LIE GROUPS AND LIE ALGEBRAS

ANDREW WOOD

1. Introduction

1.1. **Examples.** An example of a Lie group is the general linear group $G = \mathrm{GL}(d, \mathbb{R})$, consisting of all invertible $d \times d$ matrices over $\mathbb{R}$. Indeed, G is closed under multiplication, and satisfies the group axioms. Further, multiplication is *smooth*, i.e., subset of $\mathbb{R}^{d\times d}$. Inversion is smooth as well.

Another example includes upper triangular matrices. One easily computes

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+x & c+xb+z \\ 0 & 1 & y+b \\ 0 & 0 & 1 \end{pmatrix}$$

to observe these matrices are indeed closed under multiplication.

Now, another example is $\mathbb{R}^3$ with $(x, y, z) * (a, b, c) = (x + a, y + b, z + c + xb)$, which is smooth on all entries.

A few others are the rotational group (i.e., SO(3)) and orthogonal transformations/group.

1.2. **Exponential map.** Let $X \in M^{d\times d}(\mathbb{R})$ be a matrix. We define $e^X = \sum_{n=0}^{\infty} \frac{1}{n!} X^n$. Observe $e^X \cdot e^{-X} = I = e^{-X} \cdot e^X$. So, $e^X \in \mathrm{GL}(d, \mathbb{R})$. Also, $e^0 = I$, where 0 is the zero matrix. Note that $x \mapsto e^X$ "close enough to the identity" is surjective, and "close enough to zero matrix" is one-to-one.

If X and Y are close to the zero matrix, e^X and e^Y are close to the identity, so $e^X e^Y = e^Z$ where $Z = X+Y+\frac{1}{2}[X, Y]+\ldots$ and $[X, Y] = XY - YX$ (commutator).

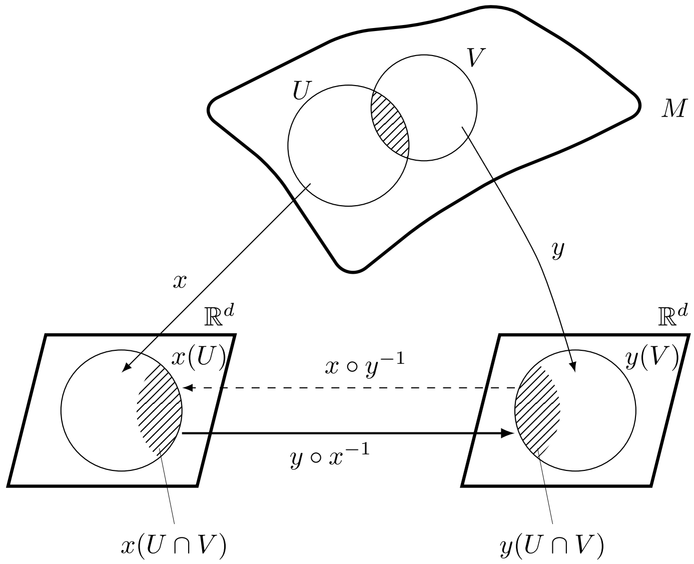
2. Differential Geometry

2.1. **Manifolds.** Fix M a Hausdorff space (one might also impose M to be second countable, but we do not), and $d \in \mathbb{N}_0$, which will be the dimension. If $d = 0$, M consists of disjoint singletons. We now provide terminology to properly define a manifold.

A *local chart* is a pair (U, x), where $U \subset M$ is open and $x : U \to \overbrace{x(U)}^{\subset \mathbb{R}^d}$ is a homeomorphism with $x(U)$ open in $\mathbb{R}^d$.

Let (U, x) and (V, y) be charts. The charts (U, x) and (V, y) are called C^∞-*related* if the maps $y \circ x^{-1} : x(U \cap V) \to \mathbb{R}^d$ and $x \circ y^{-1} : y(U \cap V) \to \mathbb{R}^d$ are both C^∞-maps (i.e., all partial derivatives exist).

A function f is called *real analytic* if every point has a nhood in which there is a power series that converges to the function value. Notation: f is called a C^ω-map. Every C^ω-function is C^∞, but converse is false. Similar definitions for two charts being C^ω-*related*. For our purposes, we refer to C^k-maps, where $k \in \{\infty, \omega\}$.



An *atlas* on M is a collection of charts $\mathscr{A} = \{(U_\alpha, x_\alpha) \mid \alpha \in I\}$ such that $\bigcup_{\alpha \in I} U_\alpha = M$. Now, let $k \in \{\infty, \omega\}$. An atlas $\mathscr{A}$ is a C^k-*atlas* if for all $\alpha, \beta \in I$ the charts (U_α, x_α) and (U_β, x_β) are C^k-related. If $\mathscr{A}$ is a C^k-atlas, then all charts (U_α, x_α) with $\alpha \in I$ are called *smooth*.

If $\mathscr{A}$ and $\mathscr{B}$ are C^k-atlases, then $\mathscr{A} \leq \mathscr{B}$ if $\mathscr{A} \subset \mathscr{B}$. This is clearly a partial order. A C^k-*structure* on M is a maximal C^k-atlas on M, with respect to this partial order.

Proposition 2.1. *Let $\mathscr{A} = \{(U_\alpha, x_\alpha) \mid \alpha \in I\}$ be a C^k-atlas on M. Define $\mathscr{B}$ as the collection of all charts on M that are C^k-related to each chart in $\mathscr{A}$. Then $\mathscr{B}$ is the unique C^k-structure on M containing $\mathscr{A}$.*

Proof. Let $\mathscr{U}$ be the collection of C^k-atlases on M which contain $\mathscr{A}$. We claim $\mathscr{B} = \bigcup \mathscr{U}$. Clearly, $\bigcup \mathscr{U} \subset \mathscr{B}$. It suffices to show $\mathscr{B}$ is a C^k-atlas to prove our claim (it is obvious $\mathscr{B}$ is an atlas containing $\mathscr{A}$). Observe once we do so, uniqueness and maximality are established by our claim, and we are done. Now, suppose (U, x) and (V, y) are charts in $\mathscr{B}$, and that $p \in U \cap V$. There exists $(W, z) \in \mathscr{A}$ such that $p \in W$, since $\mathscr{A}$ is an atlas. By definition, (U, x) and (V, y) are both C^k-related to (W, z). Hence, it follows that $x \circ y^{-1}\big|_{y(U \cap V \cap W)} = x \circ z^{-1} \circ z \circ y^{-1}$ is C^k (implying $x \circ y^{-1}$ is C^k at $y(p)$), since z is a homeomorphism (in particular, a bijection) and both $x \circ z^{-1}$, $z \circ y^{-1}$ are C^k-maps (recall composition of C^k-maps are C^k). It follows that $x \circ y^{-1}$ is C^k, and similarly $y \circ x^{-1}$ (and thus the result follows). $\square$

Example 2.2. The *torus* $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ is the circle in complex plane (not to be confused with $S^1 \times S^1$). Let $U = (-\pi, \pi)$, $x(t) = e^{it}$. Can get everything except -1 on circle. Could take $V = (36\pi, 38\pi)$ and $y(t) = e^{it}$ to again get everything except a single point. We then take maximal atlas generated by (U, x) and (V, y) to get structure on $\mathbb{T}$.

A C^k-*manifold* $(M, \mathscr{A})$ is a Hausdorff space M together with a C^k-structure $\mathscr{A}$ (i.e., maximal C^k-atlas). Shortly: M is a manifold. We say d is the *dimension*.

Example 2.3. $M = \mathbb{R}$, $U = \mathbb{R}$ and $x(t) = t$. (U, x) is a chart. $\mathscr{A} = \{(U, x)\}$ is a C^k-atlas. $\widetilde{\mathscr{A}}$ is the structure generated by $\mathscr{A}$.

$V = \mathbb{R}$, $y(t) = t^3$. y is a homeomorphism, so (V, y) is a chart. $\mathscr{B} = \{(V, y)\}$ is a C^k-atlas. $\widetilde{\mathscr{B}}$ structure generated by $\mathscr{B}$.

$\left(\mathbb{R}, \widetilde{\mathscr{A}}\right)$ is a C^k-manifold and $\left(\mathbb{R}, \widetilde{\mathscr{B}}\right)$ is a C^k-manifold. These are different manifolds, since (U, x) and (V, y) are not C^∞-related. For $x \circ y^{-1}$ is not C^∞ (not differentiable at zero).

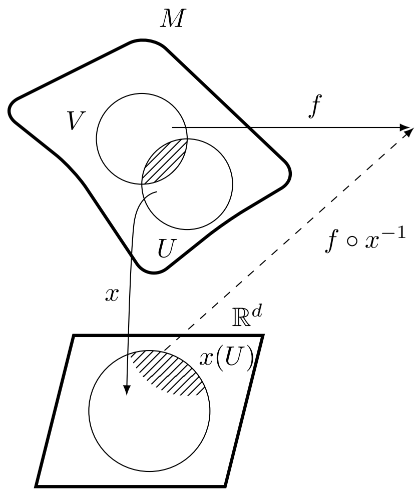
Let $(M, \mathscr{A})$ be a manifold. Let Ω be an open subset of M. The *induced atlas* on Ω is the structure on Ω generated by the atlas $\{(U \cap \Omega, x\,|_{U\cap\Omega}) \mid (U, x) \in \mathscr{A}\}$, so Ω becomes a manifold.

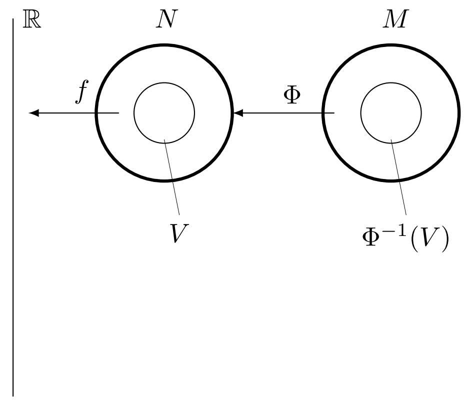
Example 2.4. $M = \mathbb{R}^d$, $U = \mathbb{R}^d$, $x : U \to \mathbb{R}^d$ by $x : p \mapsto p$. (U, x) chart. Take structure generated by $\{(U, x)\}$. We always consider the manifold $\mathbb{R}^d$ with this structure.

2.2. **Differentiable functions.** Let $(M, \mathscr{A})$ be a manifold. Let $V \subset M$ open and $f : V \to \mathbb{R}$, a function. Then f is called a C^k*-function* if for all $(U, x) \in \mathscr{A}$ one has that the function $f \circ x^{-1} : x(U \cap V) \to \mathbb{R}$ is a C^k-function.

Lemma 2.5. *Suppose $\mathscr{A}$ is generated by C^k-atlas $\mathscr{B}$. Then f is a C^k-function iff for all $(U, x) \in \mathscr{B}$ the function $f \circ x^{-1} : x(U \cap V) \to \mathbb{R}$ is C^k.*

Proof. The forward direction is trivial. Conversely, suppose for each $(U, x) \in \mathscr{B}$ the function $f \circ x^{-1} : x(U \cap V) \to \mathbb{R}$ is C^k. Suppose $(U, x) \in \mathscr{A}$, and let $p \in U$. Then there exists $(W, y) \in \mathscr{B}$ such that $p \in W$. Since $f \circ y^{-1} : y(V \cap W) \to \mathbb{R}$ and $y \circ x^{-1} : x(U \cap W) \to \mathbb{R}^d$ are C^k-maps, $f \circ x^{-1}\,|_{x(U\cap V\cap W)} = f \circ y^{-1} \circ y \circ x^{-1}$ is a C^k-map. That is to say, $f \circ x^{-1}$ is C^k at point p; therefore, $f \circ x^{-1}$ is clearly a C^k-function. $\square$





If $f : V \to \mathbb{R}$ is a C^k-map and $p \in V$, then f is called a C^k*-map about* p. Notation: all C^k-maps about p are $\mathcal{F}_p = \mathcal{F}_p M = \mathcal{F}_p(M) = \mathcal{F}(M, p)$. Two C^k maps about p; we can add, scalar multiplication, constant function on full manifold is zero, but not a vector space due to the domain problem for inverse (always restrict domain when composing, but zero is defined on full manifold).

Example 2.6. Let (U, x) be a smooth chart. Let $i \in \{1, \ldots, d\}$. Define projection map $\pi_i : \mathbb{R}^d \to \mathbb{R}$ by $\pi_i(p_1, \ldots, p_d) = p_i$. Then $\pi_i \circ x : U \to \mathbb{R}$ is a C^k-map.

Note. In terms of Lie groups want $g \mapsto g^{-1}$ to be smooth (where $G \to G$).

Let $(M, \mathscr{A})$ and $(N, \mathscr{B})$ be two manifolds. Let $\Phi : M \to N$ be a function. Then Φ is called a C^k*-map* if for every C^k-function $f : V \to \mathbb{R}$, with $V \subset N$ open in N the composition $f \circ \Phi : \Phi^{-1}(V) \to \mathbb{R}$ is a C^k-function. Since we only defined C^k-functions on open sets, $\Phi^{-1}(V)$ must be open. In particular, this implies Φ is continuous.

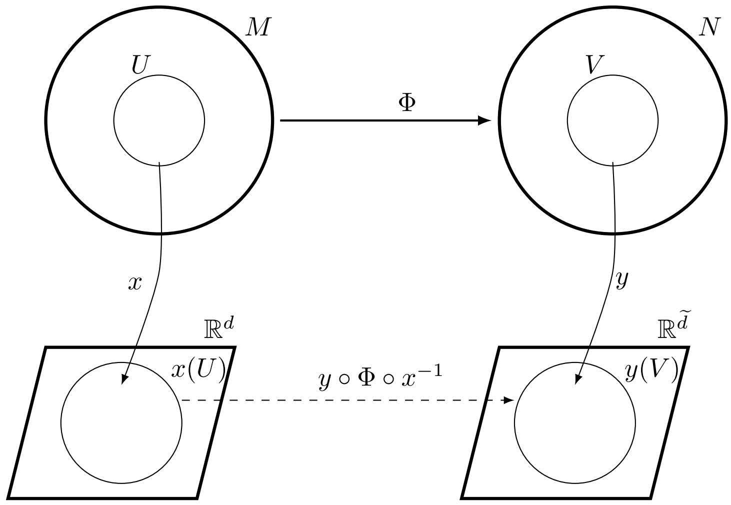
Note. A C^k-map is a continuous map from M into N.

Lemma 2.7. *Let $\Phi : M \to N$. The following are equivalent.*

(1) *Φ is a C^k-map.*
(2) *Φ is continuous and for all $(U, x) \in \mathscr{A}$ and $(V, y) \in \mathscr{B}$, the map $y \circ \Phi \circ x^{-1} : x\big(U \cap \Phi^{-1}(V)\big) \to y(V)$ is C^k (suffices for generating atlas).*

Proof. ((1) $\Longrightarrow$ (2)). Suppose Φ is a C^k-map. Suppose V is open in N, and that $f : V \to \mathbb{R}$ is a C^k-function. Then $f \circ \Phi : \Phi^{-1}(V) \to \mathbb{R}$ is a C^k-function, implying $\Phi^{-1}(V)$ is open in M. Hence, Φ is continuous. Now, suppose $(U, x) \in \mathscr{A}$ and $(V, y) \in \mathscr{B}$. Then $\pi_i \circ y : V \to \mathbb{R}$ is a C^k-map, and consequently $\pi_i \circ y \circ \Phi : \Phi^{-1}(V) \to \mathbb{R}$ is a C^k-function. Hence, $\pi_i \circ y \circ \Phi \circ x^{-1}$ is a C^k-function. Since $y \circ \Phi \circ x^{-1}$ is coordinate-wise C^k, it follows that it is C^k.

((2) $\Longrightarrow$ (1)). Now, suppose $f : V \to \mathbb{R}$ is a C^k-function, with V open subset of N. Consider $f \circ \Phi : \Phi^{-1}(V) \to \mathbb{R}$. We aim to show $f \circ \Phi$ is a C^k-function. To this end, suppose $p \in \Phi^{-1}(V)$. Since $\mathscr{A}$ is an atlas, there exists chart (U, x) such that $p \in U$. We may assume without loss of generality $U \subseteq \Phi^{-1}(V)$. Since $\mathscr{B}$ is an atlas, there exists chart (W, y) such that $\Phi(p) \in W$. Since f is a C^k-function, $f \circ y^{-1} : y(W) \to \mathbb{R}$ is C^k. Also, $y \circ \Phi \circ x^{-1}$ is C^k by assumption; so their composition $f \circ \Phi \circ x^{-1}$ is C^k around a nhood of p. Hence, $f \circ \Phi \circ x^{-1}$ is C^k. Thus, $f \circ \Phi$ is a C^k-function, which proves Φ is a C^k-map. $\square$



Let M and N be two manifolds. A C^k-*diffeomorphism* is a bijection $\Phi : M \to N$ such that both Φ and Φ^{-1} are C^k-maps. We then say M and N are *diffeomorphic*.

2.3. **Tangent space.** A *tangent vector* at point p of a C^k-manifold M is a function $v : \mathcal{F}_p \to \mathbb{R}$ such that

$$\begin{aligned} v(f+g) &= v(f) + v(g) \\ v(\lambda f) &= \lambda v(f) \\ v(f \cdot g) &= v(f) \cdot g(p) + f(p) \cdot v(g), \end{aligned}$$

for each $f, g \in \mathcal{F}_p$ and $\lambda \in \mathbb{R}$. The final property is called the *Leibniz rule/property.* Let T_pM be the vector space of all tangent vectors at p, which is called the *tangent space.* Let $U \subset M$ open such that $p \in U$. Then the smooth function $1_U : U \to \mathbb{R}$ defined by $1_U(q) = 1$ for each $q \in U$ is in $\mathcal{F}_p$. Observe

$$v(1_U) = v(1_U \cdot 1_U) = v(1_U) \cdot 1_U(p) + 1_U(p) \cdot v(1_U) = 2v(1_U),$$

implying $v(1_U) = 0$.

Let $f : V \to \mathbb{R}$ function with $V \subset M$ open and $f \in \mathcal{F}_p$. Then

$$v(f)_{U\cap V} = v(f \cdot 1_U) = v(f)\overbrace{1_U(p)}^{1} + f(p)\overbrace{v(1_U)}^{0} = v(f),$$

so restricting f to smaller open sets does not change anything w.r.t. tangent.

Let $I \subset \mathbb{R}$ open interval and $t_0 \in I$. Let $\gamma : I \to M$ be a C^k-map. Define $\dot\gamma(t_0) : \mathcal{F}_p \to \mathbb{R}$,

$$(\dot\gamma(t_0))(f) = \frac{d}{dt}(f \circ \gamma)(t) \mid_{t=t_0} .$$

Then $\dot\gamma(t_0)$ is a tangent vector. Clearly linear, and satisfies Leibniz rule due to product rule.

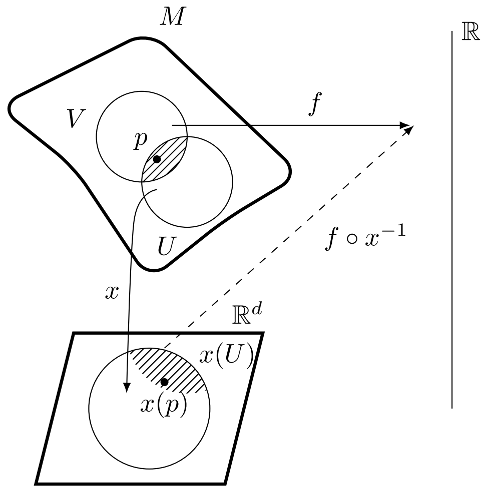
Let (U, x) be a smooth chart on M. Let $v \in \{1, \dots, d\}$ and $p \in U$. Define

$$\left.\frac{\partial}{\partial x^i}\right|_p : \mathcal{F}_p \to \mathbb{R}$$

by

$$\begin{aligned}\left.\frac{\partial}{\partial x^i}\right|_p (f) &= D_i\big(f \circ x^{-1}\big) \mid_{x(p)} \\ &= \left.\frac{d}{dt}\right|_{t=0} \big(f \circ x^{-1}\big)(x(p) + te_i) \\ &= \left.\frac{d}{dt}\right|_0 (f \circ \gamma)(t),\end{aligned}$$

where $\gamma(t) = x^{-1}(x(p) + te_i)$ and e_i standard basis vector for $\mathbb{R}^d$. Then $\gamma(0) = x^{-1}(x(p)) = p$. D_i ordinary partial derivative on $\mathbb{R}^d$, and $\frac{\partial}{\partial x^i} \mid_p$ tangent vector.



Theorem 2.8. *Let (U, x) be a smooth chart with $p \in U$. Then $\left\{\frac{\partial}{\partial x^1} \mid_p, \dots, \frac{\partial}{\partial x^d} \mid_p\right\}$ is a basis for T_pM. So $\dim T_pM = d$.*

Proof. First show linear independence. Let $i, j \in \{1, \dots, d\}$. Then

$$\begin{aligned}\left.\frac{\partial}{\partial x^i}\right|_p (\pi_j \circ x) &= D_i\big((\pi_j \circ x) \circ x^{-1}\big) \mid_{x(p)} \\ &= D_i(\pi_j) \mid_{x(p)} = \begin{cases} 0 & \text{if } i \neq j; \\ 1 & \text{if } i = j; \end{cases}\end{aligned}$$

which establishes independence. Remains to show spanning.

Let $v \in T_pM$. Define $\lambda_i = v(\pi_i \circ x) \in \mathbb{R}$ for all $i \in \{1, \ldots, d\}$. We claim $v = \sum_{i=1}^d \lambda_i \frac{\partial}{\partial x^i}\Big|_p$. Proof of claim: Let $f \in \mathcal{F}_p$; $f : V \to \mathbb{R}$, V open and $p \in V$. W.l.o.g., $V = U$, $x(p) = 0$ and $x(U) = B(0, \epsilon)$ for some $\epsilon > 0$. Let $\xi \in B(0, \epsilon)$. Then

$$\int_0^1 \frac{d}{dt}\big(f \circ x^{-1}\big)(t\xi)dt = \big(f \circ x^{-1}\big)(\xi) - \big(f \circ x^{-1}\big)(0),$$

so

$$\begin{aligned}
\big(f \circ x^{-1}\big)(\xi) &= \big(f \circ x^{-1}\big)(0) + \int_0^1 \frac{d}{dt}\big(f \circ x^{-1}\big)(t\xi)dt \\
&= f(p) + \int_0^1 \sum_{i=1}^d \xi_i \big(D_i\big(f \circ x^{-1}\big)\big)(t\xi)dt \\
&= f(p) + \sum_{i=1}^d \pi_i(\xi) \int_0^1 \big(D_i\big(f \circ x^{-1}\big)\big)(t\xi)dt,
\end{aligned}$$

and we let $h_i : B(0, \epsilon) \to \mathbb{R}$ be the C^k-function defined by $h_i(\xi) = \int_0^1 \big(D_i\big(f \circ x^{-1}\big)\big)(t\xi)dt$. Hence, $f \circ x^{-1} = f(p) + \sum_{i=1}^d \pi_i \cdot h_i$. That is,

$$f = f(p) + \sum_i^d (\pi_i \circ x) \cdot (h_i \circ x).$$

Define $g_i = h_i \circ x : U \to \mathbb{R}$. Then $g_i \in \mathcal{F}_p$. Then $f = f(p) + \sum_i^d (\pi_i \circ x) \cdot g_i$. Let $j \in \{1, \ldots, d\}$. Then

$$\frac{\partial}{\partial x^j}\bigg|_p (f) = \frac{\partial}{\partial x^j}\bigg|_p (f(p)1_U) + \sum_{i=1}^d \frac{\partial}{\partial x^j}\bigg|_p (\pi_i \circ x) \cdot g_i(p) + \sum_{i=1}^d (\pi_i \circ x)(p) \frac{\partial}{\partial x^j}\bigg|_p (g_i),$$

and thus

$$\frac{\partial}{\partial x^j}\bigg|_p (f) = g_j(p),$$

since $\frac{\partial}{\partial x^j}\big|_p (f(p)1_U) = f(p) \frac{\partial}{\partial x^j}\big|_p (1_U) = f(p) \cdot 0 = 0$, $\frac{\partial}{\partial x^j}\big|_p (\pi_i \circ x) \cdot g_i(p) = \delta_{ij} \cdot g_i$, and $(\pi_i \circ x)(p) = 0$ because $x(p) = 0$. Now,

$$v(f) = v(f(p)1_U) + \sum_{i=1}^d v(\pi_i \circ x) \cdot g_i(p) + \sum_{i=1}^d (\pi_i \circ x)(p) \cdot v(g_i) = \sum_{i=1}^d \lambda_i \frac{\partial}{\partial x^i}\bigg|_p (f).$$

□

Note. We will use $v = \sum_{i=1}^d v(\pi_i \circ x) \cdot \frac{\partial}{\partial x^i}\big|_p$ numerous times. We also consider $w = \sum_i w(\pi_i \circ y) \frac{\partial}{\partial y_i} |_q$, with reference to below.

Let M, N be C^k-manifolds. Let $\Phi : M \to N$ be a C^k-map. Let $p \in M$. Define $\Phi_{*p} : T_pM \to T_{\Phi(p)}N$, $v \in T_pM$ by $(\Phi_{*p}(v))(f) := v(f \circ \Phi)$. Then Φ_{*p}, which is called the *differential* of Φ at p, is linear. Let (U, x) be a smooth chart, with $p \in U$; and let (V, y) be a smooth chart, with $\Phi(p) \in V$. By Theorem 2.8, $\big\{\frac{\partial}{\partial x^1} |_p, \ldots, \frac{\partial}{\partial x^d} |_p\big\}$ is a basis for T_pM and $\Big\{\frac{\partial}{\partial y^1} |_{\Phi(p)}, \ldots, \frac{\partial}{\partial y^d} |_{\Phi(p)}\Big\}$ is a basis for $T_{\Phi(p)}N$. Then the matrix element at place i, j of Φ_{*p} with respect to the two given bases is given by,

$$\left(\Phi_{*p}\left(\frac{\partial}{\partial x^j}\bigg|_p\right)\right)(\pi_i \circ y) = \frac{\partial}{\partial x^j}\bigg|_p (\pi_i \circ y \circ \Phi) = D_j\big(\pi_i \circ y \circ \Phi \circ x^{-1}\big) |_{\Phi(p)} .$$

Suppose $M \overset{\Phi}{\to} N \overset{\Psi}{\to} L$ and $p \in M$. Observe $(\Psi \circ \Phi)_{*p} = \Psi_{*\Phi(p)} \circ \Phi_{*p}$, since

$$(\Psi \circ \Phi)_{*p}(v)(f) = v(f \circ \Psi \circ \Phi) = (\Phi_{*p}(v))(f \circ \Psi) = \big(\Psi_{*\Phi(p)}(\Phi_{*p}(v))\big)(f) = \big(\big(\Psi_{*\Phi(p)} \circ \Phi_{*p}\big)(v)\big)(f).$$

Let M be a C^k-manifold. $\Omega \subset M$ open; Ω induced manifold. Let $p \in \Omega$. Clearly, $\mathcal{F}_p\Omega \subset \mathcal{F}_pM$. If $f \in \mathcal{F}_pM$, $f : U \to \mathbb{R}$ (where U open in M), then $f\mid_\Omega = f\mid_{U\cap\Omega} \in \mathcal{F}_p\Omega$. Let $v \in T_pM$. Then $v : \mathcal{F}_pM \to \mathbb{R}$. Hence, $f \mapsto v(f)$ by $\mathcal{F}_p\Omega \to \mathbb{R}$ is an element of $T_p\Omega$.

Let $w \in T_p(\Omega)$. Now, $w : \mathcal{F}_p\Omega \to \mathbb{R}$. Then $\mathcal{F}_pM \to \mathbb{R}$ defined $f \mapsto w(f\mid_p)$ is an element of T_pM. We identify v with w.

2.4. **Vector Fields.** Let M be a C^k-manifold. Let $U \subset M$ open. A *vector field* on U is a function $X : U \to \bigcup_{p\in U} T_pM$ such that $X(p) \in T_pM$ for all $p \in U$. Notation: $X_p = X(p)$.

Let $V \subset M$ open, and $f : V \to \mathbb{R}$ a C^k-function. Define $Xf : U \cap V \to \mathbb{R}$ by $(Xf)(p) = X_p(f)$ for all $p \in U \cap V$.

X is called a *smooth vector field on* U if Xf is a smooth function for every C^k-function f.

Let $g : U \to \mathbb{R}$ be a function. Define vector field, notation gX on U by $(gX)(p) = g(p)X_p$, $p \in U$.

If both g and X are smooth, then gX is smooth (product of two smooth functions are smooth).

Example 2.9. Let (U, x) be a smooth chart, and $i \in \{1, \ldots, d\}$. Define $X : U \to \bigcup_{p\in U} T_pM$ by $X_p = \frac{\partial}{\partial x^i}\big|_p$. Then X is a smooth vector field. Notation: $\frac{\partial}{\partial x^i} = X$.

Let $X : M \to \bigcup_{p\in M} T_pM$ be a smooth vector field. Note $\bigcup_{p\in M} \mathcal{F}_pM$ is the set of all smooth functions defined on an open subset of M. Notation: $\operatorname{dom}(X)$ is the domain of f. Then:

- (i) If $f \in \bigcup_{p\in M} \mathcal{F}_pM$, then $Xf \in \bigcup \mathcal{F}_pM$ and $\operatorname{dom}(Xf) = \operatorname{dom}(f)$.
- (ii) $X(f + g) = X(f) + X(g)$.
- (iii) $X(\lambda f) = \lambda X(f)$.
- (iv) $X(fg) = (Xf) \cdot g + f \cdot Xg$.

$f \mapsto Xf$ where $\bigcup_{p\in M} \mathcal{F}_pM \to \bigcup \mathcal{F}_pM$.

Lemma 2.10. *Let* $D : \bigcup_{p\in M} \mathcal{F}_pM \to \bigcup_{p\in M} \mathcal{F}_pM$ *be a map such that*

1. $\operatorname{dom}(Df) = \operatorname{dom}(f)$
2. $D(f + g) = Df + Dg$
3. $D(\lambda f) = \lambda D(f)$
4. $D(fg) = (Df)g + f \cdot Dg$

for each $f, g \in \bigcup \mathcal{F}_pM$, *for all* $\lambda \in \mathbb{R}$. *Then there exists a unique* C^k*-vector field* X *on* M *such that* $Df = Xf$ *for all* $f \in \bigcup \mathcal{F}_pM$.

Proof. For all $p \in M$ define $X_p : \mathcal{F}_pM \to \mathbb{R}$ by $X_p(f) = (Df)(p)$. Then $X_p \in T_pM$. Certainly (i)-(iv) hold, and X_p is smooth. □

Notation: $\mathcal{D}$ is the vector space of all smooth vector fields on M.

