lie subalgebra of $\mathfrak{gl}(V)$ such that every element of L is nilpotent (i.e. $T^n = 0$ for all $T \in L$). Then there exists a basis B of V with respect to which every $T \in L$ is represented by a strictly upper-triangular matrix, i.e.*

$$L \cong \{[T]_B : T \in L\} \subseteq \mathfrak{u}(n,F).$$

In particular, L is a nilpotent Lie algebra.

Idea of the proof. Consider the case where L is 1-dimensional, so $L = \mathrm{Span}(T)$, where $T\colon V \to V$ is nilpotent. The conclusion is that there exists a basis B of V such that

$$[T]_B = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix},$$

(and the same for any λT). This is true by the Jordan Canonical Form theorem.

How do we prove the Jordan Canonical Form theorem for nilpotent T?

(a) There is an eigenvector $0 \neq u \in V$ for T, so $T(u) = 0$.
(b) Let $U = \mathrm{Span}(u)$. Then T induces a linear map $\overline{T}\colon \frac{V}{U} \to \frac{V}{U}$ by

$$\overline{T}(v+U) = T(v) + U \text{ for all } v \in V.$$

It is well-defined, since $T(u) = 0 \in U$. Observe that since T is nilpotent, $\overline{T}$ is nilpotent, and $\dim \frac{V}{U} = n-1$.
By induction (on n), there exists a basis $\overline{B} = \{v_1 + U, \dots, v_{n-1} + U\}$ of $\frac{V}{U}$ such that

$$[\overline{T}]_{\overline{B}} = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

(an $(n-1)\times(n-1)$ matrix). Then $B = \{u, v_1, \dots, v_{n-1}\}$ is a basis of V, and

$$[T]_B = \begin{pmatrix} 0 & * & \dots & * \\ 0 & 0 & & * \\ \vdots & & \ddots & \\ 0 & 0 & & 0 \end{pmatrix}$$

(the first row is $T(u)$), which is strictly upper-triangular.

The proof of Engel's theorem runs along the same lines, replacing $\mathrm{Span}(T)$ by a whole Lie algebra L. The hard step is (a): showing that there exists a common eigenvector for all $T \in L$.

We will now work towards proving this.

Lemma 4.2. *Let L be a Lie subalgebra of $\mathfrak{gl}(V)$. Let $x \in L$, and suppose $x\colon V \to V$ is nilpotent. Then $\mathrm{ad}\, x\colon L \to L$ is also nilpotent.*

Proof. Let $y \in L$. Then
$$(\operatorname{ad} x)(y) = [xy] = xy - yx$$
and hence
$$(\operatorname{ad} x)^2(y) = (\operatorname{ad} x)(xy - yx) = [x, xy - yx] = x(xy - yx) - (xy - yx)x = x^2y - 2xyx + yx^2.$$
Similarly, $(\operatorname{ad} x)^3$ is a linear combination of terms of the form $x^3y, x^2yx, xyx^2, xyx^2$, and in general, $(\operatorname{ad} x)^m(y)$ is a linear combination of terms of the form $x^j yx^{m-j}$ where $0 \le j \le m$.[2]

By hypothesis, $x^n = 0$ for some n. Then for $m = 2n$, either x^j or x^{m-j} is 0, for all $0 \le j \le m$. This shows that $(\operatorname{ad} x)^{2n} = 0$, hence $\operatorname{ad} x$ is nilpotent. □

Lemma 4.3. *Let L be a Lie subalgebra of $\mathfrak{gl}(V)$, and I an ideal of L. Define*
$$W = \{v \in V : T(v) = 0 \text{ for all } T \in I\}.$$
*Then W is L-*invariant*, i.e. $S(W) \subseteq W$ for all $S \in L$.*

Proof. Let $S \in L$ and $w \in W$. For $T \in I$, note $[TS] = TS - ST$, so
$$TS(w) = ST(w) + [TS](w) = 0 + 0$$
since both T and $[TS]$ are in I. Therefore, $S(w) \in W$. □

The following proposition is a key step in the proof of Engel's Theorem 4.1.

Proposition 4.4. *Let L be a Lie subalgebra of $\mathfrak{gl}(V)$ consisting of nilpotent elements. Then there exists $0 \neq v \in V$ such that $T(v) = 0$ for all $T \in L$.*

Proof. We proceed by induction on $\dim L$. The base case $\dim L = 1$ is clear since then $L = \operatorname{Span}(T)$ with $T\colon V \to V$ nilpotent and T has an eigenvector v, so $T(v) = 0$.

Now, let L be as in the statement. Let A be a maximal Lie subalgebra of L, i.e. if $A \le B < L$ with B a subalgebra, then $A = B$.

We claim that A is an ideal of L; in fact, there exists $y \in L \setminus A$ such that
$$L = A \oplus \operatorname{Span}(y)$$
and $[A, y] \subseteq A$. To show this, let $\overline{L} = \frac{L}{A}$ be the quotient vector space (not a Lie algebra). Define $\varphi\colon A \to \mathfrak{gl}(\overline{L})$ by
$$\varphi(a)(x + A) = [ax] + A \text{ for all } a \in A, x \in L.$$

We check that $\varphi(a)$ is well-defined: if $x + A = x' + A$ then $x - x' \in A$, so $[a(x - x')] \in A$, so $[ax] + A = [ax'] + A$.

We check that φ is a Lie homomorphism: φ is linear, and for $a, b \in A$ we have:
$$\begin{aligned}
[\varphi(a), \varphi(b)](x + A) &= (\varphi(a)\varphi(b) - \varphi(b)\varphi(a))(x + A) \\
&= [a[bx]] - [b[ax]] + A \\
&= [[ab]x] + A \qquad \text{by the Jacobi identity} \\
&= \varphi([ab])(x + A)
\end{aligned}$$

[2] Formally, we can prove this by induction on m.

Consider the image $\varphi(A)$, a Lie subalgebra of $\mathfrak{gl}(\overline{L})$. For $a \in A$, $a\colon V \to V$ is nilpotent, and hence $\operatorname{ad} a$ is nilpotent by Lemma 4.2, and therefore $\varphi(a) \in \mathfrak{gl}(\overline{L})$ is also nilpotent. Thus $\varphi(A)$ is a Lie subalgebra of $\mathfrak{gl}(\overline{L})$, of dimension less than $\dim L$ (as $\dim A < \dim L$), and consisting of nilpotent elements. So we can apply the inductive hypothesis to $\varphi(A)$: there exists a non-zero $y + A \in \overline{L}$ such that

$$[ay] + A = \varphi(a)(y + A) = A \text{ for all } a \in A$$

i.e.

$$[ay] \in A \text{ for all } a \in A.$$

Let $A_1 = A \oplus \operatorname{Span}(y)$. This is a subalgebra of L (as $[A, y] \subseteq A$), so as A is a maximal subalgebra of L, $L = A_1$. We have shown that

$$L = A \oplus \operatorname{Span}(y).$$

Also, $[A, y] \subseteq A$ implies that A is an ideal of L.

By the inductive hypothesis applied to A, there exists $0 \neq w \in V$ such that $a(w) = 0$ for all $a \in A$. Let

$$W = \{v \in V : a(v) = 0 \text{ for all } a \in A\}.$$

Then $w \in W$, so $W \neq 0$. By Lemma 4.3, W is L-invariant. In particular, $y(W) \subseteq W$. Now, $y\colon V \to V$ is nilpotent, and restricts to a nilpotent map $W \to W$. Therefore, there exists $0 \neq v \in W$ such that $y(v) = 0$. We can write any $T \in L$ as $T = a + \lambda y$ for some $a \in A$, $\lambda \in F$, and

$$T(v) = (a + \lambda y)(v) = 0.$$

Hence $T(v) = 0$ for any $T \in L$. □

Proof of Engel's Theorem 4.1. Let L be a Lie subalgebra of $\mathfrak{gl}(V)$ consisting of nilpotent elements. We aim to show that there exists a basis B such that $\{[T]_B : T \in L\} \subseteq \mathfrak{u}(n, F)$. Let $n = \dim V$. We proceed by induction on n. For $n = 1$, the claim is clear. By Proposition 4.4, there exists $0 \neq v \in V$ such that $T(v) = 0$ for all $T \in L$. Let $U = \operatorname{Span}(v)$ and $\overline{V} = \frac{V}{U}$. Any $T \in L$ induces $\overline{T}\colon \overline{V} \to \overline{V}$ by

$$\overline{T}(x + U) = T(x) + U \text{ for all } x \in V.$$

Moreover, since T is nilpotent, $\overline{T}$ is nilpotent. The map $T \mapsto \overline{T}$ from $L \to \mathfrak{gl}(\overline{V})$ is a Lie homomorphism, and its image is a Lie subalgebra of $\mathfrak{gl}(\overline{V})$ consisting of nilpotent elements. As $\dim \overline{V} = n - 1$, we can apply the inductive hypothesis to obtain a basis

$$\overline{B} = \{v_1 + U, \ldots, v_{n-1} + U\}$$

of $\overline{V}$ such that

$$\left\{ [\overline{T}]_{\overline{B}} : T \in L \right\} \subseteq \mathfrak{u}(n-1, F).$$

Then

$$B = \{v, v_1, \ldots, v_{n-1}\}$$

is a basis of V, and

$$\{[T]_B : T \in L\} \subseteq \mathfrak{u}(n, F),$$

which completes the proof. □

Engel's Theorem 4.1 is about subalgebras of $\mathfrak{gl}(V)$. There is a second version of Engel's theorem that is about *abstract* Lie algebras.

Theorem 4.5 (Engel's Theorem, version II). *A (finite-dimensional) Lie algebra L is* nilpotent *if and only if* $\operatorname{ad} x\colon L \to L$ *is nilpotent for all $x \in L$.*

Proof. For the *only if* implication, recall that L nilpotent means $L^m = 0$ for some m, so

$$[x_0[x_1[\ldots[x_{m-2},[x_{m-1}x_m]]]\ldots]] = 0$$

for all $x_i \in L$. Hencee $(\operatorname{ad} x)^m = 0$.

For the *if* implication, suppose that $\operatorname{ad} x\colon L \to L$ is nilpotent for all $x \in L$. Recall that $\operatorname{ad}\colon L \to \mathfrak{gl}(L)$ is a Lie homomorphism, so its image $\overline{L} = \operatorname{ad}(L)$ is a Lie subalgebra of $\mathfrak{gl}(L)$ consisting of nilpotent elements $\operatorname{ad} x$. By Engel's Theorem 4.1, $\overline{L}$ is nilpotent. Also, $\operatorname{Ker}(\operatorname{ad}) = Z(L)$, so

$$\overline{L} \cong \frac{L}{Z(L)}.$$

Suppose $\overline{L}^m = 0$. Then $0 = \overline{L}^m = \frac{L^m + Z(L)}{Z(L)}$, hence $L^m \leq Z(L)$. So $L^{m+1} = [L, L^m] \subseteq [L, Z(L)] = 0$. Hence L is nilpotent. ☐

5. Lie's Theorem

Recall that $\mathfrak{t}(n, F)$ is the Lie algebra of upper-triangular matrices, and is a soluble Lie algebra (Sheet 1, Question 14).

Theorem 5.1 (Lie's Theorem). *Let V be an n-dimensional vector space over $\mathbb{C}$ and let L be a soluble Lie subalgebra of $\mathfrak{gl}(V)$. Then there exists a basis B of V such that*

$$L \cong \{[T]_B : T \in L\} \subseteq \mathfrak{t}(n, \mathbb{C}).$$

Notes.

1. The theorem is false for fields of prime characteristic (Sheet 2).
2. When $\dim L = 1$, $L = \operatorname{Span}(T)$, then the theorem says that there exists a basis B such that $[T]_B$ is upper-triangular. This is easily proved by induction on $\dim V$:
   (a) There exists an eigenvector $0 \neq v \in V$ for T (since the field is $\mathbb{C}$).
   (b) Let $U = \operatorname{Span}(v)$. Then T induces $\overline{T}\colon \frac{V}{U} \to \frac{V}{U}$ by $\overline{T}(v + U) = T(v) + U$ and by the inductive hypothesis, there exists a basis $\overline{B} = (v_1 + U, \ldots, v_n + U)$ of $\frac{V}{U}$ such that $\left[\overline{T}\right]_{\overline{B}}$ is upper-triangular. Then $B = \{v, v_1, \ldots, v_{n-1}\}$ is a basis of V, and $[T]_B$ is upper-triangular.

   As for the proof of Engel's Theorem 4.1, the main part of the proof of Lie's Theorem 5.1 is Step (a): finding a common eigenvector for all T in the Lie algebra L.

Weight spaces. Let V be a vector space over F. Let A be a Lie subalgebra of $\mathfrak{gl}(V)$, and suppose there exists $0 \neq v \in V$ such that v is a common eigenvector for all $a \in A$. So

$$a(v) = \lambda(a)v, \text{ for all } a \in A,$$

where $\lambda(a) \in F$. The function $\lambda\colon A \to F$ is linear, since

$$\lambda(\alpha a + \beta b)v = (\alpha a + \beta b)(v) = \alpha a(v) + \beta b(v) = (\alpha\lambda(a) + \beta\lambda(b))v \quad \text{for } a, b \in A,\ \alpha, \beta \in F.$$

Definition. A *weight* for a Lie subalgebra A of $\mathfrak{gl}(V)$ is a linear map $\lambda\colon A \to F$ such that

$$V_\lambda = \{v \in V : a(v) = \lambda(a)v \text{ for all } a \in A\} \neq 0.$$

Such a non-zero subspace V_λ is the *weight space* for the weight λ.

Example. Let $L = d(n, F)$ be the set of diagonal matrices. Let $e_1, \ldots, e_n$ be the standard basis of F^n. Then for

$$x = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \in L,$$

we have $x(e_i) = \alpha_i e_i$. So $\mathrm{Span}(e_i)$ is a weight space for L, with weight $\lambda_i\colon L \to F$, where

$$\lambda_i \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} = \alpha_i$$

Proposition 5.2. *Let V be a finite-dimensional vector space over F, where* $\mathrm{char}(F) = 0$. *Let L be a Lie subalgebra of $\mathfrak{gl}(V)$ and A be an ideal of L. Suppose $\lambda\colon A \to F$ is a weight for A, with weight space*

$$V_\lambda = \{v \in V : a(v) = \lambda(a)v \text{ for all } a \in A\} \neq 0.$$

Then V_λ is L-invariant (i.e. $l(V_\lambda) \subseteq V_\lambda$ for all $l \in L$).

Note. If λ is the zero weight ($\lambda(a) = 0$ for all $a \in A$), Proposition 5.2 is just Lemma 4.3.

Proof of Proposition 5.2. Let $y \in L$, $w \in V_\lambda$. We want to show that $y(w) \in V_\lambda$, i.e.

$$ay(w) = \lambda(a)y(w).$$

Now, $[ay] \in A$, since A is an ideal, and $ay = ya + [ay]$, so

$$ay(w) = (ya + [ay])(w) = \lambda(a)y(w) + [ay](w) = \lambda(a)y(w) + \lambda([ay])w.$$

We need to prove

$$\lambda([ay]) = 0. \tag{1}$$

Let

$$U = \mathrm{Span}(w, y(w), y^2(w), \ldots).$$

As U is finite-dimensional, we can choose the minimal m such that w, $y(w)$, $\ldots$, $y^m(w)$ are linearly dependent. Then

$$B = \{w, y(w), \ldots, y^{m-1}(w)\}$$

is a basis of U. We claim that if $z \in A$ then

(i) $z(U) \subseteq U$
(ii) if $z_{|U}\colon U \to U$ is the restriction of z to U then

$$[z_U]_B = \begin{pmatrix} \lambda(z) & & * \\ & \ddots & \\ 0 & & \lambda(z) \end{pmatrix}$$

We need to show that for $0 \leq r \leq m-1$,

$$zy^r(w) = \lambda(z)y^r(w) + \sum_{i \leq r-1} \gamma_i y^i(w) \qquad (\text{for } \gamma_i \in F). \tag{2}$$

We prove this by induction on r. This is true for $r = 0$, as $z(w) = \lambda(z)w$ (because $z \in A$). For the inductive step, assume that the statement is true for $r-1$. Observe

$$zy^r(w) = zyy^{r-1}(w) = (yz + [zy])y^{r-1}(w). \tag{3}$$

By the inductive hypothesis,

$$zy^{r-1}(w) = \lambda(z)y^{r-1}(w) + \sum_{i \leq r-2} \gamma_i y^i(w),$$

so

$$yzy^{r-1}(w) = \lambda(z)y^r(w) + \sum_{i \leq r-1} \gamma_i y^i(w). \tag{4}$$

Also $[zy] \in A$, as $z \in A$ (because A is an ideal), so by the inductive hypothesis

$$[zy]y^{r-1}(w) = \sum_{i \leq r-1} \delta_i y^i(w). \tag{5}$$

Since, by equation (3), we have $zy^r = yzy^{r-1}(w) + [zy]y^{r-1}(w)$, we add equations (4) and (5) together to obtain the required equality (2). This completes the induction, and hence proves (i) and (ii).

Let $z = [ay] \in A$. By (i), $z(U) \subseteq (U)$, and by (ii)

$$\mathrm{Tr}(z_U) = m\lambda(z) \quad (m = \dim U).$$

Also,

$$z = [ay] = ay - ya,$$

and $a(U) \subseteq U$ by (i), $y(U) \subseteq U$ by definition of U. Thus the restrictions $a_{|U}, y_{|U} \colon U \to U$ exist, and

$$z_{|U} = a_{|U}y_{|U} - y_{|U}a_{|U}.$$

Hence

$$\mathrm{Tr}(z_{|U}) = \mathrm{Tr}(a_{|U}y_{|U}) - \mathrm{Tr}(y_{|U}a_{|U}) = 0.$$

Therefore

$$m\lambda(z) = 0.$$

Since $\mathrm{char}(F) = 0$, $m \neq 0$ in F, and hence

$$\lambda(z) = \lambda([ay]) = 0.$$

This proves equation (1), completing the proof. □

Proposition 5.3. *Let V be a finite-dimensional vector space over $\mathbb{C}$. Let L be a soluble Lie subalgebra of $\mathfrak{gl}(V)$. Then there exists $0 \neq v \in V$ such that v is a common eigenvector of all $x \in L$ (i.e. $x(v) \in \mathrm{Span}(v)$ for all $x \in L$).*

Proof. We proceed by induction on $\dim L$. The statement is true for $\dim L = 1$, i.e. $L = \mathrm{Span}(T)$, as T has an eigenvector (since the field is $\mathbb{C}$).