Let X and Y be smooth vector fields on M. Then $X + Y$ is a smooth vector field on M. But $f \mapsto XYf$ does not satisfy (4) in the above Lemma (in general); so, D is not a vector field. Nevertheless: $f \mapsto X(Yf) - Y(Xf)$ satisfies (1)-(4) in the Lemma. For example, satisfies (4):

$$\begin{aligned} X(Y(fg)) &= X((Yf)g + f\cdot Yg) \\ &= (XYf)g + (Yf)Xg + (Xf)(Yg) + fXYg \\ \implies \quad XY(fg) - YX(fg) &= (XYf)g + f(XYg) - (YXf)g - f(YXg) \\ &= ((XY-YX)f)g + f(XYg - YXg). \end{aligned}$$

Define smooth vector field, named $[X,Y]$ by $[X,Y]f = XYf - YXf$. $[X,Y]$ is linear in both entries. Also, $[X,Y] = -[Y,X]$ (antisymmetric) and

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$

which is called the *Jacobi identity.* If it satisfied these three properties, called a Lie algebra. More formally, we define a Lie algebra below.

A *Lie algebra* is a pair $(V,[\cdot,\cdot])$ where V is a vector space and $[\cdot,\cdot] : V\times V \to V$, called the *Lie bracket*, such that $[\cdot,\cdot]$ is *bilinear* (linear in both entries), antisymmetric (i.e., $[X,Y] = -[Y,X]$ for each $X,Y\in V$) and satisfies the Jacobi identity (i.e., $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ for each $X,Y,Z\in V$).

$[X,Y]$ is called the *commutator* of X and Y. If the commutator is clear, we simply say V is a Lie algebra. Fix $X\in V$. Then the map $V\to V$ defined $Y\mapsto [X,Y]$ is linear. We call this map $\mathrm{ad}\,(X)$, the *adjoint representation.* So, $(\mathrm{ad}\,(X))(Y) = [X,Y]$. For example,

$$[X,[X,[X,[X,Y]]]] = (\mathrm{ad}\,(X))^4(Y)$$

Example 2.11. $\mathcal{D}$ vector space of all smooth vector fields on M, with $[X,Y]f = XYf - YXf$ is an example. Another one is by taking V vector space, and $\mathcal{L}(V)$ the vector space of all linear maps from V into V. Then $[A,B] = A\circ B - B\circ A$ (for each $A,B\in\mathcal{L}(V)$) defines Lie algebra $(\mathcal{L}(V),[\cdot,\cdot])$.

Proposition 2.12. *Let $X,Y\in\mathcal{D}$ and let (U,x) be a smooth chart on M. For each $i\in\{1,\dots,d\}$ define $x^i : U\to\mathbb{R}$ by $x^i = \pi_i\circ x$ (x^i is index, not exponent). Then*

$$X_p = \sum_{i=1}^{d} \big(Xx^i\big)(p)\, \frac{\partial}{\partial x^i}\bigg|_p$$

and

$$\begin{aligned} [X,Y]_p &= \sum_{i=1}^{d} \left[\big(X\big(Yx^i\big)\big)(p) - \big(Y\big(Xx^i\big)\big)(p)\right] \frac{\partial}{\partial x^i}\bigg|_p \\ &= \sum_{i,j=1}^{d} \left[\big(Xx^j\big)(p)\, \frac{\partial}{\partial x^j}\bigg|_p \big(Yx^i\big) - \big(Yx^j\big)(p)\, \frac{\partial}{\partial x^j}\bigg|_p \big(Xx^i\big)\right] \frac{\partial}{\partial x^i}\bigg|_p \end{aligned}$$

for each $p\in U$.

Proof. □

2.5. **Direct Product.** Let $(M,\mathscr{A})$ and $(N,\mathscr{B})$ be two C^k-manifolds. Say $\dim M = d_1$ and $\dim N = d_2$. With: to make $M\times N$ a manifold, clearly product of two Hausdorff spaces is Hausdorff (so only need to define C^k-structure on $M\times N$). Let $(U,x)\in\mathscr{A}$ and $(V,y)\in\mathscr{B}$ where $x : U\to\mathbb{R}^{d_1}$ and $y : V\to\mathbb{R}^{d_2}$ are homeomorphisms. Define map $x\times y : U\times V\to\mathbb{R}^{d_1+d_2}\cong\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}$ by $(x\times y)(p,q) = (x(p),y(q))$. Then $(U\times V, x\times y)$ is a chart on $M\times N$. Also, $\{(U\times V, x\times y)\mid (U,x)\in\mathscr{A},(V,y)\in\mathscr{B}\}$ is an atlas on $M\times N$. In fact, it is a C^k-atlas. Of course, we can then take our C^k-structure on $M\times N$ to be the one generated by this C^k-atlas.

Proposition 2.13. *Let $p \in M$ and $q \in N$. Then $T_pM \times T_qN \cong T_{(p,q)}(M \times N)$ (vector spaces isomorphic).*

Proof. Define $\varphi : T_pM \times T_qN \to T_{(p,q)}(M \times N)$ by $(\varphi(v,w))(f) := v(f \circ i_1) + w(f \circ i_2)$, where $v \in T_pM$ and $w \in T_qN$ (observe φ is linear). Define $i_1 : M \to M \times N$ by $i_1(a) = (a,q)$ and $i_2 : N \to M \times N$ by $i_2(p,b)$ (both i_1 and i_2 are smooth). Need $f \in \mathcal{F}_{(p,q)}(M \times N)$ where $f : M \times N \to \mathbb{R}$. $M \xrightarrow{i_1} M \times N \xrightarrow{f} \mathbb{R}$.

Define $\psi : T_{(p,q)}(M \times N) \to T_pM \times T_qN$ by $\psi(u) = (v,w)$ where $u \in T_{(p,q)}(M \times N)$, $v(f) = u(f \otimes 1_N)$ and $w(g) = u(1_M \otimes g)$. Also, $f \in \mathcal{F}_pM$, $g \in \mathcal{F}_qN$. $(f \oplus g)(a,b) = f(a) \cdot g(b)$. Bijective, so vector spaces are isomorphic. □

Note. By convention, we identify $T_pM \times T_qN$ with $T_{(p,q)}(M \times N)$ via the map φ.

Let L be a manifold and $\theta : M \times N \to L$ be a C^k-map. Let $p \in M$, $q \in N$, $v \in T_pM$, $w \in T_qN$. $\theta_{*(p,q)}(v,w) \in T_{\theta(p,q)}L$. $\theta^{(1)} := \theta \circ i_1$ and $\theta^{(2)} := \theta \circ i_2$. Let f be a smooth function on L, i.e., $f \in \mathcal{F}_{(p,q)}L$. Then

$$\begin{aligned}
\big(\theta_{*(p,q)}(v,w)\big)(f) &= \big(\theta_{*(p,q)}\varphi(v,w)\big)(f) \\
&= \varphi(v,w)(f \circ \theta) \\
&= v(f \circ \theta \circ i_1) + w(f \circ \theta \circ i_2) \\
&= v\Big(f \circ \theta^{(1)}\Big) + w\Big(f \circ \theta^{(2)}\Big) \\
&= \Big(\theta^{(1)}_{*p}(v)\Big)(f) + \Big(\theta^{(2)}_{*q}(w)\Big)(f) \\
&= \Big(\theta^{(1)}_{*p}(v) + \theta^{(2)}_{*q}(w)\Big)(f).
\end{aligned}$$

Hence, $\theta_{*(p,q)}(v,w) = \theta^{(1)}_{*p}(v) + \theta^{(2)}_{*q}(w)$.

Proposition 2.14. *Let M, N be C^k-manifolds and $f : M \times N \to \mathbb{R}$ a C^k-function. Fix $q \in N$ and $v \in T_qN$. For all $p \in M$ define $i_p : N \to M \times N$ by $i_p(b) = (p,b)$. Define $g : M \to \mathbb{R}$ by $g(p) = v(f \circ i_p)$. Then g is a C^k-function on M.*

Proof. □

2.6. **Vector Spaces.** Let V be a finite dimensional vector space over $\mathbb{R}$ with $d = \dim V \in \mathbb{N}$ (i.e., $d \neq 0$). Let $E = \{e_1, \ldots, e_d\}$ be a basis for V. Define norm $\|\cdot\|_E$ on V by $\left\|\sum_{i=1}^d \lambda_i e_i\right\|_E = \sqrt{\lambda_1^2 + \ldots + \lambda_d^2}$. Let $\widetilde{E}$ be another basis on V. Then $\|\cdot\|_E$ and $\|\cdot\|_{\widetilde{E}}$ are equivalent. Hence, open sets independent of norm chosen. Define chart (V, x_E) by $x_E : V \to \mathbb{R}^d$ with $x_E\Big(\sum_{i=1}^d \lambda_i e_i\Big) = (\lambda_1, \ldots, \lambda_d)$, which is indeed a homeomorphism. Then (V, x_E) and $\big(V, x_{\widetilde{E}}\big)$ are C^ω-related. So, $\{(V, x_E) \mid E \text{ basis for } V\}$ is a C^ω-atlas on V. Therefore, generates a structure.

Fix $p \in V$. Let $q \in V$. Define $\gamma_q : \mathbb{R} \to V$ by $\gamma_q(t) = p + tq$ for each $t \in \mathbb{R}$. Then γ_q is a C^ω-map. So, $\dot{\gamma}_q(0) \in T_pV$. Let $E = \{e_1, \ldots, e_d\}$ be a basis for V. Let

$q = \sum_{i=1}^d \lambda_i e_i \in V$. Let $f \in \mathcal{F}_p V$. Then

$$\begin{aligned}
(\dot{\gamma}_q(0))(f) &= \frac{d}{dt}\bigg|_0 (f \circ \gamma_q)(t) \,|_{t=0} \\
&= \frac{d}{dt}\bigg|_0 \big(f \circ x_E^{-1}\big) \overbrace{(x_E(p+tq))}^{x_E(p)+x_E(q)=x_E(p)+(\lambda_1 t,\ldots,\lambda_d t)} \\
&= \sum_{i=1}^d \lambda_i D_i \big(f \circ x_E^{-1}\big) \,|_{x_E(p)} \\
&= \sum_{i=1}^d \lambda_i \, \frac{\partial}{\partial x_E^i}\bigg|_p .
\end{aligned}$$

Define $\mathscr{I} : V \to T_pV$ by $\mathscr{I}(q) = \dot{\gamma}_q(0)$; $(\mathscr{I}(q))(f)$. So,

$$\mathscr{I}\left(\sum_i \lambda_i e_i\right) = \sum_{i=1}^d \lambda_i \, \frac{\partial}{\partial x_E^i}\bigg|_p$$

for each $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$. So, $\mathscr{I}$ is an isomorphism, i.e., $V \cong T_pV$.

3. Lie Groups

A C^k*-Lie group* G is a group G, which is also a C^k manifold such that the multiplication $(g,h) \mapsto gh$ $(G \times G \to G)$ is a C^k map, and inversion $g \mapsto g^{-1}$ $(G \to G)$ is a C^k-map. Furthermore, $\dim G$ is equal to the dimension of the manifold.

Note. Let $(G, \mathscr{A})$ be a C^∞-Lie group. Then there is a C^ω-structure $\mathscr{B}$ on G such that $\mathscr{B} \subset \mathscr{A}$ and $(G, \mathscr{B})$ is a C^ω-Lie group. In fact, would only require $(G, \mathscr{A})$ to be a C^1-Lie group (which was one of Hilbert's problems).

Proposition 3.1. *If G, H are C^k-Lie groups, then $G \times H$ is a C^k-Lie group.*

Proof. □

Example 3.2. $(\mathbb{R}^d, +)$ is a Lie group (in fact, a C^ω-Lie group) of dimension d.

We claim the torus $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ is a group under multiplication, and has C^k-atlas generated by (U, x) and (V, y) from our introduction of the torus, yielding a C^ω-Lie group of dimension 1. We will now show in detail why the torus is indeed a Lie group.

Example 3.3. $\mathrm{GL}\,(d, \mathbb{R}) = \left\{A \in M^{d\times d}(\mathbb{R}) \;\middle|\; \det A \neq 0\right\}$ is the group of all invertible $d \times d$ matrices over $\mathbb{R}$. Observe $M^{d\times d}(\mathbb{R}) \cong \mathbb{R}^{d^2}$, where $\mathrm{GL}\,(d, \mathbb{R})$ homeomorphic to open subset in $\mathbb{R}^{d^2}$; hence, we obtain the induced manifold for $\mathrm{GL}\,(d, \mathbb{R})$, which has same dimension d^2. Both inversion and multiplication smooth/C^ω, so a Lie group. Indeed, as in our introduction section, the upper triangular matrices is an example. Also $\mathbb{H} = \mathbb{R}^3$ as a manifold, and $(\mathbb{H}, *)$ is a group with multiplication $(x, y, z) * (a, b, c) = (x + a, y + b, z + c + xb)$. Observe $*$ is a C^ω-map; also, $\mathbb{H}$ has identity element $e = (0, 0, 0)$ and inverse $(x, y, z)^{-1} = (-x, -y, -z + xy)$. So, C^ω-Lie group, which is called the *Heisenberg group*: note $(\mathbb{H}, *)$ is not commutative (however, it is in the first two coordinates).

Example 3.4. Let $a > 0$ and $b \in \mathbb{R}$. Define $\tau_{(a,b)} : \mathbb{R} \to \mathbb{R}$ by $\tau_{(a,b)}(x) = ax + b$. The map is affine. $c > 0$ and $d \in \mathbb{R}$.

$$\begin{aligned}\big(\tau_{(a,b)} \circ \tau_{(c,d)}\big)(x) &= \tau_{(a,b)}(cx + d)\\ &= a(cx + d) + b\\ &= acx + b + ad\\ &= \tau_{(ac,b+ad)}(x).\end{aligned}$$

Hence, $\tau_{(a,b)} \circ \tau_{(c,d)} = \tau_{(ac,b+ad)}$.

Now, define $G = (0, \infty) \times \mathbb{R}$ (which is connected; if one takes $\mathbb{R} \setminus \{0\} \times \mathbb{R}$, not connected). Observe $\dim G = 2$. Define multiplication by $(a, b) * (c, d) = (ac, b + ad)$. Then $(G, *)$ is a group where identity $e = (1, 0)$ and inverse $(a, b)^{-1} = \left(\frac{1}{a}, -\frac{b}{a}\right)$. Indeed, a C^ω-Lie group, called the $(ax + b)$-group.

Example 3.5. Consider matrices of the form $\begin{pmatrix}\cos\theta & -\sin\theta & x\\ \sin\theta & \cos\theta & y\\ 0 & 0 & 1\end{pmatrix}$. $G = \mathbb{R}^3$ as a manifold with multiplication

$$(\theta, x, y) * (\eta, a, b) = (\theta + \eta, x + a\cos\theta - b\sin\theta, y + a\sin\theta + b\cos\theta).$$

$(G, *)$ is a group, and indeed a C^ω-Lie group called the covering group of the Euclidean motion group. The identity element is $(0, 0, 0)$ and

$$(\theta, x, y)^{-1} = (-\theta, -x\cos\theta - y\sin\theta, x\sin\theta - y\cos\theta).$$

Consider $\mathbb{T} \times \mathbb{R}^2$ (which is connected), and multiplication

$$(z, x, y) * (w, a, b) = (zw, x + a\cos\theta - b\sin\theta, y + a\sin\theta + b\cos\theta),$$

and $z = e^{i\theta}$. Indeed, a C^ω-Lie group, called Euclidean motion group. $\mathbb{T}$ is not simply connected; hence covering topological space makes a covering space ($\mathbb{R}^3$ is simply connected).

Let $g \in G$, where G is a C^k-Lie group. Define $L_g : G \to G$ (i.e., left translation) by $L_g h := gh$. Let $h \in G$. Define $R_h : G \to G$ (i.e., right translation) by $R_h(g) = gh^{-1}$. L_g is a C^k-diffeomorphism; also,

$$(L_{g_1} \circ L_{g_2})(h) = L_{g_1}(g_2 h) = g_1 g_2 h = L_{g_1 g_2} h$$

and

$$(R_{h_1} \circ R_{h_2})(g) = R_{h_1}\big(gh_2^{-1}\big) = gh_2^{-1}h_1^{-1} = g(h_1 h_2)^{-1} = R_{h_1 h_2}(g).$$

Hence, $R_{h_1} \circ R_{h_2} = R_{h_1 h_2}$.

Let X be a vector field on G. Then X is called *left invariant* if for each $g, h \in G$,

$$(L_g)_{*h} X_h = X_{L_g h},$$

where $X_h \in T_h G$, $(L_g)_{*h} X_h$ and $X_{L_g h}$ both members of $T_{L_g h} G$. Similarly, *right invariant* if

$$(R_g)_{*h} X_h = X_{R_g h}$$

for each $g, h \in G$.

Proposition 3.6. *Every left invariant vector field is a C^∞ vector field.*

Proof. □

Lemma 3.7. *X is left invariant if and only if $(Xf) \circ L_g = X(f \circ L_g)$ for all $f \in \bigcup_{g \in G} \mathcal{F}_g G$ and each $g \in G$.*

Proof. Let $g, h \in G$, and $f \in \mathcal{F}_{gh}G$ (smooth function about gh). Then

$$X_{L_g h} f = (Xf)(gh) = ((Xf) \circ L_g)(h)$$

and

$$\big((L_g)_{*h}(X_h)\big) f = X_h(f \circ L_g) = (X(f \circ L_g))(h).$$

□

Before we typically let v denote our tangent vectors. From now on: $X \in T_eG$.

Lemma 3.8. *For all $X \in T_eG$, there exists a unique left invariant vector field $\widetilde{X}$ on G such that $\widetilde{X}_e = X$.*

Proof. For all $g \in G$, define $\widetilde{X}_g = (L_g)_{*e}(X)$. Then $\widetilde{X}$ is left invariant. Observe

$$\left(\widetilde{X} f\right)(g) = \widetilde{X}_g f = \big((L_g)_{*e}(X)\big)(f) = X(f \circ L_g).$$

Indeed, let $g \in G$ and f smooth. Then for all $h \in G$,

$$\left(\left(\widetilde{X} f\right) \circ L_g\right)(h) = \left(\widetilde{X} f\right)(gh) = X(f \circ L_{gh})$$

and

$$\left(\widetilde{X}(f \circ L_g)\right)(h) = X(f \circ L_g \circ L_h) = X(f \circ L_{gh}).$$

Hence, by previous Lemma, result easily follows. □

Lemma 3.9. *Let $\widetilde{X}$ and $\widetilde{Y}$ be left invariant vector fields on G. Then $[\widetilde{X}, \widetilde{Y}]$ is left invariant.*

Proof. Let $g \in G$ and f smooth. Then we want to show that

$$\left([\widetilde{X}, \widetilde{Y}] f\right) \circ L_g = [\widetilde{X}, \widetilde{Y}](f \circ L_g).$$

Observe

$$\begin{aligned}
\left([\widetilde{X}, \widetilde{Y}] f\right) \circ L_g &= \left(\widetilde{X}\widetilde{Y} f - \widetilde{Y}\widetilde{X} f\right) \circ L_g \\
&= \left(\widetilde{X}\left(\widetilde{Y} f\right)\right) \circ L_g - \left(\widetilde{Y}\left(\widetilde{X} f\right)\right) \circ L_g \\
&= \left(\widetilde{X}\left(\widetilde{Y} f\right) \circ L_g\right) - \left(\widetilde{Y}\left(\widetilde{X} f\right) \circ L_g\right) \\
&= \widetilde{X}\widetilde{Y}(f \circ L_g) - \widetilde{Y}\widetilde{X}(f \circ L_g) \\
&= [\widetilde{X}, \widetilde{Y}](f \circ L_g).
\end{aligned}$$

□

The *Lie algebra* $\mathfrak{g}$ of G is the vector space $\mathfrak{g} = T_eG$ and Lie bracket $[X, Y] := [\widetilde{X}, \widetilde{Y}]_e \in T_eG = \mathfrak{g}$. The dimension of the Lie algebra is $\dim \mathfrak{g} = d = \dim G$.

Example 3.10. Heisenberg group $\mathbb{H} = \mathbb{R}^3$ as manifold, multiplication as $(a, b, c) * (x, y, z) = (a + x, b + y, c + z + ay)$. Charts: $U = \mathbb{H}$ and $x : U \to \mathbb{R}^3$ defined $x(a, b, c) = (a, b, c)$. Identity element is $e = (0, 0, 0)$. Let

$$X_1 = \left.\frac{\partial}{\partial x^1}\right|_e \quad X_2 = \left.\frac{\partial}{\partial x^2}\right|_e \quad X_3 = \left.\frac{\partial}{\partial x^3}\right|_e .$$

Then $\{X_1, X_2, X_3\}$ is a Hamel basis for $\mathfrak{g} = T_e\mathbb{H}$ (recall $X \in T_pG$ can be written $X = \sum_{i=1}^d X(\pi_i \circ x) \left.\frac{\partial}{\partial x^i}\right|_p$). Let $k \in \{1, 2, 3\}$. Let $(a, b, c) \in \mathbb{H}$. Then

$$\widetilde{X}_k \mid_{(a,b,c)} = \sum_{i=1}^{3} \widetilde{X}_k \mid_{(a,b,c)} (\pi_i \circ x) \left.\frac{\partial}{\partial x^i}\right|_{(a,b,c)} .$$

Observe

$$\begin{aligned}
\widetilde{X}_k\ |_{(a,b,c)}\ (\pi \circ x) &= X_k\big(\pi_i \circ x \circ L_{(a,b,c)}\big) \\
&= \frac{\partial}{\partial x^k}\bigg|_{(0,0,0)} \big(\pi_i \circ x \circ L_{(a,b,c)}\big) \\
&= D_k\big(\pi_i \circ x \circ L_{(a,b,c)} x^{-1}\big)\ |_{x(0,0,0)} \\
&= \frac{d}{dt}\bigg|_{t=0} \big(\pi_i \circ x \circ L_{(a,b,c)} \circ x^{-1}\big)(te_k) \\
&= \frac{d}{dt}\bigg|_{t=0} \pi_i((a,b,c) * (te_k)).
\end{aligned}$$

Now,

$$\begin{aligned}
(a,b,c) * (t,0,0) &= (a+t,b,c) \\
(a,b,c) * (0,t,0) &= (a,b+t,c+at) \\
(a,b,c) * (0,0,t) &= (a,b,c+t).
\end{aligned}$$

Therefore,

$$\begin{aligned}
k=1 \quad \widetilde{X}_1\ |_{(a,b,c)} &= \frac{\partial}{\partial x^1}\bigg|_{(a,b,c)} \\
k=2 \quad \widetilde{X}_2\ |_{(a,b,c)} &= \frac{\partial}{\partial x^2}\bigg|_{(a,b,c)} + a\ \frac{\partial}{\partial x^3}\bigg|_{(a,b,c)} \\
k=3 \quad \widetilde{X}_3\ |_{(a,b,c)} &= \frac{\partial}{\partial x^3}\bigg|_{(a,b,c)}.
\end{aligned}$$

Now, we wish to calculate $[\widetilde{X}_1, \widetilde{X}_2]$. Notice

$$\Big(\widetilde{X}_1 f\Big)(a,b,c) = \left(\frac{\partial}{\partial x^1} f\right)(a,b,c) = D_1\big(f \circ x^{-1}\big)\big|_{x(a,b,c)} = (D_1 f)(a,b,c)$$

and

$$\Big(\widetilde{X}_2 f\Big)(a,b,c) = (D_2 f)(a,b,c) + a(D_3 f)(a,b,c).$$

Hence,

$$\Big(\widetilde{X}_1 \widetilde{X}_2 f\Big)(a,b,c) = (D_1 D_2 f)(a,b,c) + (D_3 f)(a,b,c) + a(D_1 D_3 f)(a,b,c)$$

and

$$\Big(\widetilde{X}_2 \widetilde{X}_1 f\Big)(a,b,c) = (D_2 D_1 f)(a,b,c) + a(D_3 D_1 f)(a,b,c).$$

As $D_1 D_2 = D_2 D_1$ etc,

$$\Big([\widetilde{X}_1, \widetilde{X}_2] f\Big)(a,b,c) = (D_3 f)(a,b,c).$$

Consequently,

$$[X_1, X_2] f = \Big([\widetilde{X}_1, \widetilde{X}_2] f\Big)(e) = (D_3 f)(e) = X_3 f.$$

So, $[X_1, X_2] = X_3$ and $[X_2, X_1] = -X_3$. All other commutators are equal to zero (for e.g., $[X_1, X_3] = 0$ etc).

Example 3.11. Let V be a vector space of $\dim(V) = d \in \mathbb{N}$ (so $d > 0$). $\mathcal{L}(V)$ is the vector space of all linear maps $A : V \to V$. $\mathrm{GL}(V) = \{A \in \mathcal{L}(V) \mid \det(A) \neq 0\}$ is a group, open in $\mathcal{L}(V)$. $\mathcal{L}(V)$ is a manifold, since $\mathcal{L}(V)$ has finite dimension d^2. $\mathrm{GL}(V)$ is also a manifold (taking the induced manifold), with multiplication and inversion smooth. Thus, $\mathrm{GL}(V)$ is a C^ω-Lie group with identity element $e = I$. Lie algebra of $\mathrm{GL}(V)$ is denoted by $\mathfrak{gl}(V) = T_I \mathrm{GL}(V) = T_I \mathcal{L}(V)$, where the final equality holds since we identify $T_I\mathrm{GL}(V)$ with $T_I\mathcal{L}(V)$. Define $\mathscr{J} : \mathcal{L}(V) \to \mathfrak{gl}(V)$ by $(\mathscr{J}(A))f = \frac{d}{dt}\big|_0 f(I+tA)$, where $f \in \mathcal{F}_I\mathrm{GL}(V)$. Then $\mathfrak{gl}(V)$ is a Lie algebra. Also, $\mathcal{L}(V)$ is a Lie algebra where $[A,B] = AB - BA$.

Theorem 3.12. *Define $\mathscr{J} : \mathcal{L}(V) \to \mathfrak{gl}(V)$ by $(\mathscr{J}(A))f = \frac{d}{dt}\big|_0 f(I+tA)$. $\mathscr{J}$ is a Lie algebra isomorphism (i.e., linear, bijective and $[\mathscr{J}(A), \mathscr{J}(B)] = \mathscr{J}([A,B])$; note $[\mathscr{J}(A), \mathscr{J}(B)] \in \mathfrak{gl}(V)$ and $[A,B] \in \mathcal{L}(V)$).*

Proof. $[\mathscr{J}(A), \mathscr{J}(B)] = [\widetilde{\mathscr{J}(A)}, \widetilde{\mathscr{J}(B)}]_{e=I}$. Chart on $\mathcal{L}(V)$. Chart on $\mathcal{L}(V)$. Fix a basis $E = \{e_1, \ldots, e_d\}$ for V. If $A \in \mathcal{L}(V)$, write a matrix of A w.r.t. basis E, $\begin{pmatrix} a_{11} & \cdots & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{d1} & \cdots & a_{dd} \end{pmatrix}$. Define $y_E^{ij} : \mathcal{L}(V) \to \mathbb{R}$ the (i,j)-matrix element of A. Chart $(\mathcal{L}(V), y_E)$ with $y_E = \left(y_E^{11}, \ldots, y_E^{dd}\right) : \mathcal{L}(V) \to \mathbb{R}^{d^2}$. Let $A, B \in \mathcal{L}(V)$. Then $[\mathscr{J}(A), \mathscr{J}(B)] = \mathscr{J}[A,B]$ iff

$$\left([\widetilde{\mathscr{J}(A)}, \widetilde{\mathscr{J}(B)}]\left(y_E^{ij}\right)\right)(I) = [\mathscr{J}(A), \mathscr{J}(B)]\left(y_E^{ij}\right) = (\mathscr{J}[A,B])\left(y_E^{ij}\right)$$

for each $i,j \in \{1, \ldots, d\}$. Let $i,j \in \{1, \ldots, d\}$. Let f be a smooth function on $\mathcal{L}(V)$. Let $C \in \mathrm{GL}(V)$. Then

$$\left(\widetilde{\mathscr{J}(A)}f\right)(C) = \mathscr{J}(A)(f \circ L_C) = \frac{d}{dt}\bigg|_{t=0} (f \circ L_C)(I+tA) = \frac{d}{dt}\bigg|_{t=0} f(C(I+tA)).$$

Therefore,

$$\begin{aligned} \left(\widetilde{\mathscr{J}(A)}\widetilde{\mathscr{J}(B)}f\right)(C) &= \frac{d}{dt}\bigg|_{t=0} \left(\widetilde{\mathscr{J}(B)}f\right)(C(I+tA)) \\ &= \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} f(C(I+tA)(I+sB)). \end{aligned}$$

So,

$$\left(\widetilde{\mathscr{J}(A)}\widetilde{\mathscr{J}(B)}y_E^{ij}\right)(I) = \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} y_E^{ij}((I+tA)(I+sB)) = y_E^{ij}(AB).$$

Hence,

$$\begin{aligned} [\mathscr{J}(A), \mathscr{J}(B)]\left(y_E^{ij}\right) &= \left(\left[\widetilde{\mathscr{J}(A)}, \widetilde{\mathscr{J}(B)}\right]\left(y_E^{ij}\right)\right)(I) \\ &= y_E^{ij}([A,B]) \\ &= \frac{d}{dt}\bigg|_0 y_E^{ij}(I + t[A,B]) \\ &= \mathscr{J}([A,B])\left(y_E^{ij}\right). \end{aligned}$$

□

Let G and H be C^k-Lie groups. Then $\Phi : G \to H$ is a *Lie group homomorphism*, if Φ is a C^k-map and a group homomorphism.

Hence, $\Phi(e) = e$ (group HM). Old: $\Phi_{*e} : \overbrace{T_eG}^{=\mathfrak{g}} \to \overbrace{T_eH}^{=\mathfrak{h}}$ is a linear map.

Note. Lie groups to Lie algebras results typically easy. Conversely, Lie algebras to Lie groups results a bit of work, and requires connectedness.

Theorem 3.13. *Φ_{*e} is a Lie algebra homomorphism.*

Proof. Let $X, Y \in \mathfrak{g}$. Want to show $[\Phi_{*e}(X), \Phi_{*e}(Y)] = \Phi_{*e}[X, Y]$. Consider $\Big(\Big(\widetilde{\Phi}_{*e}(X)f\Big) \circ \Phi\Big)(g)$, f is a smooth function on H, $g \in G$. Observe

$$\begin{aligned}\Big(\Big(\widetilde{\Phi}_{*e}(X)f\Big) \circ \Phi\Big)(g) &= \Big(\widetilde{\Phi}_{*e}(X)f\Big)(\Phi(g)) \\ &= \Phi_{*e}(X)\big(f \circ L_{\Phi(g)}\big) \\ &= X\big(f \circ L_{\Phi(g)} \circ \Phi\big) \\ &= X((f \circ \Phi) \circ L_g) \\ &= \Big(\widetilde{X}(f \circ \Phi)\Big)(g).\end{aligned}$$

For $L_{\Phi(g)} \circ \Phi = \Phi \circ L_g$, since

$$\big(L_{\Phi(g)} \circ \Phi\big)(g') = L_{\Phi(g)}\Phi(g') = \Phi(g)\Phi(g') = \Phi(gg') = (\Phi \circ L_g)(g').$$

So, $\Big(\widetilde{\Phi}_{*e}(X)f\Big) \circ \Phi = \widetilde{X}(f \circ \Phi)$. Hence,

$$[X, Y](f \circ \Phi) = [\Phi_{*e}(X), \Phi_{*e}(Y)](f) = \Big([\widetilde{\Phi}_{*e}(X), \widetilde{\Phi}_{*e}(Y)]f\Big) \circ \overbrace{\Phi(e)}^{e} = \Phi_{*e}[X, Y](f).$$

□

Lemma 3.14. *Let $X, Y \in T_eG$. Let X^R and Y^R be the unique right invariant vector fields on G such that $X^R_e = X$ and $Y^R_e = Y$. Then $[\widetilde{X}, \widetilde{Y}]_e = -[X^r, Y^R]_e$.*

4. Exponential map

$\widetilde{X}_g = (L_g)_{*e}(X)$. $e^A \in G$, A matrix element. $e^{tA}e^{sA} = e^{(t+s)A}$. $\gamma : \mathbb{R} \to G$ by $t \mapsto e^{tA}$ homomorphism. $\dot{\gamma}(0) = A$, $e^A = \gamma(1)$ derivative.