Now, assume that $\dim L > 1$. As L is soluble, $L' \subset L$. Choose a subspace A of L such that

$$L' \subseteq A \text{ and } \dim A = \dim L - 1.$$

Let $z \in L \setminus A$, so

$$L = A \oplus \mathrm{Span}(z) \quad \text{(as vector spaces).}$$

Then $[AA] \subseteq L' \subseteq A$, and $[A, z] \subseteq L' \subseteq A$, so A is an ideal of L. By the inductive hypothesis, there exists $0 \neq w \in V$ such that w is a common eigenvector for all $a \in A$. So

$$a(w) = \lambda(a)w \text{ for all } a \in A.$$

Let

$$V_\lambda = \{v \in V : a(v) = \lambda(a)v \text{ for all } a \in A\}.$$

Then $w \in V_\lambda$, so $V_\lambda \neq 0$, and hence λ is a weight of A. By Proposition 5.2, V_λ is L-invariant, so

$$z(V_\lambda) \subseteq V_\lambda.$$

Therefore, there exists $0 \neq v \in V_\lambda$ such that v is an eigenvector for z; say $z(v) = \beta v$. Then for any $a + \alpha z \in L$ where $a \in A$, $\alpha \in \mathbb{C}$, we have

$$(a + \alpha z)(v) = \lambda(a)v + \alpha\beta v.$$

Hence v is a common eigenvector for all $x \in L$. □

The proof of Lie's Theorem 5.1 is completed using Proposition 5.3 and induction on $\dim V$, in the same way as the completion of Engel's Theorem 4.1. (Sheet 2, Question 4)

6. Representations

Definition. Let L be a Lie algebra over F. A finite-dimensional *representation* of L is a Lie homomorphism

$$\rho\colon L \to \mathfrak{gl}(V)$$

where V is a finite-dimensional vector space over F.

A representation ρ is *faithful* if $\mathrm{Ker}\,\rho = 0$, in which case $L \cong \mathrm{Im}(\rho)$, a Lie subalgebra of $\mathfrak{gl}(V)$.

If we fix a basis B of V, the map

$$x \mapsto [\rho(x)]_B \quad (x \in L)$$

from L to $\mathfrak{gl}(n, F)$ (where $n = \dim V$) is a *matrix representation* of L.

Examples.

(1) If L is any Lie algebra, the map

$$\mathrm{ad}\colon L \to \mathfrak{gl}(L)$$

sending x to $\mathrm{ad}\,x$ is a representation, called the *adjoint representation* of L. Its kernel is $Z(L)$.

(2) Let $L = \mathfrak{sl}(2, F)$ with basis B

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

and structure constants $[ef] = h$, $[eh] = -2e$, $[fh] = 2f$. With respect to this basis, the adjoint representation of L gives representation of L, sending

$$e \mapsto (\operatorname{ad} e)_B = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$h \mapsto (\operatorname{ad} h)_B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

$$f \mapsto (\operatorname{ad} f)_B = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

Now, $Z(L) = 0$, provided $\operatorname{char}(F) \neq 2$, so L is isomorphic to the Lie subalgebra of $\mathfrak{gl}(3, F)$, spanned by these 3 matrices.

(3) Let L be the 2-dimensional non-abelian Lie algebra over F, with basis x, y and structure constant $[xy] = x$. Define $\rho\colon L \to \mathfrak{gl}(2, F)$ by

$$\rho(x) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; \rho(y) = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$$

and extend it linearly to L, i.e. for all $\alpha, \beta \in F$

$$\rho(\alpha x + \beta y) = \begin{pmatrix} -\beta & \alpha + \beta \\ 0 & 0 \end{pmatrix}.$$

One can easily check that ρ is a representation of L. (It suffices to check that $[\rho(x), \rho(y)] = \rho[xy] = \rho(x)$, which is clear when we multiply out the matrices.)

(4) If L is a Lie subalgebra of $\mathfrak{gl}(V)$ for some vector space V, the inclusion map sending $l \mapsto l$ for all $l \in L$ is a representation of L, called the *natural representation* of L.

(5) For any Lie algebra L over F, the zero map $L \to F$ sending $l \mapsto 0$ for all $l \in L$ is the *trivial* (1-dimensional) representation of L.

Modules. If $\rho\colon L \to \mathfrak{gl}(V)$ is a representation, it is notationally convenient to *drop* the ρ and instead of $\rho(l)v$ write lv. This defines a map $L \times V \to V$ sending $(l, v) \mapsto lv (= \rho(l)(v))$, satisfying

(1) $(l_1 + l_2)v = l_1 v + l_2 v$,
(2) $l(v_1 + v_2) = lv_1 + lv_2$,
(3) $\lambda(lv) = l(\lambda v) = (\lambda l)v$,
(4) $[l_1 l_2]v = l_1(l_2 v) - l_2(l_1 v)$,

for all $l_i, l \in L$, $v_i, v \in V$, $\lambda \in F$. All these statements are clear. For example, for (1), we check that

$$(l_1 + l_2)v = \rho(l_1 + l_2)(v) = (\rho(l_1) + \rho(l_2))(v) = \rho(l_1)(v) + \rho(l_2)(v) = l_1 v + l_2 v,$$

and for (4), we note that $\rho[l_1 l_2] = [\rho(l_1), \rho(l_2)]$ implies that

$$[l_1 l_2]v = \rho([l_1 l_2])(v) = \rho(l_1)\rho(l_2)(v) - \rho(l_2)\rho(l_1)(v) = l_1(l_2 v) - l_2(l_1 v).$$

Definition. Let L be a Lie algebra over F. A vector space V over F is an *L-module* if there exists a map $L \times V \to V$ sending $(l, v) \mapsto lv$, satisfying (1)–(4) above.

As we have seen, any representation of L gives an L-module. In fact, L-modules and representations of L are equivalenct concepts:

(a) Given a representation $\rho \colon L \to \mathfrak{gl}(V)$, defining $lv = \rho(l)(v)$ for $l \in L$, $v \in V$ makes V an L-module.

(b) Given an L-module V, we can define $\rho \colon L \to \mathfrak{gl}(V)$ by

$$\rho(l)(v) = lv \text{ for all } l \in L,\ v \in V.$$

Then (1)–(4) imply that ρ is a Lie homomorphism, hence a representation of L.

Examples.

(1) Let L be a 2-dimensional Lie algebra, with basis x, y and $[xy] = x$. We have a matrix representation $\rho \colon L \to \mathfrak{gl}(2, F)$ sending

$$x \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ y \mapsto \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$$

The corresponding L-module is $V = F^2$ with standard basis e_1, e_2 and multiplication

$$xe_1 = 0,\ xe_2 = e_1,$$

$$ye_1 = -e_1,\ ye_2 = e_1.$$

(2) Let L be any Lie algebra. We have the *adjoint representation* $\mathrm{ad} \colon L \to \mathfrak{gl}(L)$. The corresponding L-module, the *adjoint module* is L itself, with multiplication

$$lv = (\mathrm{ad}\, l)(v) = [lv] \text{ for all } l \in L, v \in L.$$

Submodules.

Definition. Let L be a Lie algebra, V an L-module. A subspace W of V is a *submodule* if W is L-invariant, i.e. $lW \subseteq W$ for all $l \in L$ (where $lW = \{lw : w \in W\}$).

Examples.

(1) Let L be the 2-dimensional Lie algebra with basis x, y and module $V = F^2$ as above. Since $xe_1 = 0$, $ye_1 = -1$, $W = \mathrm{Span}(e_1)$ is a submodule.

(2) Let L be a Lie algebra. Then W is a submodule of the adjoint module L if and only if $LW \subseteq W$, i.e. $[LW] \subseteq W$, which is equivalent to W being an ideal of L.

Proposition 6.1. *Let L be a soluble (finite-dimensional) Lie algebra over $\mathbb{C}$, and let V be a (finite-dimensional) non-zero L-module. Then V has a 1-dimensional submodule.*

In example (1) above, $\mathrm{Span}(e_1)$ was such a submodule.

Proof. Let $\rho\colon L \to \mathfrak{gl}(V)$ be the corresponding representation of L (so $\rho(l)(v) = lv$ for all $l \in L$, $v \in V$). Then

$$\frac{L}{\ker\rho} \cong \operatorname{Im}\rho \subseteq \mathfrak{gl}(V).$$

Now, $\operatorname{Im}\rho$ is a soluble Lie subalgebra of $\mathfrak{gl}(V)$ by Proposition 3.4, so by Proposition 5.3, there exists a common eigenvector $0 \neq v \in V$ for all $\rho(l) \in \operatorname{Im}\rho$. So

$$lv = \rho(l)(v) \in \operatorname{Span}(v) \text{ for all } l \in L.$$

Hence $\operatorname{Span}(v)$ is a submodule of dimension 1. □

Irreducible modules.

Definition. An L-module is *irreducible* if it is nonzero, and the only submodules are 0 and V. We then also say that the corresponding representation $\rho\colon L \to \mathfrak{gl}(V)$ is *irreducible*.

Examples.

(1) Any 1-dimensional L-module is irreducible.
(2) If L is the adjoint module, we have seen that the submodules are the ideals of L. So the adjoint L-module is irreducible if and only if L is a simple Lie algebra (or $\dim L = 1$). For example, the adjoint representation $\mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(3,\mathbb{C})$ is irreducible.

Proposition 6.2. *Let L be a soluble Lie algebra over $\mathbb{C}$. Then every irreducible L-module is 1-dimensional.*

Proof. This is immediate from Proposition 6.1. □

Quotient modules. Let L be a Lie algebra, V and L-module, and W a submodule. We can make $\frac{V}{W}$ into an L-module, called the *quotient module*, by defining

$$l(v + W) = lv + W \text{ for } l \in L, v \in V.$$

We check that the multiplication is well-defined; indeed, if $v_1+W = v_2+W$ then $v_1-v_2 \in W$, so $lv_1-lv_2 = l(v_1-v_2) \in W$ since W is a submodule, and hence $lv_1+W = lv_2+W$. Moreover, it is easy to check that V/W satisfies properties (1)–(4), and hence is well-defined.

Composition series. Let L be a Lie algebra and V a (finite-dimensional) nonzero L-module. Choose a non-zero submodule $V_1 \subset V$ of minimal dimension. Then V_1 is irreducible. Form the quotient $\frac{V}{V_1} \neq 0$. If $\frac{V}{V_1} \neq 0$, find a submodule $\frac{V_2}{V_1}$ of minimal dimension, where $V_2 \subset V$ is a submodule $V_1 \subset V_2$. Then $\frac{V_2}{V_1}$ is irreducible. Continuing, the sequence

$$0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_r = V,$$

where $\frac{V_i}{V_{i-1}}$ are irreducible, terminates, because V is finite-dimensional. This is a *composition series* of V.

Direct sums. Let V be an L-module. Suppose $V = U_1 \oplus \cdots \oplus U_r$, i.e. $V = U_1 + \cdots + U_r$ and $U_i \cap \left(\sum_{j \neq i} U_j\right) = 0$ for all i, where each U_i is a submodule. Then we say V is the *direct sum* of the submodules $U_1, \ldots, U_r$. If $V = U_1 \oplus \cdots \oplus U_r$ with every U_i irreducible, we call V a *completely reducible* L-module.

Examples.

(1) Let $L = d(2, F) = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} : \alpha, \beta \in F \right\}$, an abelian Lie algebra. Let $V = F^2$, the natural L-module with standard basis e_1, e_2. Then $V_1 = \mathrm{Span}(e_1)$, $V_2 = \mathrm{Span}(e_2)$ are submodules, as
$$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} e_1 = \alpha e_1, \ \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} e_2 = \beta e_2,$$
and $V = V_1 \oplus V_2$. Each V_i is irreducible, so V is completely reducible.

(2) Let $L = t(2, F)$, $V = F^2$ be the natural L-module. The only nontrivial (not 0 and not V) submodule of V is $\mathrm{Span}(e_1)$, so V is not completely reducible.

Homomorphisms.

Definition. Let L be a Lie algebra, V and W be L-modules. A linear map $\varphi\colon V \to W$ is a *homomorphism* if
$$\varphi(lv) = l\varphi(v) \text{ for all } l \in L, v \in V.$$
An *isomorphism* is a bijective homomorphism.

We want to interpret the notion of homomorphism in terms of representations. Let $\rho_V\colon L \to \mathfrak{gl}(V)$ and $\rho_W\colon L \to \mathfrak{gl}(V)$ be the corresponding representations of L. Then φ is a homomorphism if
$$\varphi(\rho_V(l)(v)) = \rho_W(l)(\varphi(v)) \text{ for all } l \in L, v \in V$$
i.e. $\varphi \circ \rho_V(l) = \rho_W(l) \circ \varphi$ for all $l \in L$. If φ is an isomorphism, this says
$$\rho_W(l) = \varphi \circ \rho_V(L) \circ \varphi^{-1}.$$
So we can choose bases B_V of V and B_W of W such that the corresponding matrices are equal:
$$[\rho_V(l)]_{B_v} = [\rho_W(l)]_{B_w} \text{ for all } l \in L,$$
i.e. the matrix representations
$$l \mapsto [\rho_V(l)]_{B_V}$$
$$l \mapsto [\rho_W(l)]_{B_W}$$
are **identical**.

This is a way to tell whether two L-modules are isomorphic.

Example. Let $L = \mathrm{Span}(x)$, a 1-dimensional Lie algebra. Then L has matrix representations
$$\rho_1\colon x \mapsto \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \ \rho_2\colon x \mapsto \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}.$$
The corresponding L-modules are isomorphic as the matrices $\rho_1(x)$, $\rho_2(x)$ are similar.

7. Representations of $\mathfrak{sl}(2,\mathbb{C})$

Let $L = \mathfrak{sl}(2,\mathbb{C})$, with basis e, f, h and structure constants

$$[ef] = h,\ [he] = 2e,\ [hf] = -2f.$$

We will classify all the irreducible L-modules (up to isomorphism), showing there is exactly one of each dimension.

Construction. Let $\mathbb{C}[X,Y]$ be the vector space of all polynomials in X, Y. For each $d \geq 0$, define

$$V_d = \mathrm{Span}(X^d, X^{d-1}Y, \dots, XY^{d-1}, Y^d),$$

the subspace of all *homogeneous* polynomials of degree d. So $\dim V_d = d+1$. Define

$$\varphi\colon L \to \mathfrak{gl}(V_d)$$

by

$$\varphi(e) = X\frac{\partial}{\partial Y},\ \varphi(f) = Y\frac{\partial}{\partial X},\ \varphi(h) = X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}.$$

So, explicitly

$$\begin{array}{ll} \varphi(e)\colon & X^aY^b \mapsto bX^{a+1}Y^{b-1}\ (b \geq 1) \\ & X^d \mapsto 0 \\[1ex] \varphi(f)\colon & X^aY^b \mapsto aX^{a-1}Y^{b+1}\ (a \geq 1) \\ & Y^d \mapsto 0 \\[1ex] \varphi(h)\colon & X^aY^b \mapsto (a-b)X^aY^b \end{array}$$

and extend φ to a linear map $L \to \mathfrak{gl}(V)$.

Proposition 7.1. *The map φ is a representation of $L = \mathfrak{sl}(2,\mathbb{C})$. Hence V_d is an L-module of dimension $d+1$.*

Proof. We need to check:

(1) $[\varphi(h)\varphi(e)] = \varphi([he]) = 2\varphi(e)$,
(2) $[\varphi(h)\varphi(f)] = -2\varphi(f)$,
(3) $[\varphi(e)\varphi(f)] = \varphi(e)$.

We check (1). For $b \geq 1$, we have

$$\begin{aligned} [\varphi(h)\varphi(e)](X^aY^b) &= \varphi(h)(bX^{a+1}Y^{b-1}) - \varphi(e)((a-b)X^aY^b) \\ &= b(a-b+1)X^{a+1}Y^{b-1} - (a-b)bX^{a+1}Y^{b-1} \\ &= 2bX^{a+1}Y^{b-1} \\ &= 2\varphi(e)(X^aY^b) \end{aligned}$$

and

$$[\varphi(h)\varphi(e)](X^d) = \varphi(h)(0) - \varphi(e)(dX^d) = 0 = 2\varphi(e)(X^d).$$

Hence (1) is verified. Similarly, one can check (2) and (3). (Exercise.) $\square$

Matrix representations. Let $B = \{X^d, X^{d-1}Y, \ldots, Y^d\}$. The corresponding matrix representation is

$$e \mapsto [\varphi(e)]_B = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 2 & & \\ & & \ddots & \ddots & \\ & & & 0 & d \\ & & & & 0 \end{pmatrix}$$

$$f \mapsto [\varphi(f)]_B = \begin{pmatrix} 0 & & & & \\ d & 0 & & & \\ & d-1 & \ddots & & \\ & & \ddots & & \\ & & & 1 & 0 \end{pmatrix}$$

$$h \mapsto [\varphi(h)]_B = \begin{pmatrix} d & & & & 0 \\ & d-2 & & & \\ & & \ddots & & \\ & & & -(d-2) & \\ 0 & & & & -d \end{pmatrix}$$

Examples. The module V_0 is the trivial representation $e, f, g \mapsto 0$. The module V_1 is given by

$$e \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ f \mapsto \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \ h \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and it is equal to the natural L-module. The module V_2 is given by:

$$e \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$h \mapsto \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

$$f \mapsto \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and is isomorphic to the adjoint L-module. (Exhibiting the isomorphism is an exercise: Sheet 3, Question 6.)

Proposition 7.2. *The module V_d is irreducible.*

Proof. Recall that

$$[\varphi(h)]_B = \begin{pmatrix} d & & & & 0 \\ & d-2 & & & \\ & & \ddots & & \\ & & & -(d-2) & \\ 0 & & & & -d \end{pmatrix}.$$

Suppose U is a non-zero submodule of V_d. Then $\varphi(e)$, $\varphi(f)$, $\varphi(h)$ send U to U. Since $\varphi(h)$ has $d+1$ distinct eigenvalues, the eigenvalues of the restriction $\varphi(h)_{|U}\colon U \to U$ are also distinct, and U contains an eigenvector for $\varphi(h)$. The eigenspaces of $\varphi(h)$ are the 1-dimensional spaces spanned by the basis vectors in B. Hence

$$X^aY^b \in U \text{ for some } a, b.$$

Apply $\varphi(e)$ to this successively to get

$$X^{a+1}Y^{b-1},\ X^{a+2}Y^{b-2},\ \ldots,\ X^d \in U,$$

and apply $\varphi(f)$ to this successively to get

$$X^{a-1}Y^{b+1},\ X^{a-2}Y^{b+2},\ \ldots,\ Y^d \in U.$$

Hence $U = V_d$. This shows V_d is irreducible. □

Theorem 7.3. *If V is an irreducible finite-dimensional L-module, then $V \cong V_d$ for some d.*

We first prove a lemma.

Notation. For an L-module V, write $e(ev) = e^2v$, $e(e^2v) = e^3v$, and so on, $e^kv = e(e^{k-1}v)$.

Lemma 7.4. *Let V be a finite-dimensional L-module.*

(1) *If $v \in V$ with $hv = \lambda v$ then*

$$h(ev) = (\lambda + 2)ev,\ h(fv) = (\lambda - 2)fv.$$

(2) *The L-module V contains an eigenvector $w \neq 0$ for h such that $ew = 0$.*

Proof. For (1), observe that

$$h(ev) = e(hv) + [he]v = e(\lambda v) + 2ev = (\lambda + 2)ev,$$

and similarly for $h(fv)$.

For (2), note that the linear map $v \mapsto hv$ has an eigenvector (since the field is $\mathbb{C}$). Say $hv = \lambda v$. Consider

$$v, ev, e^2v, \ldots.$$

If all of them are non-zero, by (1), these are eigenvectors for h with distinct eigenvalues, hence they are linearly independent. Hence, as V is finite-dimensional, there exists k such that $e^kv \neq 0$, but $e^{k+1}v = 0$. Put $w = e^kv$ to complete the proof. □

Proof of Theorem 7.3. Let V be an irreducible finite-dimensional L-module. By Lemma 7.4 (2), there exists $w \neq 0$ such that

$$hw = \lambda w,\ ew = 0.$$

By the proof of Lemma 7.4, there exists d such that

$$f^dw \neq 0,\ f^{d+1}w = 0.$$

(A) The elements $w, fw, \ldots, f^dw$ form a basis of V, consisting of h-eigenvectors with eigenvalues $\lambda, \lambda - 2, \ldots, \lambda - 2d$.

Indeed, by Lemma 7.4, these are eigenvectors for h with the given eigenvalues, hence they are linearly independent. To show they span V, we set

$$U = \mathrm{Span}(w, fw, \ldots, f^dw).$$

We show U is a submodule. Well, clearly

$$fU \subseteq U,$$

$$hU \subseteq U.$$

We will show that $ef^k w \in \mathrm{Span}(w, fw, \dots, f^{k-1}w)$ for all $k \leq d$ by induction on k. This is clearly true for $k = 0$, as $ew = 0$. Assume this is true for $k-1$. Then

$$\begin{aligned} ef^k w &= e(f(f^{k-1}w)) \\ &= (fe + [ef])(f^{k-1}w) \\ &= (fe + h)f^{k-1}w. \end{aligned}$$

By the inductive hypothesis, $ef^{k-1}w \in \mathrm{Span}(w, \dots, f^{k-2}w)$, and hence

$$fef^{k-1}w \in \mathrm{Span}(w, \dots, f^{k-1}w),$$

and so is $hf^{k-1}w$. This completes the induction and shows that $eU \subseteq U$. We have hence shown that U is a submodule of V and, as V is irreducible, we have $U = V$.

(B) If B is the basis in (A), then

$$[h]_B = \begin{pmatrix} \lambda & & & \\ & \lambda - 2 & & \\ & & \ddots & \\ & & & \lambda - 2d \end{pmatrix}.$$

Also, $h = [ef] \in L'$, so $\mathrm{Tr}[h]_B = 0$ (by Sheet 1, Question 4). Hence

$$\lambda + \lambda - 2 + \cdots + \lambda - 2d = 0.$$

Thus $(d+1)\lambda = d(d+1)$, which shows that $\lambda = d$.

(C) We have that $V \cong V_d$.

The L-module V has a basis $w, fw, \dots, f^d w$, and the L-module V_d has a basis $X^d, fX^d, \dots, f^d X^d$. Both bases consist of h-eigenvectors with eigenvalues $d, d-2, \dots, -d$. Define $\varphi \colon V \to V_d$ by

$$\varphi(f^k w) = f^k X^d \text{ for } 0 \leq k \leq d.$$

We show that φ is an isomorphism of L-modules. We need to show that

$$\varphi(lv) = l\varphi(v) \text{ for all } v \in V$$

for $l = e, f, h$. For f:

$$f\varphi(f^k w) = f(f^k X^d) = f^{k+1} X^d = \varphi(f^{k+1} w).$$

For h:

$$h\varphi(f^k w) = h(f^k X^d) = (d - 2k) f^k X^d = \varphi(h(f^k w)).$$

For e, we show that

$$e\varphi(f^k w) = \varphi(ef^k w)$$

by induction on k. For $k = 0$, we have

$$e\varphi(w) = eX^d = 0 = \varphi(ew)$$

as $ew = 0$. We assume the equation holds for $k-1$. Then

$$\begin{aligned}\varphi(ef^k w) &= \varphi((fe+h)f^k w)\\ &= f\varphi(ef^{k-1}w) + h\varphi(f^{k-1}w)\\ &= fe\varphi(f^{k-1}w) + h\varphi(f^{k-1}w) \quad \text{by the inductive hypothesis}\\ &= ef\varphi(f^{k-1}w)\\ &= e\varphi(f^k w)\end{aligned}$$

which completes the induction and shows that $V \cong V_d$.

This completes the proof. □

8. The Killing form

Let L be a Lie algebra over $\mathbb{C}$ and $\mathrm{ad}\colon L \to \mathfrak{gl}(L)$ be the adjoint representation. The *Killing form* is the map $K\colon L \times L \to \mathbb{C}$ given by

$$(x, y) \mapsto \mathrm{Tr}(\mathrm{ad}(x)\,\mathrm{ad}(y)).$$

Remark. This is a symmetric bilinear form. (Symmetric because $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, and bilinear because ad is linear.)

Jordan decomposition. If $f\colon V \to V$ is a linear transformation of a complex vector space V, then there is a basis B of V in which f is given by the matrix $[f]_B$ which is a direct sum of Jordan blocks:

$$\left(\begin{array}{c|c|c|c} J_{r_1}(\lambda_1) & 0 & 0 & 0 \\ \hline 0 & J_{r_2}(\lambda_2) & 0 & 0 \\ \hline 0 & 0 & \ddots & 0 \\ \hline 0 & 0 & 0 & J_{r_k}(\lambda_k) \end{array}\right)$$

where

$$J_r(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & & & 0 \\ & \lambda & 1 & 0 & & \\ & & \lambda & 1 & & \\ & & & \ddots & & \\ & & & & \lambda & 1 \\ 0 & & & & & \lambda \end{pmatrix}$$

Rewrite this matrix as $D + N$ where

$$D = \mathrm{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_k, \dots, \lambda_k)$$

and

$$N = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$

is strictly upper triangular. Clearly, N is a nilpotent matrix ($N^l = 0$ for some l). We have $DN = ND$.

The Jordan normal form theorem says that any linear transformation $f\colon V \to V$ is uniquely written as $f = d + n$ (the *Jordan decomposition*), where $d\colon V \to V$ is a diagonalisable linear transformation and $n\colon V \to V$ is a nilpotent linear transformation and such that $dn = nd$.

A *Jordan basis* B of V is a basis in which $[d]_B = D$ and $[n]_B = N$.

Lemma 8.1. *Let $x\colon V \to V$ be a linear transformation with Jordan decomposition $x = d+n$.*

(1) *Then there is a polynomial $p(X) \in \mathbb{C}[X]$ such that $p(x) = d$.*
(2) *Define a linear transformation $\overline{d}\colon V \to V$ by $[\overline{d}]_B = \overline{D} = \bigoplus_i \overline{\lambda_i} I_{r_i}$ (the complex conjugate of D). Then there is a polynomial $q(X) \in \mathbb{C}[X]$ such that $q(x) = \overline{d}$.*

Proof. Let $\lambda_1, \dots, \lambda_k$ be eigenvalues of the linear transformation x. Let a_i be the size of the largest Jordan block with eigenvalue λ_i. Let $V_i \subseteq V$ be the subspace $V_i = \operatorname{Ker}(x - \lambda_i I)^{a_i}$. In terms of Jordan basis, V_i is spanned by the basis elements associated to λ_i. Then $V = V_1 \oplus \cdots \oplus V_k$. Then d acts on V_i as $\lambda_i I_{v_i}$. Therefore, $d = \bigoplus_{i=1}^{k} \lambda_i I_{v_i}$.

Observe that $(X - \lambda_i)^{a_i}$, for $i = 1, \dots, k$, are pairwise coprime. The Chinese Remainder Theorem[3] says that the natural map

$$\mathbb{C}[X] \to \bigoplus_{i=1}^{k} \mathbb{C}[X] \big/ \big((X - \lambda_i)^{a_i}\big)$$

is surjective.

Hence, for $\lambda_1, \dots, \lambda_k$, we can find a polynomial $p(X)$ such that

$$p(X) \equiv \lambda_i \bmod (X - \lambda_i)^{a_i}, \text{ for } i = 1, 2, \dots, k.$$

Then we have that

$$p(X) = \lambda_i + \varphi_i(X)(X - \lambda_i)^{a_i}, \text{ for some } \varphi_i(X) \in \mathbb{C}[X],$$

and hence for $v \in V_i$ we have $p(x)v = \lambda_i v + \varphi(x)(x - \lambda_i I)^{a_i} v = \lambda_i v$. Therefore,

$$p(x) = \bigoplus_{i=1}^{k} \lambda_i I_{V_i} = d.$$

This proves (1).

To prove (2), define $q(X)$ to be the polynomial equivalent to $\overline{\lambda_i}$ modulo $(X - \lambda_i)^{a_i}$, for $i = 1, \dots, k$. The same argument now shows that $q(x) = \overline{d}$. □

Lemma 8.2. *Suppose that $x\colon V \to V$ is a linear transformation with Jordan decomposition $x = d + n$. Consider $\operatorname{ad}(x)\colon \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ (recall that $\operatorname{ad}(x)$ sends y to $[x, y] = xy - yx$.) Then $\operatorname{ad}(x)$ has Jordan decomposition $\operatorname{ad}(d) + \operatorname{ad}(n)$.*

[3] Recall the proof of the Chinese Remainder Theorem. For $k = 2$, if $f(X)$, $g(X)$ are coprime polynomials, then there exist polynomials $p(X)$ and $q(X)$ such that $1 = p(X)f(X) + q(X)g(X)$. Then the polynomial $b(X)p(X)f(X) + a(X)q(X)g(X)$ is congruent to $a(X)$ modulo $f(X)$ to $b(X)$ modulo $g(X)$. The general case $k \geq 2$ is done by induction.

Proof. By linearity of ad, we have $\mathrm{ad}(x) = \mathrm{ad}(d) + \mathrm{ad}(n)$. Next, by Lemma 4.2, $\mathrm{ad}(n)$ is nilpotent, because n is nilpotent. We claim that $\mathrm{ad}(d)$ is diagonalisable. Consider the basis in which d is diagonal. Recall that E_{ij} (the matrix with 1 in the ij entry and 0 everywhere else) is the standard basis of $\mathfrak{gl}(V)$. For this basis of $\mathfrak{gl}(V)$, $\mathrm{ad}(d)$ is diagonal. Finally, $[\mathrm{ad}(d), \mathrm{ad}(n)] = \mathrm{ad}([dn]) = 0$. By the Jordan canonical form theorem, $\mathrm{ad}(d)$ and $\mathrm{ad}(n)$ are the diagonalisable and the nilpotent parts of $\mathrm{ad}(x)$, respectively. □

Recall that a Lie algebra L is *soluble* if $L^{(N)} = 0$ for some N, where $L^{(i)}$ is the *derived series* of L, i.e. $L^{(0)} = L$, $L^{(1)} = [L, L]$, ..., $L^{(n+1)} = [L^{(n)}, L^{(n)}]$, and we have

$$L \supseteq L^{(1)} \supseteq \cdots \supseteq L^{(n)} \supseteq \cdots \supseteq L^{(N)} = 0.$$

Remark. We have that $\mathrm{ad}(L) = L/Z(L)$, where $Z(L)$ is the center of L. Since $Z(L)$ is abelian, it is always soluble, and hence L is soluble if and only if $\mathrm{ad}(L)$ is soluble.

Theorem 8.3 (Cartan's first criterion)**.** *Let L be a Lie algebra over $\mathbb{C}$. Then L is soluble if and only if $K(x, y) = 0$ for any $x \in L$, $y \in L'$.*

We will work towards proving this theorem.

Proposition 8.4. *Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $L \subseteq \mathfrak{gl}(V)$ be a soluble Lie subalgebra. Then $\mathrm{Tr}(xy) = 0$ for $x \in L$, $y \in L'$.*

Proof. By Lie's Theorem 5.1, there is a basis B of V such that $[x]_B$ is an upper-triangular matrix for all $x \in L$. Then $[x_1]_B[x_2]_B - [x_2]_B[x_1]_B$ is strictly upper-triangular for all $x_1, x_2 \in L$. Hence for any $y \in L'$, $[y]_B$ is strictly upper-triangular, hence $[x]_B[y]_B$ is strictly upper-triangular. Therefore, $\mathrm{Tr}(xy) = \mathrm{Tr}([x]_B[y]_B) = 0$. □

This allows us to prove on implication of Theorem 8.3: if L is soluble then $K(x, y) = 0$ for $x \in L$, $y \in L'$.

Proof of the "only if" implication in Theorem 8.3. Consider $\mathrm{ad}\colon L \to \mathfrak{gl}(L)$. Now, $\mathrm{ad}(L)$ is a Lie subalgebra of $\mathfrak{gl}(V)$. By the remark above, the solubility of L is equivalent to the solubility of $\mathrm{ad}(L)$. We are given that L is soluble, hence $\mathrm{ad}(L) \subseteq \mathfrak{gl}(V)$ is soluble. Thus $\mathrm{Tr}(\mathrm{ad}(x)\,\mathrm{ad}(y)) = 0$ for any $x \in L$ and any $y \in L'$ by Proposition 8.4. □

To prove the other implication of the theorem, we first prove a lemma.

Lemma 8.5. *Let V be a finite-dimensional vector space over $\mathbb{C}$. If $x, y, z\colon V \to V$ are linear maps, then $\mathrm{Tr}([xy]z) = \mathrm{Tr}(x[yz])$.*

Note that this implies that $K([xy], z) = K(x, [yz])$, which will be useful later on.

Proof. We have that $\mathrm{Tr}([xy]z) = \mathrm{Tr}(xyz - yxz) = \mathrm{Tr}(xyz) - \mathrm{Tr}(yxz)$ and $\mathrm{Tr}(x[yz]) = \mathrm{Tr}(xyz - xzy) = \mathrm{Tr}(xyz) - \mathrm{Tr}(xzy)$, so we only have to check that $\mathrm{Tr}(yxz) = \mathrm{Tr}(xzy)$. But this is clear, since $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$:

$$\mathrm{Tr}(yxz) = \mathrm{Tr}(y(xz)) = \mathrm{Tr}((xz)y) = \mathrm{Tr}(xzy),$$

which completes the proof. □

Proposition 8.6. *Let V be a finite-dimensional vector space over $\mathbb{C}$. Let L be a Lie subalgebra of $\mathfrak{gl}(V)$. Suppose $\mathrm{Tr}(xy) = 0$ for $x \in L$, $y \in L$. Then L is soluble.*

Proof. The idea is to show that every $x \in L'$ is nilpotent. Then Engel's Theorem 4.1 implies that L' is a nilpotent algebra. Hence L' is soluble, therefore L is soluble.

Let $x \in L'$. Consider the Jordan decomposition $x = d + n$. There exists a basis B of V with respect to which $d = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ and $[n]_B$ is strictly upper-triangular. We need to prove that $d = 0$, i.e. $\lambda_i = 0$ for all i. This will follow if we show that $\sum_{i=1}^{n} \lambda_i \overline{\lambda_i} = 0$. Define $\overline{d} \colon V \to V$ by $[\overline{d}]_B = \mathrm{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$. Let us note that

$$\mathrm{Tr}(\overline{d}x) = \sum_{i=1}^{n} \lambda_i \overline{\lambda_i}.$$

It is enough to prove that

$$\mathrm{Tr}(\overline{d}[yz]) = 0$$

for any $y, z \in L$, because $x \in L' = \mathrm{Span}\{[yz] : y, z \in L\}$. By Lemma 8.5,

$$\mathrm{Tr}(\overline{d}[yz]) = \mathrm{Tr}([\overline{d}y]z).$$

By Lemma 8.2, the Jordan decomposition of $\mathrm{ad}(x)$ is $\mathrm{ad}(d) + \mathrm{ad}(n)$. By Lemma 8.1 (2), there is a polynomial $q(X) \in \mathbb{C}[X]$ such that $\overline{\mathrm{ad}(d)} = q(\mathrm{ad}(x))$. But it is clear that $\overline{\mathrm{ad}(d)} = \mathrm{ad}(\overline{d})$. Therefore, $\mathrm{ad}(\overline{d})$ maps L to L (because this is a polynomial in $\mathrm{ad}(x)$ and $\mathrm{ad}(x) \colon L \to L$). In particular, $\mathrm{ad}(\overline{d})(y) = [\overline{d}y] \in L$. Since $\mathrm{Tr}(xy) = 0$ for any $x, y \in L$, we conclude that $\mathrm{Tr}([\overline{d}y]z) = 0$. We have seen above that this implies $\lambda_1 = \dots = \lambda_n = 0$, so that x is nilpotent. □

Finally, we complete the proof of Theorem 8.3.

Proof of the "if" implication in Theorem 8.3. We are given that $\mathrm{Tr}(\mathrm{ad}(x)\,\mathrm{ad}(y)) = 0$ for any $x \in L$, $y \in L'$. By Proposition 8.6, we obtain that $L^{(2)}$ is soluble. But this implies that L' is soluble so L is soluble. □

A digression into linear algebra. Let V be a vector space over $\mathbb{C}$ of dimension $\dim V = n$ ($\mathbb{C}$ can be replaced by any field). Define the *dual vector space* V^* as the set of linear maps $\alpha \colon V \to \mathbb{C}$. This a vector space over $\mathbb{C}$. If $v_1, \dots, v_n$ is a basis of V, then there is a natural basis $f_1, \dots, f_n$ of V^* defined by the condition

$$f_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases},$$

the *Krönecker delta*. This basis $\{f_1, \dots, f_n\}$ is called the dual basis. In particular, $\dim V^* = \dim V$.