Lemma 4.1. *Let $\delta > 0$ and $\varphi : (-2\delta, 2\delta) \to G$ be a function, such that $\varphi(s)\varphi(t) = \varphi(s + t)$ for all $s, t \in (-\delta, \delta)$. Then there exists a unique homomorphism $\gamma : \mathbb{R} \to G$ such that $\varphi(s) = \gamma(s)$ for all $s \in (-2\delta, 2\delta)$. Furthermore, if φ is continuous (C^k-map), then γ is continuous (C^k-map).*

Proof. If γ exists, $\gamma(t) = \big(\gamma\big(\frac{t}{n}\big)\big)^n$ for each $n \in \mathbb{N}$. So restrict to $(-\delta, \delta)$ to get $\gamma(t) = \big(\varphi\big(\frac{t}{n}\big)\big)^n$. Let $t \in \mathbb{R}$, $n, m \in \mathbb{N}$. Suppose $\frac{t}{n} \in (-\delta, \delta)$ and $\frac{t}{m} \in (-\delta, \delta)$. Then

$$\left(\varphi\left(\frac{t}{n}\right)\right)^n = \left(\left(\varphi\left(\frac{t}{nm}\right)\right)^m\right)^n = \left(\left(\varphi\left(\frac{t}{nm}\right)\right)^n\right)^m = \left(\varphi\left(\frac{t}{m}\right)\right)^m.$$

Define $\gamma : \mathbb{R} \to G$ by $\gamma(t) = \varphi\big(\frac{t}{n}\big)^n$ if n large enough. □

Lemma 4.2. *There exists open $U' \subset G$ with $e \in U$ and $\epsilon > 0$ and C^k-function $\Phi : (-\epsilon, \epsilon) \times U' \to G$ such that $\Phi(0, g) = g$ for all $g \in U'$ and, for all $g \in U'$, if one writes $\alpha(t) = \Phi(t, g)$ (where $\alpha : (-\epsilon, \epsilon) \to G$) then $\dot{\alpha}(t) = \widetilde{X}_{\alpha(t)}$ for all $t \in (-\epsilon, \epsilon)$. This is for existence.*

Moreover, let $g \in U'$, $\epsilon' \in (0, \epsilon)$ and $\beta : (-\epsilon', \epsilon') \to G$ a C^k-function such that $\beta(0) = g$ and also $\dot{\beta}(t) = \widetilde{X}_{\beta(t)}$ for all $t \in (-\epsilon', \epsilon')$ then $\beta(t) = \alpha(t)$ for all $t \in (-\epsilon', \epsilon')$. This is for uniqueness.

Theorem 4.3. *Let $X \in \mathfrak{g}$. Then there exists a unique Lie algebra homomorphism $\gamma : \mathbb{R} \to G$, such that $\dot{\gamma}(0) = X$. Also: $\dot{\gamma}(t) = \widetilde{X}_{\gamma(t)}$ for all $t \in \mathbb{R}$.*

Proof. Suppose $\gamma : \mathbb{R} \to G$ is a Lie group homomorphism such that $\dot{\gamma}(0) = X$. Then $\gamma(s + t) = \gamma(s)\gamma(t)$ for all $s, t \in \mathbb{R}$. Let $s \in \mathbb{R}$ fixed. Let f smooth function. Then

$$\dot{\gamma}(s)f = \frac{d}{dt}\bigg|_{t=0} f(\gamma(s + t)) = \frac{d}{dt}\bigg|_{t=0} f(\gamma(s)\gamma(t)) = \frac{d}{dt}\bigg|_{t=0} \big(f \circ L_{\gamma(s)}\big)(\gamma(t)) = \dot{\gamma}(0)\big(f \circ L_{\gamma(s)}\big).$$

So,

$$\dot{\gamma}(s)f = X\big(f \circ L_{\gamma(s)}\big) = \big(\big(L_{\gamma(s)}\big)_{*e}(X)\big)(f) = \widetilde{X}_{\gamma(s)}(f).$$

Hence, $\dot{\gamma}(s) = \widetilde{X}_{\gamma(s)}$ for each s in $\mathbb{R}$.

Let (U, x) be a chart with $e \in U$. For $|t|$ small, $\dot{\gamma}(t) = \widetilde{X}_{\gamma(t)}$ if and only if $\dot{\gamma}(t)(\pi_i \circ x) = \widetilde{X}_{\gamma(1)}(\pi_i \circ x)$ for each $i \leq d$ and $|t|$ small (recall $v = \sum v(\pi_i \circ x)\frac{\partial}{\partial x^i}$). But this happens if and only if

$$\frac{d}{dt}\pi_i \circ (x \circ \gamma)(t) = \Big(\widetilde{X}(\pi_i \circ x)\Big)(\gamma(t)) = \Big(\Big(\widetilde{X}(\pi_i \circ x)\Big)x^{-1}\Big)(x \circ \gamma)(t).$$

Let $F_i = \Big(\widetilde{X}(\pi_i \circ x)\Big)x^{-1}$, where $F_i : x(U) \to \mathbb{R}$. Also let $F : x(U) \to \mathbb{R}^d$, and $F = (F_1, \ldots, F_d)$. Then we want $F : \frac{d}{dt}(x \circ \gamma)(t) = F((x \circ \gamma)(t))$ for all $|t|$ small. F is smooth C^k defined on $x(U)$ open nhood of 0 in $\mathbb{R}^d$.

Define $\gamma : (-\epsilon, \epsilon) \to G$, $\varphi(t) = \Phi(t, e)$. Then φ is continuous, so there is a $\delta > 0$ such that $\varphi(t) \in U'$ for all $t \in (-2\delta, 2\delta)$, where $2\delta < \epsilon$. Let $s \in (-\delta, \delta)$. Want $\varphi(s)\varphi(t) = \varphi(s+t)$. Define $h_1, h_2 : (-\delta, \delta) \to G$ by $h_1(t) = \varphi(s+t)$ and $h_2(t) = \varphi(s)\varphi(t)$. $h_1(0) = \varphi(s) = \varphi(s) \cdot e = h_2(0)$. Let $t \in (-\delta, \delta)$. Then $\dot{h}_1(t) = \dot{\varphi}(s+t) = \widetilde{X}_{\varphi(s+t)} = \widetilde{X}_{h_1(t)}$. $\dot{h}_2(t) = \ldots = \widetilde{X}_{h_2}(t)$. By uniqueness, we get $h_1(t) = \Phi(t, \varphi(s)) = h_2(t)$ for each $t \in (-\delta, \delta)$. So $\varphi(s+t) = \varphi(s)\varphi(t)$ for all $s, t \in (-\delta, \delta)$.

By Lemma 4.1, there exists Lie group homomorphism $\gamma : \mathbb{R} \to G$ such that $\gamma(t) = \varphi(t)$ for all $t \in (-\delta, \delta)$. Then $\dot{\gamma}(0) = \dot{\varphi}(0) = X$, yielding existence.

Let $\widetilde{\gamma} : \mathbb{R} \to G$ be a Lie group homomorphism with $\dot{\widetilde{\gamma}}(0) = X$. Then $\dot{\widetilde{\gamma}}(t) = \widetilde{X}_{\widetilde{\gamma}(t)}$ for all $t \in \mathbb{R}$, in particular for small $|t| < \epsilon$. Also, $\widetilde{\gamma}(0) = e$. Uniqueness: $\widetilde{\gamma}(t) = \varphi(t) = \gamma(t)$ for all $|t| < \delta$. By Lemma 4.1, $\widetilde{\gamma}(t) = \gamma(t)$ for all $t \in \mathbb{R}$. □

Define $\exp : \mathfrak{g} \to G$ by $\exp X = \gamma(1)$, where $\gamma : \mathbb{R} \to G$ is the unique Lie group homomorphism, such that $\dot{\gamma}(0) = X$. Theory of ordinary differential equations: there is an open set $V \subset \mathfrak{g}$ with $0 \in V$ such that $\exp|_V : V \to G$ is a C^k-map.

We now recall some definitions: $\Big(\widetilde{X}f\Big)(g) = X(f \circ L_g)$, and for each $X \in \mathfrak{g}$ there exists a unique Lie group homomorphism $\gamma_X : \mathbb{R} \to G$, $\dot{\gamma}_X(0) = X$. There exists open nhood $V \subseteq \mathfrak{g}$ of 0 and $\delta > 0$ such that the map $V \times B(0, \delta) \to G$ defined $(x, t) \mapsto \gamma_X(t)$ is a C^k-map. Also, $\dot{\gamma}_X(t) = \widetilde{X}_{\gamma_X(t)}$ for each $t \in \mathbb{R}$, $\exp(X) = \gamma_X(1)$ where $\exp : \mathfrak{g} \to G$.

Example 4.4. $G = \mathbb{T}$ torus, i.e., unit circle in complex plane. Then 1 is the identity; $U = \{e^{i\theta} \mid \theta \in (-\pi, \pi)\}$, $x : U \to (-\pi, \pi) \subset \mathbb{R}$ defined $x(e^{i\theta}) = \theta$. Then (U, x) is a smooth chart. $X = \frac{\partial}{\partial x^1}\big|_e \in \mathfrak{g}$, $d = 1$ dimension.

Let $s \in \mathbb{R}$ be fixed. We want to find/calculate $\exp(sX)$. Claim: $\exp(sX) = e^{is} \in G$.

Proof. Define $\gamma : \mathbb{R} \to G$, $\gamma(t) = e^{ist}$. Clearly, γ is a C^ω-map. Let $t_1, t_2 \in \mathbb{R}$. Then

$$\gamma(t_1)\gamma(t_2) = e^{ist_1}e^{ist_2} = e^{is(t_1+t_2)} = \gamma(t_2 + t_2),$$

so γ is a homomorphism. Hence, γ is a Lie group homomorphism.

Let $f \in \mathcal{F}_e G$. Then

$$\dot{\gamma}(0)f = \frac{d}{dt}\bigg|_{t=0} f(\gamma(t)) = \frac{d}{dt}\bigg|_{t=0} f\big(e^{ist}\big) = s\,\frac{d}{dt'}\bigg|_{t'=0} f\Big(e^{it'}\Big) = s\,\frac{\partial}{\partial x^1}\bigg|_e f = (sX)(f),$$

where $t' = st$. Hence, $\dot{\gamma}(0) = sX$. So, by definition: $\exp(sX) = \gamma(1) = e^{is}$. □

Lemma 4.5. $\exp(tX) = \gamma_X(t)$, *where* $X \in \mathfrak{g}$ *and* $t \in \mathbb{R}$ *arbitrary.*

Proof. Define $\omega : \mathbb{R} \to G$ by $\omega(s) = \gamma(st)$, $s \in \mathbb{R}$ and $\gamma = \gamma_X$. Then ω is a homomorphism, and C^k-map, so Lie group homomorphism. Also, $\dot{\omega}(0) = t\dot{\gamma}(0) = tX$. So, $\exp(tX) = \omega(1) = \gamma(t)$. □

Example 4.6. $G = \mathbb{H}$, $U = \mathbb{R}$, $x(a,b,c) = (a,b,c)$ and $e = (0,0,0)$, where

$$(a,b,c) * (x,y,z) = (a+x, b+y, c+z+ay).$$

$X_1 = \frac{\partial}{\partial x^1}\big|_e$, $X_2 = \frac{\partial}{\partial x^2}\big|_e$, and $X_3 = \frac{\partial}{\partial x^3}\big|_e$. Let $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$. Claim:

$$\exp(\lambda_1 X_1 + \lambda_2 X_2 + \lambda_3 X_3) = \left(\lambda_1, \lambda_2, \lambda_3 + \frac{1}{2}\lambda_1\lambda_2\right).$$

Proof. Define $\gamma : \mathbb{R} \to \mathbb{H}$ by $\gamma(t) = \left(\lambda_1 t, \lambda_2 t, \lambda_3 t + \frac{1}{2}t^2\lambda_1\lambda_2\right)$. Clearly, γ is a C^ω-map. Let $s, t \in \mathbb{R}$. Then

$$\begin{aligned}
\gamma(t) * \gamma(s) &= \left(\lambda_1 t, \lambda_2 t, \lambda_3 t + \frac{1}{2}t^2\lambda_1\lambda_2\right) * \left(\lambda_1 s, \lambda_2 s, \lambda_3 s + \frac{1}{2}s^2\lambda_1\lambda_2\right) \\
&= \left(\lambda_1 t + \lambda_1 s, \lambda_2 t + \lambda_2 s, \lambda_3 t + \frac{1}{2}t^2\lambda_1\lambda_2 + \lambda_3 s + \frac{1}{2}s^2\lambda_1\lambda_2 + \lambda_1 t\lambda_2 s\right) \\
&= \left(\lambda_1(t+s), \lambda_2(t+s), \lambda_3(t+s) + \frac{1}{2}\lambda_1\lambda_2(t+s)^2\right) \\
&= \gamma(t+s).
\end{aligned}$$

So, γ is a Lie group homomorphism. Let $f \in \mathcal{F}_e\mathbb{H}$. Then

$$\begin{aligned}
\dot{\gamma}(0)f &= \frac{d}{dt}\bigg|_{t=0} f(\gamma(t)) \\
&= \frac{d}{dt}\bigg|_{t=0} f\left(\lambda_1 t, \lambda_2 t, \lambda_3 t + \frac{1}{2}t^2\lambda_1\lambda_2\right) \\
&= \lambda_1(D_1 f)(0) + \lambda_2(D_2 f)(0) + \lambda_3(D_3 f)(0) + 0 \\
&= (\lambda_1 X_1 + \lambda_2 X_2 + \lambda_3 X_3)f.
\end{aligned}$$

So, $\dot{\gamma}(0) = \lambda_1 X_1 + \lambda_2 X_2 + \lambda_3 X_3$. Hence,

$$\exp(\lambda_1 X_1 + \lambda_2 X_2 + \lambda_3 X_3) = \gamma(1) = \left(\lambda_1, \lambda_2, \lambda_3 + \frac{1}{2}\lambda_1\lambda_2\right).$$

□

Proposition 4.7. *Let $t \in \mathbb{R}$, $X \in \mathfrak{g}$, $f \in \mathcal{F}_{\exp(tX)}G$. Then*

$$\left(\widetilde{X} f\right)(\exp(tX)) = \frac{d}{dt} f(\exp(tX)).$$

Proof. Observe

$$\left(\widetilde{X} f\right)(\exp(tX)) = \widetilde{X}_{\gamma(t)} f = \dot{\gamma}(t)(f) = \frac{d}{dt}\bigg|_t (f \circ \gamma)(t) = \frac{d}{dt} f(\exp(tX)).$$

So, $Xf = \left(\widetilde{X} f\right)(e) = \frac{d}{dt}\big|_{t=0} f(\exp(tX))$. □

Proposition 4.8. *Let $X \in \mathfrak{g}$, $n \in \mathbb{N}$, $f \in \mathcal{F}_g G$. Then*

$$\left(\widetilde{X}^n f\right)(g) = \frac{d^n}{dt^n} f(g\exp(tX)).$$

Proof. For $n = 1$,

$$\left(\widetilde{X} f\right)(g) = X(f \circ L_g) = \frac{d}{dt}\bigg|_{t=0} (f \circ L_g)(\exp(tX)) = \frac{d}{dt}\bigg|_{t=0} f(g\exp(tX)).$$

Suppose valid for n. Then

$$\begin{aligned}\left(\widetilde{X}^{n+1}f\right)(g) &= \left(\widetilde{X}\left(\widetilde{X}^n f\right)\right)(g)\\ &= \left.\frac{d}{dt}\right|_{t=0}\left(\widetilde{X}^n f\right)(g\exp(tX))\\ &= \left.\frac{d}{dt}\right|_0 \left.\frac{d^n}{ds^n}\right|_{s=0} f\left(g\overbrace{\exp(tX)\exp(sX)}^{=\exp((t+s)X)}\right)\\ &= \left.\frac{d^{n+1}}{dr^{n+1}}\right|_{r=0} f(g\exp(rX)).\end{aligned}$$

□

Theorem 4.9. *The map* $\exp : \mathfrak{g} \to G$ *is a* C^k*-map.*

Proof. Fix $t = \frac{1}{2}\delta$. Then $V \to G$ defined $X \mapsto \gamma_X(t)$ is a C^k-map. But $\gamma_X(t) = \exp(tX) = \exp\left(\frac{1}{2}\delta X\right)$. Define $W = \left\{\frac{1}{2}\delta X \;\middle|\; X \in V\right\}$ open in $\mathfrak{g}$, and $0 \in W$. Then $W \to G$ defined $Y \mapsto \exp(Y)$ is a C^k-map. Let $X \in \mathfrak{g}$. There is $n \in \mathbb{N}$ such that $\frac{1}{n}X \in W$. Then $\exp(X) = \left(\exp\frac{1}{n}X\right)^n$. Since each $\exp\frac{1}{n}X$ is C^k, $\exp X$ is a C^k-map (for sufficiently large n). □

Theorem 4.10. *The map* $\exp_{*0} : T_0\mathfrak{g} \to T_eG$ *is bijective.*

Proof. Recall $\mathscr{J} : \mathfrak{g} \to T_0\mathfrak{g}$ is bijective, where $\mathscr{J}(X)f = \frac{d}{dt}\big|_{t=0} f(0 + tX)$. Now, let $X \in \mathfrak{g}$ and $f \in \mathcal{F}_eG$. Then

$$\begin{aligned}((\exp_{*0}\circ\mathscr{J})(X))(f) &= \mathfrak{g}(X)(f\circ\exp)\\ &= \left.\frac{d}{dt}\right|_{t=0}(f\circ\exp)(tX)\\ &= \left.\frac{d}{dt}\right|_{t=0} f(\exp(tX))\\ &= Xf.\end{aligned}$$

So, $\exp_{*0}\circ\mathscr{J}$ is the identity. Thus, $\exp_{*0}$ is bijective. □

Let G be a Lie group of a Lie algebra. Then $\exp : \mathfrak{g} \to G$ is a C^k-map. For all $X \in \mathfrak{g}$, one has $\gamma : t \mapsto \exp(tX)$ is a Lie group homomorphism such that $\dot\gamma(0) = X$. This characterises $\exp(X) = \gamma(1)$.

Theorem 4.11. *There exists open* $U \subset G$ *and open* $V \subset \mathfrak{g}$ *such that* $0 \in V$, $e \in U$ *and* $\exp|_V : V \to U$ *is a* C^k*-diffeomorphism.*

Proof. Let $\{X_1, \ldots, X_d\}$ be a Hamel basis for $\mathfrak{g}$. Define $\alpha : \mathbb{R}^d \to \mathfrak{g}$ by $\alpha(\xi_1, \ldots, \xi_d) = \xi_1X_1 + \ldots + \xi_dX_d$. Then $\alpha^{-1}\circ(\exp|_V)^{-1} : U \to W$ is a bijective C^ω-map, where $W = \alpha^{-1}(V)$. So, (U, x) is a chart on G (called a normal chart; also exponential coordinates of the first kind), and that $e \in U$, where $x = \alpha^{-1}\circ(\exp|_V)^{-1}$. □

Note. Name: log for $\alpha^{-1}\circ(\exp|_V)^{-1}$.

Lemma 4.12. *Let* $m \in \mathbb{N}$, $\mathfrak{g}_1, \ldots, \mathfrak{g}_m$ *subspaces of* $\mathfrak{g}$. *Suppose that* $\mathfrak{g} = \oplus_{i=1}^m\mathfrak{g}_i$. *Define* $\Phi : \oplus_{i=1}^m\mathfrak{g}_i \to G$ *by* $\Phi(X_1, \ldots, X_m) = (\exp X_1)\ldots(\exp X_m)$. *Then* Φ *is a* C^k*-map, and* Φ_{*0} *is bijective.*

Proposition 4.13. *Let* $\{X_1, \ldots, X_d\}$ *be a basis for* $\mathfrak{g}$ *Define* $\alpha : \mathbb{R}^d \to G$ *by* $\alpha(\xi_1, \ldots, \xi_d) = \exp(\xi_1X_1)\ldots\exp(\xi_dX_d)$. *Then* α *is clearly* C^k*-map, and there exists open* $U \subset G$ *and open* $V \subset \mathbb{R}^d$ *such that* $e \in U$, $0 \in V$ *and* $\alpha|_V : V \to U$ *is a* C^k*-diffeomorphism.*

Proof. α is a C^k-map and $\alpha(0) = e$. By the Inverse Function theorem, the result follows. □

Note. Name: $\Big(U, (\alpha\,|_V)^{-1}\Big)$ is called exponential coordinates of the second kind. Recall first kind: $\exp(\xi_1 X_1 + \ldots + \xi_d X_d)$. Second kind: $\exp(\xi_1 X_1)\ldots\exp(\xi_d X_d)$.

Proposition 4.14. *Let G, H be C^k-Lie groups, with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $\Phi : G \to H$ be a Lie group homomorphism. Then $\Phi_{*e} : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. Moreover,*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\Phi_{*e}} & \mathfrak{h} \\ {\scriptstyle \exp_G}\downarrow & & \downarrow{\scriptstyle \exp_H} \\ G & \xrightarrow[\Phi]{} & H \end{array}$$

is a commutative diagram, i.e., if $X \in \mathfrak{g}$, then $\Phi(\exp_G X) = \exp_H(\Phi_{*e}(X))$.

Proof. Define $\gamma : \mathbb{R} \to H$ by $\gamma(t) = \Phi(\exp_G(tX))$. Then γ is a C^k-map. Also, if $t, s \in \mathbb{R}$, then

$$\begin{aligned} \gamma(s)\gamma(t) &= \Phi(\exp_G(sX))\Phi(\exp_G(tX)) \\ &= \Phi(\exp_G(sX)\exp_G(tX)) \\ &= \Phi(\exp_G((s+t)X)) \\ &= \gamma(s+t). \end{aligned}$$

So, γ is a Lie group homomorphism. Let $f \in \mathcal{F}_e H$. Then

$$\dot{\gamma}(0)f = \frac{d}{dt}\bigg|_{t=0} f(\gamma(t)) = \frac{d}{dt}\bigg|_{t=0} (f \circ \Phi)(\exp_G tX) = X(f \circ \Phi) = (\Phi_{*e}(X))(f).$$

So, $\dot{\gamma}(0) = \Phi_{*e}(X)$. Then by definition, $\exp_H(\Phi_{*e}(X)) = \Phi(\exp_G X)$. □

Lemma 4.15. *Let G be a C^k-Lie group. Let $\gamma : \mathbb{R} \to G$ be a continuous homomorphism. Then γ is a Lie group homomorphism, that is, γ is a C^k-map.*

Proof. There are open $U \subset G$ and $r > 0$ such that $e \in U$ and $\exp|_{B(0,r)}: B(0,r) \to U$ is a C^k-diffeomorphism. U is open, γ is continuous, so there is a $\delta > 0$ such that $\gamma(t) \in \exp\big(B\big(0, \frac{1}{2}r\big)\big)$ for all $t \in [-\delta, \delta]$. Define $f : [-\delta, \delta] \to B(0, \delta) \subset \mathfrak{g}$ by $\exp f(t) = \gamma(t)$. Let $t \in \mathbb{R}$ and suppose $|2t| \leq \delta$. Then

$$\exp(2f(t)) = \exp f(t) \exp f(t) = \gamma(t)\gamma(t) = \gamma(2t) = \exp f(2t).$$

So, $2f(t) = f(2t)$, since map is injective. Now, let $t \in \mathbb{R}$ with $|4t| \leq \delta$. Then

$$f(4t) = f(2 \cdot 2t) = 2(f(2t)) = 4f(t),$$

so by induction $f(2^n t) = 2^n f(t)$ if $n \in \mathbb{N}$, $t \in \mathbb{R}$, $|2^n t| \leq \delta$. So, $f(2^{-n}t) = 2^{-n}f(t)$ for all $n \in \mathbb{N}$, $t \in \mathbb{R}$ with $|t| \leq \delta$ (since $t := 2^n t$). Write $X = f(\delta) \in \mathfrak{g}$. Then

$$\gamma\big(2^{-n}\delta\big) = \exp\big(f\big(2^{-n}\delta\big)\big) = \exp\big(2^{-n}f(\delta)\big) = \exp 2^{-n}X.$$

So, $\gamma(2^{-n}\delta) = \exp(2^{-n}X)$ for each $n \in \mathbb{N}$. So,

$$\gamma\big(m2^{-n}\delta\big) = \exp\big(m \cdot 2^{-n}X\big)$$

for each $n \in \mathbb{N}$ and $m \in \mathbb{Z}$. As the dyadic rationals are dense in $\mathbb{R}$ and γ is continuous, so $\gamma(t\delta) = \exp(tX)$ for all $t \in \mathbb{R}$. $\gamma(t) = \exp\big(t\delta^{-1}X\big)$ for each $t \in \mathbb{R}$, where $t \mapsto \exp\big(t\big(\delta^{-1}X\big)\big)$ is a C^k-map. Thus, γ is C^k. □

Theorem 4.16. *Let G and H be Lie groups. Let $\Phi : G \to H$ be a continuous homomorphism. Then Φ is a C^k-map, so it is a Lie group homomorphism.*

Proof. Let $X_1, \dots, X_d$ be a basis for $\mathfrak{g}$. Let $\alpha : \mathbb{R}^d \to G$ be defined $\alpha(\xi_1, \dots, \xi_d) = \exp(\xi_1 X_1) \dots \exp(\xi_d X_d)$, and $U \subset G$ and $V \subset \mathbb{R}^d$ be open nhoods of e and 0 (respectively), where $\alpha|_V \colon V \to U$ is a C^k-diffeomorphism. Define $\beta : \mathbb{R}^d \to H$ by $\beta(\xi_1, \dots, \xi_d) = \Phi(\exp(\xi_1 X_1)) \dots \Phi(\exp(\xi_d X_d))$. Then β is a C^k-map by the above Lemma. For $\xi_i \mapsto \Phi(\exp(\xi_i X_i))$ is a continuous homomorphism, so C^k. Also, multiplication is smooth, implying β is smooth. But also,

$$\begin{aligned}\beta(\xi_1, \dots, \xi_d) &= \Phi(\exp(\xi_1 X_1)) \dots \Phi(\exp(\xi_d X_d)) \\ &= \Phi(\exp(\xi_1 X_1) \dots \exp(\xi_d X_d)) \\ &= p\Phi \circ \alpha(\xi_1, \dots, \xi_d),\end{aligned}$$

since Φ is a homomorphism. So, $\Phi \circ \alpha : \mathbb{R}^d \to H$ is C^k, $\Phi \circ \alpha|_V \colon V \to H$ is C^k. $\Phi|_U = \Phi \circ \alpha|_V \circ (\alpha|_V)^{-1} : U \to H$ is C^k.

Now, let $g_0 \in G$. Wish: Φ is C^k on open nhood of g_0. Let $g \in g_0 U = \{g_0 u \mid u \in U\}$ (where $g_0 U$ is open, contains g_0). On the other hand, if $g = g_0 u$ where $u \in U$, then $u = g_0^{-1} g$ implying

$$\Phi(g) = \Phi(g_0)\Phi(u) = \Phi(g_0)\Phi\big(g_0^{-1} g\big) = \left(L^{(H)}_{\Phi(g_0)} \circ \Phi \circ \left(L^{(G)}_{g_0^{-1}}\right)\right)(g).$$

□

Corollary 4.17. *Let G be an abstract group, together with topology. Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be two C^k-structures on G such that $(G, \mathscr{A}_1)$ and $(G, \mathscr{A}_2)$ are Lie groups. Then $\mathscr{A}_1 = \mathscr{A}_2$.*

Proof. Consider identity map $\mathrm{id} : (G, \mathscr{A}_1) \to (G, \mathscr{A}_2)$, which is clearly continuous homeomorphism and hence C^k. So, $\mathscr{A}_2 \subseteq \mathscr{A}_1$. Similarly, $\mathscr{A}_1 \subseteq \mathscr{A}_2$, and $\mathscr{A}_1 = \mathscr{A}_2$. □

Note. $\mathbb{R}$ with usual topology is a Lie group, and $\mathbb{R}$ with discrete topology is a Lie group. As they have different C^k-structures, this shows G must have same topology for this result to hold.

Example 4.18. Let V be a finite dimensional vector space with $\dim(V) = d \in \mathbb{N}$. Let $G = \mathrm{GL}(V) \subset \mathcal{L}(V)$ (vector space). Also, $\mathfrak{g} = \mathfrak{gl}(V) = T_I \mathrm{GL}(V) = T_I \mathcal{L}(V)$, where $\mathcal{L}(V) \overset{\mathscr{J}}{\to} T_I \mathcal{L}(V)$. Recall

$$(\mathscr{J}(A))(f) = \left.\frac{d}{dt}\right|_0 f(I + tA),$$

where $A \in \mathcal{L}(V)$ and $f \in \mathcal{F}_I G$.

Now, let $A \in \mathcal{L}(V)$, $v \in V$, $n \in \mathbb{N}$. Then $\|Av\| \le \|A\|\|v\|$, so $\|A^n v\| \le \|A\|^n \|v\|$. Hence,

$$\sum_{n=0}^{\infty} \frac{\|A^n v\|}{n!} \le \sum_{n=0}^{\infty} \frac{\|A\|^n}{n!} \|v\| = e^{\|A\|}\|v\|.$$

So, $\sum \frac{1}{n!} A^n v$ is convergent in V (as it is absolutely convergent). Write $e^A v = \sum_{n=0}^{\infty} \frac{1}{n!} A^n v$. The map e^A is linear. $\left\|e^A\right\| \le e^{\|A\|}$. If $t, s \in \mathbb{R}$, then $e^{tA} \circ e^{sA} = e^{(t+s)A}$. So, $t \mapsto e^{tA}$ is a homomorphism from $\mathbb{R}$ into $\mathrm{GL}(V)$. We now show it is smooth. Fix $t_0 \in \mathbb{R}$. Then

$$e^{tA} = e^{(t-t_0)A} e^{t_0 A} = \sum_{n=0}^{\infty} \frac{1}{n!} (t - t_0)^n A^n e^{t_0 A}.$$

So, $t \mapsto e^{tA}$ is a C^ω-map. Write $\gamma : \mathbb{R} \to \mathrm{GL}(V)$ by $\gamma(t) = e^{tA}$. Then γ is a Lie group homomorphism. Question: What is $\dot\gamma(0) =: X$? Then $\gamma(t) = \exp(tX)$, for all $t \in \mathbb{R}$. We have $X = \mathscr{J}(A)$.