A *bilinear* form $V \times V \to \mathbb{C}$ is a function (u, v) which is linear in each argument. A bilinear form $(\ ,\ ) \colon V \times V \to \mathbb{C}$ is *symmetric* if $(u, v) = (v, u)$ for all $u, v \in V$. Let $W \subseteq V$ be a vector subspace. Then

$$W^{\perp} = \{v \in V \mid (v, x) = 0 \text{ for all } x \in W\}.$$

Remark. The set $W^\perp$ is a vector subspace.

Definition. A bilinear form $(\ ,\ )\colon V \times V \to \mathbb{C}$ is called *non-degenerate* if $V^\perp = 0$.

If we choose a basis $v_1, \ldots, v_n$ of V, then $(\ ,\ )$ is given by a $n \times n$ matrix $A = (a_{ij})$, where $a_{ij} = (v_i, v_j)$ for all $1 \le i, j \le n$. Then for $v = \sum \lambda_i v_i$, $w = \sum \mu_j v_j$, we have

$$(v, w) = \sum a_{ij} \lambda_i \mu_j.$$

Exercise. A bilinear form $(\ ,\ )$ is non-degenerate if and only if A is invertible.

Proposition 8.7. *Let $(\ ,\ )$ be a non-degenerate bilinear form V. For $u \in V$, define $f_u \in V^*$ by the rule*

$$f_u(x) = (u, x) \text{ for all } x \in V.$$

Then the map $V \to V^*$ given by $u \mapsto f_u$ is an isomorphism of vector spaces.

Proof. Linearity follows from the linearity of $(\ ,\ )$ in the first argument. (That $f_u \in V^*$ follows from the linearity of $(\ ,\ )$ in the second argument.)

Since $\dim V^* = \dim V$, it is enough to show that the kernel is 0. But the kernel of this map is $V^\perp$, and this is 0 since $(\ ,\ )$ is non-degenerate. □

Proposition 8.8. *Let $(\ ,\ )$ be a non-degenerate bilinear form on V. Let $W \subseteq V$ be a vector subspace. Then*

(1) $\dim W + \dim W^\perp = \dim V$,
(2) *if $W \cap W^\perp = \{0\}$, then $V = W \oplus W^\perp$.*

Proof. For (1), choose a basis $v_1, \ldots, v_r$ of W, and then extend it to a basis $v_1, \ldots, v_n$ of V. By definition, $u \in W^\perp$ means that $f_u(x) = (u, x) = 0$ for all $x \in W$, which is equivalent to $f_u(v_i) = 0$ for $i = 1, \ldots, r$. This happens if and only if $f_u \in \mathrm{Span}(f_{r+1}, \ldots, f_n)$ (where we recall that $f_i(v_j) = \delta_{ij}$). Hence the image of $W^\perp$ in V^* has dimension $n - r$. But by Proposition 8.7, the map $V \to V^*$ that sends u to f_u is an isomorphism. Thus

$$\dim W^\perp = n - r = \dim V - \dim W.$$

This proves (1).

If we have subspaces $V_1 \subseteq V$ and $V_2 \subseteq V$ such that $\dim V_1 + \dim V_2 = \dim V$ and $V_1 \cap V_2 = \{0\}$, then $V = V_1 \oplus V_2$. We apply this to $V_1 = W$ and $V_2 = W^\perp$ to get (2). □

Definition. If $W \cap W^\perp = \{0\}$, let us call W a *non-degenerate* subspace of V.

Let L be a Lie algebra over $\mathbb{C}$. Recall that we defined the bilinear Killing form $K\colon L \times L \to \mathbb{C}$ by $K(x, y) = \mathrm{Tr}(\mathrm{ad}(x)\,\mathrm{ad}(y))$.

We recall Theorem 8.3.

Theorem 8.9. *A Lie algebra L over $\mathbb{C}$ is soluble if and only if then $K(L, L') = 0$.*

Example. Let L be a Lie algebra such that $\dim L = 2$ and L has a vector space basis $B = \{x, y\}$ with $[xy] = x$. Recall that $\operatorname{ad} x$ is a linear transformation $L \to L$ sending z to $[xz]$. Then

$$[\operatorname{ad} x]_B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ [\operatorname{ad} y]_B = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Let us write down the matrix of the Killing form of L. We have

$$K(x, x) = \operatorname{Tr}(\operatorname{ad}(x)\operatorname{ad}(x)) = \operatorname{Tr} 0 = 0,$$

$$K(x, y) = \operatorname{Tr}(\operatorname{ad}(x)\operatorname{ad}(y)) = \operatorname{Tr} 0 = 0,$$

$$K(y, y) = \operatorname{Tr}(\operatorname{ad}(y)\operatorname{ad}(y)) = \operatorname{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1.$$

Therefore, the matrix associated to K is

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

We have that $K(x, x) = K(y, x) = 0$, so $K(L, L') = 0$ (since $L' = \operatorname{Span}(x)$), and hence Cartan's criterion 8.3 says that the Lie algebra is soluble. This can be confirmed by noting that $L' = \operatorname{Span}(x)$, so $L^{(2)} = 0$.

Ideals. Let L be a Lie algebra over F. For an ideal L of L, let K_I be the Killing form of I (as a Lie algebra itself). For $x, y \in I$, what is the relationship between $K_I(x, y)$ and $K(x, y)$?

Lemma 8.10. *Let I be an ideal of L.*

(1) *For $x, y \in I$, $K_I(x, y) = K(x, y)$.*
(2) *Define*

$$I^{\perp} = \{x \in L : K(x, i) = 0 \text{ for all } i \in I\}.$$

Then $I^{\perp}$ is an ideal of L.

Proof. For (1), let B be a basis of I, and $x, y \in I$. Let

$$[\operatorname{ad} x]_B = M_x, \ [\operatorname{ad} y]_B = M_y.$$

So $K_I(x, y) = \operatorname{Tr}(M_x M_y)$ by definition. Extend B to a basis B' of L. As I is an ideal, $\operatorname{ad} x$ maps $L \to I$, so

$$[\operatorname{ad} x]_{B'} = \begin{pmatrix} M_x & N_x \\ 0 & 0 \end{pmatrix}, \ [\operatorname{ad} y]_{B'} = \begin{pmatrix} M_y & N_y \\ 0 & 0 \end{pmatrix}.$$

So

$$[(\operatorname{ad} x)(\operatorname{ad} y)]_{B'} = \begin{pmatrix} M_x M_y & M_x N_y \\ 0 & 0 \end{pmatrix}$$

and hence $K(x, y) = \operatorname{Tr}((\operatorname{ad} x)(\operatorname{ad} y)) = \operatorname{Tr}(M_x M_y) = K_I(x, y)$.

For (2), let $x \in I^{\perp}$. Then

$$0 = K(x, i) = \operatorname{Tr}((\operatorname{ad} x)(\operatorname{ad} i)) \text{ for all } i \in I.$$

Let $y \in L$, $i \in I$. Then

$$\begin{aligned} K([xy], i) &= \mathrm{Tr}(\mathrm{ad}[xy]\,\mathrm{ad}\, i) \\ &= \mathrm{Tr}([\mathrm{ad}\, x, \mathrm{ad}\, y]\,\mathrm{ad}\, i) \\ &= \mathrm{Tr}(\mathrm{ad}\, x[\mathrm{ad}\, y, \mathrm{ad}\, i]) && \text{by Lemma 8.5} \\ &= \mathrm{Tr}(\mathrm{ad}\, x\,\mathrm{ad}[yi]) \\ &= K(x, [yi]) \\ &= 0 && \text{since } [yi] \in I \end{aligned}$$

Hence $[xy] \in I^\perp$ for all $x \in I^\perp$, $y \in L$. So $I^\perp$ is an ideal. □

Semisimplicity. Recall that a Lie algebra L is *semisimple* if $\mathrm{Rad}(L) = 0$, i.e. L has no nonzero soluble ideals.

Theorem 8.11. *A finite dimensional Lie algebra L over $\mathbb{C}$ is semisimple if and only if its Killing form is non-degenerate.*

Proof. Suppose L is semisimple. By Lemma 8.10, $L^\perp$ is an ideal of L. Also,

$$K(x, y) = 0 \text{ for all } x, y \in L^\perp.$$

This implies that $L^\perp$ is soluble by Theorem 8.9. Therefore, $L^\perp = 0$, which means that K is non-degenerate.

For the converse implication, we show that if L is non-semisimple, then K is degenerate (i.e. $L^\perp \neq 0$). Suppose $R = \mathrm{Rad}(L) \neq 0$. Let the derived series of R be

$$R \supset R^{(1)} \supset R^{(2)} \supset \cdots \supset R^{(t)} = 0.$$

Then $A = R^{(t-1)}$ is a nonzero abelian ideal of L. We will show that $A \subseteq L^\perp$.

We claim that the map $(\mathrm{ad}\, a)(\mathrm{ad}\, x)\colon L \to L$ is nilpotent for any $a \in A$, $x \in L$. For $l \in L$, the composition $(\mathrm{ad}\, a)(\mathrm{ad}\, x)(\mathrm{ad}\, a)$ sends

$$l \mapsto \underbrace{[al]}_{\in A} \mapsto \underbrace{[x[al]]}_{\in A} \mapsto 0$$

as A is abelian. Therefore, $(\mathrm{ad}\, a)(\mathrm{ad}\, x)(\mathrm{ad}\, a) = 0$, so $((\mathrm{ad}\, a)(\mathrm{ad}\, x))^2 = 0$.

Therefore, $\mathrm{Tr}((\mathrm{ad}\, a)(\mathrm{ad}\, x)) = 0$. Hence

$$K(a, x) = \mathrm{Tr}((\mathrm{ad}\, a)(\mathrm{ad}\, x)) = 0 \text{ for all } a \in A,\ x \in L,$$

which shows $A \subseteq L^\perp$, hence $L^\perp \neq 0$. □

Example. The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is semisimple (because it is simple), and the matrix of its Killing form with respect to the basis e, h, f is

$$\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

(Checking this is left as an exercise.)

Structure of semisimple Lie algebras.

Theorem 8.12. *Let L be a finite dimensional Lie algebra over $\mathbb{C}$. Then L is semisimple if and only if*
$$L = L_1 \oplus \cdots \oplus L_r,$$
a direct sum of simple ideals L_i.

Note. This means

(1) $L = L_1 \oplus \cdots \oplus L_r$ as vector spaces,
(2) $[L_i, L_j] \subseteq L_i \cap L_j = 0$ for all $i \neq j$.

Note. For any finite dimensional Lie algebra L over $\mathbb{C}$, this means

- $\mathrm{Rad}(L)$ is the maximal soluble ideal of L,
- $L/\mathrm{Rad}(L)$ is semisimple, so it is the direct sum of simple ideals.

Lemma 8.13. *Let L be a semisimple Lie algebra over $\mathbb{C}$. Suppose I is a non-zero ideal. Then*
$$L = I \oplus I^\perp \quad (\textit{perp. with respect to } K)$$
and I is itself semisimple.

Proof. Since $K(x, y) = 0$ for any $x, y \in I \cap I^\perp$, Theorem 8.9 shows that $I \cap I^\perp$ is soluble. As L is semisimple, $I \cap I^\perp = 0$. Hence, as K is non-degenerate by Theorem 8.11, $L = I \oplus I^\perp$ by Proposition 8.8.

As $I \cap I^\perp = 0$, the restriction of K to I is non-degenerate. Hence so is K_I, the Killing form of I, by Lemma 8.10. Therefore, I is semisimple by Theorem 8.11. □

Proof of Theorem 8.12. We first proceed by induction on $\dim L$ to show that if L is semisimple, then L is a direct sum of simple ideals. The statement is trivial when L is simple (in particular, when $\dim L = 1$). So let $0 \neq I \subsetneq L$ be an ideal. Then by Lemma 8.13, $L = I \oplus I^\perp$ and the ideals I, $I^\perp$ are semisimple. By the inductive hypothesis:
$$I = L_1 \oplus \cdots \oplus L_m, \ I^\perp = M_1 \oplus \cdots \oplus M_n$$
where L_i, M_i are simple ideals of I, $I^\perp$, respectively. Then
$$[L_i, L] = [L_i, I + I^\perp] = [L_i, I] + [L_i, I^\perp] = [L, I] \subseteq L_i$$
since $[L_i, I^\perp] \subseteq [I, I^\perp] \subseteq I \cap I^\perp = 0$. Hence all L_i are ideals of L, and similarly, M_i are ideals of L. Hence
$$L = L_1 \oplus \cdots \oplus L_m \oplus M_1 \oplus \cdots \oplus M_n$$
is a direct sum of simple ideals.

Conversely, let $L = L_1 \oplus \cdots \oplus L_r$, where L_i are simple ideals. Let I be a soluble ideal of L. Then
$$[IL_i] \subseteq I \cap L_i,$$
so $[IL_i] = 0$, since L_i is simple and I is soluble. Hence
$$[IL] = [IL_1] \oplus \cdots \oplus [IL_r] = 0$$

and $I \subseteq Z(L) = \bigoplus_i Z(L_i) = 0$. $\square$

9. Jordan decomposition in semisimple Lie algebras

Let L be a Lie subalgebra of $\mathfrak{gl}(V)$. Each $x \in L$ has a Jordan decomposition $x = d + n$ where $d\colon V \to V$ is diagonalizable, $n\colon V \to V$ is nilpotent, and $dn = nd$.

Note that d, n need not be in L.

Example. If $L = \mathrm{Span}(x)$, an abelian 1-dimensional Lie algebra ($x \in \mathfrak{gl}(V)$), then $d, n \notin L$ (unless $x = d$ or $x = n$).

But: if L is semisimple, then d, n **do** lie in L.

Theorem 9.1. *Let L be a finite dimensional semisimple Lie algebra over $\mathbb{C}$. Then every $x \in L$ can be expressed uniquely as*

$$x = d + n$$

where

(1) $d, n \in L$,
(2) $\operatorname{ad} d\colon L \to L$ *is diagonalizable, and* $\operatorname{ad} n\colon L \to L$ *is nilpotent,*
(3) $[dn] = 0$.

Moreover, for $y \in L$

$$[xy] = 0 \text{ implies that } [dy] = [ny] = 0.$$

Definition. We call $x = d + n$ the *Jordan decomposition* of x. We call d a *semisimple* element of L and n a *nilpotent* element.

The next proposition shows that this agree with the *old* notion of Jordan decomposition when $L \subseteq \mathfrak{gl}(V)$.

Proposition 9.2. *Let $L \subseteq \mathfrak{gl}(V)$ be a semisimple Lie subalgebra over $\mathbb{C}$. Then the Jordan decomposition*

$$x = d + n$$

in Theorem 9.1 is the same as the Jordan decomposition of x as a linear map $V \to V$.

Proof. Let

$$x = d' + n'$$

be the Jordan decomposition of $x\colon V \to V$. So $d'\colon V \to V$ is diagonalizable, $n'\colon V \to V$ is nilpotent, and $d'n' = n'd'$. By Lemma 8.2,

$$\mathrm{ad}(x) = \mathrm{ad}(d') + \mathrm{ad}(n')$$

is the Jordan decomposition of $\operatorname{ad} x\colon L \to L$. If $x = d + n$ as in Theorem 9.1, then by (1)–(3)

$$\mathrm{ad}(x) = \mathrm{ad}(d) + \mathrm{ad}(n)$$

is also the Jordan decomposition of $\operatorname{ad} x\colon L \to L$. By uniqueness of Jordan decomposition,

$$\mathrm{ad}(d') = \mathrm{ad}(d), \ \mathrm{ad}(n') = \mathrm{ad}(n).$$

As L is semisimple, $Z(L) = 0$, so $\mathrm{ad}\colon L \to \mathfrak{gl}(V)$ is injective. Therefore, $d' = d$, $n' = n$. $\square$

Proposition 9.3. *Let L be a semisimple Lie algebra over $\mathbb{C}$, and let $\rho\colon L \to \mathfrak{gl}(V)$ be a representation of L. Suppose $x \in L$ has Jordan decomposition*

$$x = d + n$$

as in Theorem 9.1. Then

$$\rho(x) = \rho(d) + \rho(n)$$

is the Jordan decomposition of $\rho(x)\colon V \to V$.

Proof. Set as an exercise on Sheet 4. □

Derivations. Recall that for a Lie algebra L, a *derivation* of L is a linear map $\delta\colon L \to L$ such that

$$\delta[xy] = [x, \delta(y)] + [\delta(x), y] \text{ for all } x, y \in L.$$

The set $\operatorname{Der} L$ of all derivations of L is a Lie subalgebra of $\mathfrak{gl}(L)$ by Proposition 1.3.

Also, $\operatorname{ad} x \in \operatorname{Der} L$ by Proposition 1.4.

Proposition 9.4. *The set* $\operatorname{ad} L = \{\operatorname{ad} x : x \in L\}$ *is an ideal of* $\operatorname{Der} L$.

Proof. Let $\delta \in \operatorname{Der} L$ and $x, y \in L$. Then

$$[\delta, \operatorname{ad} x](y) = \delta[xy] - [x, \delta(y)] = [x, \delta(y)] + [\delta(x), y] - [x, \delta(y)] = [\delta(x), y] = (\operatorname{ad}(\delta(x))(y).$$

Hence

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

so $\operatorname{ad} L$ is an ideal. □

Proposition 9.5. *If L is a semisimple Lie algebra over $\mathbb{C}$, then*

$$\operatorname{ad} L = \operatorname{Der} L.$$

Proof. Let $M = \operatorname{ad} L$. As $Z(L) = 0$, $\operatorname{ad}\colon L \to M$ is an isomorphism, so M is also semisimple, and M is an ideal of $\operatorname{Der} L$.

Let K be the Killing form of $\operatorname{Der} L$. We claim that $M^\perp = 0$. By Lemma 8.10, the Killing form K_M of M is the restriction of K to M. By Theorem 8.11, K_M is non-degenerate, so $M \cap M^\perp = 0$. Note that

$$[M, M^\perp] \subseteq M \cap M^\perp = 0,$$

and hence for $\delta \in M^\perp$ and $\operatorname{ad} x \in M$, we have $[\delta, \operatorname{ad} x] = 0$. By the proof of Proposition 9.4,

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

so $\delta(x) = 0$ for all $x \in L$, i.e. $\delta = 0$. Hence we have shown that $M^\perp = 0$.