Theorem 4.19. *Let $A \in \mathcal{L}(V)$. Then $e^A = \exp(\mathscr{I}(A))$.*

Proof. We claim $\dot\gamma(0) = \mathscr{I}(A)$, where $\gamma : \mathbb{R} \to \mathrm{GL}(V)$ is given by $\gamma(t) = e^{tA}$. Fix basis $\{e_1, \dots, e_d\}$ for V. $y_E^{ij} : \mathcal{L}(V) \to \mathbb{R}$ is i,j matrix element for $i, j \in \{1, \dots, d\}$. It suffices to show that $\dot\gamma(0)\left(y_E^{ij}\right) = (\mathscr{I}(A))\left(y_E^{ij}\right)$. Now,

$$\begin{aligned}\dot\gamma(0)\left(y_E^{ij}\right) &= \left.\frac{d}{dt}\right|_{t=0} y_E^{ij}(\gamma(t)) \\ &= \left.\frac{d}{dt}\right|_{t=0} y_E^{ij}\left(\sum_{n=0}^{\infty} \frac{1}{n!} t^n A^n\right) \\ &= \left.\frac{d}{dt}\right|_{t=0} \sum_{n=0}^{\infty} \frac{1}{n!} t^n y_E^{ij}(A^n) \\ &= y_E^{ij}(A).\end{aligned}$$

On the other hand,

$$(\mathscr{I}(A))\left(y_E^{ij}\right) = \left.\frac{d}{dt}\right|_0 y_E^{ij}(I + tA) = y_E^{ij}(A).$$

□

Lemma 4.20. *Let G be a C^k-Lie group and $X \in \mathfrak{g}$. Let f be a C^k-function on G such that $\exp(tX) \in D(f)$ for all $t \in [0,1]$. Then for $n \in \mathbb{N}$,*

$$\begin{aligned}f(\exp(X)) = f(e) + \left(\widetilde{X} f\right)(e) + \frac{1}{2}\left(\widetilde{X}^2 f\right)(e) + \dots + \frac{1}{(n-1)!}\left(\widetilde{X}^{n-1} f\right)(e) \\ + \int_0^1 \cdots \int_0^1 u_1^{n-1} u_2^{n-2} \dots u_{n-1} \left(\widetilde{X}^n f\right)(\exp(u_1 \dots u_n X)) du_1 \dots du_n.\end{aligned}$$

Proof. Observe

$$\int_0^1 \left(\widetilde{X} f\right)(\exp uX) du = \int_0^1 \frac{d}{du} f(\exp uX) du = [f(\exp uX)]_{u=0}^{u=1} = f(\exp X) - f(e),$$

so by simple rearrangement we get base case for $n = 1$. By induction: Have shown $g(\exp Y) = g(e) + \int_0^1 \left(\widetilde{Y} g\right) \exp(uY) du$. Let $n \in \mathbb{N}$. Choose $g = \widetilde{X}^n f$, $Y = u_1 \dots u_n X$. Then

$$\left(\widetilde{X}^n f\right) \exp(u_1 \dots u_n X) = \left(\widetilde{X}^n f\right)(e) + \int_0^1 u_1 \dots u_n \left(\widetilde{X} \widetilde{X}^n f\right) \exp(u_{n+1} u_1 \dots u_n X) du_{n+1},$$

and result follows by computation. □

Lemma 4.21. *Let $X, Y \in \mathfrak{g}$ and $f \in \mathcal{F}_e G$. Then the second order Taylor polynomial of $(s,t) \mapsto f(\exp tX \exp sY \exp(-tX))$ and $(t,s) \mapsto f(\exp(sY + st[X,Y]))$ about $(0,0)$ are equal.*

Proof. Let $Z \in \mathfrak{g}$. Let $n, m, \ell \in \mathbb{N}_0$. Then

$$\begin{aligned}\left(\widetilde{X}^n \widetilde{Y}^m \widetilde{Z}^\ell f\right)(e) &= \left.\frac{d^n}{dt^n}\right|_{t=0} \left(\widetilde{Y}^m \widetilde{Z}^\ell f\right)(\exp tX) \\ &= \left.\frac{d^n}{dt^n}\right|_{t=0} \left.\frac{d^m}{ds^m}\right|_{m=0} \left.\frac{d^\ell}{du^\ell}\right|_{u=0} f(\exp tX \exp sY \exp uZ).\end{aligned}$$

Let $N \in \mathbb{N}$. Then the N-th Taylor polynomial about $(0,0,0)$ of $(t,s,u) \mapsto f(\exp tX \exp sY \exp uZ)$ is

$$\sum_{n,m,\ell \in \mathbb{N}_0 : n+m+\ell \le N} \frac{1}{n!m!\ell!} t^n s^m u^\ell \left(\widetilde{X}^n \widetilde{Y}^m \widetilde{Z}^\ell f\right)(e).$$

Take $N = 2$, and $u = -t$, and $Z = X$. Then the second order Taylor polynomial of $(t, s) \mapsto f(\exp tX \exp sY \exp(-tX))$ is

$$\sum_{n,m,\ell \in \mathbb{N}_0 : n+m+\ell \leq 2} \frac{1}{n!m!\ell!} t^n s^m (-t)^\ell \left(\widetilde{X}^n \widetilde{Y}^m \widetilde{X}^\ell f\right)(e)$$

which is equal to

$$\begin{aligned} & f(e) + t\left(\widetilde{X} f\right)(e) + s\left(\widetilde{Y} f\right)(e) - t\left(\widetilde{X} f\right)(e) \\ & + \frac{1}{2} t^2 \left(\widetilde{X}^2 f\right)(e) + \frac{1}{2} s^2 \left(\widetilde{Y}^2 f\right)(e) + \frac{1}{2} t^2 \left(\widetilde{X}^2 f\right)(e) \\ & + ts\left(\widetilde{X}\widetilde{Y} f\right)(e) - t^2\left(\widetilde{X}^2 f\right)(e) - st\left(\widetilde{Y}\widetilde{X} f\right)(e) \\ & = f(e) + s\left(\widetilde{Y} f\right)(e) + \frac{1}{2} s^2 \left(\widetilde{Y}^2 f\right)(e) + ts\left([\widetilde{X}, \widetilde{Y}] f\right)(e). \end{aligned}$$

But also,

$$\begin{aligned} f(\exp(sY + st[X, Y])) &= f(e) + \left(\left(s\widetilde{Y} + st[\widetilde{X}, \widetilde{Y}]\right) f\right)(e) \\ &\quad + \frac{1}{2}\left(\left(s\widetilde{Y} + st[\widetilde{X}, \widetilde{Y}]\right)^2 f\right)(e) \\ &\quad + \text{ higher order terms} \\ &= f(e) + s\left(\widetilde{Y} f\right)(e) + st\left([\widetilde{X}, \widetilde{Y}] f\right)(e) + \frac{1}{2} s^2 \left(\widetilde{Y}^2 f\right)(e) + \text{h.o.t.} \end{aligned}$$

□

5. The adjoint map

Let $g \in G$. Define $\gamma_g : G \to G$ by $\gamma_g(h) = ghg^{-1}$ (that is, conjugation). Then γ_g is a Lie group homomorphism. Define $\mathrm{Ad}(g) = (\gamma_g)_{*e}$ (not to be confused with $(\mathrm{ad}(X))(Y) = [X, Y]$).

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\mathrm{Ad}(g)} & \mathfrak{g} \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow[\gamma_g]{} & G \end{array}$$

Let $X \in \mathfrak{g}$. Then $\exp(\mathrm{Ad}(g)X) = \gamma_g(\exp X) = g(\exp X)g^{-1}$.

Let V be a finite dimensional vector space $\dim V \in \mathbb{N}$. $G = \mathrm{GL}(V)$. $\mathfrak{g} = \mathfrak{gl}(V) = T_I G \cong T_I \mathcal{L}(V) \xleftarrow{\mathscr{J}} \mathcal{L}(V)$. Let $A \in \mathrm{GL}(V)$, $B \in \mathcal{L}(V)$. Then $\mathscr{J}(B) \in \mathfrak{gl}(V)$.

Theorem 5.1. *Then* $Ad(A)\mathscr{J}(B) = \mathscr{J}\left(ABA^{-1}\right)$.

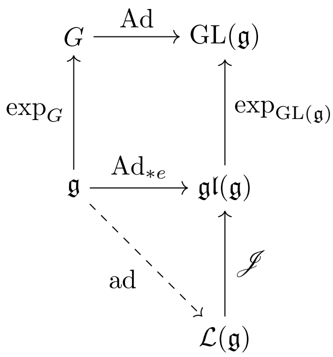
Proof. If $X, Y \in \mathfrak{g}$, then $X = Y$ iff for each $t \in \mathbb{R}$, $\exp tX = \exp tY$ (useful!). Let $t \in \mathbb{R}$. Then

$$\begin{aligned} \exp(t\mathrm{Ad}(A)\mathscr{J}(B)) &= A(\exp t\mathscr{J}(B))A^{-1} \\ &= A(\exp \mathscr{J}(tB))A^{-1} \\ &= Ae^{tB}A^{-1} \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} AB^n A^{-1} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left(ABA^{-1}\right)^n \\ &= e^{tABA^{-1}} = \exp\left(t\mathscr{J}\left(ABA^{-1}\right)\right). \end{aligned}$$

□

Now, $\gamma_{g_1} \circ \gamma_{g_2} = \gamma_{g_1 g_2}$. So, $(\gamma_{g_1})_{*e} \circ (\gamma_{g_2})_{*e} = (\gamma_{g_1 g_2})_{*e}$. $\mathrm{Ad}(g_1) \circ \mathrm{Ad}(g_2) = \mathrm{Ad}(g_1 g_2)$. What about $\mathrm{Ad}(e)$? Well, γ_e is the identity map. So, $(\gamma_e)_{*e} = I_{\mathfrak{g}}$, i.e., identity map on the Lie algebra. So, $\mathrm{Ad}(g) \circ \mathrm{Ad}\big(g^{-1}\big) = \mathrm{Ad}(e) = I_{\mathfrak{g}}$. Thus, $\mathrm{Ad}(g)$ is invertible for all $g \in G$.

Let $g \in G$. Then $\mathrm{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ is linear (also invertible). So, $\mathrm{Ad}(g) \in \mathrm{GL}(\mathfrak{g})$. Also, $g \mapsto \mathrm{Ad}(g)$ from $G \to \mathrm{GL}(\mathfrak{g})$ is a homomorphism. Note $(g, h) \mapsto \gamma_g(h) = ghg^{-1}$ from $G \times G \to G$ is a C^k-map. With exercise (assignment 1 question 2), $g \mapsto \mathrm{Ad}(g)$ is a C^k-map. So, $G \to \mathrm{GL}(\mathfrak{g})$ defined $g \mapsto \mathrm{Ad}(g)$ is a Lie group homomorphism (note Ad is called the Adjoint map/adjoint representation).



Theorem 5.2. *Let $X \in \mathfrak{g}$, $B \in \mathcal{L}(\mathfrak{g})$, and suppose $\mathscr{J}(B) = Ad_{*e}X$. Then $B = adX$.*

Proof. Let $Y \in \mathfrak{g}$. Let $f \in \mathcal{F}_e G$. Let $s, t \in \mathbb{R}$ (small). Consider

$$f(\exp tX \exp_G sY \exp(-tX)) = f(\exp_G(s(\mathrm{Ad}(\exp tX))Y)).$$

So,

$$\begin{aligned}\left.\frac{d}{ds}\right|_{s=0} f(\exp tX \exp_G sY \exp(-tX)) &= ((\mathrm{Ad}(\exp_G(tX)))Y)f \\ &= \Big(\Big(\exp_{\mathrm{GL}(\mathfrak{g})} t(\mathrm{Ad}_{*e}(X))\Big)Y\Big)f \\ &= F\Big(\exp_{\mathrm{GL}(\mathfrak{g})}(t\mathrm{Ad}_{*e}(X))\Big),\end{aligned}$$

where $F(A) = (A(Y))f$ $(A \in \mathrm{GL}(\mathfrak{g}))$. So,

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=0} F\Big(\exp_{\mathrm{GL}(\mathfrak{g})}(t\mathrm{Ad}_{*e}X)\Big) &= \mathrm{Ad}_{*e}(X)F = (\mathscr{J}(B))F \\ &= \left.\frac{d}{du}\right|_{u=0} F(I + uB) \\ &= \left.\frac{d}{du}\right|_{u=0} ((I + uB)(Y))f \quad = (B(Y))f.\end{aligned}$$

Hence,

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=0} \left.\frac{d}{ds}\right|_{s=0} f(\exp tX \exp sY \exp(-tX)) &= \left.\frac{d}{dt}\right|_{t=0} \left.\frac{d}{ds}\right|_{s=0} f(\exp(s(Y + t[X, Y]))) \\ &= \left.\frac{d}{dt}\right|_{t=0} (Y + t[X, Y])f = [X, Y]f.\end{aligned}$$

Thus, $(B(Y))f = [X, Y]f$. So, $B(Y) = [X, Y] = (\mathrm{ad}X)(Y)$. Hence, $B = \mathrm{ad}X$. $\square$

Corollary 5.3. *Let $X, Y \in \mathfrak{g}$. Then $Ad(\exp X) = e^{adX}$ and $\exp X \exp Y \exp -X = \exp\big(e^{adX} Y\big)$.*

Proof. Observe

$$\mathrm{Ad}(\exp X) = \exp \mathrm{Ad}_{*e}(X) = \exp\left(\mathscr{J}(\mathrm{ad}X)\right) = e^{\mathrm{ad}X}$$

and

$$\exp X \exp Y \exp -X = \exp X \exp Y (\exp X)^{-1} = \exp\left(\mathrm{Ad}(\exp X)Y\right) = \exp\left(e^{\mathrm{ad}X}Y\right).$$

□

Let G be a C^k-Lie group ($k \in \{\infty, \omega\}$), $\exp : \mathfrak{g} \to G$. Exists open nhood $V \in \mathfrak{g}$ of 0 (identity of Lie algebra) and open nhood $U \subset G$ of 0 (identity of group), such that $\exp|_V \colon V \to U$ is a C^k-diffeomorphism. Choose norm $\|\cdot\|$ on $\mathfrak{g}$ (it is finite dimensional vector space, so for example could take the Euclidean norm as they are equivalent). Let $X, Y \in \mathfrak{g}$. Then $\exp X \exp Y \in G$. If $\|X\|$ an $\|Y\|$ are both small enough, then $(\exp X)(\exp Y) \in U$. Write $(\exp X)(\exp Y) = \exp\left(M(X,Y)\right)$. Clearly, if $\delta > 0$ is small enough, then $(X,Y) \mapsto M(X,Y)$ from $B(0,\delta) \times B(0,\delta) \to V \subset \mathfrak{g}$ is a C^∞-map.

Theorem 5.4. *This map is C^ω, even in a C^∞-lie group.*

Proof. In a matrix group, $e^A e^B = e^C$. What is C as function of A and B (where A, B close to zero)? Well, $C = C_1 + C_2 + \ldots$, and

$$e^A e^B = \left(I + A + \frac{1}{2}A^2 + \ldots\right)\left(I + B + \frac{1}{2}B^2 + \ldots\right)$$

which is equal to

$$I + A + B + \frac{1}{2}A^2 + AB + \frac{1}{2}B^2 + \ldots.$$

As $e^C = I + C + \frac{1}{2}C^2 + \ldots = e^A e^B$, and $e^C = I + C_1 + C_2 + \ldots + \frac{1}{2}(C_1 + C_2 + \ldots)^2 + \ldots$, we take $C_1 = A + B$, $C_2 = \frac{1}{2}A^2 + AB + \frac{1}{2}B^2 = C_2 + \frac{1}{2}C_1^2$. So, $C_2 = \frac{1}{2}A^2 + AB + \frac{1}{2}B^2 - \frac{1}{2}C_1^2$. Therefore, get

$$\frac{1}{2}A^2 + AB + \frac{1}{2}B^2 - \frac{1}{2}A^2 - \frac{1}{2}AB - \frac{1}{2}BA - \frac{1}{2}B^2 = \frac{1}{2}AB - \frac{1}{2}BA = \frac{1}{2}[A,B].$$

□

Theorem 5.5. (Cambell-Baker-Hausdorff formula). *$M(X,Y)$ can be expressed as the sum of $X+Y$, and a power series in higher orders (e.g., $[X,[X,Y]]$) commutators of X and Y. If $\delta > 0$ is small enough and $\|X\| < \delta$ and $\|Y\| < \delta$, then the power series converges absolutely. $M(X,Y) = X + Y + \frac{1}{2}[X,Y] + h.o.t.$.*

$\exp X \exp Y = \exp M(X,Y)$. Since M is real analytic, even in C^∞-Lie group, it is C^ω.

Theorem 5.6. *Let $(G, \mathscr{A})$ be a C^∞-Lie group. Then there exists a C^ω-atlas $\mathscr{B}$ on G such that $\mathscr{B} \subset \mathscr{A}$ and also $(G, \mathscr{G}, \mathscr{B})$ is a C^ω-Lie group.*

Note. The following on $\mathbb{R}$

$$\begin{cases} e^{-\frac{1}{x}} & \text{if } x > 0; \\ 0 & \text{if } x \leq 0; \end{cases}$$

is a chart not real analytic.

Proposition 5.7. *Let $X, Y \in \mathfrak{g}$. Then* $\exp[X,Y] = \lim_{n\to\infty}\left(\exp \frac{1}{n}X \exp \frac{1}{n}Y \exp -\frac{1}{n}X \exp -\frac{1}{n}Y\right)^{n^2}$ *in G.*

Proof. Let $t \in \mathbb{R}$ small. Then

$$\exp tX \exp tY = \exp\left(tX + tY + \frac{1}{2}t^2[X,Y] + O\big(t^3\big)\right)$$

and

$$\exp -tX \exp -tY = \exp\left(-tX - tY + \frac{1}{2}t^2[X,Y] + O\big(t^3\big)\right).$$

Hence, $(\exp tX \exp tY)(\exp -tX \exp -tY)$ is equal to

$$\begin{aligned}
&\left(\exp\left(tX + tY + \frac{1}{2}t^2[X,Y] + O\big(t^3\big)\right)\right)\left(\exp\left(-tX - tY + \frac{1}{2}t^2[X,Y] + O\big(t^3\big)\right)\right)\\
&= \exp\left(\left(tX + tY + \frac{1}{2}t^2[X,Y]\right) + \left(-tX - tY + \frac{1}{2}t^2[X,Y]\right) + \frac{1}{2}[\ldots,\ldots] + O\big(t^3\big)\right)\\
&= \exp\left(t^2[X,Y] - \frac{t^2}{2}[X+Y, X+Y] + O\big(t^3\big)\right)\\
&= \exp\big(t^2[X,Y] + O\big(t^3\big)\big)\\
&= \exp\big(t^2[X,Y] + t^3F(t,X,Y)\big),
\end{aligned}$$

where

$$[\ldots,\ldots] = [t(X+Y) + \frac{1}{2}t^2[X,Y] + O\big(t^3\big), -t(X+Y) + \frac{1}{2}t^2[X,Y] + O\big(t^3\big)].$$

Let $n \in \mathbb{N}$. Then

$$\begin{aligned}
\left(\exp\frac{1}{n}X \exp\frac{1}{n}Y \exp -\frac{1}{n}X \exp -\frac{1}{n}Y\right)^{n^2} &= \left(\exp\left(\frac{1}{n^2}[X,Y] + \frac{1}{n^3}F\left(\frac{1}{n}, X, Y\right)\right)\right)^{n^2}\\
&= \exp\left([X,Y] + \frac{1}{n}F\left(\frac{1}{n}, X, Y\right)\right).
\end{aligned}$$

Taking limits on both sides, yields the desired result. □

6. Lie subgroups

Let G be a C^ω-Lie group. Then a Lie group H is called a *Lie subgroup* of G if

(1) algebraically, H is a subgroup of G;
(2) H is a Lie group with its own topology and manifold structure;
(3) the inclusion map $\iota : H \to G$ is continuous.

Lemma 6.1. *The inclusion map $\iota : H \to G$ is a Lie group homomorphism (so it is a C^ω-map). Furthermore, $\iota_{*e} : \mathfrak{h} \to \mathfrak{g}$ is injective.*

Proof. Since $H \leq G$, and ι is continuous, ι is a Lie group homomorphism. Now, let $Y \in \mathfrak{h}$, and suppose $\iota_{*e}(Y) = 0 \in \mathfrak{g}$.

$$\begin{array}{ccc}
\mathfrak{h} & \xrightarrow{\iota_{*e}} & \mathfrak{g}\\
{\scriptstyle \exp_H}\big\downarrow & & \big\downarrow{\scriptstyle \exp_G}\\
H & \xrightarrow[\iota]{} & G
\end{array}$$

Let $t \in \mathbb{R}$. Then

$$\iota(\exp_H tY) = \exp_G\left(\iota_{*e}(tY)\right) = \exp_G 0 = e.$$

Hence, $\exp_H tY = e = \exp t \cdot 0$ for each $t \in \mathbb{R}$. Hence, $Y = 0$ (kernel is trivial). Thus, ι_{*e} is injective. □

Let $(V, [\cdot,\cdot])$ be a Lie algebra. Let $W \subset V$ be a set. Then W is called a *subalgebra* of $(V, [\cdot,\cdot])$ if W is a linear subspace of V, and for all $X, Y \in W$ one has $[X,Y] \in W$.

Corollary 6.2. *$\iota_{*e}(\mathfrak{h})$ is a subalgebra of $\mathfrak{g}$.*

Note. In literature: $\mathfrak{h}$ is identified with $\iota_{*e}(\mathfrak{h}) \in \mathfrak{g}$. Also, the circle group is compact.

Proof. $[\iota_{*e}(Y_1), \iota_{*e}(Y_2)] = \iota_{*e}([Y_1, Y_2])$. □

Lemma 6.3. *Let H be a Lie subgroup of G, and $\gamma : \mathbb{R} \to H$ be defined $\gamma(t) = \exp_G(t, X)$. Then*

$$\iota_{*e}(\mathfrak{h}) = \{X \in \mathfrak{g} \mid \exp_G(tX) \in H \textit{ for all } t \in \mathbb{R}, \textit{ and the map } \gamma \textit{ is continuous}\}.$$

Proof. Let $Y \in \mathfrak{h}$. Then $\exp_G(t\iota_{*e}(Y)) = \iota(\exp_H tY) = \exp_H tY \in H$ for all $t \in \mathbb{R}$. Simply set $X = \iota_{*e}(Y)$ for $\subset$.

For the other inclusion, $\supset$, let $X \in \mathfrak{g}$. Suppose $\exp_G(tX) \in H$ for all $t \in \mathbb{R}$ and $\gamma : t \mapsto \exp_G tX$ is continuous from $\mathbb{R}$ into H. Then γ is a Lie group homomorphism on H. Let $Y = \dot{\gamma}(0) \in \mathfrak{h}$. Also, for all $t \in \mathbb{R}$,

$$\gamma(t) = \exp_H tY = \iota(\exp_H tY) = \exp_G t\iota_{*e}(Y).$$

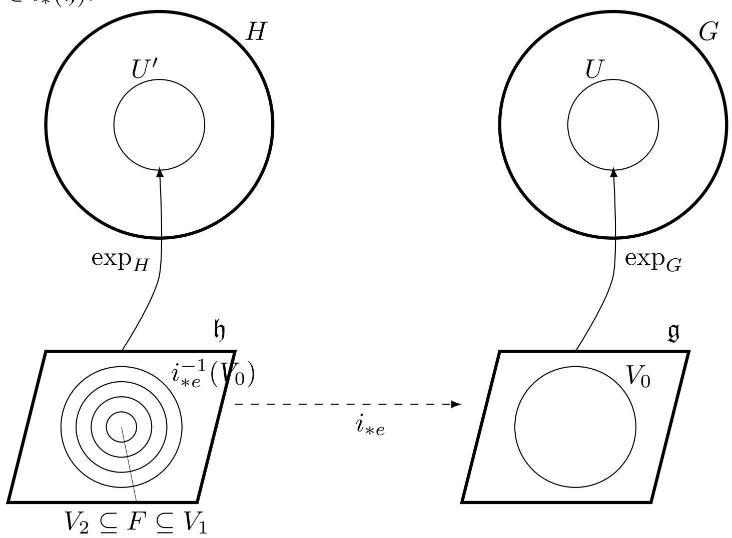
But also, $\gamma(t) = \exp_G tX$. Hence, $X = \iota_{*e}(Y) \in \iota_{*e}(\mathfrak{h})$. □

It continuity of γ necessary? Well, removing it completely with no extra assumptions, yes (we now provide an example). It is worth noting, continuity can be replaced by H being connected/path-connected. Take $G = \mathbb{R}^2$ and $\mathbb{R}_d$ group $\mathbb{R}$ with the discrete topology (otherwise use Euclidean). Then $\mathbb{R}_d \times \mathbb{R}$ is a Lie group, $H = \mathbb{R}_d \times \mathbb{R}$. Observe $\mathbb{R}_d \times \mathbb{R}$ is a Lie subgroup of $\mathbb{R}^2$. But $\mathbb{R}_d \times \mathbb{R}$ has dimension 1 (since $\dim \mathbb{R}_d = 0$) and $\mathbb{R}^2$ has dimension 2. Also, $t \mapsto (t, 0)$ from $\mathbb{R} \to \mathbb{R}_d \times \mathbb{R}$ is not continuous. Take $X = \frac{\partial}{\partial x^1}\big|_{(0,0)}$. Then $(t, 0) = \exp_{\mathbb{R}^2}(tX)$.

Proposition 6.4. *Let H be a Lie subgroup of G. Suppose H is (path) connected in H. Then $i_{*e}(\mathfrak{h}) = \{X \in \mathfrak{g} \mid \exp_G tX \in H$ for all $t \in \mathbb{R}\}$.*

Proof. ($\subset$) is trivial.

($\supset$). Let $X \in \mathfrak{g}$, and suppose $\exp_G tX \in H$ for all $t \in \mathbb{R}$. If $X = 0$, then $X \in i_*(\mathfrak{h})$.



Suppose $X \neq 0$. There exists open $V_0 \subset \mathfrak{g}, U \subset G$ such that $0 \in V_0$ and $\exp_G|_{V_0}$: $V_0 \to U$ is a C^ω-diffeomorphism. Then $i_{*e}^{-1}(V_0)$ is open in $\mathfrak{h}$ and $0 \in i_{*e}^{-1}(V_0)$. There is an open $V_1 \subset \mathfrak{h}$ and open $U' \subset H$ such that $0 \in V_1 \subset i_{*e}^{-1}(V_0)$ and $\exp_H|_{V_1}$: $V_1 \to U'$ is a C^ω-diffeomorphism. There are open $V_2 \subset \mathfrak{h}$ and closed bounded $F \subset \mathfrak{h}$ such that $0 \in V_2 \subset F \subset V_1$ and V_2 is symmetric (i.e., $Y \in V_2$ iff $-Y \in V_2$). There is a $\delta > 0$ such that for all $t \in [-2\delta, 2\delta]$ one has $tX \in V_0$. By assignment 1 question 5, $\exp_H V_2$ open in H where $e \in \exp_H$, so $\bigcup_{m=1}^{\infty} (\exp_H V_2)^m = H$. Also,

$H = \bigcup_{m=1}^{\infty} (\exp_H V_2)^m \subset \bigcup_{m=1}^{\infty} (\exp_H F)^m$. For each $t \in [-\delta, \delta]$, $\exp_G tX \in H$. There is an $m \in \mathbb{N}$ such that $\{t \in [-\delta, \delta] \mid \exp_G tX \in (\exp_H F)^m\}$ is infinite. Then there are different $t_1, t_2, \ldots \in [-\delta, \delta]$ such that $\exp_G t_n X \in (\exp_H F)^m$ for all $n \in \mathbb{N}$. For all $n \in \mathbb{N}$, there exists $Y_{n1}, \ldots, Y_{nm} \in F$ such that $\exp_G t_n X = \exp_H Y_{n1} \ldots \exp_H Y_{nm}$. Without loss of generality, $\lim_{n\to\infty} Y_{nj} =: Y_j$ exists in $\mathfrak{h}$ for all $j \in \{1, \ldots, m\}$. $\lim_{n\to\infty} t_n =: s$ in $\mathbb{R}$. Without loss of generality, $s \neq t_n$ for each $n \in \mathbb{N}$. We claim

(1) $\lim_{n\to\infty} \exp_G t_n X = \exp_G sX$ in H;
(2) $\lim_{n\to\infty} \exp_G t_n X$ exists in H; and
(3) $\lim_{n\to\infty} \exp_G t_n X = \exp_G sX$ in G.

We now prove claim (1), assuming claims (2) and (3). Observe

$$\begin{aligned}\lim_{n\to\infty} \exp t_n X &= \lim_{n\to\infty} \exp_H Y_{n1} \ldots \exp_H Y_{nm} \\ &= \exp_H Y_1 \ldots \exp_H Y_m \quad \text{in } H.\end{aligned}$$

The inclusion map i is continuous. So,

$$\exp sX = \lim_{n\to\infty} i(\exp_G t_n X) = i(\exp_H Y_1 \ldots \exp_H Y_m) \quad \text{in } G.$$

So, $\exp_G sX = i(\exp_H Y_1 \ldots \exp_H Y_m)$ and

$$\lim_{n\to\infty} \exp_H t_n X = \exp_H Y_1 \ldots \exp_H Y_m = \exp_G sX.$$

Hence, $\lim_{n\to\infty} (\exp_G t_n X)(\exp_G sX)^{-1} = e$ in H, and this is equal to $\lim_{n\to\infty} \exp_G \overbrace{(t_n - s)X}^{\in H}$. So there are $n \in \mathbb{N}$ and $Y \in V_1$ such that $\exp_G \overbrace{(t_n - s)X}^{\in V_0} = \exp_H Y$ (and $t_n - s \in [-2\delta, 2\delta]$). Also,

$$\exp_H Y = i(\exp_H Y) = \exp_G \overbrace{(i_{*e}(Y))}^{\in V_0}.$$

So, $(t_n - s)X = i_{*e}(Y)$, since diffeo/bijection on V_0, so $X = \frac{1}{t_n - s} i_{*e}(Y) \in i_{*e}(\mathfrak{h})$. □

Theorem 6.5. *If Ω is a small enough open nhood of $0 \in \mathfrak{g}$, then there is a real analytic $M : \Omega \times \Omega \to \mathfrak{g}$ such that $\exp X \exp Y = \exp M(X,Y)$ for all $X, Y \in \Omega$ and $M(X,Y) = X + Y + \sum c_N(X,Y)$ converges absolutely (where $c_N(X,Y)$ are sums of commutators of order N).*

Theorem 6.6. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let $\underline{\mathfrak{h}}$ be a subalgebra of $\mathfrak{g}$. Then there exists a unique path connected Lie subgroup $\overline{H}$ of G such that $i_{*e}(\mathfrak{h}) = \underline{\mathfrak{h}}$, where $i : H \to G$ is the inclusion map and $\mathfrak{h}$ is the Lie algebra of H.*

Proof. We provide a sketch proof. Uses the Campbell-Baker-Hausdorff formula. Recall $\exp X \exp Y = \exp M(X,Y)$, where $M(X,Y) = X + Y + \sum c_N \in \mathfrak{h}$. Also, $\langle\{\exp X \mid X \in \mathfrak{h}\}\rangle$. If $X, Y \in \underline{\mathfrak{h}} \cap \Omega$ with $M(X,Y) \in \underline{\mathfrak{h}}$. Charts close to the identity. □

Theorem 6.7. *There is a unique Lie subgroup $\widehat{H}$ of G such that $i_{*e}(\mathfrak{h}) = \underline{\mathfrak{h}}$, where $i : \widehat{H} \to G$ inclusion map and $\mathfrak{h}$ is the Lie algebra of $\widehat{H}$.*

Theorem 6.8. *Let G be a Lie group and $H \leq G$ be closed in G. Then there exists a C^ω-structure on H such that H is a Lie subgroup of G.*

Proof. Define $\underline{\mathfrak{h}} := \{X \in \mathfrak{g} \mid \exp_G tX \in H \text{ for all } t \in \mathbb{R}\}$. Clearly, $\underline{\mathfrak{h}}$ closed under scalar multiplication. Next, let $X, Y \in \underline{h}$. Let $n \in \mathbb{N}$. Then

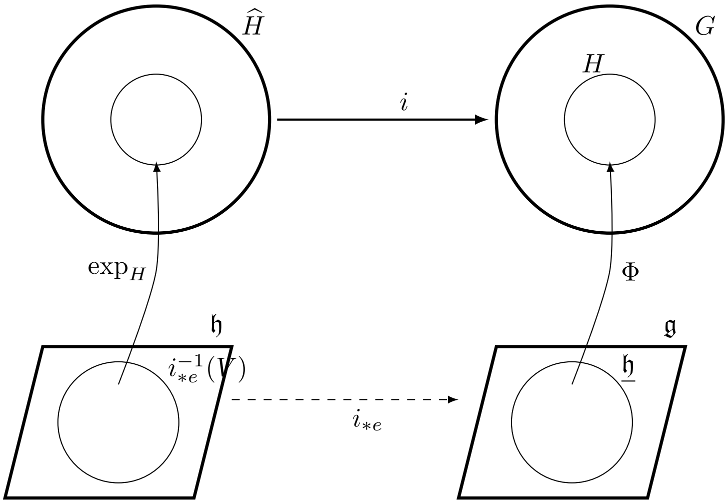
$$\lim_{n\to\infty} \left(\exp \frac{1}{n} X \exp \frac{1}{n} Y \exp -\frac{1}{n} X \exp -\frac{1}{n} Y\right)^{n^2} \in H,$$

since H is closed (limit exists since we showed it equals $\exp[X,Y]$). So,

$$\lim_{n\to\infty}\left(\exp\frac{1}{n}tX\exp\frac{1}{n}Y\exp-\frac{1}{n}tX\exp-\frac{1}{n}Y\right)^{n^2}\in H,$$

where limit is $\exp t[X,Y]$ for all $t\in\mathbb{R}$, so $[X,Y]\in\underline{\mathfrak{h}}$. Also, $X+Y\in\underline{\mathfrak{h}}$. So, $\underline{\mathfrak{h}}$ is a subalgebra of $\mathfrak{g}$.