This simplies that

$$(\operatorname{Der} L)^\perp \subseteq M^\perp = 0,$$

and therefore K is non-degenerate, so by Proposition 8.8,

$$\dim(\operatorname{Der} L) = \dim M + \dim M^\perp = \dim M,$$

as $\dim M^\perp = 0$, which implies that $M = \operatorname{Der} L$. □

Proposition 9.6. *Let L be a Lie algebra over $\mathbb{C}$. Let $\delta \in \operatorname{Der} L$ have Jordan decomposition (as a linear map $L \to L$)*

$$\delta = \sigma + \nu$$

where $\sigma\colon L \to L$ is diagonal, $\nu\colon L \to L$ is nilpotent, and $\sigma\nu = \nu\sigma$. Then $\sigma, \nu \in \operatorname{Der} L$.

Proof. Let $\lambda_1, \ldots, \lambda_r$ be distinct eigenvalues of $\delta\colon L \to L$, and m_i be the size of the largest λ_i-Jordan block. Then

$$L = \bigoplus_{i=1}^{r} L_{\lambda_i}$$

where

$$L_{\lambda_i} = \operatorname{Ker}(\delta - \lambda_i I)^{m_i}.$$

On each L_{λ_i}, σ acts as $\lambda_i I$, and ν as a strictly upper-triangular matrix. For each $\lambda \in \mathbb{C}$, define

$$L_\lambda = \{x \in L : (\delta - \lambda I)^k(x) = 0 \text{ for some } k\} = \begin{cases} L_{\lambda_i} & \text{if } \lambda = \lambda_i, \\ 0 & \text{if } \lambda \neq \lambda_i. \end{cases}$$

We claim that

$$[L_\lambda, L_\mu] \subseteq L_{\lambda+\mu}.$$

We show that for any $n \in \mathbb{N}$,

$$(\delta - (\lambda+\mu)I)^n[xy] = \sum_{k=0}^{n} \binom{n}{k} [(\delta - \lambda I)^k(x), (\delta - \mu I)^{n-k}(y)] \tag{1}$$

by induction on n. The base case $n = 1$ is clear:

$$\begin{aligned} \text{RHS} &= [x, (\delta - \mu)(y)] + [(\delta - \lambda)(x), y] \\ &= [x, \delta y] + [\delta x, y] - (\lambda + \mu)[xy] \\ &= \delta([xy]) - (\lambda + \mu)[xy] \\ &= \text{LHS} \end{aligned}$$

The inductive step is left as an exercise.

By equation (1), for $x \in L_\lambda$, $y \in L_\mu$,

$$(\delta - (\lambda+\mu)I)^n[xy] = 0$$

for sufficiently large n. Hence $[xy] \in L_{\lambda+\mu}$, which shows that $[L_\lambda, L_\mu] \subseteq L_{\lambda+\mu}$.

Now, we use this to show that $\sigma \in \operatorname{Der} L$. Recall that L_λ is a λ-eigenspace for σ. For $x \in L_\lambda$, $y \in L_\mu$, $[xy] \in L_{\lambda+\mu}$, so

$$\sigma(x) = \lambda x,\ \ \sigma(y) = \mu y \text{ and } \sigma[xy] = (\lambda+\mu)[xy].$$

Hence

$$[\sigma(x), y] + [x, \sigma(y)] = \lambda[xy] + \mu[xy] = (\lambda+\mu)[xy] = \sigma[xy],$$

showing that $\sigma \in \operatorname{Der} L$. Hence $\nu = \delta - \sigma \in \operatorname{Der} L$. ∎

We can finally prove Theorem 9.1.

Proof of Theorem 9.1. Let L be a semisimple Lie algebra over $\mathbb{C}$ and let $x \in L$. Then $\operatorname{ad} x \in \operatorname{Der} L$. Let the Jordan decomposition of $\operatorname{ad} x \colon L \to L$ be

$$\operatorname{ad} x = \sigma + \mu$$

(for σ diagonalizable, ν nilpotent, $\sigma\nu = \nu\sigma$). By Proposition 9.6, $\sigma, \nu \in \operatorname{Der} L$. By Proposition 9.5, $\operatorname{Der} L = \operatorname{ad} L$, so there exist $d, n \in L$ such that

$$\sigma = \operatorname{ad}(d), \ \nu = \operatorname{ad}(n).$$

Hence

$$\operatorname{ad} x = \operatorname{ad}(d) + \operatorname{ad}(n) = \operatorname{ad}(d+n).$$

As $Z(L) = 0$, $\operatorname{ad} \colon L \to \operatorname{ad} L = \operatorname{Der} L$ is an isomorphism, so

$$x = d + n.$$

As $d, n \in L$, part (1) of Theorem 9.1 holds. As $\sigma = \operatorname{ad} d$, $\nu = \operatorname{ad} n$, part (2) holds.

Also

$$\operatorname{ad}[dn] = [\operatorname{ad}(d), \operatorname{ad}(n)] = [\sigma, \nu] = 0,$$

hence $[dn] = 0$, so part (3) holds.

As σ and ν are unique (by uniqueness of Jordan decomposition), so are d and n.

We finally prove the last part of Theorem 9.1. Let $y \in L$ with $[xy] = 0$. By Lemma 8.1, σ and $\nu = \sigma - \operatorname{ad} x$ are polynomials in $\operatorname{ad} x$. Say

$$\nu = c_r(\operatorname{ad} x)^r + \cdots + c_1 \operatorname{ad} x + c_0 I.$$

As $(\operatorname{ad}(x))(y) = [xy] = 0$, this means

$$\nu(y) = c_0 y.$$

Since ν is nilpotent, $c_0 = 0$, so $\nu(y) = 0$. Hence

$$0 = \nu(y) = (\operatorname{ad} n)(y) = [ny],$$

so $[ny] = 0$, and then $[dy] = 0$ as $[xy] = 0$. □

10. Cartan subalgebras and root spaces

We now work towards the classification of simple Lie algebras over $\mathbb{C}$.

We have seen one such: $\mathfrak{sl}(2, \mathbb{C})$. Features: basis e, f, h with:

- $\operatorname{ad} h$ is diagonalizable — i.e. h is a *semisimple* element
- e, f, h are a basis of eigenvectors for $\operatorname{ad} h$.

For $\mathfrak{sl}(3, \mathbb{C})$, replace h by

$$H = \{\text{diagonal matrices}\} = \left\{ h = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} : \sum a_i = 0 \right\},$$

a 2-dimensional abelian subalgebra.

Let $E_{ij} \in \mathfrak{sl}(2, \mathbb{C})$ be the matrix with 1 in the ij-entry and 0 elsewhere ($i \neq j$). Check that

$$(\operatorname{ad} h)(E_{ij}) = [hE_{ij}] = (a_i - a_j)E_{ij}.$$

Hence

$$\mathfrak{sl}(3, \mathbb{C}) = H \oplus \bigoplus_{i+j} \operatorname{Span}(E_{ij})$$

a direct sum of weight spaces for $\operatorname{ad} H$. The weights of H corresponding to each of these weight spaces are

space	weight
H	0
$\operatorname{Span}(E_{ij})$	$\epsilon_i - \epsilon_j$, where $\epsilon_i \colon H \to \mathbb{C}$, $h \mapsto a_i$.

Note also: H consists of **semisimple**[4] elements.

The general strategy to understand structure of a simple Lie algebra L is:

(1) find an abelian subalgebra H consisting of semisimple elements,
(2) decompose L into a direct sum of weight spaces for $\operatorname{ad} H$,
(3) use decomposition to pin down structure constants.

First steps. Suppose L is a Lie algebra over $\mathbb{C}$, and H is an abelian subalgebra of L consisting of semisimple elements.

Lemma 10.1. *The Lie algebra L has a basis of common eigenvectors for all $\operatorname{ad} h$, $h \in H$.*

Proof. Let $h_1, \dots, h_r$ be a basis for H and $\alpha_i = \operatorname{ad} h_i \colon L \to L$. Then

$$[\alpha_i, \alpha_j] = \operatorname{ad}[h_i h_j] = 0$$

since H is abelian. So $\alpha_1, \dots, \alpha_r$ are commuting, diagonalizable linear maps $L \to L$. Such maps have a basis of common eigenvectors. (Standard linear algebra fact, see Sheet 4, Question 5). $\square$

Let $x \in L$ be a common eigenvector for $\operatorname{ad} H$. Define weight $\alpha \colon H \to \mathbb{C}$ (i.e. $\alpha \in H^*$) by

$$(\operatorname{ad} h)(x) = [hx] = \alpha(h)x \qquad (h \in H).$$

The *weight space* of α is

$$L_\alpha = \{x \in L : [hx] = \alpha(h)x \text{ for all } h \in H\}$$

By Lemma 10.1, L is the direct sum of these weight spaces. One of them is the 0-*weight space*

$$L_0 = \{x \in L : [hx] = 0 \text{ for all } h \in H\}.$$

Note that $H \subseteq L_0$. Define

$$\Phi = \text{set of nonzero weights } \alpha \in H^* \text{ for which } L_\alpha \neq 0.$$

Then

$$L = L_0 \oplus \bigoplus_{\alpha \in \Phi} L_\alpha \tag{1}$$

[4] Recall that $x \in L$ is *semisimple* if $\operatorname{ad} x$ is diagonalizable

Note: Φ is a finite set, since L is finite dimensional.

Let K be the Killing form of L (i.e. $K(x,y) = \mathrm{Tr}((\mathrm{ad}\, x)(\mathrm{ad}\, y))$).

Proposition 10.2. *Let $\alpha, \beta \in H^*$.*

(1) $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$
(2) *If $\alpha + \beta \neq 0$ then $K(L_\alpha, L_\beta) = 0$.*
(3) *Suppose L is semisimple. Then the restriction of K to L_0 is non-degenerate (i.e. $L_0 \cap L_0^\perp = 0$).*

Proof. For (1), let $x \in L_\alpha$, $y \in L_\beta$. For $h \in H$,

$$\begin{aligned} h[xy] &= [[hx]y] + [x[hy]] && \text{by the Jacobi identity} \\ &= \alpha(h)[xy] + \beta(h)[xy] \\ &= (\alpha+\beta)(h)[xy] \end{aligned}$$

and hence $[xy] \in L_{\alpha+\beta}$.

For (2), let $h \in H$ with $(\alpha+\beta)(h) \neq 0$. For $x \in L_\alpha$, $y \in L_\beta$,

$$\begin{aligned} \alpha(h)K(x,y) &= K([hx], y) && \text{as } K \text{ is bilinear} \\ &= -K([xh], y) && \text{as } K \text{ is bilinear} \\ &= -K(x, [hy]) && \text{by Lemma 8.5} \\ &= -\beta(h)K(x,y) && \text{as } K \text{ is bilinear} \end{aligned}$$

Hence $(\alpha+\beta)(h)K(x,y) = 0$, so $K(x,y) = 0$.

For (3), let $y \in L_0 \cap L_0^\perp$. Then $K(L_0, y) = 0$. For $x \in L$, use the weight space decomposition (1), to write

$$x = x_0 + \sum_{\alpha \in \Phi} x_\alpha \qquad (x_\alpha \in L_\alpha).$$

By (2), $K(L_0, L_\alpha) = 0$ if $\alpha \neq 0$. So $K(x_\alpha, y) = 0$ for all $\alpha \in \Phi$. Also, $K(x_0, y) = 0$ by assumption. Hence

$$K(x,y) = 0 \text{ for all } x \in L.$$

So $y \in L^\perp$. As L is semisimple, K is non-degenerate, so $L^\perp = 0$. Hence $y = 0$. $\square$

Corollary 10.3. *If $x \in L_\alpha$, where $\alpha \neq 0$, $\alpha \in H^*$, then $\mathrm{ad}\, x$ is nilpotent.*

Proof. For any weight $\beta \in \Phi \cup \{0\}$, we have

$$(\mathrm{ad}\, x)(L_\beta) \subseteq L_{\alpha+\beta},$$

$$(\mathrm{ad}\, x)^2(L_\beta) \subseteq L_{2\alpha+\beta},$$

$$\vdots$$

$$(\mathrm{ad}\, x)^r(L_\beta) \subseteq L_{r\alpha+\beta}$$

$$\vdots$$

by Proposition 10.2 (1). But Φ is finite, so for some r we will have $r\alpha + \beta \notin \Phi$. This means that $(\mathrm{ad}\, x)^r(L_\beta) = 0$.

Once again, since Φ is finite and for each $\beta \in \Phi \cup \{0\}$, we can find r such that $(\operatorname{ad} x)^r(L_\beta) = 0$, we can take the maximum of these r to obtain one r such that

$$(\operatorname{ad} x)^r(L_\beta) = 0$$

for any $\beta \in \Phi \cup \{0\}$.

Since

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

(equation (1)), we see that $(\operatorname{ad} x)^r$ is the zero transformation $L \to L$. □

Cartan subalgebras. Each choice of an abelian subalgebra $H \subseteq L$ gives rise to a direct sum decomposition

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

Definition. A subalgebra H of a Lie algebra L is called a *Cartan subalgebra* if

(1) H is abelian,
(2) every element $h \in H$ is semisimple,
(3) H is maximal among subalgebras $H \subseteq L$ satisfying properties (1) and (2).

Proposition 10.4. *Let L be a finite dimensional semisimple Lie algebra over $\mathbb{C}$. Then L has a Cartan subalgebra.*

Proof. Recall that each $x \in L$ has a Jordan decomposition $x = s + n$ where s is semisimple and n is nilpotent, $s, n \in L$ and $[s, n] = 0$.

Suppose that the semisimple part of any $x \in L$ is zero. Then all elements of L are nilpotent. Then $\operatorname{ad} x$ is nilpotent for any $x \in L$. Engel's Theorem 4.1 implies now that L is a nilpotent Lie algebra. This contradicts the fact that L is semisimple.

Hence there exists $x \in L$ whose semisimple part is nonzero. So semisimple elements exist. Let $s \in L$ be a semisimple element. Consider $\operatorname{Span}(s)$. It is an abelian subalgebra of L. Hence subalgebras $H \subseteq L$ satisfying properties (1) and (2) exist. A maximal dimensional such subalgebra is a Cartan subalgebra. □

Definition. For a subset $X \subseteq L$, define the *centralizer* of X in L by the formula

$$C_L(X) = \{l \in L \mid [lx] = 0 \text{ for all } x \in X\}.$$

Exercise. $C_L(X)$ is a Lie subalgebra of L.

Lemma 10.5. *Suppose H is a Lie subalgebra of L such that*

(1) *H consists of semisimple elements,*
(2) *$C_L(H) = H$.*

Then H is a Cartan subalgebra in L.

Proof. Note that clearly H is abelian if and only if $H \subseteq C_L(H)$. Thus H is abelian.

We claim that H is the maximal subalgebra with properties (1) and (2). Otherwise, there exists subalgebra H_1 such that $H \subsetneqq H_1$. Then $H_1 \subseteq C_L(H)$. But we are given that $C_L(H) = H$, so $H_1 = H$, a contradiction. □

Example. Let $L = \mathfrak{sl}(n, \mathbb{C}) = \{n \times n \text{ matrices with trace } 0\}$. Consider

$$H = \{\text{diagonal } n \times n \text{ matrices with trace } 0\}.$$

Then $H \subseteq L$ is an abelian subalgebra. Then

$$\mathfrak{sl}(n, \mathbb{C}) = H \oplus \bigoplus_{\substack{1 \le i,j \le n \\ i \ne j}} \mathbb{C}E_{ij}.$$

One can check that $C_L(H) = H$ (Sheet 4). Hence H is a Cartan subalgebra.

On Sheet 4, we will find a Cartan subalgebra in $\mathfrak{so}(n, \mathbb{C})$.

Theorem 10.6. *Let L be a semisimple Lie algebra over $\mathbb{C}$. Let H be a Cartan subalgebra in L. Then $C_L(H) = H$.*

Consequence. In the decomposition

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

we know that

$$L_0 = \{x \in L \mid [h, x] = 0 \text{ for all } h \in H\} = C_L(H) = H$$

if H is a Cartan subalgebra.

The elements $\alpha \in \Phi$ are called *roots*. It turns out that semisimple Lie algebras will be classified according to so-called *root systems*.

Proof of Theorem 10.6. Let L be semisimple with Cartan subalgebra H.

Step 1. Choose $h \in H$ such that $\dim C_L(h)$ is minimal. We will show that

$$C_L(h) = C_L(H).$$

Suppose this is false, i.e. $C_L(h) \ne C_L(H)$ so there exists $s \in H$ such that $C_L(h) \not\subseteq C_L(s)$. So

$$C_L(h) \cap C_L(s) \subsetneqq C_L(h).$$

(The aim is to find linear combinations of h, s with smaller centralizer than h.)

Choose basis $c_1, \ldots, c_n$ of $C_L(h) \cap C_L(s)$. Since $s \in H \subseteq C_L(H)$ and s is semisimple, we can extend to a basis

$$c_1, \ldots, c_n, x_1, \ldots, x_p \text{ of } C_L(h)$$

consisting of eigenvectors for $\operatorname{ad}(s)$. Also, we can extend to a basis

$$c_1, \ldots, c_n, y_1, \ldots, y_q \text{ of } C_L(s)$$

of eigenvectors for $\operatorname{ad}(h)$. So

$$c_1, \ldots, c_n, x_1, \ldots, x_p, y_1, \ldots, y_q$$

is a basis of $C_L(h) + C_L(s)$. Also, $\operatorname{ad} h$ and $\operatorname{ad} s$ commute (since $[hs] = 0$), so we can extend to a basis of L:

$$c_1, \dots, c_n, x_1, \dots, x_p, y_1, \dots, y_q, z_1, \dots, z_r$$

consisting of simultaneous $\operatorname{ad}(h)$ and $\operatorname{ad}(s)$ eigenvectors.

Note $[sx_i] \neq 0$, since $x_i \notin C_L(s)$, and similarly $[hy_i] \neq 0$. Let

$$[hz_i] = \alpha_i z_i, \ [sz_i] = \beta_i z_i$$

where $\alpha_i, \beta_i \neq 0$.

We have the following table of eigenvalues.

	c_i	x_i	y_i	z_i
$\operatorname{ad}(s)$	0	$\neq 0$	0	β_i
$\operatorname{ad}(h)$	0	0	$\neq 0$	α_i
$\operatorname{ad}(s) + \lambda \operatorname{ad}(h)$ $(\lambda \neq 0)$	0	$\neq 0$	$\neq 0$	$\beta_i + \lambda\alpha_i$

Choose $\lambda \in \mathbb{C}$ such that $\beta_i + \lambda\alpha_i \neq 0$ for all i. Then $C_L(s + \lambda h) = C_L(s) \cap C_L(h)$ which has smaller dimension than $C_L(h)$, a contradiction.