There is a vector subspace $\mathfrak{m}$ such that $\mathfrak{g}=\underline{\mathfrak{h}}\oplus\mathfrak{m}$. Define $\Phi:\underline{\mathfrak{h}}\oplus\mathfrak{m}\to G$ by $\Phi(X,Y)=\exp X\exp Y$.



There are open convex $V\subset\underline{\mathfrak{h}}$ and open $W\subset\mathfrak{m}$ and open $U\subset G$ such that $0\in V$, $0\in W$, $e\in U$ and $\Phi\mid_{V\times W}: V\times W\to U$ is a C^ω-diffeomorphism. Without loss of generality, if V small enough, $\exp_{\widehat{H}}\mid_{i_{*e}^{-1}(V)}: i_{*e}^{-1}(V)\to\exp_{\widehat{H}}\big(\big(i_{*e}^{-1}(V)\big)\big)$ is a C^ω-diffeomorphism. Let $V'\subset\underline{\mathfrak{h}}$ open, with $0\in V'\subset V$.

Claim: there exists a $\delta>0$ such that $H\cap U'=\exp_G V'$ where $U'=\Phi(V'\times B_\delta(0))$ where $B_\delta(0)\subset\mathfrak{m}$. Proof of claim: ($\supset$) is trivial. ($\subset$) Suppose not. Then for all $n\in\mathbb{N}$, there exists $X_n\in V'$, $Y_n\in\mathfrak{m}$ such that $\Phi(X_n,Y_n)\in H$, $\|Y_n\|<\frac{1}{n}$ and $\Phi(X_n,Y_n)\notin V'$. So, $Y_n\neq 0$. Define $Z_n=\frac{1}{\|Y_n\|}Y_n\in\mathfrak{m}$, $\|Z_n\|=1$. Without loss of generality, (Z_n) is convergent. There exists $Z\in\mathfrak{m}$ such that $Z_n\to Z$. Write $\lambda_n=\frac{1}{\|Y_n\|}$. Let $t\in\mathbb{R}$. Then

$$t-\|Y_n\|=\frac{t\lambda_n-1}{\lambda_n}\leq\frac{\lfloor t\lambda_n\rfloor}{\lambda_n}\leq\frac{t\lambda_n}{\lambda_n}=t,$$

so by Squeeze theorem, $\lim_{n\to\infty}\frac{\lfloor t\lambda_n\rfloor}{\lambda_n}=t$. Also, $(\exp X_n\exp Y_n)\in H$, so $(\exp Y_n)=(\exp -X_n)(\exp X_n\exp Y_n)\in H$. Then

$$\exp\frac{\lfloor t\lambda_n\rfloor}{\lambda_n}Z_n=\exp\left(\lfloor t\lambda_n\rfloor Y_n\right)=(\exp Y_n)^{\lfloor t\lambda_n\rfloor}\in H.$$

Converges to $\exp tZ$, so $\exp tZ\in H$ for all $t\in\mathbb{R}$. So $Z\in\underline{\mathfrak{h}}$, $\|Z\|=1$, $Z\in\underline{\mathfrak{h}}\cap\mathfrak{m}=\{0\}$, which proves the claim.

Define H_0 as the path connected component of e in H. Claim: $H_0=\widehat{H}$ as sets. Proof of claim: Consider $\exp_G V=i(\exp_G V)=i\big(\exp_{\widehat{H}}\big(i_{*e}^{-1}(V)\big)\big)$ is open in $\widehat{H}$, and contains e. $\widehat{H}$ is path connected (assignment 1). So $\widehat{H}$ is generated by $\exp_G V$ as group. $\exp_G V$ is path connected, since V is convex. $\exp_G V\subset H$, so $\exp_G V\subset H_0$. Choose $V'=V$. First claim: there exists open $U'\subset G$, $e\in U'$, $\exp_G V=H\cap U'$, so $\exp_G V$ is open in H (relative topology). $\exp_G V$ is open in H_0 (relative topology). Hence, H_0 is generated by $\exp_G V$ (also H_0 open in H. Conclusion: $H_0=\widehat{H}$ as sets.

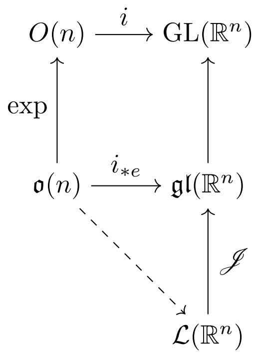
Claim: The topologies on H_0 and $\widehat{H}$ are equal. Proof of claim: Let $A \subset H_0 = \widehat{H}$. Suppose A is open in H_0. Then there is an open $U_0 \subset G$ such that $A = H \cap U_0$. Then $i^{-1}(U_0)$ is open in $\widehat{H}$, where $i^{-1}(U_0) = U_0 \cap \widehat{H} = U_0 \cap H_0 = A$. Suppose A is open in $\widehat{H}$. W.l.o.g., $A \subset \exp_G V$. Then, $A = \exp_G V'$ for some open $V' \subset V$. Claim: there exists open $U' \subset G$ such that $H \cap U' = \exp_G V' = A$. So A is open in H. Also, $A \subset H_0$, so A is open in H_0.

So, $H_0 = \widehat{H}$ as topological spaces. So, H_0 is a Lie group. Take H to be the disjoint union translates of H_0, so H is a Lie subgroup. □

Note. Also: $\underline{\mathfrak{h}} = \{X \in \mathfrak{g} \mid \exp_G tX \in H \text{ for all } t \in \mathbb{R}\} = i_{*e}(\mathfrak{h})$, and so it suffices for H to be closed in G to remove condition of continuity (instead of connectedness).

Corollary 6.9. *Let H be a closed Lie subgroup of a Lie group G. Then $i_{*e}(\mathfrak{h}) = \{X \in \mathfrak{g} \mid \exp_G tX \in H$ for all $t \in \mathbb{R}\}$.*

Example 6.10. Let $n \in \mathbb{N}$. Define $O(n) = \{U \in M^{n\times n}(\mathbb{R}) \mid U^T U = I\}$ (i.e., orthogonal matrices). Then $O(n)$ is a Lie group. Note $O(n)$ is closed in $\mathrm{GL}(\mathbb{R}^n)$. By previous theorem, $O(n)$ is a Lie subgroup of $\mathrm{GL}(\mathbb{R}^n)$.



Question: What is $i_{*e}(\mathfrak{o}(n))$? $\mathscr{J}(A) \in i_{*e}(\mathfrak{o}(n))$? Let $A \in \mathcal{L}(\mathbb{R}^n)$. Then $\mathscr{J}(A) \in i_{*e}(\mathfrak{o}(n))$ iff for all $t \in \mathbb{R}$, $\exp t\mathscr{J}(A) \in O(n)$. However, $e^{tA} = \exp t\mathscr{J}(A) = I$, implying $A^T + A = 0$. So, $e^{tA^T} = e^{-tA}$ implying $\left(e^{tA}\right)^T e^{tA} = e^{-tA}e^{tA} = I$, so converse holds (implying iff $A^T = -A$).

Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$. Then $\mathfrak{h}$ is called an *ideal* in $\mathfrak{g}$ if $\mathfrak{h}$ is a linear subspace of $\mathfrak{g}$, and $[X, Y] \in \mathfrak{h}$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$.

Theorem 6.11. *Let H be a Lie subgroup of a Lie group G.*

1. *If H is normal in G, then $i_{*e}(\mathfrak{h})$ is an ideal in $\mathfrak{g}$.*
2. *If both H and G are path connected, and $i_{*e}(\mathfrak{h})$ is an ideal in $\mathfrak{g}$, then H is normal in G.*

Proof. Without loss of generality, H is path connected. Let $Y \in \mathfrak{h}$ and $g \in G$. Then $g(\exp_H Y)g^{-1} \in H$, since $H \lhd G$. Also,

$$gi(\exp_H Y)g^{-1} = g(\exp_G(i_{*e}(Y)))g^{-1} = \exp(\mathrm{Ad}(g)i_{*e}(Y)) \in H.$$

Let $t \in \mathbb{R}$. Then

$$gi(\exp_H tY)g^{-1} = g(\exp_G t(i_{*e}(Y)))g^{-1} = \exp(t\mathrm{Ad}(g)i_{*e}(Y)) \in H.$$

So, $\mathrm{Ad}(g)i_{*e}(Y) \in i_{*e}(\mathfrak{h})$, since H is path connected Lie subgroup, so $i_*(\mathfrak{h}) = \{X \in \mathfrak{g} \mid \exp_G tX \in H \text{ for all } t \in \mathbb{R}\}$. Let $X \in \mathfrak{g}$. Then for all $t \in \mathbb{R}$,

$$e^{t\mathrm{ad}(X)}i_{*e}(Y) = e^{\mathrm{ad}(tX)}i_{*e}(Y) = \mathrm{Ad}(\exp_G tX)i_{*e}(Y) \in i_{*e}(\mathfrak{h}).$$

So,

$$[X, i_{*e}(Y)] = (\mathrm{ad}X)i_{*e}(Y) \in i_{*e}(\mathfrak{h}),$$

and therefore $i_{*e}(\mathfrak{h})$ is an ideal (proving (1)).

Let $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$. Then

$$\begin{aligned}\exp_G X \exp_H Y \exp_G -X &= (\exp_G X)\overbrace{i(\exp_H Y)}^{\exp_G i_{*e}(Y)}(\exp_G X)^{-1}\\ &= \exp\overbrace{\left(e^{\mathrm{ad}X} i_{*e}(Y)\right)}^{\in i_{*e}(\mathfrak{h})}\\ &= \exp_G\left(i_{*e}(Z)\right) = i(\exp_H Z) = \exp_H Z \in H,\end{aligned}$$

where $Z \in \mathfrak{h}$ and write $e^{\mathrm{ad}X} i_{*e}(Y) = i_{*e}(Z)$. So, $\exp_G X \exp_H Y (\exp_G X)^{-1} \in H$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$. Fix $Y \in \mathfrak{h}$. Define $G_0 = \left\{g \in G \mid g(\exp_H Y)g^{-1} \in H\right\}$. Then $\exp \mathfrak{g} \subset G_0$. Also, G_0 is a subgroup of G algebraically. G is path connected, so G is generated by every open nhood of e. Hence, $G = G_0$ (there exists open nhood $V \subset \mathfrak{g}$ of 0 and open nhood $U \subset G$ of e and $\exp_G \mid V : V \to U$ is C^ω-diffeomorphism, so $U = \exp V \subset \exp \mathfrak{g}$). So, $g \exp_H Y g^{-1} \in H$ for all $Y \in \mathfrak{h}$ and $g \in G$. Define $H_0 = \left\{h \in H \mid ghg^{-1} \in H \text{ for all } g \in G\right\}$ (the core of H in G). Then $\exp_H \mathfrak{h} \subset H_0$. Also, H_0 is a subgroup of H. H is path connected, so H is generated by $\exp_H \mathfrak{h}$. So, $H = H_0$. Then $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$ (proving (2)). □

7. Introduction to Lie algebras

7.1. **Why study Lie algebras?** Lie algebras arise as the tangent space of Lie groups T_1G, namely infinitesimal component of a Lie group G.

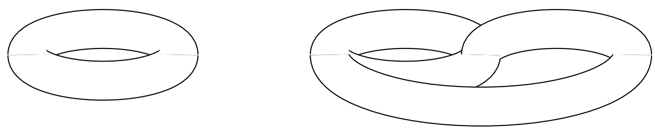
Theorem 7.1. (Lie's Third Theorem). *Every real finite dimensional Lie algebra is isomorphic to a Lie algebra of a Lie group.*

The functor $G \to T_1G$ proceeds by linearising the group structure near the unit $1 \in G$, by differentiating along smooth curves $\gamma : [-\epsilon, \epsilon] \to G$, $\gamma(0) = 1$.

Conversely, we can integrate a Lie algebra to a (local) Lie group by exponentiation.

The Baker-Campbell-Hausdorff formula describes the group multiplication near $1 \in G$ in terms of the Lie bracket.

In fact, $\exp : \mathfrak{g} \to G$ is a local diffeomorphism of unit neighbourhoods. It does not capture the global structure (i.e., topology) of G in general.



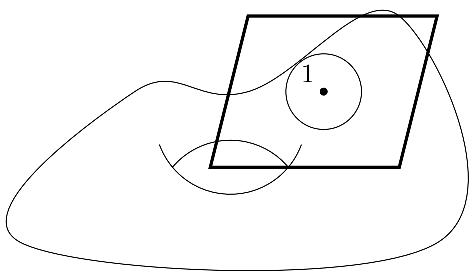
Since $\mathfrak{g}$ is connected, $\exp(\mathfrak{g})$ must be connected and cannot "see" disconnected components (for example, the above is disconnected). $\exp : \mathbb{R} \to S^1$ is surjective, but is not globally injective ($\theta \mapsto e^{2\pi i\theta}$, yet θ and $\theta + n$ where $n \in \mathbb{Z}$ mapped to same point). However, exp is a global diffeomorphism for connected, simply connected, nilpotent Lie groups (follows by the BCH formula).

Proposition 7.2. *If G is a compact, connected, real Lie group, then the exponential map is surjective.*

Note.

- One can show that $\exp : \mathfrak{gl}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ is also onto, while $\exp : \mathfrak{sl}_n(\mathbb{C}) \to \mathrm{SL}_n(\mathbb{C})$ is not. For instance, $\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C})$ does not lie in $\exp(\mathfrak{sl}_2(\mathbb{C}))$.

- SU(2) (the *spin group*) and SO(3) have isomorphic Lie algebras, but are not isomorphic as Lie groups (since their fundamental groups differ). For $\pi_1(\mathrm{SU}(2)) = \{1\}$, while $\pi_1(\mathrm{SO}(3)) = \mathbb{Z}_2$. Issue? $\exp : \mathfrak{so}(3) \to \mathrm{SO}(3)$ is not a local diffeomorphism at all points (consider the antipodal map).
- If G is a connected Lie group, then a nhood of $1 \in G$ generates all of G. Thus, every $g \in G$ is a finite product of elements of the form e^X, $X \in T_1G$. Note $X \in \mathfrak{g} \cong T_1G$, G non-compact ($h = e^{X_1}e^{X_2}\dots e^{X_n}$ for $X_1, \dots, X_n \in \mathfrak{g}$).



Lie algebras have fundamental connections to finite simple groups, algebraic groups, coxeter systems, reflection groups, and so on.

7.2. **Basic Theory.**

Definition 7.3. A finite-dimensional complex *Lie algebra* is a finite dimensional $\mathbb{C}$-vector space $\mathfrak{g}$ with an anti-symmetric bilinear map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ that satisfies the Jacobi identity

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0.$$

Note. The *adjoint map* $\mathrm{ad}_X = [X, -]$ is defined $\mathrm{ad}_X(Y) = [X,Y]$, and $\mathrm{ad}_X([Y,Z]) = [\mathrm{ad}_X(Y), Z] + [Y, \mathrm{ad}_X(Z)]$, which is the Jacobi identity and a form of the Leibniz rule. One can consider Lie algebras in infinite dimension and other fields, but we will restrict attention to the finite dimensional complex case. An example of infinite dimension is $D : C^\infty\big(S^1, \mathbb{C}\big) \to C^\infty\big(S^1, \mathbb{C}\big)$ (called derivation) and $D(f \cdot g) = (Df) \cdot g + f \cdot D(g)$ (Leibniz rule). One considers $X = -\frac{id}{d\theta}$. The Lie algebra of vector fields on S^1 is an infinite dimensional Lie algebra, called the *Witt algebra*. The Lie bracket is not associative, $[X,Y] = -[Y,X]$, non-commutative, and non-unital (i.e., no identity). For g_1, g_2 near the identity, can consider $g_1g_2g_1^{-1}g_2^{-1}$ which has $\frac{d}{dt}\big(e^{tX}e^{tY}e^{-tX}e^{-tY}\big)\big|_{t=0} = [X,Y]$. Aside: the *genus* is the number of holes (e.g., torus has $g = 1$).

Proposition 7.4. *Suppose* $(\mathfrak{g}, [\cdot,\cdot])$ *is a Lie algebra. Then anti-symmetry (i.e.,* $[X,Y] = -[Y,X]$ *for each* $X, Y \in \mathfrak{g}$*) is equivalent to alternativity (i.e.,* $[X,X] = 0$ *for all* $X \in \mathfrak{g}$*).*

Note. If $[X,Y] \neq 0$, then X and Y are linearly independent.

Proof. Suppose $[X,Y] = -[Y,X]$ for each $X, Y \in \mathfrak{g}$. Then $[X,X] = -[X,X]$, which clearly implies $[X,X] = 0$. Conversely, suppose $[X,X] = 0$ for each $X \in \mathfrak{g}$. Then $[X+Y, X+Y] = 0$, and so by bilinearity, it follows $[X,Y] = -[Y,X]$ (for each $X, Y \in \mathfrak{g}$). □

Example 7.5. A vector space V with the zero Lie bracket $[X,Y] = 0$ for all $X, Y \in V$ is an *abelian* Lie algebra. In particular, $\mathbb{C}$ is a 1-dimensional abelian Lie algebra.

$(\mathbb{R}^3, \times)$ where $\times$ is the cross product is a *real* Lie algebra, since $v \times v = 0$ and using $(u \times v) \times w = (u \cdot w)v - (v \cdot w)u$ one verifies the Jacobi identity:

$$\begin{aligned}&(u \times v) \times w + (v \times w) \times u + (w \times u) \times v \\ &= ((u \cdot w)v - (v \cdot w)u) + ((v \cdot u)w - (w \cdot u)v) + ((w \cdot v)u - (u \cdot v)w) = 0.\end{aligned}$$

Let $(A, \cdot)$ be a complex associative algebra, namely a $\mathbb{C}$-vector space with multiplication $\cdot : A \times A \to A$ that is bilinear:

(1) $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$; and
(2) $\lambda(a \cdot b) = (\lambda a) \cdot b = a \cdot (\lambda b)$,

for each $a, b, c \in A$ and $\lambda \in \mathbb{C}$, and associative:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c)$$

for all $a, b, c \in A$. The bracket $[\cdot, \cdot] : A \times A \to A$ defined $[a, b] = a \cdot b - b \cdot a$ makes $(A, [\cdot, \cdot])$ into a Lie algebra.

Example 7.6. Main example: Let $A = \operatorname{End}(V)$ be the associative algebra of linear transformations of a $\mathbb{C}$-vector space V. The corresponding Lie algebra is called the general linear Lie algebra, denoted $\mathfrak{gl}(V)$.

- A choice of basis in V fixes an isomorphism $V \cong \mathbb{C}^n$, $n = \dim(V)$. One then usually writes $\operatorname{End}(\mathbb{C}^n) = M_n(\mathbb{C})$ and $\mathfrak{gl}(\mathbb{C}^n) = \mathfrak{gl}_n(\mathbb{C})$.
- $\dim(\mathfrak{gl}_n(\mathbb{C})) = n^2$. Basis: $(e_{ij})_{i,j=1}^n$, where $(e_{ij})_{k\ell} = \delta_{ik}\delta_{j\ell}$.

Note that all finite dimensional complex Lie algebras are isomorphic to a subalgebra of this.

Proposition 7.7. $[e_{ij}, e_{k\ell}] = \delta_{jk}e_{i\ell} - \delta_{i\ell}e_{kj}$, *where*

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j; \end{cases}$$

is the Kronecker delta matrix.

Proof. Observe

$$[e_{ij}, e_{k\ell}] = e_{ij} \cdot e_{k\ell} - e_{k\ell} \cdot e_{ij},$$

and $e_{ij} \cdot e_{k\ell} = \delta_{jk}e_{i\ell}$, since $e_{ij} \cdot e_{k\ell} \neq 0$ iff $j = k$. □

Definition 7.8. A *Lie subalgebra* of $\mathfrak{g}$ is a vector subspace $\mathfrak{k}$ of $\mathfrak{g}$ that is closed under the Lie bracket (i.e., $[X, Y] \in \mathfrak{k}$ for all $X, Y \in \mathfrak{k}$).

Definition 7.9. An *ideal* of $\mathfrak{g}$ is a vector subspace I of $\mathfrak{g}$ such that $[I, \mathfrak{g}] \subseteq I$ (i.e., $[X, Y] \in I$ for all $X \in \mathfrak{g}$ and $Y \in I$).

Note. An ideal is always a subalgebra, but converse need not be true. Since $[X, Y] = -[Y, X]$, there is no need to distinguish between left and right ideals.

Example 7.10. $\{0\}$ and $\mathfrak{g}$ are ideals of $\mathfrak{g}$.

The *center* of $\mathfrak{g}$ is an ideal of $\mathfrak{g}$:

$$Z(\mathfrak{g}) := \{X \in \mathfrak{g} \mid [X, Y] = 0 \text{ for all } Y \in \mathfrak{g}\}.$$

The *special linear algebra*

$$\mathfrak{sl}(V) := \{X \in \mathfrak{gl}(V) \mid \operatorname{Tr}(X) = 0\},$$

where $\operatorname{Tr}(X)$ is the trace of X, is an ideal of $\mathfrak{gl}(V)$. Fix $V \cong \mathbb{C}^n$. Then $\dim(\mathfrak{sl}_n(\mathbb{C})) = n^2 - 1$. Basis:

$$\begin{cases} e_{ii} - e_{i+1,i+1}, & 1 \leq i < n; \\ (e_{ij})_{i,j=1}^n, & i \neq j. \end{cases}$$

Another example is

$$\mathfrak{b}_n(\mathbb{C}) = \{X \in \mathfrak{gl}_n(\mathbb{C}) \mid X \text{ is upper triangular (i.e., } X_{ij} = 0 \text{ for all } i > j)\}$$

and

$$\mathfrak{n}_n(\mathbb{C}) = \{X \in \mathfrak{gl}_n(\mathbb{C}) \mid X \text{ is strictly upper triangular (i.e., } X_{ij} = 0 \text{ for all } i \geq j)\}.$$

$\mathfrak{b}_n(\mathbb{C}) \subset \mathfrak{gl}_n(\mathbb{C})$ is a subalgebra, but not an ideal for $n \geq 2$, since $e_{11} \in \mathfrak{b}_n(\mathbb{C})$, $e_{21} \in \mathfrak{gl}_n(\mathbb{C})$ but $[e_{21}, e_{11}] = e_{21} \notin \mathfrak{b}_n(\mathbb{C})$.

Proposition 7.11. *$\mathfrak{n}_n(\mathbb{C}) \subset \mathfrak{b}_n(\mathbb{C})$ is an ideal.*

Proof. Clearly, it is a vector subspace. If $X \in \mathfrak{n}_n(\mathbb{C})$ and $Y \in \mathfrak{b}_n(\mathbb{C})$, then $(X \cdot Y)_{ii} = 0 = (Y \cdot X)_{ii}$ for all i, so clearly closed under Lie bracket. □

Example 7.12. We determine the center of $\mathfrak{sl}(2, \mathbb{F})$, where $\mathbb{F}$ is an arbitrary field (which depends on the characteristic of $\mathbb{F}$). Clearly, aI_2 lies in the centre, for each $a \in \mathbb{F}$ such that $2a = 0$. Since the characteristic is prime or infinite, follows that if the characteristic of $\mathbb{F}$ is 2, then $2a = 0$ for all $a \in \mathbb{F}$. Otherwise, if the characteristic is not 2, then aI_2 lies in the centre iff aI_2 is the zero matrix. Since if X lies in the centre we have $aX_{12} = bX_{21}$ for all $a, b \in \mathbb{F}$, follows that $X_{12} = X_{21} = 0$. Furthermore, $X_{11} = X_{22}$ can be easily shown, so indeed we get X is of the form aI_2.

Proposition 7.13. *If $\mathbb{F}$ has characteristic 2, then $Z(\mathfrak{sl}(2, \mathbb{F})) = \{aI_2 \mid a \in \mathbb{F}\}$. Otherwise, the centre of $\mathfrak{sl}(2, \mathbb{F})$ is trivial.*

Definition 7.14. A *homomorphism* of Lie algebras $\varphi : \mathfrak{g} \to \mathfrak{k}$ is a linear map that preserves the brackets:

$$\varphi([X, Y]_{\mathfrak{g}}) = [\varphi(X), \varphi(Y)]_{\mathfrak{k}}.$$

Definition 7.15. Let $\mathfrak{m} \subseteq \mathfrak{gl}(V)$ be a matrix subalgebra. A homomorphism $\varphi : \mathfrak{g} \to \mathfrak{m}$ is called a *representation* of $\mathfrak{g}$.

Proposition 7.16. $\mathrm{Tr} : \mathfrak{gl}_n(\mathbb{C}) \to \mathbb{C}$ *is a homomorphism.*

Proof. Suppose $X, Y \in \mathfrak{gl}_n(\mathbb{C})$. Clearly, Tr is a linear map. Furthermore, $[X, Y] = XY - YX$ has trace zero, for each $X, Y \in \mathfrak{gl}_n(\mathbb{C})$. Hence, result easily follows. □

Definition 7.17. A bijective homomorphism $\varphi : \mathfrak{g} \to \mathfrak{k}$ is an *isomorphism.*

Proposition 7.18. *If $\varphi : \mathfrak{g} \to \mathfrak{k}$ is a homomorphism, then* $\ker(\varphi) \subseteq \mathfrak{g}$ *is an ideal, and* $\mathrm{im}(\varphi) \subseteq \mathfrak{k}$ *is a subalgebra.*

Proof. $\ker(\varphi) \subseteq \mathfrak{g}$ is a subspace, and for each $X \in \ker(\varphi)$, $Y \in \mathfrak{g}$ we have

$$\varphi([X, Y]) = [\overbrace{\varphi(X)}^{=0}, \varphi(Y)] = 0.$$

So, $[X, Y] \in \ker(\varphi)$. Thus, $\ker(\varphi)$ is an ideal.

$\mathrm{im}(\varphi) \subseteq \mathfrak{k}$ is a subspace, and for each $X, Y \in \mathfrak{g}$,

$$[\varphi(X), \varphi(Y)] = \varphi([X, Y]) \in \mathrm{im}(\varphi).$$

Hence, $\mathrm{im}(\varphi)$ is a Lie subalgebra. □

Example 7.19. (Important Example). The adjoint homomorphism (representation) $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$, defined $X \mapsto \mathrm{ad}_X = [X, -]$. Linearity follows by bilinearity of the Lie bracket. Preservation of Lie brackets,

$$\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y] = \mathrm{ad}_X \cdot \mathrm{ad}_Y - \mathrm{ad}_Y \cdot \mathrm{ad}_X$$

follows by the Jacobi identity:

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

In fact, an interpretation of the Jacobi identity is that the adjoint map is a Lie homomorphism.

Note. $\ker(\mathrm{ad}) = Z(\mathfrak{g})$ is an ideal, and $\ker(\mathrm{Tr} : \mathfrak{gl}_n(\mathbb{C}) \to \mathbb{C}) = \mathfrak{sl}_n(\mathbb{C})$ is an ideal.

Definition 7.20. Let A be an $\mathbb{F}$-algebra. A linear map $D : A \to A$ is a *derivation* if $D(x \cdot y) = (Dx) \cdot y + x \cdot Dy$, for all $x, y \in A$.

Example 7.21.
(1) $A = C^\infty(\mathbb{R})$, $D = \frac{d}{dx}$.
(2) $\mathfrak{g}$ a Lie algebra, i.e., a non-associative $\mathbb{C}$-algebra with product $[\cdot,\cdot]$. For every $X \in \mathfrak{g}$, $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is a derivation. In fact, this is a re-statement of the Jacobi identity:

$$\mathrm{ad}_X([Y,Z]) = [\mathrm{ad}_X(Y), Z] + [Y, \mathrm{ad}_X(Z)].$$

If $\mathrm{Der}(\mathfrak{g}) = \{\text{all derivations of } \mathfrak{g}\}$, then we have

$$\mathrm{ad}(\mathfrak{g}) \subseteq \mathrm{Der}(\mathfrak{g}) \subseteq \mathfrak{gl}(\mathfrak{g}),$$

where $\mathrm{ad}(\mathfrak{g}) = \{\mathrm{ad}_X \mid X \in \mathfrak{g}\}$ are called *inner derivations.*

Let $\mathfrak{g}$ be a Lie algebra and let $\{e_i\}_{i=1}^{\dim(\mathfrak{g})}$ be a basis of $\mathfrak{g}$. Then

$$[e_i, e_j] = \sum_{k=1}^{\dim(\mathfrak{g})} C_{ijk} e_k.$$

The constants C_{ijk} are called the *structure constants* of $\mathfrak{g}$ with respect to the basis $\{e_i\}_{i=1}^{\dim(\mathfrak{g})}$. Skew-symmetry of the Lie bracket implies that $C_{ijk} = -C_{jik}$. Also, by the Jacobi identity,

$$[[e_i, e_j], e_\ell] + [[e_j, e_\ell], e_i] + [[e_\ell, e_i], e_j] = 0$$

implies

$$\sum_{k=1}^{\dim(\mathfrak{g})} C_{ijk}[e_k, e_\ell] + C_{j\ell k}[e_k, e_i] + C_{\ell ik}[e_k, e_j] = 0.$$

Example 7.22. Let e_1, e_2, e_3 be the standard basis for $\mathbb{R}^3$. The Lie algebra $(\mathbb{R}^3, \times)$ has the structure constants

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

and $C_{112} = 0$, $C_{123} = 1$, $C_{231} = 1$, etc, with respect to the standard basis.

Example 7.23. $\mathfrak{sl}_2(\mathbb{C})$ has the standard basis $e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. We will compute $[e,f]$, $[e,h]$, and $[f,h]$. All other structure constants will follow from these. Observe

$$[e,f] = [e_{12}, e_{21}] = e_{11} - e_{22} = h,$$

$$[e,h] = [e_{12}, e_{11} - e_{22}] = [e_{12}, e_{11}] - [e_{12}, e_{22}] = -e_{12} - e_{12} = -2e, \text{ and}$$

$$[f,h] = [e_{21}, e_{11} - e_{22}] = e_{21} + e_{21} = 2f.$$

One can show that two Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are isomorphic iff there exists bases B_1 of $\mathfrak{g}_1$ and B_2 of $\mathfrak{g}_2$ with respect to which the structure constants are equal.

We can write any homomorphism $\varphi : \mathfrak{g} \to \mathfrak{h}$ as a short exact sequence (image of one map equals the kernel of next map):

$$0 \to \ker(\varphi) \overset{i}{\hookrightarrow} \mathfrak{g} \overset{\varphi}{\to} \operatorname{im}(\varphi) \to 0$$

meaning that the inclusion map i is injective and φ is surjective onto its image and $\operatorname{im}(\varphi) \subseteq \mathfrak{k}$ (clearly, $\operatorname{im}(\varphi) = \ker(\varphi)$). In particular,

$$0 \to \overbrace{Z(\mathfrak{g})}^{=\ker(\mathrm{ad})} \hookrightarrow \mathfrak{g} \overset{\mathrm{ad}}{\to} \overbrace{\mathrm{ad}(\mathfrak{g})}^{=\operatorname{im}(\mathrm{ad})} \to 0.$$

A representation $\varphi : \mathfrak{g} \to \mathfrak{gl}(V)$ is called *faithful* if $\ker(\varphi) = 0$. We see that $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is faithful iff $Z(\mathfrak{g}) = 0$. This is not true for general Lie algebras, for example, if $\mathfrak{g}$ is abelian then $\mathfrak{g} = Z(\mathfrak{g})$.

A deep theorem by Ado proves that any finite dimensional Lie algebra admits a faithful representation $\varphi : \mathfrak{g} \to \mathfrak{gl}(V)$. Consequently, any Lie algebra can be viewed as a matrix Lie algebra, i.e., the Lie bracket $[X, Y]$ is a commutator under the injection φ:

$$[\varphi(X), \varphi(Y)] = \varphi(X) \circ \varphi(Y) - \varphi(Y) \circ \varphi(X).$$

We will not prove nor use Ado's theorem.

7.3. **Construction of ideals.** We have seen that the kernel $\ker(\varphi)$ of a Lie homomorphism $\varphi : \mathfrak{g} \to \mathfrak{k}$ is an ideal in $\mathfrak{g}$.

Proposition 7.24. *Given two ideals $I, J \subseteq \mathfrak{g}$, the following operations yield new ideals:*

(1) $I \cap J$;
(2) $I + J := \{X + Y \mid X \in I, Y \in J\}$; *and*
(3) $[I, J] := \operatorname{span}\{[X, Y] \mid X \in I, Y \in J\}$.

Proof. Suppose $\mathscr{I}$ is a collection of ideals in $\mathfrak{g}$. Then $\bigcap \mathscr{I}$ is a vector subspace of $\mathfrak{g}$. Furthermore,

$$[\bigcap \mathscr{I}, \mathfrak{g}] \subseteq \bigcap_{I \in \mathscr{I}} [I, \mathfrak{g}] \subseteq \bigcap \mathscr{I},$$

verifying (1).

Now, suppose I, J are ideals of $\mathfrak{g}$. Then $I + J$ clearly a vector subspace. If $X+Y \in I+J$ (where $X \in I$ and $Y \in J$) and $Z \in \mathfrak{g}$, then $[X+Y, Z] = [X, Z]+[Y, Z]$, where $[X, Z] \in I$ and $[Y, Z] \in J$ (implying $[X + Y, Z] \in I + J$). This verifies (2).

For (3), clearly a vector subspace. Also, if $X \in [I, J]$ and $Y \in \mathfrak{g}$, then $X = \sum_{i=1}^n [X_i, Y_i]$ for some $X_i \in I$ and $Y_i \in J$, so by Jacobi

$$[X, Y] = \sum_{i=1}^n [[X_i, Y_i], Y] = \sum_{i=1}^n [X_i, [Y_i, Y]] - [[Y, X_i], Y_i] \in [I, J].$$

□

Note. For (3), one must take the span of the commutators of elements of I and J. The set of commutators may not even be a vector space, and certainly not an ideal of $\mathfrak{g}$.

An important special case is when $I = J = \mathfrak{g}$. Then the ideal $\mathfrak{g}' := [\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{g}$ is called the *derived algebra* of $\mathfrak{g}$. Note that $\mathfrak{g}$ is abelian iff $\mathfrak{g}' = \{0\}$.

Proposition 7.25. $\mathfrak{sl}_2(\mathbb{C})' = \mathfrak{sl}_2(\mathbb{C})$.

Note. This fails for $\mathfrak{sl}_2(\mathbb{F})$ if characteristic of $\mathbb{F}$ is 2. For the centre is aI_2 where $a \in \mathbb{F}$.

Proof. Recall we computed $[e, f] = h$, $[e, h] = -2e$ and $[f, h] = 2f$. From this, result easily follows. □

8. Quotient algebras

Let V be a vector space and $W \subseteq V$ a subspace. A coset of W is a set of the form $v + W := \{v + w \mid w \in W\}$, where $v \in V$. The element v is a representative for the coset. Two cosets $v + W$ and $v' + W$ are the same iff $v - v' \in W$. The quotient space V/W is the set of all cosets of W. It becomes a vector space with

- $0 + W = W$ as the zero element;
- $(v + W) + (v' + W) = (v + v') + W$ as addition; and
- $\lambda(v + W) = \lambda v + W$ as scalar multiplication.

Example 8.1. For example, $V = \mathbb{R}^2$, $W = \operatorname{span}_{\mathbb{R}}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$. The coset space $V/W = \mathbb{R}^2/W$ is the vector space of all translations of the line W. Any line through the origin that is transverse to the line W picks out a set of representations for $\mathbb{R}^2/W$, which gives an identification $\mathbb{R}^2/W \cong \mathbb{R}$.

Let $v_1, \ldots, v_k \in V$ be a collection of vectors such that $v_1 + W, \ldots, v_k + W$ is a basis for V/W. Then $v_1, \ldots, v_k$ together with any basis for W forms a basis for V. Thus, $\dim(V) = \dim(W) + \dim(V/W)$.

We can think of the quotient space as a short exact sequence of vector spaces:

$$0 \to W \hookrightarrow V \xrightarrow{q} V/W \to 0.$$

Recall that this means that inclusion map i is injective, q is surjective and $\operatorname{im}(i) = \ker(q)$ ($q : V \to V/W$ defined $v \mapsto v + W$ is quotient map). A choice of basis of V/W determines a splitting of the short exact sequence:

$$V \cong W \oplus V/W.$$

Let $\mathfrak{g}$ be a Lie algebra and $I \subseteq \mathfrak{g}$ an ideal. Define the Lie bracket

$$[X + I, Y + I] := [X, Y] + I$$

on the quotient space $\mathfrak{g}/I$.

Proposition 8.2. *$\mathfrak{g}/I$ is a Lie algebra with bracket defined above.*

Proof. We check $[X + I, Y + I] := [X, Y] + I$ is well-defined. To this end, suppose $X_1 + I = X_2 + I$, and $Y_1 + I = Y_2 + I$. Hence, $X_1 - X_2 \in I$ and $Y_1 - Y_2 \in I$, and we have

$$[(X_1 - X_2), Y_1] + [X_2, (Y_1 - Y_2)] \in I,$$

since I is an ideal. It follows that $[X_1, Y_1] - [X_2, Y_2] \in I$, which means that $[X_1, Y_1] + I = [X_2, Y_2] + I$. Bilinearity, skew-symmetry and Jacobi identity follows since the original bracket has these properties. □

Note. $\mathfrak{g}/I$ is a Lie algebra iff $I \subseteq \mathfrak{g}$ is an ideal. Thus, ideals are the Lie algebraic analogues of normal subgroups.

Theorem 8.3. (First Isomorphism Theorem). *Let $\varphi : \mathfrak{g} \to \mathfrak{k}$ be a Lie homomorphism. Then* $\operatorname{im}(\varphi) \cong \frac{\mathfrak{g}}{\ker(\varphi)}$.

Note. Recall that we have a SES

$$0 \to \ker(\varphi) \hookrightarrow \mathfrak{g} \xrightarrow{\varphi} \operatorname{im}(\varphi) \to 0.$$

This theorem is simply saying that φ restricted to the quotient $\frac{\mathfrak{g}}{\iota(\ker(\varphi))}$ is injective and surjective onto its image $\operatorname{im}(\varphi)$ as a homomorphism of Lie algebras, so $\mathfrak{g}/\ker(\varphi) \cong \operatorname{im}(\varphi)$.

In particular, for $\varphi = \mathrm{ad}$, we have

$$\mathrm{ad}(\mathfrak{g}) \cong \frac{\mathfrak{g}}{Z(\mathfrak{g})}$$

called the *adjoint* of the Lie algebra $\mathfrak{g}$.

We have that the SES (i.e., $\ker(f_i) = \mathrm{im}(f_{i-1})$)

$$0 \xrightarrow{f_1} A \overset{f_2}{\hookrightarrow} B \overset{f_3}{\twoheadrightarrow} C \xrightarrow{f_4} 0$$

if, and only if, $C \cong B/A = B/f_2(A)$. We ask, does $B = B/A \oplus A = C \oplus A$? In general, no. Indeed, $\mathfrak{g} \ncong \mathrm{im}(\varphi) \oplus \ker(\varphi)$ in general, as Lie algebras. For example, consider $\mathbb{R} \to M \to S^1$. We have this sequence splits (i.e., $\mathbb{R} \to M \leftrightarrows S^1$, which means $M = S^1 \times \mathbb{R}$) when $M \cong S^1 \times \mathbb{R}$ (corresponding to a cylinder collapsing the real line). The second (and only other) case is when M is homeomorphic to the Mobius strip.

Proof. Set $I = \ker(\varphi)$ and define $\alpha : \frac{\mathfrak{g}}{I} \to \mathrm{im}(\varphi)$, $\alpha(x + I) = \varphi(x)$ for all $x \in \mathfrak{g}$. It is straightforward to check that α is well-defined, bijective and a Lie homomorphism. □

Theorem 8.4. (Second Isomorphism Theorem). *If I, J are ideals of $\mathfrak{g}$, then*

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

Proof. Define a map $\tau : I \to \frac{I+J}{J}$ by $\tau(x) = x + J$, for all $x \in I$. This is a surjective Lie homomorphism with kernel

$$\begin{aligned}\ker(\tau) &= \{x \in I \mid x + J = J\}\\ &= I \cap J.\end{aligned}$$

Result follows by First Isomorphism Theorem. □

Theorem 8.5. (Correspondence Theorem). *Suppose I is an ideal of a Lie algebra $\mathfrak{g}$. Then the ideals of $\mathfrak{g}/I$ are all of the form J/I, where $I \subseteq J$ and $J \subseteq \mathfrak{g}$ is an ideal.*

Proof. Let K be an ideal of $\mathfrak{g}/I$. Define $J = \{x \in \mathfrak{g} \mid x + I \in K\}$. Clearly, $I \subseteq J$, since $0 + I$ is the zero element in $K \subseteq \mathfrak{g}/I$, and $K = J/I$. Indeed, J is an ideal of $\mathfrak{g}$ (since it is clearly a vector subspace, and $X + I \in K$, $Y + I \in \mathfrak{g}/I$ implies $[X, Y] + I \in K$ so that $[X, Y] \in J$). Conversely, if J is an ideal of $\mathfrak{g}$ and $I \subseteq J$ is an ideal, then $J/I \subseteq \mathfrak{g}/I$ is an ideal. □

Note. If I is an ideal of $\mathfrak{g}$ and $I \subseteq \mathfrak{h} \subseteq \mathfrak{g}$, where $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then $\mathfrak{h}/I$ is a subalgebra of $\mathfrak{g}/I$. For $\mathfrak{h}/I$ is well-defined (since I is an ideal of $\mathfrak{h}$), non-empty, and $\mathfrak{h}/I \subseteq \mathfrak{g}/I$. Conversely, if $\mathfrak{s}$ is a subalgebra of $\mathfrak{g}/I$, then $\mathfrak{s} = \mathfrak{h}/I$ for some subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ where $I \subseteq \mathfrak{h} \subseteq \mathfrak{g}$. One defines $\mathfrak{h} = \{x \in \mathfrak{g} \mid x + I \in \mathfrak{s}\}$. Clearly, $I \subseteq \mathfrak{h}$, since $0 + I \in \mathfrak{s}$. Furthermore, $\mathfrak{h}$ is a vector subspace of $\mathfrak{g}$, and is closed under the Lie bracket. Clearly, $\mathfrak{s} = \mathfrak{h}/I$.

We define the *outer derivations* by $\mathrm{Out}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g})/\mathrm{ad}(\mathfrak{g})$. We will see later that $\mathfrak{g}$ semi-simple implies $\mathrm{Out}(\mathfrak{g}) = \{0\}$, and $\mathfrak{g}$ nilpotent implies $\mathrm{Out}(\mathfrak{g})$ could be non-trivial. Recall $\mathrm{ad}(\mathfrak{g}) \cong \mathfrak{g}/Z(\mathfrak{g})$. However, $\mathfrak{g} \ncong \mathrm{ad}(\mathfrak{g}) \oplus Z(\mathfrak{g})$ in general. Note $\mathfrak{n}_3(\mathbb{C}) \ncong \mathrm{ad}(\mathfrak{n}_3(\mathbb{C})) \oplus Z(\mathfrak{n}_3(\mathbb{C}))$, i.e., the SES

$$0 \to Z(\mathfrak{n}_3(\mathbb{C})) \hookrightarrow \mathfrak{n}_3(\mathbb{C}) \to \mathrm{ad}(\mathfrak{n}_3(\mathbb{C})) \to 0$$

does not split.

9. Simple Lie algebras

Definition 9.1. A Lie algebra $\mathfrak{g}$ is *simple* if it is non-abelian and $\{0\}$ and $\mathfrak{g}$ are its only ideals.

Note. Recall the SES
$$0 \to Z(\mathfrak{g}) \hookrightarrow \mathfrak{g} \overset{\text{ad}}{\to} \text{ad}(\mathfrak{g}) \to 0.$$
If $\mathfrak{g}$ is simple, then $Z(\mathfrak{g}) = \{0\}$. Hence, $\text{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a faithful representation.

Proposition 9.2. *Any 3-dimensional simple Lie algebra (over $\mathbb{C}$) is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.*

Proof. Suppose $\mathfrak{g}$ is a 3-dimensional simple Lie algebra (over $\mathbb{C}$), with basis X_1, X_2, X_3. Since $\mathfrak{g}$ is simple, $\text{ad}(\mathfrak{g}) \cong \mathfrak{g}$. Indeed, this implies $\mathfrak{g}$ is perfect, and that we may take $[X, Y] = XY - YX$. May assume $[X_1, X_2]$ does not lie in the span of X_1, X_2, so redefine $[X_1, X_2] = X_1X_2 - X_2X_1 =: X_3$. By structure constant restraint, one can easily deduce
$$-C_{232}X_3 + C_{233}[X_3, X_1] + C_{311}X_3 + C_{313}[X_3, X_2] = 0,$$
implying
$$C_{311}X_3 - C_{232}X_3 + C_{233}(C_{311}X_1 + C_{312}X_2 + C_{313}X_3) + C_{313}(C_{321}X_1 + C_{322}X_2 + C_{323}X_3) = 0.$$
Hence, $C_{311} - C_{232} + C_{233}C_{313} + C_{313}C_{323} = 0$. However, $C_{233}C_{313} = -C_{313}C_{233}$, so $C_{311} = C_{232}$. One may scale to obtain $C_{311} = 2 = C_{232}$. So,
$$C_{233}(2X_1 + C_{312}X_2 + C_{313}X_3) + C_{313}(C_{321}X_1 - 2X_2 + C_{323}X_3) = 0.$$
That is, $C_{313}C_{231} = 2C_{233}$ and $C_{233}C_{312} = 2C_{313}$. Assuming $C_{313} \neq 0$, each of these must be non-zero, so one easily obtains $C_{312}C_{231} = 4$. However, one easily obtains a contradiction (deduce $[X_3, X_1]$ and $[X_2, X_3]$ linearly dependent). One finds the structure constants coincide with $\mathfrak{sl}_2(\mathbb{C})$, so an isomorphism is evident. □

In fact, one can classify all non-abelian Lie algebras of dimension less than or equal to 3 "by hand". The idea is to organise the study via two ideals: the centre $Z(\mathfrak{g})$ and the derived algebra $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$.

Note in the above proposition, $Z(\mathfrak{g}) = \{0\}$ and $\mathfrak{g}' = \mathfrak{g}$. If one drops the "simple" condition, then $Z(\mathfrak{g})$ could be non-zero and $\dim(\mathfrak{g}') \leq \dim(\mathfrak{g})$.

Note that abelian Lie algebras of the same dimension are all isomorphic, so there is a unique abelian Lie algebra of dimension n.

- $\dim(\mathfrak{g}) = 1$: Any Lie algebra is abelian, since $[X, X] = 0$, for all $X \in \mathfrak{g}$.
- $\dim(\mathfrak{g}) = 2$: There is a unique abelian Lie algebra.
- $\dim(\mathfrak{g})$: If $\dim(\mathfrak{g}') = 3$, then $\mathfrak{g}' = \mathfrak{g}$ and $\mathfrak{g}$ must be simple. Hence, $\mathfrak{g} \cong \mathfrak{sl}_2(\mathbb{C})$ by previous proposition.

Proposition 9.3. *There exists a unique non-abelian Lie algebra of dimension 2 and it has a basis $X, Y \in \mathfrak{g}$, with Lie bracket $[X, Y] = X$, which determines all the structure constants (this is in fact true over any field $\mathbb{F}$).*

Note. $\mathfrak{g} \subset \mathfrak{n}_2(\mathbb{C})$ (upper triangular 2×2 matrices).

Proof. Consider $\left\{e, -\frac{1}{2}h\right\}$ as a basis. Observe $[e, -\frac{1}{2}h] = e$. Indeed, this Lie algebra is of dimension 2 with desired properties. Uniqueness is clear (just define the obvious isomorphism). □

Definition 9.4. *Direct sum* of Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ is the vector space $\mathfrak{g} \oplus \mathfrak{h}$ with Lie bracket
$$[(X_1, Y_1), (X_2, Y_2)] := ([X_1, X_2]_{\mathfrak{g}}, [Y_1, Y_2]_{\mathfrak{h}}).$$

Example 9.5. We will classify all the other 3-dimensional Lie algebras by considering $\dim(\mathfrak{g}') = 1$ and $\dim(\mathfrak{g}') = 2$ and two cases for each:

(1) $\mathfrak{g}' \subseteq Z(\mathfrak{g})$;
(2) $\mathfrak{g}' \not\subseteq Z(\mathfrak{g})$.

This will yield three distinct unique Lie algebras (up to isomorphism) and one infinite family of Lie algebras:

- $\dim(\mathfrak{g}') = 1$, $\mathfrak{g}' \subseteq Z(\mathfrak{g})$, there exists a unique Lie algebra $\mathfrak{g} \cong \mathfrak{n}_3(\mathbb{C})$ (strictly upper triangular 3×3 matrices) called the *Heisenberg algebra*.
- $\dim(\mathfrak{g}') = 1$, $\mathfrak{g}' \subsetneq Z(\mathfrak{g})$, there exists a unique Lie algebra, $\mathfrak{g} \cong \mathfrak{b}_2(\mathbb{C}) \oplus \mathbb{C}$.
- $\dim(\mathfrak{g}') = 2$, $\mathfrak{g}' \subsetneq Z(\mathfrak{g})$, there exists a unique Lie algebra $\mathfrak{g} = \text{span}_{\mathbb{C}}\{X, Y, Z\}$ satisfying $[X, Y] = Y$, $[X, Z] = Y + Z$, $[Y, Z] = 0$.
- $\dim(\mathfrak{g}') = 2$, $\mathfrak{g}' \subseteq Z(\mathfrak{g})$: For each $\mu \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$ there exists a basis $\mathfrak{g}_\mu = \text{span}_{\mathbb{C}}\{X, Y, Z\}$ such that $[X, Y] = Y$, $[X, Z] = \mu Z$, $[Y, Z] = 0$. Moreover, $\mathfrak{g}_\mu \cong \mathfrak{g}_{\widetilde{\mu}}$ iff $\mu = \widetilde{\mu} \pm 1$.

This approach is not feasible for higher dimensional Lie algebras. However, the idea of using the ideal $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ to factorise the Lie algebras $\mathfrak{g}$ into smaller pieces will crucial when classifying all finite complex Lie algebras.

The other key component for the classification will be understanding representations of abelian Lie algebras and the smallest simple Lie algebra $\mathfrak{sl}_2(\mathbb{C})$.

In other words, we need to classify Lie homomorphisms:

- $\varphi : \mathfrak{a} \to \mathfrak{gl}(V)$ where $\mathfrak{a} \cong \mathbb{C}^n$ is the unique abelian Lie algebra,
- $\varphi : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V)$.

10. Solvable Lie algebras

Our aim is to classify finite dimensional complex Lie algebras up to isomorphism.

- If $\mathfrak{g}$ is Abelian, then $\mathfrak{g}' = 0$ and $\mathfrak{g} = Z(\mathfrak{g}) \cong \mathbb{C}^n$.
- If $\mathfrak{g}$ is non-Abelian, then $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{g}$ is a non-trivial ideal and we have a SES:
$$0 \to [\mathfrak{g}, \mathfrak{g}] \hookrightarrow \mathfrak{g} \to \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] \to 0.$$
The quotient algebra $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is the maximal abelian quotient of $\mathfrak{g}$, known as the *abelianisation* of $\mathfrak{g}$.

Now, if $[\mathfrak{g}, \mathfrak{g}]$ was always abelian, then the classification problem would reduce to understanding all possible extensions of abelian Lie algebras (i.e., SES where $[\mathfrak{g}, \mathfrak{g}]$ and $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ are abelian).

However, $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ is not abelian in general (it is an ideal, though), which leads us to consider the derived series of a Lie algebra.

Definition 10.1. The *derived series* of a Lie algebra $\mathfrak{g}$ is a descending series of ideals:
$$\mathfrak{g} \supset \mathfrak{g}^{(1)} \supset \mathfrak{g}^{(2)} \supset \dots,$$
defined by $\mathfrak{g}^{(k)} := \mathfrak{g}^{(k-1)\prime} = [\mathfrak{g}^{(k-1)}, \mathfrak{g}^{(k-1)}]$.

By the following proposition, it follows that $\mathfrak{g}^{(k)}/\mathfrak{g}^{(k+1)}$ is abelian.

Proposition 10.2. *If $I \subset \mathfrak{g}$ is an ideal, then $\mathfrak{g}/I$ is abelian iff $\mathfrak{g}' \subseteq I$.*

Proof. $\mathfrak{g}/I$ is abelian iff $[x + I, y + I] = I$, or equivalently, $[x, y] + I = I$ for all $x, y \in \mathfrak{g}$. This implies that $[x, y] \in I$, for all $x, y \in \mathfrak{g}$, which means $\mathfrak{g}' \subseteq I$. □

Note. This means the short exact sequences
$$0 \to \mathfrak{g}^{(k)} \to \mathfrak{g}^{(k-1)} \to \mathfrak{g}^{(k-1)}/\mathfrak{g}^{(k)} \to 0$$
have $\mathfrak{g}^{(k-1)}/\mathfrak{g}^{(k)}$.

Definition 10.3. A Lie algebra $\mathfrak{g}$ is *solvable* if $\mathfrak{g}^{(k)} = 0$ for some $k \in \mathbb{N}$.

Example 10.4. Any Abelian Lie algebra is solvable, since $\mathfrak{g}^{(1)} = 0$.

If $\dim(\mathfrak{g}) = 2$, then $\dim \mathfrak{g}' \leq 1$, so $\mathfrak{g}^{(2)} = 0$ and $\mathfrak{g}$ is solvable.

If $\dim(\mathfrak{g}) = 3$ and $\dim(\mathfrak{g}') \leq 2$, then by previous example, it follows $\mathfrak{g}^{(3)} = 0$ so all those Lie algebras are solvable. If $\dim(\mathfrak{g}') = 3$, then $\mathfrak{g} = \mathfrak{g}'$ and thus $\mathfrak{g}^{(k)} = \mathfrak{g}$ for all $k \in \mathbb{N}$ so $\mathfrak{g}$ is not solvable (e.g., $\mathfrak{sl}_2(\mathbb{C})$ is simple).

Being solvable is the complementary to being simple. The intuition is that solvable Lie algebras are not abelian, but "approximately abelian" in the sense that the k-th derived algebra is abelian for some $k \in \mathbb{N}$. Alternatively, a solvable Lie algebra is a nested series of abelian Lie algebras (extensions = SES).

On the other hand, a simple Lie algebra is "maximally non-abelian", since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$.

Definition 10.5. *Semi-direct sum* of Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ is the vector space $\mathfrak{g} \oplus_s \mathfrak{h}$ with Lie bracket

$$[(X_1, Y_1), (X_2, Y_2)] := ([X_1, X_2]_{\mathfrak{g}}, [Y_1, Y_2]_{\mathfrak{h}} + \varphi(Y_1)(X_2) - \varphi(Y_2)(X_1)),$$

where $\varphi : \mathfrak{g}_2 \to \mathfrak{gl}(\mathfrak{g}_1)$.

As we will see, every Lie algebra is a semi-direct sum of solvable Lie algebra with a direct sum of finitely many simple Lie algebras ($\mathfrak{g} \cong S \oplus_s \mathfrak{g}_{ss}$ where S solvable and $\mathfrak{g}$ semisimple ($\mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_k$, where $\mathfrak{g}_i$ simple)).

Proposition 10.6. *Let $\mathfrak{g}$ be a Lie algebra.*

(1) *If $\mathfrak{g}$ is solvable, then all subalgebras of $\mathfrak{g}$ and quotient algebras $\mathfrak{g}/I$ are solvable.*
(2) *If I is an ideal, and both I and $\mathfrak{g}/I$ are solvable, then $\mathfrak{g}$ is solvable.*
(3) *If I, J are solvable ideals of $\mathfrak{g}$, then $I + J$ is solvable.*

Proof. Suppose $\mathfrak{g}^{(k)} = 0$. If $\mathfrak{k} \subseteq \mathfrak{g}$ is a subalgebra, then $\mathfrak{k}^{(k)} = 0$ hence solvable. If $I \subseteq \mathfrak{g}$ is an ideal, then

$$(\mathfrak{g}/I)^{(m)} = \frac{\mathfrak{g}^{(m)} + I}{I}$$

and therefore $(\mathfrak{g}/I)^{(k)} = 0$. To see why we have the above, observe

$$\begin{aligned}
(\mathfrak{g}/I)^{(1)} &= \operatorname{span}\{[[x], [y]] \mid [x], [y] \in \mathfrak{g}/I\} \\
&= \operatorname{span}\{[[x], [y]] \mid x, y \in \mathfrak{g}\} + I \\
&= \operatorname{span}\{[[x, y]] \mid x, y \in \mathfrak{g}\} + I \\
&= (\operatorname{span}\{[x, y] \mid x, y \in \mathfrak{g}\} + I) + I = \frac{\mathfrak{g}^{(1)} + I}{I}.
\end{aligned}$$

This proves (1).

For (2), note this is the same as the saying the SES

$$0 \to I \to \mathfrak{g} \to \mathfrak{g}/I \to 0$$

has I, $\mathfrak{g}/I$ solvable implies $\mathfrak{g}$ is solvable. Now, suppose $I^{(m)} = 0$ and $(\mathfrak{g}/I)^{(n)} = 0$, then

$$\left(\frac{\mathfrak{g}}{I}\right)^{(n)} = \frac{\mathfrak{g}^{(n)} + I}{I} = 0.$$

So, $\mathfrak{g}^{()} \subseteq I$. We have

$$\left(\mathfrak{g}^{(n)}\right)^{(m)} = \mathfrak{g}^{(n+m)} \subseteq I^{(m)} = 0,$$

hence $\mathfrak{g}$ is solvable (proving (2)).

Suppose I, J are solvable ideals. By the Second Isomorphism Theorem,

$$\frac{I+J}{I} \cong I/I \cap J.$$

By (1), $I/I \cap J$ is solvable, hence $\frac{I+J}{I}$ is solvable. By (2), it follows that $I + J$ is solvable. □

Proposition 10.7. *If* $\mathfrak{g} = \mathfrak{g}'$ *and* $\dim(\mathfrak{g}) = 3$*, then* $\mathfrak{g}$ *must be simple.*

Proof. To derive a contradiction, suppose $\{0\} \neq I \subsetneq \mathfrak{g}$ is an ideal. Then I has dimension 1 or 2. But also, $\mathfrak{g}/I$ has dimension 1 or 2. This implies I and $\mathfrak{g}/I$ are solvable, and so by the previous proposition (1), this implies $\mathfrak{g}$ is solvable (a contradiction). □

Corollary 10.8. *Let* $\mathfrak{g}$ *be a finite dimensional Lie algebra. Then* $\mathfrak{g}$ *has a unique maximal solvable ideal (i.e., it contains all solvable ideals of* $\mathfrak{g}$*).*

Proof. Let R be a solvable ideal of $\mathfrak{g}$ of maximal dimension. If I is any solvable ideal, then $I + R$ is solvable ideal. However, $\dim(I + R) \leq \dim(R)$ by the choice of R, hence $I + R = R$. Thus, $I \subseteq R$. □

11. Semi-simple Lie algebras

Definition 11.1. The unique maximal solvable ideal of $\mathfrak{g}$ is called the *radical* of $\mathfrak{g}$, written $\mathrm{rad}(\mathfrak{g})$.