We have hence concluded Step 1, showing that

$$C_L(h) = C_L(H).$$

Step 2. $C_L(h)$ is nilpotent.

Let $x \in C_L(h)$, with Jordan decomposition $x = d + n$. By Theorem 9.1,

$$[hx] = 0 \text{ implies that } [hd] = [hn] = 0.$$

So $d, n \in C_L(h)$. Also, $d \in H$: this is because by Step 1, $[d, H] = 0$, so $H + \operatorname{Span}(d)$ is abelian and consists of semisimple elements, so $d \in H$ as H is a Cartan subalgebra (maximality condition).

Now, the restriction $\operatorname{ad} d\colon C_L(h) \to C_L(h)$ is the zero map (as $d \in H$), so the restriction $\operatorname{ad} x\colon C_L(h) \to C_L(h)$ is the restriction of $\operatorname{ad} n$ to $C_L(h)$, so is nilpotent. Hence $\operatorname{ad} C_L(h)$ consists of nilpotent linear maps $C_L(h) \to C_L(h)$, hence by Engel's Theorem 4.5, $C_L(h)$ is nilpotent. This completes Step 2.

Step 3. $C_L(h) \subseteq H$.

By Step 2, $C_L(h)$ is nilpotent, hence soluble, so by Lie's Theorem 5.1, there exists a basis B of $C_L(h)$ such that

$$\{(\operatorname{ad} x)_B : x \in C_L(h)\} \subseteq \mathfrak{t}(n, \mathbb{C})$$

(where $n = \dim C_L(h)$) consists of upper-triangular matrices. Let $x \in C_L(h)$, with Jordan decomposition

$$x = d + n$$

where $d \in H$ (by proof of Step 2), and $n \in C_L(h)$ with $\operatorname{ad} n$ nilpotent. Then

$$(\operatorname{ad} n)_B \in \mathfrak{u}(n, \mathbb{C})$$

is strictly upper-triangular.

Let K be the Killing form of L. For all $y \in C_L(h)$,

$$K(n, y) = \mathrm{Tr}(\mathrm{ad}(n)\,\mathrm{ad}(y)) = 0.$$

By Proposition 10.2 (3), the restriction of K to $L_0 = C_L(H)$ is non-degenerate. Hence $n = 0$, and so

$$x = d \in H.$$

Therefore, $C_L(h) = H$, finishing Step 3.

By Steps 1 and 3, $C_L(H) = H$. □

11. $\mathfrak{sl}(2)$-subalgebras

Let L be a semisimple Lie algebra over $\mathbb{C}$ and H be a Cartan subalgebra.

We have the *root space decomposition*

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

where

$$L_\alpha = \{x \in L : [hx] = \alpha(h)x \text{ for all } h \in H\}$$

and $\Phi \subseteq H^* \setminus \{0\}$, the set of *roots* of L (with respect to H).

Aim. To pin down the structure constants; so far we know that

$$[H, H] = 0,$$

$$[H, L_\alpha]\text{: structure constants } \alpha(h),$$

$$[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}.$$

There is still a long way to go.

Proposition 11.1. *Let $\alpha \in \Phi$. Then*

(1) $-\alpha \in \Phi$,
(2) *Let $0 \neq x \in L_\alpha$. Then there exists $y \in L_{-\alpha}$ such that*

$$\mathrm{Span}(x, y, [xy])$$

is a subalgebra of L isomorphic to $\mathfrak{sl}(2, \mathbb{C})$.

Proof. For (1), let $0 \neq x \in L_\alpha$. As K is non-degenerate, there exists $y \in L$ such that $K(x, y) \neq 0$. Write

$$y = y_0 + \sum_{\beta \in \Phi} y_\beta \quad (y_0 \in H,\ y_\beta \in L_\beta).$$

By Proposition 10.2 (2), $K(L_\alpha, L_\beta) = 0$, unless $\alpha + \beta = 0$. Therefore, $y_{-\alpha} \neq 0$, and so $-\alpha \in \Phi$.

For (2), let $x \in L_\alpha$, $y \in L_{-\alpha}$ satisfy

$$K(x, y) \neq 0.$$

We claim that $[xy] \in H$ and $[xy] \neq 0$. First, $[xy] \in [L_\alpha, L_{-\alpha}] \subseteq L_0 = H$ by Proposition 10.2 (1). As $\alpha \neq 0$, there exists $u \in H$ such that $\alpha(u) \neq 0$. Then

$$\begin{aligned} K(u,[xy]) &= K([ux],y) \qquad \text{by Lemma 8.5} \\ &= \alpha(u)K(x,y) \\ &\neq 0 \end{aligned}$$

Therefore $[xy] \neq 0$.

Now, we claim that $S = \mathrm{Span}(x, y, [xy])$ is a subalgebra of L. As $[xy] \in H$, we have that

$$[[xy]x] = \alpha([xy])x,$$

$$[[xy]y] = -\alpha([xy])y,$$

and hence S is indeed a subalgebra.

Finally, we claim that $S \cong \mathfrak{sl}(2,\mathbb{C})$. To show this, we will show that $S' = S$, and refer to the classification of Lie algebras of small dimension from Chapter 2. Let $h = [xy] \in H \setminus \{0\}$. Suppose for a contradiction that $\alpha(h) = 0$, so $[hx] = [hy] = 0$. Then $\dim S' = 1$, so S is soluble. By Lie's Theorem 5.1, there exists a basis B of L such that

$$\{(\mathrm{ad}(s))_B : s \in S\} \subseteq \mathfrak{t}(3,\mathbb{C})$$

are upper-triangular matrices. Then

$$(\mathrm{ad}\, h)_B = (\mathrm{ad}[xy])_B \in \mathfrak{u}(3,\mathbb{C}),$$

so $\mathrm{ad}\, h$ is nilpotent. As h is semisimple, $\mathrm{ad}\, h$ is diagonalizable, hence $h = 0$, a contradiction.

Hence $\alpha(h) \neq 0$, so $[hx], [hy] \neq 0$, and so S' is 3-dimensional. Now, $S \cong \mathfrak{sl}(2,\mathbb{C})$ by Proposition 2.7. $\square$

Notation. Rescale x, y to take

$$S = \mathrm{Span}(e_\alpha, e_{-\alpha}, h_\alpha)$$

where

$$e_\alpha \in L_\alpha,\ e_{-\alpha} \in L_{-\alpha}$$

$$h_\alpha = [e_\alpha, e_{-\alpha}],\ [h_\alpha e_\alpha] = 2e_\alpha,\ [h_\alpha e_{-\alpha}] = -2e_{-\alpha}$$

(i.e. $\alpha(h_\alpha) = 2$). Define

$$S = \mathfrak{sl}(\alpha) \cong \mathfrak{sl}(2,\mathbb{C}).$$

Example. Take $L = \mathfrak{sl}(3,\mathbb{C})$. We have the root space decomposition

$$L = H \oplus \bigoplus_{i \neq j} \mathrm{Span}(E_{ij})$$

with roots $\epsilon_i - \epsilon_j$ where

$$\epsilon_i : \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \mapsto a_i.$$

For $\alpha = \epsilon_i - \epsilon_j$, we have that

$$\mathfrak{sl}(\alpha) = \mathrm{Span}(E_{ij}, E_{ji}, h_{ij})$$

where $h_{ij} = E_{ii} - E_{jj}$.

We can regard L as an $\mathfrak{sl}(\alpha)$-module with multiplication

$$sl = [sl] \qquad (s \in S, l \in L).$$

Proposition 11.2. *Let $\alpha \in \Phi$ and $\beta \in \Phi \cup \{0\}$. Define*

$$M = \sum_{\substack{c \in \mathbb{Z} \\ \beta + c\alpha \in \Phi}} L_{\beta + c\alpha}.$$

Then M is an $\mathfrak{sl}(\alpha)$-submodule of L.

Proof. By Proposition 10.2 (1),

$$[L_{\beta+c\alpha}, L_{\pm\alpha}] \subseteq L_{\beta+(c\pm1)\alpha}$$

and $[L_{\beta+c\alpha}, H] \subseteq L_{\beta+c\alpha}$. □

Definition. We call the set of roots

$$\{\beta + c\alpha : c \in \mathbb{Z}\} \cap \Phi$$

the *α-string through β*.

Example. Let $L = \mathfrak{sl}(3, \mathbb{C})$. Let $\alpha = \epsilon_2 - \epsilon_3$, $\beta = \epsilon_1 - \epsilon_2$. The α-string though β is

$$\beta, \beta + \alpha.$$

The corresponding submodule is

$$M = \mathrm{Span}(E_{12}, E_{13}) = \left\{ \begin{pmatrix} 0 & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}$$

and

$$\mathfrak{sl}(\alpha) = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix} \right\}.$$

Proposition 11.3. *For any $\alpha \in \Phi$:*

(1) $\dim L_\alpha = 1$,
(2) *if $n \in \mathbb{Z} \setminus \{0\}$ and $n\alpha \in \Phi$ then $n = \pm 1$.*

Proof. Define

$$W = \mathrm{Span}(e_{-\alpha}, H, L_{n\alpha} : n \in \mathbb{N}).$$

Then W is invariant under $\mathrm{ad}\, e_\alpha$, $\mathrm{ad}\, e_{-\alpha}$, and $\mathrm{ad}\, H$. So it is an $\mathfrak{sl}(\alpha)$-submodule. Write $(\mathrm{ad}\, e_\alpha)_W$ for the restriction of $\mathrm{ad}\, e_\alpha$ to W. Now

$$\mathrm{Tr}([(\mathrm{ad}\, e_\alpha)_W, (\mathrm{ad}\, e_{-\alpha})_W]) = 0$$

and is also equal to

$$\mathrm{Tr}(\mathrm{ad}[e_\alpha e_{-\alpha}])_W = \mathrm{Tr}(\mathrm{ad}\, h_\alpha)_W.$$

Now

$$(\operatorname{ad} h_\alpha))_W = \left(\begin{array}{c|c|c} -\alpha(h) & & \\ \hline & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} & \\ \hline & & \begin{matrix} n\alpha(h) & & & \\ & \ddots & & \\ & & n\alpha(h) & \\ & & & \ddots \end{matrix} \end{array}\right)$$

writing $h = h_\alpha$ so that $\alpha(h) = 2$ (the first square, $-\alpha(h)$, coming from $e_{-\alpha}$, the second square, 0, coming from H, and the third square, $n\alpha(h)I$, coming from $L_{n\alpha}$). Taking the trace, we obtain

$$0 = -\alpha(h) + \sum_{n\geq 1} n\alpha(h) \dim L_{n\alpha}.$$

Therefore,

$$\sum_{n\geq 1} n \dim L_{n\alpha} = 1.$$

Hence $\dim L_\alpha = 1$ and $L_{n\alpha} = 0$ for $n > 1$.

Similarly, $\dim L_{-\alpha} = 1$ and $L_{-n\alpha} = 0$ for $n > 1$. □

The next aim towards studying the *root system* $\Phi \subseteq H^* \setminus \{0\}$ is the following proposition.

Proposition 11.4. *Let $S = \operatorname{Span}(e, f, h) \cong \mathfrak{sl}(2, \mathbb{C})$, and let V be a finite dimensional S-module. Then every eigenvalue of the linear map*

$$v \mapsto hv \quad (v \in V)$$

is an integer.

Proof. Recall that the irreducible S-modules are V_d, and on V_d, h has eigenvalues

$$d, d-2, \dots, -d \in \mathbb{Z}.$$

Now take a composition series

$$V = W_0 \supset W_1 \supset \cdots \supset W_r = 0$$

where W_i are submodules and each quotient W_i/W_{i+1} is irreducible. So $\frac{W_i}{W_{i+1}} \cong V_{d_i}$ for some d_i, and with respect to a suitable basis of V, h acts as

$$\begin{pmatrix} d_0 & & & & & & \\ & \ddots & & & 0 & & \\ & & -d_0 & & & & \\ & & & d_1 & & & \\ & \bigstar & & & \ddots & & \\ & & & & & -d_1 & \\ & & & & & & \ddots \end{pmatrix}$$

Hence the eigenvalues are integers. □

We prove some more facts about Φ.

Proposition 11.5. *Let $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$.*

(1) $\beta(h_\alpha) \in \mathbb{Z}$.
(2) *The α-root string through β is*

$$\beta - r\alpha, \beta - (r-1)\beta, \ldots, \beta, \beta+\alpha, \ldots, \beta + q\alpha$$

where $q, r \geq 0$ and $\beta(h_\alpha) = r - q$.
(3) *If $\alpha + \beta \in \Phi$, then*

$$[e_\alpha e_\beta] = \lambda e_{\alpha+\beta} \quad (\lambda \neq 0).$$

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

Proof. Let

$$M = \sum_{c \in \mathbb{Z}} L_{\beta + c\alpha}.$$

This is a $\mathfrak{sl}(\alpha)$-submodule by Proposition 11.2.

Part (1) is clear: since $\beta(h_\alpha)$ is an eigenvalue of h_α on L_β, $\beta(h_\alpha) \in \mathbb{Z}$ by Proposition 11.4.

Then by Proposition 11.3, $\dim L_{\beta+c\alpha}$ is either 0 or 1 for all $c \in \mathbb{Z}$. So the eigenspaces of $\mathrm{ad}(h_\alpha)$ on M are all 1-dimensional. The eigenvalues are $(\beta + c\alpha)(h_\alpha) = \beta(h_\alpha) + 2c$, so they are either all even or all odd. Therefore, M is an irreducible $\mathfrak{sl}(\alpha)$-module (if not, then the eigenvalue 0 or 1 has multiplicity greater than 1).

So $M \cong V_d$ for some d. Hence the eigenvalues are

$$\{\beta(h_\alpha) + 2c : \beta + c\alpha \in \Phi\} = \{d, d-2, \ldots, -d\}.$$

Hence (2) holds, taking

$$d = \beta(h_\alpha) + 2q, \; -d = \beta(h_\alpha) - 2r$$

(so, in particular, we obtain $\beta(h_\alpha) = r - q$ by taking the sum of these equations).

For (3), we recall that the action of e_α on $M \cong V_d$ is given by the matrix

$$\begin{pmatrix} 0 & 1 & & & \\ & 0 & 2 & & \\ & & \ddots & \ddots & \\ & & & 0 & d \\ & & & & 0 \end{pmatrix}$$

So if $[e_\alpha e_\beta] = 0$ then $[h_\alpha e_\beta] = de_\beta$, so $\beta(h_\alpha) = d$, hence $q = 0$. But $\alpha + \beta \in \Phi$ by assumption, and hence $q \geq 1$, a contradiction. Therefore, $[e_\alpha e_\beta] \neq 0$, showing (3).

Finally, (4) follows from (2). $\square$

12. Cartan subalgebras as inner product spaces

As in the previous chapters, we assume that L is semisimple and H is a Cartan subalgebra.

Proposition 12.1. *For $h_1, h_2 \in H$,*

$$K(h_1, h_2) = \sum_{\alpha \in \Phi} \alpha(h_1)\alpha(h_2).$$

Proof. Let B be a basis of L:

$$B = (\text{basis of } H) \cup \{e_\alpha : \alpha \in \Phi\}.$$

For $h \in H$,

$$(\operatorname{ad} h)_B = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & \alpha(h) & \\ & & & & \ddots \end{pmatrix},$$

the first part, 0, corresponding to H, and the second part, $\alpha(h)I$, corresponding to

$$\{e_\alpha : \alpha \in \Phi\}.$$

Hence

$$K(h_1, h_2) = \operatorname{Tr}((\operatorname{ad} h_1)(\operatorname{ad} h_2)) = \sum_{\alpha \in \Phi} \alpha(h_1)\alpha(h_2),$$

completing the proof. □

Proposition 12.2. *We have that* $\operatorname{Span}(\Phi) = H^*$*, the dual space of H.*

Proof. Suppose for a contradiction that

$$\operatorname{Span}(\Phi) = W \subsetneqq H^*.$$

Then

$$\operatorname{Ann}_H(W) = \{h \in H : f(h) = 0 \text{ for all } f \in W\}$$

is nonzero (as it has dimension $\dim H^* - \dim W$).

Hence there exists $0 \neq h \in H$ such that

$$\alpha(h) = 0 \text{ for all } \alpha \in \Phi.$$

Then

$$K(h, H) = 0 \text{ by Proposition 12.1}$$
$$K(h, L_\alpha) = 0 \text{ for all } \alpha, \text{ by Proposition 10.2 (2)}$$

Hence $h \in L^\perp = 0$, a contradiction. □

By Proposition 10.2 (3), the restriction K_H of K to H is non-degenerate. For $h \in H$, define $\theta_h \in H^*$ by

$$\theta_h(x) = K(h, x).$$

Now, K_H non-degenerate implies that the map $h \mapsto \theta_h$ is an injective map $H \to H^*$. Since $\dim H = \dim H^*$, the map $h \mapsto \theta_h$ is an isomorphism $H \to H^*$.

We hence obtain the following result.

Proposition 12.3. *For each $\alpha \in \Phi$, there exists a unique $t_\alpha \in H$ such that*

$$\alpha(x) = K(t_\alpha, x) \text{ for all } x \in H.$$

Recall that $L_\alpha = \mathrm{Span}(e_\alpha)$ and $[e_\alpha, e_{-\alpha}] = h_\alpha \in H$.

Proposition 12.4. *Let $\alpha \in \Phi$ and $x \in L_\alpha$, $y \in L_{-\alpha}$. Then*

$$[xy] = K(x, y)t_\alpha.$$

Proof. For $h \in H$,

$$\begin{aligned} K(h, [xy]) &= K([hx], y) && \text{by Lemma 8.5} \\ &= \alpha(h)K(x, y) \\ &= K(t_\alpha, h)K(x, y) \\ &= K(h, K(x, y)y_\alpha) \end{aligned}$$

Finally, K_H is non-degenerate, so $[xy] = K(x, y)t_\alpha$. □

Proposition 12.5.

(1) $t_\alpha = \frac{h_\alpha}{K(e_\alpha, e_{-\alpha})}$ *and* $h_\alpha = \frac{2t_\alpha}{K(t_\alpha, t_\alpha)}$
(2) $K(t_\alpha, t_\alpha)K(h_\alpha, h_\alpha) = 4$

Proof. For (1), note that $[e_\alpha, e_{-\alpha}] = h_\alpha$ implies the formula for t_α, using Proposition 12.4. Also, $\alpha(h_\alpha) = 2$, so using the formula for t_α

$$2 = K(t_\alpha, h_\alpha) = K(e_\alpha, e_{-\alpha})K(t_\alpha, t_\alpha).$$

Hence we get the formula for h_α, using Proposition 12.4.

Finally, (2) follows from (1). □

Proposition 12.6. *If $\alpha, \beta \in \Phi$ then*

(1) $K(h_\alpha, h_\beta) \in \mathbb{Z}$,
(2) $K(t_\alpha, t_\beta) \in \mathbb{Q}$.

Proof. For (1), recall that by Proposition 12.1,

$$K(h_\alpha, h_\beta) = \sum_{\gamma \in \Phi} \gamma(h_\alpha)\gamma(h_\beta)$$

which is an integer by Proposition 11.5.

To show (2), note that by Proposition 12.5,

$$K(t_\alpha, t_\beta) = K(h_\alpha, h_\beta) \cdot \frac{K(t_\alpha, t_\alpha)}{2} \frac{K(t_\beta, t_\beta)}{2} \in \mathbb{Q},$$

using part (1). □

We define a bilinear form on H^*: for $\theta_1, \theta_2 \in H^*$ we have $\theta_i = \theta_{h_i}$ for some $h_i \in H$, and we define

$$(\theta_1, \theta_2) = K(h_1, h_2).$$

As K_H is non-degenerate, so is $(\ , \ )$ on H^*.

For $\alpha, \beta \in \Phi$,

$$(\alpha, \beta) = K(t_\alpha, t_\beta).$$

By Proposition 12.2, $\text{Span}(\Phi) = H^*$, so there exists a basis $\alpha_1, \dots, \alpha_k$ of H^* with all $\alpha_i \in \Phi$.

Proposition 12.7. *If $\beta \in \Phi$ then*

$$\beta = \sum_{i=1}^{k} r_i \alpha_i \textit{ for } r_i \in \mathbb{Q}.$$

Proof. Let $\beta = \sum_{i=1}^{k} r_i \alpha_i$ for $r_i \in \mathbb{C}$. Then

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

In matrix form,

$$\begin{pmatrix} (\beta, \alpha_1) \\ \vdots \\ (\beta, \alpha_k) \end{pmatrix} = \underbrace{\begin{pmatrix} (\alpha_1, \alpha_1) & \dots & (\alpha_1, \alpha_k) \\ & \vdots & \\ (\alpha_k, \alpha_1) & \dots & (\alpha_k, \alpha_k) \end{pmatrix}}_{A} \begin{pmatrix} r_1 \\ \vdots \\ r_k \end{pmatrix}$$

As K_H is non-degenerate, A is invertible, and all entries of A are in $\mathbb{Q}$ by Proposition 12.6. Hence all the entries of A^{-1} are in $\mathbb{Q}$. Also all $(\beta, \alpha_j) \in \mathbb{Q}$. Hence all $r_i \in \mathbb{Q}$. □

Definition. Let E be the real span of $\alpha_1, \dots, \alpha_k \in H^*$, i.e.

$$E = \left\{ \sum_{i=1}^{k} r_i \alpha_i : r_i \in \mathbb{R} \right\}.$$

By Proposition 12.7,

(1) E does not depend on the choice of basis $\alpha_1, \dots, \alpha_k$,
(2) $\Phi \subseteq E$,
(3) $E = \text{Span}_{\mathbb{R}}(\Phi)$.

Proposition 12.8. *The bilinear form $(\ , \ )$ is a real-valued inner product on the vector space E.*

Proof. For $\alpha, \beta \in \Phi$, $(\alpha, \beta) = K(t_\alpha, t_\beta) \in \mathbb{R}$. Hence $(\ , \ )$ is real-valued on E.

Let $\theta \in E$, so $\theta = \theta_h$ for some $h \in H$. Then

$$\begin{aligned} (\theta_h, \theta_h) &= K(h,h) && \\ &= \sum_{\gamma \in \Phi} \gamma(h)^2 && \text{by Proposition 12.1} \\ &= \sum_{\gamma \in \Phi} K(t_\gamma, h)^2 && \text{by definition of } t_\gamma \\ &= \sum_{\gamma \in \Phi} (\gamma, \theta_h)^2 && \text{by definition of } (\ ,\ ) \end{aligned}$$

As $(\gamma, \theta_h) \in \mathbb{R}$, this shows that $(\theta_h, \theta_h) \geq 0$.

If $(\theta_h, \theta_h) = 0$ then $\gamma(h) = 0$ for all $\gamma \in \Phi$. Hence $h = 0$, so $\theta = 0$. □

13. Root systems

Let E be a finite dimensional real vector space with an inner product $(\ ,\ )$. For $0 \neq v \in E$, the *reflection* $s_v \colon E \to E$ is defined by

$$s_v(x) = x - \frac{2(v,x)}{(v,v)} v \text{ for all } x \in E.$$

Note that s_v sends $v \mapsto -v$ and fixes every vector in $v^\perp$.

Note. The reflection s_v preserves the inner product, i.e.

$$(s_v(x), s_v(y)) = (x,y) \text{ for all } x, y \in E.$$

(Showing this is left as an exercise.)

Notation. Write $\langle x, v \rangle = \frac{2(x,v)}{(v,v)}$ (linear in x, not in v).

Definition. A subset R of E is a *root system* if

(1) R is finite, $0 \notin R$, and $\mathrm{Span}(R) = E$,
(2) for $\alpha \in R$, the only scalar multiples of α in R are $\pm\alpha$,
(3) for $\alpha \in R$, the reflection s_α sends R to R, i.e. permutes the set R,
(4) for $\alpha, \beta \in R$,

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

We call the elements of R *roots*, and $\dim E$ the rank of R.

Example. The only root system of rank 1 is

Proposition 13.1. *Let L be a semisimple Lie algebra over $\mathbb{C}$, H a Cartan subalgebra, and root decomposition*

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

Let $E = \mathrm{Span}_{\mathbb{R}}(\Phi) \subseteq H^*$, with inner product $(\ ,\ )$ as in Chapter 12. Then Φ is a root system in E.

Proof. Axiom (1) is clear. Axiom (2) is Proposition 11.3 (2), and Sheet 5, Question 2.

To check axiom (3), let $\alpha, \beta \in \Phi$. Then

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha.$$

We claim that $\langle \beta, \alpha \rangle = \beta(h_\alpha)$. Indeed:

$$\begin{aligned} \beta(h_\alpha) &= K(t_\beta, h_\alpha) && \text{by definition} \\ &= K\left(t_\beta, \tfrac{2t_\alpha}{K(t_\alpha,t_\alpha)}\right) && \text{by Proposition 12.5} \\ &= \tfrac{2(\beta,\alpha)}{(\alpha,\alpha)} && \text{by definition} \\ &= \langle \beta, \alpha \rangle && \text{by definition} \end{aligned}$$

Hence $s_\alpha(\beta) = \beta - \beta(h_\alpha)\alpha \in \Phi$ by Proposition 11.5 (4). Finally, (4) follows from $\langle \beta, \alpha \rangle = \beta(h_\alpha) \in \mathbb{Z}$ by Proposition 11.5 (1). □

Examples.

(1) Let $L = \mathfrak{sl}(n, \mathbb{C})$. The Cartan subalgebra H consists of diagonal matrices in L, the root spaces are

$$\mathrm{Span}(E_{ij}) \text{ with roots } \epsilon_i - \epsilon_j \quad (i \neq j),$$

where

$$\epsilon_i \colon \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \mapsto a_i.$$

Here

$$E = \mathrm{Span}_{\mathbb{R}}(\Phi) = \left\{ \sum_{i=1}^n \lambda_i \epsilon_i : \sum_{i=1}^n \lambda_i = 0 \right\}$$

of dimension $n-1$. Inner product (rescaling) is the usual

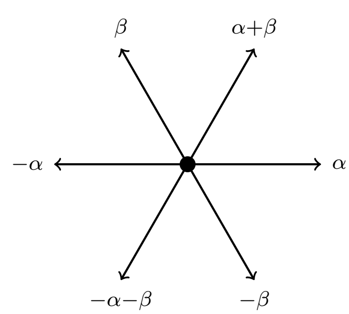
$$\left(\sum \lambda_i \epsilon_i, \sum \mu_i \epsilon_i \right) = \sum \lambda_i \mu_i.$$

(2) **Some rank 2 root systems.**

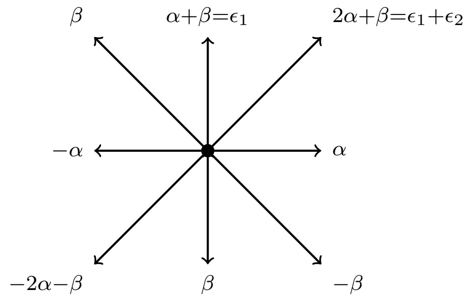
(a) Let $L = \mathfrak{sl}(3, \mathbb{C})$. Let $\alpha = \epsilon_1 - \epsilon_2$, $\beta = \epsilon_2 - \epsilon_3$. Then the root system

$$\Phi = \{\alpha, \beta, \alpha+\beta, -\alpha, -\beta, -\alpha-\beta\}.$$

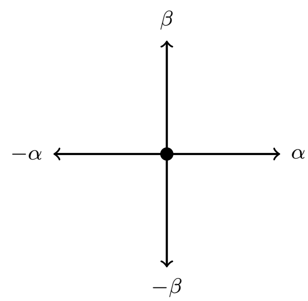
Also $(\alpha, \beta) = -1$, $(\alpha, \alpha) = 2$. So the angle between α and β is $\cos^{-1}\left(\frac{1}{2}\right) = \frac{2\pi}{3}$.



(b) Another rank 2 root system. Again, let ϵ_1, ϵ_2 be standard unit vectors in $\mathbb{R}^2$, and let $\alpha = \epsilon_2$, $\beta = \epsilon_1 - \epsilon_2$. The angle between α and β is $\cos^{-1}\left(\frac{-1}{\sqrt{2}}\right) = \frac{3\pi}{4}$.



(c) Another one



Definition. We say that root systems $R \subseteq E$, $R' \subseteq E'$ are *isomorphic* if there exists a vector space isomorphism $\phi\colon E \to E'$ such that

(1) $\phi(R) = R'$,
(2) $(\phi(\alpha), \phi(\beta)) = (\alpha, \beta)$ for any $\alpha, \beta \in R$.

Definition. A root system $R \subseteq E$ is *reducible* if $R = R_1 \cup R_2$ where $R_i \neq \emptyset$ and

$$(\alpha, \beta) = 0 \text{ for any } \alpha \in R_1, \beta \in R_2.$$

Otherwise, R is *irreducible.*

Example. In the examples above, (2)(c) is reducible, but (2)(a) and (2)(b) are irreducible.

Proposition 13.2. *Let L be a semisimple Lie algebra over $\mathbb{C}$ with root system of Φ. If Φ is irreducible, then L is simple.*

Proof. We have

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

Suppose L is not simple, so it has an ideal $I \neq 0, L$. Now, $[HI] \subseteq I$, and $\operatorname{ad} H$ is simultaneously diagonalizable on L, hence also on I.

Therefore, I has a basis of common eigenvectors for $\operatorname{ad} H$, so

$$I = H_1 \oplus \bigoplus_{\alpha \in \Phi_1} L_\alpha$$

where $H_1 \subseteq H$, $\Phi_1 \subseteq \Phi$. Similarly,

$$I^\perp = H_2 \oplus \bigoplus_{\alpha \in \Phi_2} L_\alpha.$$

As $L = I \oplus I^\perp$,

$$H = H_1 \oplus H_2,\ \Phi = \Phi_1 \cup \Phi_2,\ \Phi_1 \cap \Phi_2 = \emptyset.$$

If $\Phi_2 = \emptyset$, then $\Phi_1 = \Phi$, so I contains all L_α ($\alpha \in \Phi$), so also all $[L_\alpha L_{-\alpha}]$, which span H. Hence $I = L$, a contradiction. Hence $\Phi_i \neq \emptyset$ for $i = 1, 2$.

Finally, for $\alpha \in \Phi_1$, $\beta \in \Phi_2$,

$$[h_\beta e_\alpha] \in I \cap I^\perp = 0,$$

so

$$0 = \alpha(h_\beta) = \langle \alpha, \beta \rangle$$

by the proof of Proposition 13.1. Hence $(\alpha, \beta) = 0$, showing that Φ is reducible. □

Classification theorems (Killing, Cartan). The root system Φ depends on the choice of Cartan subalgebra H. However, we have the following theorem.

Theorem 13.3. *Let L be a semisimple Lie algebra over $\mathbb{C}$ with Cartan subalgebras H_1, H_2 and corresponding root systems Φ_1, Φ_2. Then $\Phi_1 \cong \Phi_2$.*

The proof is based on the fact that all Cartan subalgebras are *conjugate*, i.e. there exists

$$g \in \mathrm{Aut}(L) = \{x \in \mathrm{GL}(L) : x([ab]) = [x(a), x(b)] \text{ for all } a, b \in L\}$$

such that $g(H_1) = H_2$.

So every semisimple Lie algebra over $\mathbb{C}$ has a unique unique root system. The converse also holds.

Theorem 13.4. *For any root system Φ, there exists a unique (up to isomorphism) semisimple Lie algebra L over $\mathbb{C}$ with root system Φ.*

Uniqueness. We can specify the structure constants of L in terms of root systems Φ: there exists a *Chevalley* basis of L with the following structure constants. Recall, for $\alpha, \beta \in \Phi$, the α-string through β is

$$\beta - r\alpha, \ldots, \beta + q\alpha.$$

Then the Chevalley basis is h_α's, e_α's, and the structure constants are

$$[h_\alpha h_\beta] = 0$$
$$[h_\alpha e_\beta] = \beta(h_\alpha) e_\beta \text{ and } \beta(h_\alpha) = r - q$$
$$[e_\alpha e_{-\alpha}] = h_\alpha$$
$$[e_\alpha e_\beta] = \begin{cases} 0 & \text{if } \alpha + \beta \notin \Phi \cup \{0\} \\ \pm(q+1) e_{\alpha+\beta} & \text{if } \alpha + \beta \in \Phi \end{cases}$$

Existence. Given an irreducible root system (these are classified), one can construct a simple Lie algebra with that root system. The simple Lie algebras as

$$\begin{array}{rl} \text{(classical)} & \mathfrak{sl},\ \mathfrak{sp},\ \mathfrak{so}, \\ \text{(exceptional)} & \mathfrak{g}_2,\ \mathfrak{f}_2,\ \mathfrak{e}_6,\ \mathfrak{e}_7,\ \mathfrak{e}_8. \end{array}$$

Alternative approach to existence: use *Serre relations*.

Conclusion. Classification of simple Lie algebras is equivalent to classification of irreducible root systems.

14. Irreducible root systems

Let $R \subseteq E$ be a root system for a real inner product space E. Recall that for $\alpha, \beta \in R$,

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

Also,

$$(\alpha, \alpha) = ||\alpha||^2$$

(where $||\alpha||$ is the length of α), and

$$(\alpha, \beta) = ||\alpha|| ||\beta|| \cos \theta$$

where θ is the angle between α and β.

Proposition 14.1. *If $\alpha, \beta \in R$ and $\beta \neq \pm\alpha$, then*

$$\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = \frac{4(\alpha, \beta)^2}{(\alpha, \alpha)(\beta, \beta)} \in \{0, 1, 2, 3\}.$$

Proof. If θ is the angle between α and β,

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

We only have to note that it is not equal to 4; otherwise, $\cos^2 \theta = 1$, and $\theta = n\pi$, so $\beta = \pm\alpha$. □

Proposition 14.2. *Let $\alpha, \beta \in R$, $\beta \neq \pm\alpha$, and assume that $(\beta, \beta) \geq (\alpha, \alpha)$. The possibilities for $\langle \alpha, \beta \rangle$, $\langle \beta, \alpha \rangle$, θ are as follows*

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	$\cos \theta$	θ	$\frac{(\beta,\beta)}{(\alpha,\alpha)} = \frac{\langle \beta,\alpha \rangle}{\langle \alpha,\beta \rangle}$
0	0	0	$\frac{\pi}{2}$	
1	1	$\frac{1}{2}$	$\frac{\pi}{2}$	1
-1	-1	$-\frac{1}{2}$	$\frac{2\pi}{3}$	1
1	2	$\frac{1}{\sqrt{2}}$	$\frac{\pi}{4}$	2
-1	-2	$-\frac{1}{\sqrt{2}}$	$\frac{3\pi}{4}$	2
1	3	$\frac{\sqrt{3}}{2}$	$\frac{\pi}{6}$	3
-1	-3	$-\frac{\sqrt{3}}{2}$	$\frac{5\pi}{6}$	3

The proof is an easy calculation.

Proposition 14.3. *Let θ be the angle between $\alpha, \beta \in R$.*

(1) *If $\theta > \frac{\pi}{2}$, then $\alpha + \beta \in R$.*
(2) *If $\theta < \frac{\pi}{2}$, then $\alpha - \beta \in R$.*

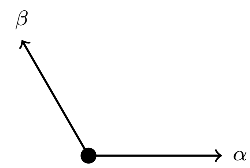
Proof. By axiom (3),

$$s_\beta = \alpha - \langle \alpha, \beta \rangle \beta \in R.$$

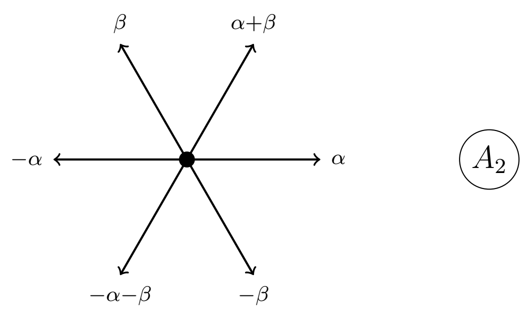
From Proposition 14.2, $\theta > \frac{\pi}{2}$ implies that $\langle \alpha, \beta \rangle = -1$, and $\theta < \frac{\pi}{2}$ implies that $\langle \alpha, \beta \rangle = 1$. □

Example (Classification of root systems of rank 2)**.** Let $R \subseteq \mathbb{R}^2$ be a rank 2 root system. Pick $\alpha, \beta \in R$ with $\beta \neq \pm\alpha$, and the angle θ between α and β as large as possible ($\theta \geq \frac{\pi}{2}$). By Proposition 14.2, the possibilities of θ are $\frac{2\pi}{3}$, $\frac{3\pi}{4}$, $\frac{5\pi}{6}$ or $\frac{\pi}{2}$.