Example 11.2.
- If $\mathfrak{g}$ is solvable, then $\mathrm{rad}(\mathfrak{g}) = \mathfrak{g}$.
- If $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$, then $\mathrm{rad}(\mathfrak{g}) = 0$.
- If $\mathfrak{g} = \mathfrak{gl}_2(\mathbb{C})$, then $\mathrm{rad}(\mathfrak{g}) = Z(\mathfrak{g}) = \{\lambda I \mid \lambda \in \mathbb{C}\}$. For $Z(\mathfrak{g})$ is clearly a solvable ideal. Furthermore, if M is the maximal solvable ideal, then $M \cap \mathfrak{sl}_2(\mathbb{C}) = \{0\}$. For $M \cap \mathfrak{sl}_2(\mathbb{C}) \subseteq \mathfrak{sl}_2(\mathbb{C})$ is an ideal, and clearly must not be equal to $\mathfrak{sl}_2(\mathbb{C})$ (since $\mathfrak{sl}_2(\mathbb{C})$ has trivial radical and is simple). Suppose M contains a non-central element X. Then there exists $Y \in \mathfrak{g}$, such that $[X, Y] \neq 0$ (since X is not in the centre of $\mathfrak{g}$). As M is an ideal, $[X, Y] \in M$. But $[X, Y]$ is non-zero, and also a member of $\mathfrak{sl}_2(\mathbb{C})$ (a contradiction), thus establishing our claim.

Definition 11.3. A Lie algebra $\mathfrak{g}$ is called *semi-simple*, if $\mathrm{rad}(\mathfrak{g}) = 0$.

Proposition 11.4. *Given a Lie algebra* $\mathfrak{g}$*, the following are equivalent.*

(1) $\mathfrak{g}$ *is semi-simple;*
(2) $\mathfrak{g}$ *does not contain any non-trivial solvable ideals;*
(3) $\mathfrak{g}$ *does not contain any non-trivial abelian ideals.*

Proof. ((1) $\Longrightarrow$ (2)). By definition.

((2) $\Longrightarrow$ (3)). Suppose $\mathfrak{g}$ contains a non-trivial abelian ideal I. Then I is a solvable ideal of $\mathfrak{g}$, so radical is non-trivial.

((3) $\Longrightarrow$ (1)). Suppose $\mathfrak{g}$ is not semi-simple. Then $\mathfrak{g}$ has a non-trivial unique maximal solvable ideal I. Following the derived series by I, there exists non-trivial ideal J such that $[J, J] = 0$ (i.e., J is non-trivial abelian ideal). □

Proposition 11.5. *If* $\mathfrak{g}$ *is simple, then* $\mathfrak{g}$ *is semi-simple.*

Proof. If $\mathfrak{g}$ is simple, then the ideal $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ equals either $\{0\}$ or $\mathfrak{g}$. Since $\mathfrak{g}$ is non-abelian, this excludes $\{0\}$, so $\mathfrak{g} = \mathfrak{g}'$, and thus $\mathfrak{g}$ is not solvable.

Similarly, the ideal $\mathrm{rad}(\mathfrak{g})$ must be either $\{0\}$ or $\mathfrak{g}$. Since $\mathfrak{g}$ is not solvable, $\mathrm{rad}(\mathfrak{g}) = 0$. □

Proposition 11.6. *Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Then $\frac{\mathfrak{g}}{rad(\mathfrak{g})}$ is semi-simple.*

Proof. We want to prove that $\operatorname{rad}\left(\frac{\mathfrak{g}}{\operatorname{rad}(\mathfrak{g})}\right) = 0$. Set $R = \operatorname{rad}(\mathfrak{g})$. Let I be a solvable ideal of $\mathfrak{g}/R$. Then, by the Correspondence Theorem, $I = J/R$ for an ideal $J \subseteq \mathfrak{g}$ containing I. Now, since I is solvable and R is solvable, J must be solvable. Hence, $J \subseteq \operatorname{rad}(\mathfrak{g})$ by maximality of the radical, and therefore $I = J/\operatorname{rad}(\mathfrak{g}) = 0$. Hence, $\mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ contains no solvable ideals, so $\operatorname{rad}(\mathfrak{g}/\operatorname{rad}(\mathfrak{g})) = 0$. □

Note. This means that there exists a SES

$$0 \to \operatorname{rad}(\mathfrak{g}) \hookrightarrow \mathfrak{g} \to \frac{\mathfrak{g}}{\operatorname{rad}(\mathfrak{g})} \to 0.$$

A deep theorem by Levi shows that this is a split SES. This means that $\mathfrak{g}$ is a semi-direct sum:

$$\mathfrak{g} \cong \overbrace{\operatorname{rad}(\mathfrak{g})}^{\text{solvable}} \oplus_s \overbrace{\left(\frac{\mathfrak{g}}{\operatorname{rad}(\mathfrak{g})}\right)}^{\text{semi-simple}}.$$

We will not prove the "Levi decomposition" theorem or define semi-direct sums in this course. It suffices to know that in order to classify finite dimensional Lie algebras, we need to understand:

- solvable Lie algebras;
- semi-simple Lie algebras.

Remarkably, over complex numbers these are completely classified.

12. Nilpotent Lie algebras

12.1. **Classification of Lie algebras.** Structure of solvable Lie algebras is given by Lie's theorem:

Theorem 12.1. (Lie's Theorem). *Every solvable Lie algebra is (essentially; actually, its adjoint) isomorphic to a subalgebra of*

$$\mathfrak{b}_n(\mathbb{C}) = \{\text{upper triangular matrices}\}$$

for some n.

By Cartan's second criterion:

Theorem 12.2. *Every semi-simple Lie algebra is isomorphic to a direct sum of finitely many simple Lie algebras $\mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_r$.*

Thus, the classification boils down to classifying simple Lie algebras. This is achieved by using some representation theory of abelian subalgebras and of $\mathfrak{sl}_2(\mathbb{C})$ to translate the problem to classifying root systems. The latter is readily doable.

Theorem 12.3. *Every finite dimensional simple Lie algebra over $\mathbb{C}$ is isomorphic to one of the classical Lie algebras $\mathfrak{sl}_n(\mathbb{C})$, $\mathfrak{so}_n(\mathbb{C})$ (often split even and odd n), $\mathfrak{sp}_{2n}(\mathbb{C})$ or one of the exceptional types:*

$$\mathfrak{g}_2, f_4, e_6, e_7, e_8.$$

Note. We make note that e_8 is exceptional among even the exceptional types.

Let $J \in \mathfrak{gl}_n(\mathbb{C})$ and define

$$\mathfrak{gl}_J(n, \mathbb{C}) := \left\{x \in \mathfrak{gl}_n(\mathbb{C}) \;\middle|\; x^t J = -Jx\right\}.$$

Orthogonal Lie algebras:

- If $n = 2\ell$, take $J = \begin{pmatrix} 0 & I_\ell \\ I_\ell & 0 \end{pmatrix}$. Then $\mathfrak{so}_{2\ell}(\mathbb{C}) = \mathfrak{gl}_J(2\ell, \mathbb{C})$. Note J^2 is identity matrix.
- If $n = 2\ell + 1$, take $J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_\ell \\ 0 & I_\ell & 0 \end{pmatrix}$. Then $\mathfrak{so}_{2\ell+1}(\mathbb{C}) = \mathfrak{gl}_J(2\ell + 1, \mathbb{C})$. Again, J^2 is identity.

Symplectic Lie algebras:

- If $n = 2\ell$, take $J = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}$. Then $\mathfrak{sp}_{2\ell}(\mathbb{C}) = \mathfrak{gl}_J(2\ell, \mathbb{C})$. Notice $J^2 = -I$ (where I is identity).

Proposition 12.4. *$\mathfrak{so}_n(\mathbb{C})$ and $\mathfrak{sp}_{2\ell}(\mathbb{C})$ are Lie subalgebras of $\mathfrak{sl}_n(\mathbb{C})$.*

Proof. Suppose $x, y \in \mathfrak{so}_n(\mathbb{C})$. Then we get that the trace of x must be 0. For the trace of $x^t J$ is equal to trace of $-Jx$. Also, routine to show it is a subalgebra. Same with $\mathfrak{sp}_{2\ell}(\mathbb{C})$. □

12.2. **Nilpotent Lie algebras.** Starting with the derived algebra, one can define a slightly larger descending series of ideals that contain the derived series.

Definition 12.5. Let $\mathfrak{g}$ be a Lie algebra. The *lower central series*

$$\mathfrak{g} \supset \mathfrak{g}^1 \supset \mathfrak{g}^2 \supset \dots,$$

is defined by $\mathfrak{g}^1 = [\mathfrak{g}, \mathfrak{g}]$, and $\mathfrak{g}^k = [\mathfrak{g}, \mathfrak{g}^{k-1}]$.

Note. Indeed, as with group theory, one can define the upper central series.

Proposition 12.6. *If $\mathfrak{g}$ is a Lie algebra, then $\mathfrak{g}^{(k)} \subseteq \mathfrak{g}^k$ for all $k \in \mathbb{N}$.*

Proof. Proceed inductively. Clearly, $\mathfrak{g}^{(1)} = \mathfrak{g}^1$. Suppose $\mathfrak{g}^{(k)} \subseteq \mathfrak{g}^k$. Then $\mathfrak{g}^{(k+1)} = [\mathfrak{g}^{(k)}, \mathfrak{g}^{(k)}] \subseteq [\mathfrak{g}, \mathfrak{g}^k] = \mathfrak{g}^{k+1}$, and we are done. □

Definition 12.7. A Lie algebra $\mathfrak{g}$ is *nilpotent* if $\mathfrak{g}^k = 0$ for some $k \in \mathbb{N}$.

Note. That the series is central means

$$\frac{\mathfrak{g}^k}{\mathfrak{g}^{k+1}} \subseteq Z\left(\frac{\mathfrak{g}}{\mathfrak{g}^{k+1}}\right).$$

Indeed,

$$\left[\frac{\mathfrak{g}^k}{\mathfrak{g}^{k+1}}, \frac{\mathfrak{g}}{\mathfrak{g}^{k+1}}\right] = \frac{[\mathfrak{g}^k, \mathfrak{g}]}{\mathfrak{g}^{k+1}} = 0.$$

Since $\mathfrak{g}^{(k)} \subseteq \mathfrak{g}^k$ for all $k \in \mathbb{N}$, it follows that nilpotent Lie algebras are solvable. Converse is not true, however: let $\mathfrak{g} \subset \mathfrak{b}_2(\mathbb{C})$ be the unique non-abelian 2-dimensional Lie algebra, which is solvable. Recall that $\mathfrak{g}$ has a basis X, Y such that $[X, Y] = X$. Thus, $\mathfrak{g}^1 = \text{span}_{\mathbb{C}}(X)$, $\mathfrak{g}^2 = \text{span}_{\mathbb{C}}(X), \dots$, so $\mathfrak{g}^k = \text{span}_{\mathbb{C}}(X)$ for all $k \in \mathbb{N}$. Hence, $\mathfrak{g} = \mathfrak{b}_2(\mathbb{C})$ is not nilpotent.

We will show that a Lie algebra $\mathfrak{g}$ is solvable if, and only if, $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ is nilpotent.

Proposition 12.8. *$\mathfrak{b}_n(\mathbb{C})$ is a solvable Lie algebra, and $\mathfrak{n}_n(\mathbb{C})$ is a nilpotent Lie algebra. In fact, $\mathfrak{b}_n(\mathbb{C})' = \mathfrak{n}_n(\mathbb{C})$.*

Proof. Clearly, $\mathfrak{b}_n(\mathbb{C})' \subseteq \mathfrak{n}_n(\mathbb{C})$. The reverse inclusion is by noting $[e_{11}, e_{1n}] = e_{1n}$, and then getting all others similarly. Now, we show $\mathfrak{n}_n(\mathbb{C})$ is nilpotent (which will imply $\mathfrak{b}_n(\mathbb{C})$ is solvable). We observe $[\mathfrak{n}_n(\mathbb{C}), \mathfrak{n}_n(\mathbb{C})]$ has the diagonal above the usual is zero, and so inductively, we get $\mathfrak{n}_n(\mathbb{C})^{n-1} = 0$, finishing the proof. □

Proposition 12.9. *Subalgebras of nilpotent Lie algebras are nilpotent.*

Proof. Suppose $\mathfrak{h}$ is a subalgebra of a nilpotent Lie algebra $\mathfrak{g}$. Then $\mathfrak{g}^k = 0$ for some $k \in \mathbb{N}$. Hence, $\mathfrak{h}^k = [\mathfrak{h}, \mathfrak{h}^{k-1}] \subseteq [\mathfrak{g}, \mathfrak{g}^{k-1}] = \mathfrak{g}^k = 0$, so clearly $\mathfrak{h}$ is nilpotent. □

12.3. **Outline.** We now provide an outline for the remainder of the notes:

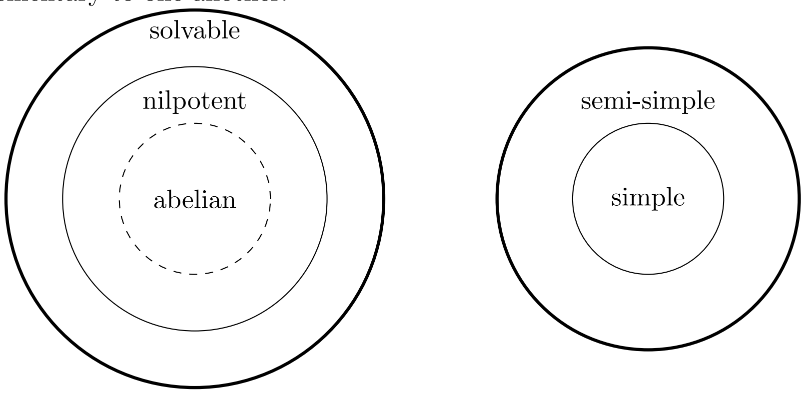
(1) Characterise nilpotent Lie algebras (Engel's Theorem);
(2) Characterise solvable Lie algebras (Lie's theorem and Cartan's first criterion);
(3) Characterise semi-simple Lie algebras (Cartan's second criterion).
(4) Classify simple Lie algebras.

The first three will require representation theory for the abelian Lie algebra $\mathbb{C}$ and introducing the Killing form.

The fourth will require studying:

- Cartan subalgebras (which are self-normalising nilpotent subalgebras),
- representation theory for $\mathfrak{sl}_2(\mathbb{C})$,
- root systems and their classification.

Below are two useful Venn diagrams to keep in mind. Note that they are complementary to one another:



- Abelian: $\mathfrak{g}' = 0$;
- Nilpotent: $\mathfrak{g}^k = 0$;
- Solvable: $\mathfrak{g}^{(k)} = 0$;
- Semi-simple implies $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}'$ and $\mathfrak{g} \cong \bigoplus_{i=1}^{r} \mathfrak{g}_i$, each $\mathfrak{g}_i$ simple (where simple means non-abelian and $\{0\}$, $\mathfrak{g}$ only ideals).

13. Characterisation of nilpotent Lie algebras

13.1. **Engel's Theorem.** Characterisation of nilpotent Lie algebras. We begin by studying nilpotent matrix Lie algebras, i.e., subalgebras of $\mathfrak{gl}(V)$, where V is a $\mathbb{C}$-vector space. First, recall that the strictly upper triangular matrices $\mathfrak{n}_n(\mathbb{C})$ is a nilpotent Lie algebra, and any subalgebra of $\mathfrak{n}_n(\mathbb{C})$ is nilpotent.

Definition 13.1. A linear map $T : V \to V$ is *nilpotent* if $T^r = 0$ for some $r \in \mathbb{N}$.

Lemma 13.2. *Elements of* $\mathfrak{n}_n(\mathbb{C})$ *are nilpotent matrices.*

Note. This justifies calling subalgebras of $\mathfrak{n}_n(\mathbb{C})$ nilpotent Lie algebras.

Proof. Let $T \in \mathfrak{n}_n(\mathbb{C})$. Its characteristic polynomial equals $\chi_T(\lambda) = \det(T - \lambda) = (-1)^n \lambda^n$, so all eigenvalues are zero. By the Cauchy-Hamilton theorem (proved below), $\chi_T(T) = (-1)^n T^n = 0$, so T is a nilpotent matrix. □

Question: If $T : V \to V$ is a nilpotent linear map, is there a basis B of V such that $[T]_B$ is strictly upper triangular?

Answer: Yes! In fact, we will show that over $\mathbb{C}$ (or any algebraically closed field $\mathbb{F}$), there is a basis B of V such that any linear transformation T is upper triangular w.r.t. B and the diagonal entries are the eigenvalues of T. If T is nilpotent, then all the eigenvalues are zero so $[T]_B$ is strictly upper triangular as required.

Note. A basis independent way of describing a linear transformation $T : V \to V$ that can be put into a strictly upper triangular form is as follows. A chain of subspaces

$$\{0\} = V_0 \subset V_1 \subset V_2 \subset \ldots \subset V_{n-1} \subset V_n = V$$

with $\dim(V_i) = i$, is called a (complete) *flag*.

The subspace V_i is *T-invariant* if $T(V_i) \subseteq V_i$. The map $T \in \mathrm{End}(V)$ has a flag of invariant subspaces iff there exists a basis B of V such that $[T]_B$ is upper triangular. The next theorem proves that over $\mathbb{C}$, this is always possible. If $T(V_i) \subseteq V_{i-1}$, then all diagonal entries are zero. This is true if T is a nilpotent map.

Theorem 13.3. *Suppose $T \in \mathrm{End}(V)$. Then there exists a basis B of V such that $[T]_B$ is upper triangular, the diagonal entries are the eigenvalues, repeated as many times as their algebraic multiplicities.*

Proof. Let $\chi_T(\lambda) = \prod_{i=1}^{p} (\lambda - \lambda_i)^{n_i}$, where $\sum_{i=1}^{p} n_i = \dim(V) =: n$, denote the characteristic polynomial of T. So, $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of T with algebraic multiplicities $n_1, n_2, \ldots, n_p$. Note that we have used the algebraic closure of $\mathbb{C}$ to ensure the existence of eigenvalues (i.e., roots of $\chi_T(\lambda)$).

Let v_1 be an eigenvector of λ_1. Then $U = \mathrm{span}(v_1)$ is a one-dimensional *invariant* subspace (i.e., $T(U) \subseteq U$). Consider the SES:

$$0 \to U \to V \to V/U \to 0.$$

T naturally defines an operator on U by restriction $T\mid_U (v) = T(v)\mid_U$ for all $v \in U$ and on the quotient space V/U by:

$$T\mid_{V/U} (v + U) = T(v) + U \quad \forall [v] \in V/U.$$

The basis $\{v_1\}$ for U can be extended to a basis $\widetilde{B} = \{v_1, \ldots, v_n\}$ of V and $\{v_2 + U, \ldots, v_n + U\}$ is a basis of V/U. These choices are equivalent to a splitting of the SES $V \cong U \oplus V/U$.

In general, a linear transformation on a direct sum can be written in block form: $T = \begin{bmatrix} T\mid_U & T_{U,V/U} \\ T_{V/U,U} & T\mid_{V/U} \end{bmatrix}$. Also, recall that the matrix of T w.r.t. the basis $\widetilde{B}$ is given by $[T]_{\widetilde{B}} = [T(v_1)\ T(v_2)\ \ldots\ T(v_n)]$. Since $T(U) \subseteq U$, it follows that

$$T(v_1) = \lambda_1 v_1 + 0 \cdot v_2 + \ldots + 0 \cdot v_n = \lambda_1 v_1.$$

Thus, $[T]_{\widetilde{B}} = \begin{bmatrix} \lambda_i & T_{U,V/U} \\ 0 & T_{V/U} \end{bmatrix}$. We have

$$\begin{aligned} \chi_T(\lambda) &= \det(T - \lambda) = (\lambda - \lambda_1) \det\left(T_{V/U} - \lambda\right) \\ &= (\lambda - \lambda_1) \chi_{T|_{V/U}}(\lambda), \end{aligned}$$

where $\chi_{T|_{V/U}}(\lambda) = (\lambda - \lambda_1)^{n_1 - 1} \cdot \prod_{i=2}^{p} (\lambda - \lambda_i)$. As $\dim(V/U) = n - 1$, we can proceed by induction to construct a basis B which satisfies the conditions in the theorem. □

Corollary 13.4. (Cayley-Hamilton Theorem). *If $T \in \mathrm{End}(V)$, then $\chi_T(T) = 0$, i.e., $T : V \to V$ is a root of its characteristic polynomial.*

Proof. Let v_1 be an eigenvector of λ_1, and set $U = \mathrm{span}(v_1)$. Arguing by induction, suppose the statement is true for $T\mid_{V/U}$. So, $\chi_{T|_{V/U}} = 0$ on V/U. Hence, $\chi_{T|_{V/U}}(T)(V) \subseteq U$. Moreover, $(T - \lambda_1)(U) = 0$,

$$\chi_T(T)(V) = (T - \lambda_1) \chi_{T|_{V/U}}(T)(V) \subseteq (T - \lambda_1)(U) = 0.$$

□

Returning to the nilpotent Lie algebra $\mathfrak{n}_n(\mathbb{C})$, we have that every element is a nilpotent matrix. Further, we have shown that any nilpotent matrix can be put into a strictly upper triangular form. Engel's theorem extends this form a single operator to an entire algebra of such operators.

Theorem 13.5. (Engel). *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra where every element $X \in \mathfrak{g}$ is nilpotent. Then there exists a basis B of V w.r.t. which $\mathfrak{g} \cong \{[X]_B \mid X \in \mathfrak{g}\} \subseteq \mathfrak{n}_n(\mathbb{C})$, so in particular, $\mathfrak{g}$ is a nilpotent Lie algebra.*

Note. The proof follows the same line of argument as proof of Theorem 13.3. The difficult step is to show that there exists a common eigenvector v_1 for all $T \in \mathfrak{g}$. We work towards this by proving by first proving two lemmas.

Lemma 13.6. *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra. If $X : V \to V \in \mathfrak{g}$ is nilpotent, then $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is nilpotent.*

Proof. If $Y \in \mathfrak{g}$, then

- $\mathrm{ad}_X(Y) = [X, Y] = XY - YX$,
- $\mathrm{ad}_X^2(Y) = X(XY - YX) - (XY - YX)X = X^2Y - 2XYX + YX^2$,

and thus $\mathrm{ad}_X^m(Y)$ is a linear combination of terms of the form X^jYX^{m-j} where $0 \leq j \leq m$. By hypothesis, $X^n = 0$ for some $n \in \mathbb{N}$, so for $m = 2n$, either X^j or X^{m-j} is zero for all $0 \leq j \leq m$. Hence $\mathrm{ad}_X^{2n} = 0$, so ad_X is nilpotent. □

Lemma 13.7. *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ and $I \subseteq \mathfrak{g}$ be an ideal. Define*

$$V_0 = \{v \in V \mid X(v) = 0 \ \forall X \in I\}.$$

Then V_0 is $\mathfrak{g}$-invariant, i.e., $Y(V_0) \subseteq V_0$ for all $Y \in \mathfrak{g}$.

Proof. For all $X \in I$, $Y \in \mathfrak{g}$ and $v \in V_0$,

$$XY(v) = \overbrace{YX(v)}^{=0} + \overbrace{[X, Y](v)}^{=0, \text{ since } [X,Y] \in I} = 0.$$

So, $Y(V_0) \subseteq V_0$, for all $Y \in \mathfrak{g}$. □

Using these lemmas, we can show that a Lie algebra of nilpotent matrices have a common eigenvector. Recall that the eigenvalue is always zero.

Proposition 13.8. *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a subalgebra consisting of nilpotent elements. Then there exists $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{g}$.*

Proof. We proceed by induction on $\dim(\mathfrak{g})$. Base case $\dim(\mathfrak{g}) = 1$ is clear since $\mathfrak{g} = \mathrm{span}\{T : V \to V\}$, T is nilpotent, has an eigenvector and $T(v) = 0$. Next, let $A \subset \mathfrak{g}$ be a maximal Lie subalgebra. We show that A must be an ideal and of codimension one, i.e., there exists $Y \in \mathfrak{g}/A$ such that $\mathfrak{g} = A \oplus \mathrm{span}\{Y\}$. First, define $\varphi : A \to \mathfrak{gl}\big(\frac{\mathfrak{g}}{A}\big)$ by $\varphi(a)(X + A) := \mathrm{ad}_a(X) + A$, for all $a \in A$ and $X \in \mathfrak{g}$.

We claim this is a well-defined Lie homomorphism. For,

$$\begin{aligned}\varphi(a + b)(X + A) &= \mathrm{ad}_{a+b}(X) + A = (\mathrm{ad}_a(X) + \mathrm{ad}_b(X)) + A \\ &= (\mathrm{ad}_a + A) + (\mathrm{ad}_b + A) \\ &= \varphi(a)(X + A) + \varphi_b(X + A),\end{aligned}$$

verifies φ is a homomorphism. It preserves the Lie bracket, since

$$\begin{aligned}[\varphi(a), \varphi(b)](X + A) &= \varphi(a) \circ \varphi(b)(X + A) - \varphi(b) \circ \varphi(a)(X + A) \\ &= \varphi(a)(\mathrm{ad}_b(X) + A) - \varphi(b)(\mathrm{ad}_a(X) + A) \\ &= (\mathrm{ad}_a \circ \mathrm{ad}_b(X) + A) - (\mathrm{ad}_b \circ \mathrm{ad}_a(X) + A) \\ &= [\mathrm{ad}_a, \mathrm{ad}_b](X) + A \\ &= \mathrm{ad}_{[a,b]}(X) + A = \varphi([a, b])(X + A),\end{aligned}$$

which verifies φ is a Lie homomorphism.

Since $a : V \to V \in A$ is nilpotent, $\mathrm{ad}_a : \mathfrak{g} \to \mathfrak{g}$ is nilpotent (by a previous lemma), and thus $\varphi(a) \in \mathfrak{gl}\big(\frac{\mathfrak{g}}{A}\big)$ is nilpotent. $\varphi(A) \subset \mathfrak{gl}\big(\frac{\mathfrak{g}}{A}\big)$ is a Lie subalgebra of dimension less than $\dim(\mathfrak{g})$, consisting of nilpotent elements. By the induction hypothesis, there exists $Y + A \in \frac{\mathfrak{g}}{A}$ such that

$$\varphi(a)(Y + A) = \mathrm{ad}_a(Y) + A = A,$$

so $\mathrm{ad}_a(Y) = [a, Y] \in A$ for all $a \in A$. Let $L = A \oplus \mathrm{span}(Y)$. This is a subalgebra of $\mathfrak{g}$ as $[A, Y] \subseteq A$, and since A is maximal, it follows that $L = \mathfrak{g}$. We have shown that $\mathfrak{g} = A \oplus \mathrm{span}\{Y\}$ and $[A, Y] \subseteq A$, so A is an ideal of $\mathfrak{g}$. By the induction hypothesis, there exists $0 \neq v \in V$ such that $a(v) = 0$ for all $a \in A$. Let

$$V_0 = \{v \in V \mid a(v) = 0 \ \forall a \in A\}.$$

Then $V_0 \neq \varnothing$ and by previous lemma, V_0 is $\mathfrak{g}$-invariant. In particular, $Y(V_0) \subseteq V_0$. Since $Y : V \to V$ is nilpotent, it restricts to a nilpotent map $V_0 \to V_0$. Thus, there exists $0 \neq v \in V_0$ such that $Y(v) = 0$. Any $X \in \mathfrak{g} = A \oplus \mathrm{span}\{Y\}$ can be written as $X = a + \lambda Y$ for some $a \in A$, $\lambda \in \mathbb{C}$, and $X(v) = (a + \lambda Y)(v) = 0$, so $X(v) = 0$ for all $X \in \mathfrak{g}$. □

Now, we provide the proof of Engel's theorem, and restate it below.

Theorem 13.9. (Engel). *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra where every element $X \in \mathfrak{g}$ is nilpotent. Then there exists a basis B of V w.r.t. which $\mathfrak{g} \cong \{[X]_B \mid X \in \mathfrak{g}\} \subseteq \mathfrak{n}_n(\mathbb{C})$, so in particular, $\mathfrak{g}$ is a nilpotent Lie algebra.*

Proof. We have a Lie subalgebra $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ consisting of nilpotent matrices. By the above proposition, they have a common eigenvector, i.e., exists $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{g}$. Define $U = \mathrm{span}\{v\}$ and proceed exactly as in the proof of Theorem 13.3 to find a basis B where $\{[X]_B \mid X \in \mathfrak{g}\} = \mathfrak{n}_n(\mathbb{C})$. □

Using Engel's Theorem for subalgebras of $\mathfrak{gl}(V)$, we can derive a second version of Engel's theorem for abstract nilpotent Lie algebras.

Theorem 13.10. (Engel, version 2). *A Lie algebra $\mathfrak{g}$ is nilpotent iff $\mathrm{ad}_X \in \mathfrak{gl}(\mathfrak{g})$ is a nilpotent operator for all $X \in \mathfrak{g}$.*

Proof. ($\Longrightarrow$) $\mathfrak{g}$ being nilpotent means that $\mathfrak{g}^k = 0$ for some $k \in \mathbb{N}$, so

$$[X_0, [X_1, [\ldots, [X_{k-2}, [X_{k-1}, X_k]]] \ldots]] = 0$$

for all $X_i \in \mathfrak{g}$. Hence, $(\mathrm{ad}_X)^k = 0$.

($\Longleftarrow$) Suppose that ad_X is nilpotent for all $X \in \mathfrak{g}$, so $\mathrm{ad}(\mathfrak{g}) = \frac{\mathfrak{g}}{Z(\mathfrak{g})}$ consists of nilpotent elements. By Engel's theorem, $\mathrm{ad}(\mathfrak{g}) \subseteq \mathfrak{gl}(\mathfrak{g})$ is a nilpotent subalgebra. Suppose $\mathrm{ad}(\mathfrak{g})^k = 0$. Then

$$0 = \mathrm{ad}(\mathfrak{g})^k = \left(\frac{\mathfrak{g}}{Z(\mathfrak{g})}\right)^k = \frac{\mathfrak{g}^k + Z(\mathfrak{g})}{Z(\mathfrak{g})},$$

so $\mathfrak{g}^k \subseteq Z(\mathfrak{g})$ which is abelian. It follows that $\mathfrak{g}^{k+1} = [\mathfrak{g}, \mathfrak{g}^k] \subseteq [\mathfrak{g}, Z(\mathfrak{g})] = 0$ so $\mathfrak{g}$ is nilpotent. □

Note.
- Engel's theorem can equivalently be stated as: "$\mathfrak{g}$ is nilpotent iff $\mathrm{ad}(\mathfrak{g})$ is nilpotent".
- The proof works over any field of arbitrary characteristic.
- If $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is a nilpotent Lie subalgebra, then by Engel's theorem, $\mathrm{ad}(\mathfrak{g}) = \{\mathrm{ad}_X \mid X \in \mathfrak{g}\}$ is a nilpotent Lie algebra and $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ are nilpotent matrices for all $X \in \mathfrak{g}$. This does not imply that $X \in \mathfrak{g}$ is a nilpotent matrix, however. For example, one needs only take the centre of $\mathfrak{gl}(V)$.

14. Characterisation of solvable Lie algebras

14.1. **Lie's Theorem.** Recall Theorem 13.3, which says that any linear transformation $T : V \to V$, where V is a $\mathbb{C}$-vector space, can be put into an upper triangular form.

In the special case when T is nilpotent, so all eigenvalues are zero, this can be extended to an entire subgalgebra of matrices, which is then necessarily form nilpotent Lie algebra.

One can ask when Theorem 13.3 can be extended to an entire Lie subalgebra of $\mathfrak{gl}(V)$, i.e., for an arbitrary collection of matrices with no constraint on the eigenvalues?

Lie's Theorem answers this affirmatively if, and only if, the subalgebra is solvable. Only then can one find a common eigenvector to all matrices in that subalgebra.