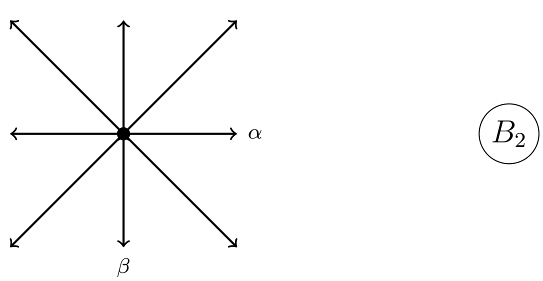
Take $\theta = \frac{2\pi}{3}$. Then α, β have the same length, and we have



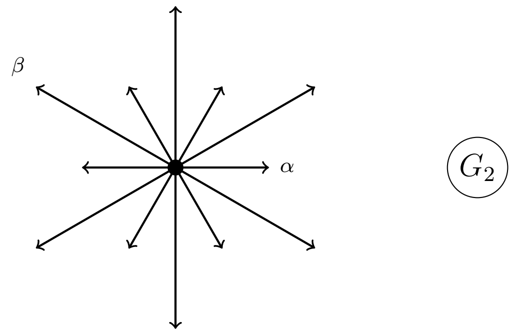
Apply reflections to get the root system



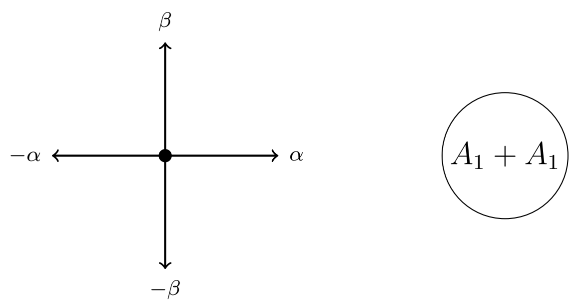
Take $\theta = \frac{3\pi}{4}$. Here, $(\beta, \beta) = 2(\alpha, \alpha)$, and we get



Take $\theta = \frac{5\pi}{6}$. Here, $(\beta, \beta) = 3(\alpha, \alpha)$, and we get



Take $\theta = \frac{\pi}{2}$. Here, we get



This root system is *reducible*.

Weyl groups. Recall that for a root system $R \subseteq E$ and a root $\alpha \in R$, we defined the reflection s_α by

$$s_\alpha(x) = x - \frac{2(\alpha, x)}{(\alpha, \alpha)}\alpha \text{ for all } x \in E.$$

Definition. The *Weyl group* $W(R)$ of a root system $R \subseteq E$ is

$$W(R) = \langle s_\alpha : \alpha \in R \rangle$$

a subgroup of $\mathrm{GL}(E)$.

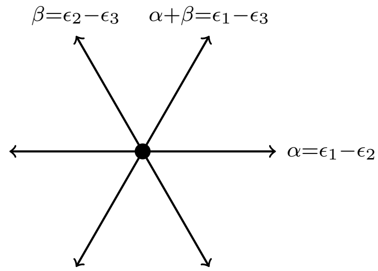
Proposition 14.4. *The Weyl group $W(R)$ is finite.*

Proof. By axioms (3), each reflection s_α gives a permutation of the finite set R. So we have a homomorphism

$$\phi \colon W(R) \to \mathrm{Sym}(R).$$

If $g \in \operatorname{Ker} \phi$, then $g(\alpha) = \alpha$ for all $\alpha \in R$, hence as $E = \mathrm{Span}(R)$, $g = 1$. Therefore, $\operatorname{Ker} \phi = 1$, so $W(R) \cong \operatorname{Im} \phi \leq \mathrm{Sym}(R)$. $\square$

Examples. $\underline{R = A_2}$



Here $W(A_2) = \langle s_\alpha, s_\beta, s_{\alpha+\beta}\rangle$. Action on basis $\epsilon_1, \epsilon_2, \epsilon_3$ of $\mathbb{R}^3$:

$$s_\alpha = s_{\epsilon_1-\epsilon_2} : \ \epsilon_1 \leftrightarrow \epsilon_2, \ \epsilon_3 \mapsto \epsilon_3$$

$$s_\beta : \epsilon_2 \leftrightarrow \epsilon_3, \ \epsilon_1 \mapsto \epsilon_1$$

$$s_{\alpha+\beta} : \epsilon_1 \leftrightarrow \epsilon_3, \epsilon_2 \mapsto \epsilon_2$$

Hence $W(A_2) \cong S_3$.

$\underline{R = A_{n-1}}$ Roots $\epsilon_i - \epsilon_j$ $(i \neq j)$ for $i, j \in \{1, \ldots, n\}$. Reflection $s_{\epsilon_i-\epsilon_j}$ sends $\epsilon_i \leftrightarrow \epsilon_j$, and fixes the other basis vectors. Hence $W(A_{n-1}) \cong S_n$.

Bases.

Definition. Let $R \subseteq E$ be a root system. We say a subset B of R is a *base* of R if

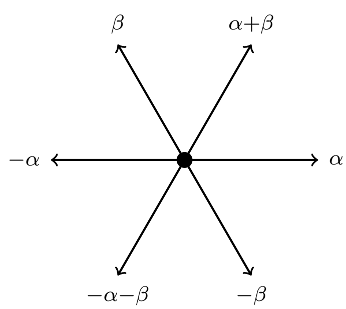
(1) B is a basis of E (as a vector space)
(2) for any $\beta \in R$,

$$\beta = \sum_{\alpha \in B} n_\alpha \alpha$$

where $n_\alpha \in \mathbb{Z}$, and either $n_\alpha \geq 0$ for all α or $n_\alpha \leq 0$ for all α.

We say $\beta \in R$ is a *positive root* (with respect to base B) if all $n_\alpha \geq 0$. Similarly, define *negative roots*.

Example. $\underline{R = A_2}$



A base is α, β (the positive roots are α, β, $\alpha+\beta$).

Another base is $\alpha, -\alpha-\beta$ (the positive roots are $\alpha, -\alpha-\beta, -\beta$).

Note, there exists $w \in W(A_2)$ sending $\{\alpha, \beta\}$ to $\{\alpha, -\alpha-\beta\}$. (Exercise)

Theorem 14.5.

(1) *Every root system has a base.*
(2) *If B, B' are bases of R, then there exists a unique $w \in W(R)$ such that $w(B) = B'$.*

Hence R has precisely $|W(R)|$ different basis.

Example. The root system A_2 has 6 different bases, all of the form $w(\{\alpha, \beta\})$ for $w \in W(A_2)$.

Proof of Theorem 14.5. Omitted. $\square$

Dynkin diagrams. Let R be a root system with a base B. Define the *Dynkin diagram* $\Delta = \Delta(R)$ of R (with respect to B) to be the following graph[5]

vertices: elements of B
edges: join α, β in B by $d_{\alpha\beta}$ edges, where $d_{\alpha\beta} = \langle \beta, \alpha \rangle \in \{0, 1, 2, 3\}$

If $d_{\alpha\beta} > 1$, then α, β have different lengths, and we draw an arrow from the longer to the shorter root.

Note. By Theorem 14.5, the diagram Δ does not depend on the choice of the base B.

Examples.

(1) A_{n-1}. Roots $\epsilon_i - \epsilon_j$ for $i \neq j$ in $\{1, \dots, n\}$. Here is a base:

$$\underbrace{\epsilon_1 - \epsilon_2}_{\alpha_1}, \ \underbrace{\epsilon_2 - \epsilon_3}_{\alpha_2}, \ \dots, \ \underbrace{\epsilon_{n-1} - \epsilon_n}_{\alpha_{n-1}}$$

This **is** a base, as for $i < j$,

$$\epsilon_i - \epsilon_j = \alpha_i + \alpha_{i+1} + \dots + \alpha_{j-1}.$$

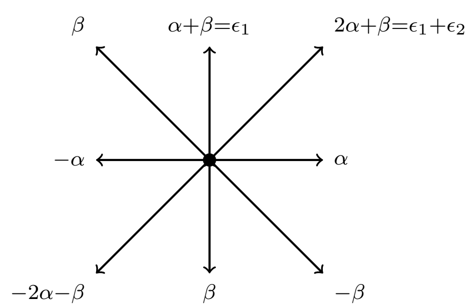
So the positive roots are $\epsilon_i - \epsilon_j$ for $i < j$ and the negative roots are $\epsilon_i - \epsilon_j$ for $i > j$.

Here,

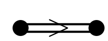
$$d_{\alpha_i \alpha_{i+1}} = 1$$
$$d_{\alpha_i \alpha_j} = 0 \text{ if } j \neq i \pm 1.$$

So the Dynkin diagram is

(2) B_2.

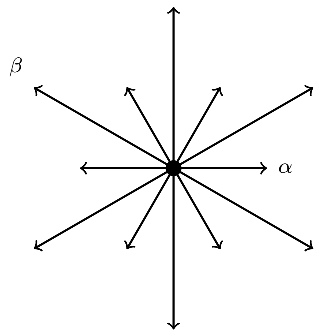


It has base α, β, and its Dynkin diagram is

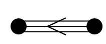


[5] A graph where we allow multiple edges between vertices.

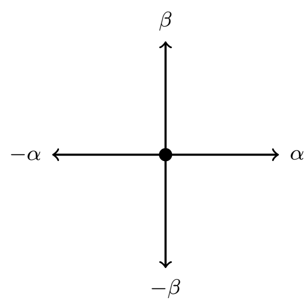
(3) G_2.



It has base α, β, and its Dynkin diagram is



(4) $A_1 + A_1$.



It has base α, β, and its Dynkin diagram is

Proposition 14.6. *A root system R is irreducible if and only if its Dynkin diagram is connected.*

Proof. Sheet 5, Question 3. $\square$

Simplicity of classical Lie algebras.

Theorem 14.7. *Let $n \geq 2$. Then the classical Lie algebras, $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{so}(n, \mathbb{C})$, $\mathfrak{sp}(n, \mathbb{C})$ (n even) are simple, apart from $\mathfrak{so}(2, \mathbb{C})$ and $\mathfrak{so}(4, \mathbb{C})$.*

Idea.

(1) Use the next proposition to show L is semisimple.
(2) Use the root system and Proposition 14.6 to show L is simple.

Proposition 14.8. *Let L be a finite-dimensional Lie algebra over $\mathbb{C}$, with $Z(L) = 0$. Assume H is a Cartan subalgebra and*

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

where $\Phi \subseteq H^* \setminus \{0\}$. Suppose

(1) $\dim L_\alpha = 1$ *for all* $\alpha \in \Phi$

(2) *if* $\alpha \in \Phi$ *then* $-\alpha \in \Phi$
(3) $[[L_\alpha, L_{-\alpha}], L_\alpha] \neq 0$ *for all* $\alpha \in \Phi$.

Then L is semisimple.

Proof. Suppose for a contradiction that L is not semisimple. Then L has a soluble ideal, so it has an abelian ideal $I \neq 0$. Now $[HI] \subseteq I$, and $\operatorname{ad} H$ acts diagonalizably on L, hence also on I.

Hence I is a sum of eigenspaces of $\operatorname{ad} H$, so

$$I = (I \cap H) \oplus \sum_{\alpha \in \Phi} (I \cap L_\alpha).$$

If $I \cap L_\alpha \neq 0$ for some α, then $L_\alpha \subseteq I$ by (1), so

$$[[L_\alpha, L_{-\alpha}], L_\alpha] \subseteq [[I, L_{-\alpha}], I] \subseteq [I, I] = 0,$$

a contradiction with (3).

Therefore, $I = I \cap H$, i.e. $I \subseteq H$. As $Z(L) = 0$, there exists $\alpha \in \Phi$ such that $[IL_\alpha] \neq 0$. But

$$[IL_\alpha] \subseteq I \cap [HL_\alpha] \subseteq I \cap L_\alpha = 0,$$

a contradiction. □

Proof of Theorem 14.7. We just do the case $L = \mathfrak{sl}(n, \mathbb{C})$. The other cases $\mathfrak{so}$, $\mathfrak{sp}$ are Sheet 5, Question 7.

We use Proposition 14.8. First, the case $Z(L) = 0$ was Sheet 3, Question 3. Otherwise, we have a Cartan subalgebra H of diagonal matrices, and root space decomposition

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

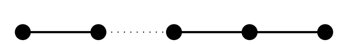
We check assumptions (1)–(3) of Proposition 14.8:

(1) the root spaces L_α are of the form $\operatorname{Span}(E_{ij})$, 1-dimensional,
(2) $\epsilon_i - \epsilon_j \in \Phi$ implies that $\epsilon_j - \epsilon_i \in \Phi$
(3) for $L_\alpha = \operatorname{Span}(E_{ij})$, we have $L_{-\alpha} = \operatorname{Span}(E_{ji})$, and check

$$[[E_{ij}, E_{ji}], E_{ij}] = 2E_{ij} \neq 0.$$

Hence L is semisimple by Proposition 14.8.

Finally, the Dynkin diagram of L is



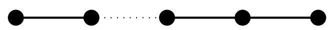
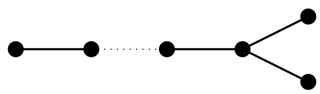
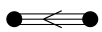
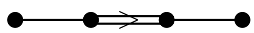
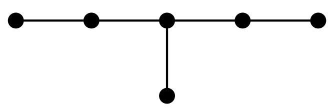
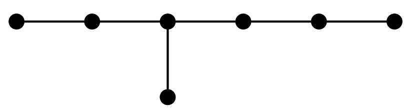
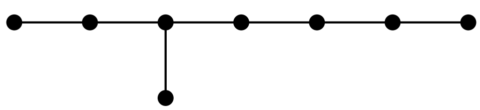
which is connected, so L is simple by Propositions 13.2 and 14.6. □

Classification of irreducible root systems.

Theorem 14.9. *Let $R \subseteq E$ and $R' \subseteq E'$ be roots, and suppose R, R' have the same Dynkin diagram. Then $R \cong R'$.*

So we just need to classify the possible Dynkin diagrams.

Theorem 14.10. *The Dynkin diagrams of the irreducible root systems are*

Dynkin diagram	*Notation*	*Corresponding simple Lie algebra*
	A_n ($n \geq 1$)	$\mathfrak{sl}(n+1, \mathbb{C})$
	B_n ($n \geq 2$)	$\mathfrak{so}(2n+1, \mathbb{C})$
	C_n ($n \geq 2$)	$\mathfrak{sp}(2n, \mathbb{C})$
	D_n ($n \geq 3$)	$\mathfrak{so}(2n, \mathbb{C})$
	G_2	$\mathfrak{g}_2$
	F_4	$\mathfrak{f}_2$
	E_6	$\mathfrak{e}_6$
	E_7	$\mathfrak{e}_7$
	E_8	$\mathfrak{e}_8$

In particular, each of these Dynkin diagram shows the existence of the Lie algebra on the right. However, these Lie algebras are actually not easy to construct.

We will prove Theorem 14.9. For that sake, we need the following proposition.

Proposition 14.11. *Let $R \subseteq E$ be a root system, with base B. Define*

$$W_0 = \langle s_\alpha : \alpha \in B \rangle \leq W(R).$$

If $\beta \in R$, then there exists $\alpha \in B$, $w \in W_0$ such that $w(\alpha) = \beta$ (i.e. $W_0(B) = R$).

Proof. Suppose $\beta \in R^+$ (positive root with respect to B), so

$$\beta = \sum_{\gamma \in B} k_\gamma \gamma, \ k_\gamma \geq 0.$$

Define

$$\mathrm{ht}(\beta) = \sum_{\gamma \in B} k_\gamma.$$

We proceed by induction on $\mathrm{ht}(\beta)$. If $\mathrm{ht}(\beta) = 1$, then $\beta \in B$ and take $\alpha = \beta$, $w = 1$.

Now, assume $\mathrm{ht}(\beta) \geq 2$, By axiom (2) of a root system, at least two k_γ's are nonzero.

We first claim that there exists $\gamma_0 \in B$ such that $(\beta, \gamma_0) > 0$. Otherwise,

$$(\beta, \beta) = \sum_{\gamma \in B} k_\gamma (\beta, \gamma) \leq 0,$$

so $(\beta, \beta) = 0$, hence $\beta = 0$, contradicting $\beta \in R$.

We now claim that $s_{\gamma_0}(\beta) \in R^+$. Well,

$$s_{\gamma_0}(\beta) = \beta - \langle \beta, \gamma_0 \rangle \gamma_0,$$

so $s_{\gamma_0}(\beta)$ has at least one coefficient $k_\gamma > 0$, and hence $s_{\gamma_0}(\beta) \in R^+$.

By the previous two claims,

$$\mathrm{ht}(s_{\gamma_0}(\beta)) = \mathrm{ht}(\beta) - \langle \beta, \gamma_0 \rangle < \mathrm{ht}(\beta).$$

By the inductive hypothesis, there exists $\alpha \in B$, $w \in W_0$ such that

$$w(\alpha) = s_{\gamma_0}(\beta).$$

Then $s_{\gamma_0} w \in W_0$ and it sends α to β. □

Proof of Theorem 14.9. We have $R \subseteq E$, $R' \subseteq E'$ with bases

$$B = \{\alpha_1, \ldots, \alpha_n\}, \ B' = \{\alpha'_1, \ldots, \alpha'_n\}$$

such that

$$\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle \text{ for all } i, j$$

(same Dynkin diagram).

Define a linear map $\phi \colon E \to E'$ by

$$\phi(\alpha_i) = \alpha'_i \text{ for all } i.$$

We need to show that $\phi(R) = R'$. (This will show that $R \cong R'$.)

We use Proposition 14.11 to obtain

$$\{w_0(\alpha) : \alpha \in B, w_0 \in W_0\} = R.$$

Now

$$\phi(s_{\alpha_i}(\alpha_j)) = \phi(\alpha_j - \langle \alpha_j, \alpha_i \rangle \alpha_i) = \alpha'_j - \langle \alpha'_j, \alpha'_i \rangle \alpha'_i = s_{\alpha'_i}(\alpha'_j) \in R'$$

by axiom (3). Hence for $w_0 \in W_0$,

$$\phi(w_0(\alpha)) \in R'.$$

So $\phi(R) \subseteq R'$.

The same argument for ϕ^{-1} gives $\phi^{-1}(R') \subseteq R$. Hence $\phi(R) = R'$. □