Note that two matrices $A, B \in \mathrm{End}(V)$ are diagonalisable iff $[A, B] = 0$. That is, if $X \in \mathrm{End}(V)$, then matrices which are diagonalisable with X are precisely those in the centralizer of X (i.e., $C_{\mathfrak{gl}(V)}(X)$).

Theorem 14.1. (Lie's Theorem). *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a solvable Lie algebra. Then there exists a basis $\mathscr{B}$ of V such that*

$$\mathfrak{g} \cong \{[X]_{\mathscr{B}} \mid X \in \mathfrak{g}\} \subseteq \mathfrak{b}_n(\mathbb{C}).$$

Note. Unlike Engel's theorem, Lie's theorem is false for fields of prime characteristic. Lie's Theorem characterises solvable matrix Lie algebras (recall by Ado's theorem, all Lie algebras are matrix Lie algebras), namely they all arise as Lie subalgebras of $\mathfrak{b}_n(\mathbb{C})$ (upper triangular matrices).

The difficult step in Lie's theorem is finding a common eigenvector for all $X \in \mathfrak{g}$, i.e., show that

$$V_\lambda = \{v \in V \mid X(v) = \lambda(v)v, \ \forall X \in \mathfrak{g}\}$$

is non-empty, where $\lambda : \mathfrak{g} \to \mathbb{C}$. The map $\lambda : \mathfrak{g} \to \mathbb{C}$ must be linear, so $\lambda \in \mathrm{Hom}(\mathfrak{g}, \mathbb{C}) =: \mathfrak{g}^*$ (called the dual of $\mathfrak{g}$), and is called a *weight* for $\mathfrak{g}$.

The space V_λ of common eigenvectors to $\mathfrak{g}$, if non-empty, is called the *weight space* of λ. For each $X \in \mathfrak{g}$, $\lambda(X) \in \mathbb{C}$ is the eigenvalue of that element and $v \in V_\lambda$ are the eigenvectors.

Example 14.2. Let $\mathfrak{d}_n(\mathbb{C}) \subseteq M_n(\mathbb{C})$ denote the subalgebra of diagonal matrices. Let $\{e_i\}_{i=1}^n$ be the standard basis of $\mathbb{C}^n$. Then for $X = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$, $Xe_i = \alpha_i e_i$, so we have that $V_{\lambda_i} = \mathrm{span}_{\mathbb{C}}\{e_j\}$ is a weight space for the weight $\lambda_i : \mathfrak{d}_n(\mathbb{C}) \to \mathbb{C}$ defined $\lambda_i(X) = \alpha_i$.

Now, we proceed to show that over $\mathbb{C}$ any solvable Lie subalgebra $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ has a weight space $V_\lambda \neq \varnothing$. Once this has been established, the proof of Lie's theorem proceeds exactly as of Theorem 13.3. Namely, let $0 \neq v \in V_\lambda$ be a common eigenvector of all $X \in \mathfrak{g}$, set $U = \mathrm{span}\{v\}$ and note that $\mathfrak{g}(U) \subseteq U$ (i.e., it is a $\mathfrak{g}$-invariant subspace), and so on.

We first need an analogue of Lemma 13.7, now with $\lambda \neq 0$.

Proposition 14.3. *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ and $I \subseteq \mathfrak{g}$ an ideal. Suppose $\lambda : I \to \mathbb{C}$ is a weight for I, with weight space*

$$V_\lambda = \{v \in V \mid X(v) = \lambda(X)v, \ \forall X \in I\} \neq \varnothing.$$

Then V_λ is $\mathfrak{g}$-invariant, i.e., $\mathfrak{g}(V_\lambda) \subseteq V_\lambda$.

Note. Lemma 13.7 is a corollary, for $\lambda = 0$.

Proof. We want to show that $Y(v) \in V_\lambda$ for all $Y \in \mathfrak{g}$ and $v \in V_\lambda$. Since I is an ideal, we have $[X, Y] \in I$ for all $X \in I$ and $Y \in \mathfrak{g}$, and $XY = YX + [X, Y]$, so

$$\begin{aligned} XY(v) &= (YX + [X, Y])(v) \\ &= \lambda(X)Y(v) + [X, Y](v) \\ &= \lambda(X)Y(v) + \lambda([X, Y])v. \end{aligned}$$

Hence, we need to show that $\lambda([X, Y]) = 0$.

Let $U = \operatorname{span}_{\mathbb{C}}\{v, Y(v), Y^2(v), \dots\} \subseteq V$. Since V is finite dimensional, there is a minimal $m \in \mathbb{N}$ such that $v, Y(v), \dots, Y^m(v)$ are linearly dependent. Then, $B = \{v, Y(v), \dots, Y^{m-1}(v)\}$ is a basis of U.

Now, we prove that for all $z \in I$:

(1) $z(U) \subseteq U$;
(2) the matrix of the restriction $z\,|_U$ with respect to the basis B is of the form

$$[z\,|_U]_B = \begin{pmatrix} \lambda(z) & & * \\ & \ddots & \\ 0 & & \lambda(z) \end{pmatrix}.$$

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

$$zY^r(v) = \lambda(z)Y^r(v) + \sum_{i \le r-1} c_i Y^i(v), \tag{1}$$

for $c_i \in \mathbb{C}$. We proceed by induction on r.

Base case: $r = 0$ is true as $z(v) = \lambda(z)v$ because $z \in I$. Next, suppose the claim is true for $r - 1$. Observe that

$$zY^r(v) = zY \cdot Y^{r-1}(v) = (Yz + [z, Y])Y^{r-1}(v). \tag{2}$$

By the inductive hypothesis,

$$zY^{r-1}(v) = \lambda(z)Y^{r-1}(v) + \sum_{i \le r-2} c_i Y^i(v),$$

so we get

$$YzY^{r-1}(v) = \lambda(z)Y^r(v) + \sum_{i \le r-1} c_i Y^i(v). \tag{3}$$

Moreover, since $[z, y] \in I$ (ideal condition), it follows by the inductive hypothesis

$$[z, Y]Y^{r-1}(v) = \sum_{i \le r-1} d_i Y^i(v) \tag{4}$$

for some $d_i \in \mathbb{C}$. Inserting Equation 3 and Equation 4 into Equation 2 yields Equation 1. This completes the induction and proves (1) and (2).

Now, let $z = [X, Y] \in I$. By (1), $z(U) \subseteq U$, and by (2), $\operatorname{Tr}(z\,|_U) = m\lambda(z)$, $m = \dim(U)$. Furthermore,

$$z = [X, Y] = XY - YX$$

and $X(U) \subseteq U$ by (1) and $Y(U) \subseteq U$ by definition of U. Thus, the restrictions $X\,|_U$, $Y\,|_U$ exist and $z\,|_U = [X\,|_U, Y\,|_U]$. Hence,

$$\operatorname{Tr}(z\,|_U) = \operatorname{Tr}([X\,|_U, Y\,|_U]) = 0,$$

so $m\lambda(z) = m\lambda([X, Y]) = 0$. Since $\operatorname{Char}(\mathbb{C}) = 0$, we have shown that $\lambda([X, Y]) = 0$ for all $X \in I$ and $Y \in \mathfrak{g}$. □

Proposition 14.4. *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a solvable Lie algebra. Then there exists a common eigenvector $0 \neq v \in V$ for all $X \in \mathfrak{g}$ (i.e., $X(v) \in \operatorname{span}\{v\}$ for all $X \in \mathfrak{g}$).*

Proof. The statement is true for $\dim(\mathfrak{g}) = 1$, since $\mathfrak{g} = \operatorname{span}\{T : V \to V\}$ and T has an eigenvector (since we are working over $\mathbb{C}$).

We proceed by induction on $\dim(\mathfrak{g})$. Assume $\dim \mathfrak{g} > 1$. Since $\mathfrak{g}$ is solvable, $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$. Choose a subspace $A \subset \mathfrak{g}$ such that $\mathfrak{g}' \subseteq A$ and $\dim(A) = \dim(\mathfrak{g}) - 1$. Let $z \in \mathfrak{g}/A$, so $\mathfrak{g} = A \oplus \operatorname{span}\{z\}$ (as vector spaces). Then, $[A, A] \subseteq \mathfrak{g}' \subseteq A$ and $[A, z] \subseteq \mathfrak{g}' \subseteq A$, so $A \subset \mathfrak{g}$ is an ideal. By the inductive hypothesis, there exists a common eigenvector $0 \neq v \in V$ for all $X \in A$. Hence, $V_\lambda = \{v \in V \mid X(v) = \lambda(X)v \ \forall X \in A\}$ is non-empty. By previous proposition, V_λ is $\mathfrak{g}$-invariant, so $z(V_\lambda) \subseteq V_\lambda$. Thus there exists $0 \neq w \in V_\lambda$ that is an eigenvector for z, i.e., $z(w) = \beta w$, where $\beta \in \mathbb{C}$. For any $Y = X + \alpha z \in \mathfrak{g}$, $\alpha \in \mathbb{C}$, we have

$$\begin{aligned} Y(w) &= (X + \alpha z)(w) \\ &= \lambda(X)w + \alpha \cdot \beta \cdot w. \end{aligned}$$

So, w is a common eigenvector for all $Y \in \mathfrak{g}$. □

Lie's theorem characterises solvable Lie subalgebras of $\mathfrak{gl}(V)$.

Proposition 14.5. *An abstract Lie algebra* $\mathfrak{g}$ *is solvable iff* $\operatorname{ad}(\mathfrak{g}) = \frac{\mathfrak{g}}{Z(\mathfrak{g})}$ *is solvable.*

Proof. ($\Longrightarrow$) Suppose $\mathfrak{g}$ is solvable. Then $\mathfrak{g}^{(k)} = 0$ for some $k \in \mathbb{N}$. Hence, $\mathfrak{g}^{(k)} + Z(\mathfrak{g}) = Z(\mathfrak{g})$, implying

$$0 = \frac{\mathfrak{g}^{(k)} + Z(\mathfrak{g})}{Z(\mathfrak{g})} = \left(\frac{\mathfrak{g}}{Z(\mathfrak{g})}\right)^{(k)} = \operatorname{ad}(\mathfrak{g})^{(k)}.$$

Thus, $\operatorname{ad}(\mathfrak{g})$ is solvable.

($\Longleftarrow$) Conversely, suppose $\operatorname{ad}(\mathfrak{g}) = \frac{\mathfrak{g}}{Z(\mathfrak{g})}$ is solvable. Then

$$\frac{\mathfrak{g}^{(k)} + Z(\mathfrak{g})}{Z(\mathfrak{g})} = \left(\frac{\mathfrak{g}}{Z(\mathfrak{g})}\right)^{(k)} = \operatorname{ad}(\mathfrak{g})^{(k)} = 0,$$

for some $k \in \mathbb{N}$. Hence, $\mathfrak{g}^{(k)} \subseteq \mathfrak{g}^{(k)} + Z(\mathfrak{g}) \subseteq Z(\mathfrak{g})$. Clearly, this implies $[\mathfrak{g}^{(k)}, \mathfrak{g}^{(k)}] = 0$, so $\mathfrak{g}^{(k+1)} = 0$. Thus, $\mathfrak{g}$ is solvable. □

Proposition 14.6. *A Lie algebra* $\mathfrak{g}$ *is solvable iff* $\mathfrak{g}'$ *is nilpotent.*

Proof. By the proposition above, it suffices to show this for $\operatorname{ad}(\mathfrak{g})$. By Lie's theorem, there is a basis B of $\mathfrak{g}$ w.r.t. which all elements $\operatorname{ad}_X \in \operatorname{ad}(\mathfrak{g})$ are upper triangular. Now, $\operatorname{ad}_{[X,Y]} = [\operatorname{ad}_X, \operatorname{ad}_Y]$ which in the basis B are all strictly upper triangular, and thus nilpotent matrices. By Engel's theorem, $\operatorname{ad}(\mathfrak{g})' = [\operatorname{ad}(\mathfrak{g}), \operatorname{ad}(\mathfrak{g})]$ is a nilpotent Lie algebra. The converse is obvious. □

Our next goal is to characterise semi-simple Lie algebras, and show that they factor into a direct sum of finitely many simple Lie algebras:

$$\overbrace{\mathfrak{g}}^{\text{semi-simple}} \cong \overbrace{\mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_k}^{\text{simple Lie algebras}} .$$

Thus far, we have used suitable ideals (e.g., $\mathfrak{g}'$ and $Z(\mathfrak{g})$) to break up and study Lie algebras. To advance further, we will also need to consider representations of abelian Lie algebras and the smallest simple Lie algebra $\mathfrak{sl}_2(\mathbb{C})$, which plays a crucial role. Recall that a representation of $\mathfrak{g}$ is a Lie homomorphism to (a subalgebra of) the general linear algebra $\mathfrak{gl}(V)$. We will only need to consider $\mathbb{C}$-vector spaces V.

15. Representations of abelian Lie algebras

We begin with the smallest abelian Lie algebra $\mathbb{C}$. A representation $\varphi : \mathbb{C} \to \mathfrak{gl}(V)$ is determined by the action of $1 \in \mathbb{C}$. Set $T = \varphi(1) : V \to V$. The representation theory of $\mathbb{C}$ thus boils down to the representation theory of a single linear map $T \in \mathrm{End}(V)$. This theory is given by the Jordan-Chevalley decomposition.

Definition 15.1. Let $T \in \mathrm{End}(V)$, where V is a $\mathbb{C}$-vector space. For each $\lambda \in \mathbb{C}$, the *generalised eigenspace* of T is defined by

$$V_\lambda := \{v \in V \mid (T-\lambda)^n v = 0, \quad \text{for some } n \in \mathbb{N}\}.$$

Proposition 15.2. *Suppose $T \in End(V)$, where V is a $\mathbb{C}$-vector space, and that V_λ is non-empty (where $\lambda \in \mathbb{C}$). Then V_λ is T-invariant, and $V_{\widetilde{\lambda}}$ are linearly independent for different $\widetilde{\lambda} \in \mathbb{C}$.*

Proof. Suppose $v \in V_\lambda$. Then, there exists $n \in \mathbb{N}$ such that $(T-\lambda)^n v = 0$. Observe

$$(T-\lambda)^{n-1}(T-\lambda)(v) = 0,$$

implying $(T-\lambda)^{n-1}(T(v)) = \lambda(T-\lambda)^{n-1}(v)$. Hence,

$$(T-\lambda)^n(T(v)) = \lambda(T-\lambda)^n(v) = 0,$$

and so $T(v) \in V_\lambda$. Thus, V_λ is T-invariant.

Now, suppose $v \neq 0 \in V_\lambda$ and $\widetilde{v} \neq 0 \in V_{\widetilde{\lambda}}$, where $\lambda \neq \widetilde{\lambda}$. Then, there exists $m, n \in \mathbb{N}$ such that $(T-\lambda)^m v = 0$, and $\left(T - \widetilde{\lambda}\right)^n \widetilde{v} = 0$. Suppose $\alpha v + \widetilde{\alpha}\widetilde{v} = 0$, that is, v and $\widetilde{v}$ are linearly dependent. One can easily prove by induction that

$$(T-\lambda)^m v = \sum_{j=0}^{m} (-1)^{m-j} \lambda^{m-j} T^j(v).$$

Choose m and n such that they are equal. Then,

$$\sum_{j=0}^{m} (-1)^{m-j} \lambda^{m-j} T^j(v) = 0 = \sum_{j=0}^{m} (-1)^{m-j} \widetilde{\lambda}^{m-j} T^j(\widetilde{v}).$$

Indeed, each of these individual terms will be linearly dependent on our assumption. Therefore, $\lambda T^{m-1}(\alpha v) + \widetilde{\lambda} T^{m-1}(\widetilde{\alpha}\widetilde{v}) = 0$. So, it follows $\left(\lambda - \widetilde{\lambda}\right) T^{m-1}(\widetilde{\alpha}\widetilde{v}) = 0$. Since $\lambda - \widetilde{\lambda} \neq 0$, we may then go down the sum and divide through the lambdas, deducing that $\left(\lambda - \widetilde{\lambda}\right)\widetilde{\alpha}\widetilde{v} = 0$. This means that $\widetilde{\alpha} = 0$, and one similarly deduces $\alpha = 0$. That is, v and $\widetilde{v}$ are linearly dependent. One continues by induction, yielding the result. □

Theorem 15.3. (Jordan-Chevalley decomposition). *Any $T \in \mathrm{End}(V)$ has a* unique *decomposition:*

$$T = T_{SS} + T_n,$$

with $[T_{SS}, T_n] = 0$, where T_{SS} is a semi-simple operator (i.e., diagonalisable) and T_n is a nilpotent operator.

Note. If T is written in Jordan normal form, then T_{SS} is the matrix with only diagonal entries and T_n the matrix with just off-diagonal terms. The Jordan-Chevalley decomposition holds under a weaker hypothesis than the existence of a Jordan normal form however. Also, unlike the latter which is basis dependent, the Jordan-Chevalley decomposition is unique.

Proof. First, we note that the generalised eigenspaces V_λ of T are non-trivial for only finitely many $\lambda \in \mathbb{C}$, namely the eigenvalues of T. Let $\sigma(T)$ denote the set of eigenvalues, i.e., the spectrum of T. Consider the quotient space $\overline{V} = V/\bigoplus_{\lambda\in\sigma(T)} V_\lambda$. The quotient transformation $T\mid_{\overline{V}}$ no longer has any spectrum, but this contradicts the fundamental theorem of algebra, applied to the characteristic polynomial of $T\mid_{\overline{V}}$. Hence, the V_λ's span V:

$$V = \bigoplus_{\lambda\in\sigma(T)} V_\lambda.$$

On each V_λ, the operator $(T-\lambda\mathrm{id})\mid_{V_\lambda}$ is nilpotent and $\lambda\mathrm{id}_{V_\lambda}$ is diagonal. Define $T_{SS} = \bigoplus_{\lambda\in\sigma(T)} \lambda\mathrm{id}_{V_\lambda}$ and $T_n = \bigoplus_{\lambda\in\sigma(T)} (T-\lambda)\mid_{V_\lambda}$, and we obtain the Jordan-Chevalley decomposition. T_{SS} is semisimple (or diagonalisable) which means that all generalised eigenspaces are actually eigenspaces. Clearly, $T_{SS}T_n = T_nT_{SS}$. To prove uniqueness, suppose we have an alternate decomposition: $T = \widetilde{T_{SS}} + \widetilde{T_n}$, where $[\widetilde{T_{SS}}, \widetilde{T_n}] = 0$. We have $T_{SS} - \widetilde{T_{SS}} = \widetilde{T_n} - T_n$. Since $\widetilde{T_{SS}}, \widetilde{T_n}$ commute with T, they also commute with T_{SS}, T_n. Now, the sum of commuting semisimple or nilpotent operators are semisimple respectively nilpotent. Since the only operator that is both semisimple and nilpotent is the zero operator, we conclude that $T_{SS} = \widetilde{T_{SS}}$ and $T_n = \widetilde{T_n}$. □

Proposition 15.4. *Let $T \in \mathrm{End}(V)$ with Jordan-Chevalley decomposition $T = T_{SS} + T_n$. There are polynomials $p, q \in \mathbb{C}[\lambda]$ such that $T_{SS} = p(T)$ and $T_n = q(T)$.*

Proof. Consider the characteristic polynomial of T

$$\chi_T(\lambda) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{n_i}.$$

The factors $(\lambda-\lambda_i)^{n_i}$ are pairwise coprime. By the Chinese Remainder theorem, the quotient map $\mathbb{C}[\lambda] \to \bigoplus_{i=1}^{k} \mathbb{C}[\lambda]/((\lambda_i-\lambda)^{n_i})$ is surjective. Thus, for $\lambda_1,\dots,\lambda_k$ we can find a polynomial such that $p(\lambda) = \lambda_i \mod (\lambda-\lambda_i)^{n_i}$ for all $i = 1, 2, \dots, k$ and $p(\lambda) = 0 \mod \lambda$. Note: if 0 is an eigenvalue of T, then the last condition is redundant; if not, then λ is relatively prime to all other factors. Set $T_{SS} = p(T)$. Since T_{SS} commutes with T, it leaves all the generalised eigenspaces V_{λ_i} invariant. As $(T-\lambda_i)^{n_i}$ restricts to zero on V_{λ_i}, by construction $T_{SS} = p(T) = \lambda_i\mathrm{id}_{V_{\lambda_i}}$ on V_{λ_i}, which is semisimple. Define $q(\lambda) = \lambda - p(\lambda)$ and set $T_n = q(T)$, which restricts to the nilpotent operator $T - \lambda_i$ on each V_{λ_i}. □

Lemma 15.5. *Let $T \in \mathrm{End}(V)$ with Jordan decomposition $T = T_{SS} + T_n$. Then $\mathrm{ad}_T \in \mathfrak{gl}(\mathrm{End}(V))$ has Jordan decomposition $\mathrm{ad}_T = \mathrm{ad}_{T_{SS}} + \mathrm{ad}_{T_n}$.*

Proof. By linearity, $\mathrm{ad}_T = \mathrm{ad}_{T_{SS}} + \mathrm{ad}_{T_n}$. Moreover, $[\mathrm{ad}_{T_{SS}}, \mathrm{ad}_{T_n}] = \mathrm{ad}_{[T_{SS},T_n]} = 0$, so $\mathrm{ad}_{T_{SS}}$ and ad_{T_n} commute. By a previous lemma, ad_{T_n} is nilpotent since T_n is nilpotent. It remains to show that $\mathrm{ad}_{T_{SS}}$ is semisimple and the result follows by the uniqueness of the Jordan-Chevalley decomposition. Consider the basis in which T_{SS} is diagonal. Then for the standard basis $\{e_{ij}\}$ of $\mathrm{End}(V)$, $\mathrm{ad}_{T_{SS}}$ is diagonal, i.e., semisimple. Namely, if $T_{SS} = \mathrm{diag}(\lambda_1,\dots,\lambda_k)$, then $\mathrm{ad}_{T_{SS}}(e_{ij}) = (\lambda_i - \lambda_j)e_{ij}$. □

Next, we consider the representations of abelian Lie algebras $\mathfrak{h} \cong \mathbb{C}^k$. In the proof of Proposition 15.4, we used the fact that if two operators commute, then the generalised eigenspaces of one operator are preserved by the other. Using this observation, we can understand the representations of an abelian Lie algebra $\mathfrak{h} \cong \mathbb{C}^k$:

$$\rho : \mathfrak{h} \to \mathfrak{gl}(V).$$

Namely, there is a finite set $\sigma(\rho) \subset \mathfrak{h}^* = \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$ of linear functionals $\lambda : \mathfrak{h} \to \mathbb{C}$ for which the generalised eigenspaces

$$V_\lambda = \{v \in V \mid (\rho(\mathfrak{h}) - \lambda)^n v = 0 \text{ for some } n \in \mathbb{N}\}$$

are non-trivial and $\mathfrak{h}$-invariant $(\rho(\mathfrak{h}) - \lambda)^n v = 0$ is short-hand notation for

$$(p(X_1) - \lambda(X_n)) \ldots (p(X_n) - \lambda(X_n))v = 0$$

for all $X_1, \ldots, X_n \in \mathfrak{h}$. Moreover, we have the decomposition $V = \bigoplus_{\lambda \in \sigma(\rho(\mathfrak{h}))} V_\lambda$. An important special case is when all the elements of $\mathfrak{h}$ act is a semisimple fashion, i.e., $\rho(X)$ is semisimple (equivalently, diagonalisable) for all $X \in \mathfrak{h}$. Then, the generalised eigenspaces V_λ are just eigenspaces or weight spaces, i.e., $\rho(x)v = \lambda(x)v$ for all $x \in \mathfrak{h}$ and $v \in V_\lambda$, so v is a weight vector with weight λ.

This will be crucial in understanding simple Lie algebras. Namely, we will find a maximal commutative subalgebra $\mathfrak{h} \subset \mathfrak{g}$ consisting of semisimple elements. Then, $\mathrm{ad}\,|_{\mathfrak{h}} \colon \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g})$ yields $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in \sigma(\mathrm{ad}(\mathfrak{h}))} \mathfrak{g}_\alpha$.

16. Characterisation of semi-simple Lie algebras

Recall that a Lie algebra $\mathfrak{g}$ is semisimple if its maximal solvable subalgebra $\mathrm{rad}(\mathfrak{g})$ is trivial. Equivalently, $\mathfrak{g}$ does not have any non-trivial solvable ideals. Recall that this implies that $\mathfrak{g}$ does not have any non-trivial abelian ideals, and the converse also holds. Namely, if $\mathfrak{g}$ has a non-trivial solvable ideal I, then every element of the derived series of I is also an ideal of $\mathfrak{g}$ and in particular $\mathfrak{g}$ will have a non-trivial abelian ideal.

Our next objective is to give two other characterisations of semisimple Lie algebras.

Definition 16.1. The *Killing form* $K : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ is a symmetric bilinear form on $\mathfrak{g}$ defined by

$$K(X, Y) := \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y).$$

Lemma 16.2. *$K(X, [Y, Z]) = K([X, Y], Z)$, for each $X, Y, Z \in \mathfrak{g}$.*

Proof. Observe

$$\begin{aligned} K(X, [Y, Z]) &= \mathrm{Tr}\big(\mathrm{ad}_X \circ \mathrm{ad}_{[Y,Z]}\big) \\ &= \mathrm{Tr}(\mathrm{ad}_X \circ (\mathrm{ad}_Y \circ \mathrm{ad}_Z - \mathrm{ad}_Z \circ \mathrm{ad}_Y)) \\ &= \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y \circ \mathrm{ad}_Z) - \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Z \circ \mathrm{ad}_Y) \\ &= \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y \circ \mathrm{ad}_Z) - \mathrm{Tr}(\mathrm{ad}_Y \circ \mathrm{ad}_X \circ \mathrm{ad}_Z) \\ &= \mathrm{Tr}((\mathrm{ad}_X \circ \mathrm{ad}_Y - \mathrm{ad}_Y \circ \mathrm{ad}_X) \circ \mathrm{ad}_Z) \\ &= \mathrm{Tr}\big(\mathrm{ad}_{[X,Y]} \circ \mathrm{ad}_Z\big) = K([X, Y], Z). \end{aligned}$$

□

Using K, we first give another characterisation of solvable Lie algebras.

We work towards proving Cartan's first criterion.

Lemma 16.3. *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a solvable Lie subalgebra. Then $\mathrm{Tr}(XY) = 0$, for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}'$.*

Proof. By Lie's theorem, there is a basis B of V such that $[X]_B \in \mathfrak{b}_n(V)$ for all $X \in \mathfrak{g}$. Then $[[X_1]_B, [X_2]_B] \in \mathfrak{n}_n(V)$ for all $X_1, X_2 \in \mathfrak{g}$. Hence $[Y]_B \in \mathfrak{n}_n(V)$ for all $Y \in \mathfrak{g}'$, and $[X]_B, [Y]_B \in \mathfrak{n}_n(V)$, so $\mathrm{Tr}(XY) = \mathrm{Tr}([X]_B[Y]_B) = 0$. □

Proposition 16.4. *Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra. Suppose $\mathrm{Tr}(XY) = 0$, for all $X, Y \in \mathfrak{g}$. Then $\mathfrak{g}$ is solvable.*

Proof. The idea is to show that if $\mathrm{Tr}(XY) = 0$ for all $X, Y \in \mathfrak{g}$, then every $Z \in \mathfrak{g}'$ is a nilpotent operator. By Engel's theorem $\mathfrak{g}'$ is then a nilpotent Lie algebra and by Proposition 14.6, $\mathfrak{g}$ is solvable.

Let $X \in \mathfrak{g}'$ and consider its Jordan decomposition $X = X_{SS} + X_n$. Note that X_{SS}, X_n are not necessarily in $\mathfrak{g}$. There is a basis B with respect to which $[X_{SS}]_B = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ and $[X_n]_B$ is strictly upper triangular.

We need to prove that $X_{SS} = 0$, i.e., $\lambda_1 = \dots = \lambda_n = 0$. This will follow if we can show that $\sum_{i=1}^n \lambda_i \overline{\lambda_i} = 0$. Define $\overline{X_{SS}} : V \to V$ by $\overline{X_{SS}} := \mathrm{diag}\big(\overline{\lambda_1}, \dots, \overline{\lambda_n}\big)$. Then $\mathrm{Tr}\big(\overline{X_{SS}}X\big) = \sum_{i=1}^n \lambda_i \overline{\lambda_i}$, so it suffices to show that $\mathrm{Tr}\big(\overline{X_{SS}}[Y,Z]\big) = 0$, for all $Y, Z \in \mathfrak{g}$ because $X \in \mathfrak{g}' = \mathrm{span}\{[Y,Z] \mid Y, Z \in \mathfrak{g}\}$. Observe

$$\begin{aligned}\mathrm{Tr}([X,Y]Z) &= \mathrm{Tr}((XY - YX)Z) = \mathrm{Tr}(XYZ - YXZ) \\ &= \mathrm{Tr}(XYZ) - \mathrm{Tr}(YXZ) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(XZY) \\ &= \mathrm{Tr}(XYZ - XZY) = \mathrm{Tr}(X(YZ - ZY)) = \mathrm{Tr}(X[Y,Z]),\end{aligned}$$

for all $X, Y, Z \in \mathfrak{gl}(V)$. It follows that

$$\mathrm{Tr}\big(\overline{X_{SS}}[Y,Z]\big) = \mathrm{Tr}\big([\overline{X_{SS}}, Y]Z\big).$$

By Lemma 15.5, the Jordan decomposition of ad_X is $\mathrm{ad}_{X_{SS}} + \mathrm{ad}_{X_n}$. By Proposition 15.4, there is a polynomial $q(X) \in \mathbb{C}[X]$ such that $\overline{\mathrm{ad}_{X_{SS}}} = q(\mathrm{ad}_X)$ and clearly $\overline{\mathrm{ad}_{X_{SS}}} = \mathrm{ad}_{\overline{X}_{SS}}$. Therefore, $\mathrm{ad}_{\overline{X}_{SS}}$ maps $\mathfrak{g}$ to $\mathfrak{g}$ (even if $\overline{X}_{SS}$ does not belong to $\mathfrak{g}$), since $\mathrm{ad}_{\overline{X}_{SS}}$ is a polynomial in ad_X. In particular, $\mathrm{ad}_{\overline{X}_{SS}}(Y) = [\overline{X}_{SS}, Y] \in \mathfrak{g}$. Since $\mathrm{Tr}(XY) = 0$ for all $X, Y \in \mathfrak{g}$, we conclude that $\mathrm{Tr}\big([\overline{X}_{SS}, Y]Z\big) = 0$, so $\lambda_i = 0$ for all $i = 1, \dots, n$ and X is nilpotent. □

Theorem 16.5. (Cartan's first criterion). *$\mathfrak{g}$ is solvable iff $\mathfrak{g} \perp_K \mathfrak{g}'$, i.e., $K(X,Y) = 0$ for all $X \in \mathfrak{g}$, $Y \in \mathfrak{g}'$.*

Note. Recall that $\mathfrak{b}_n(\mathbb{C})' = \mathfrak{n}_n(\mathbb{C})$, so $\mathfrak{b}_n(\mathbb{C})$ is K-orthogonal to $\mathfrak{n}_n(\mathbb{C})$.

Proof. ($\Longrightarrow$) Suppose $\mathfrak{g}$ is solvable, so equivalently $\mathrm{ad}(\mathfrak{g}) \subseteq \mathfrak{gl}(\mathfrak{g})$ is solvable. By Lemma 16.3, $\mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y) = K(X,Y) = 0$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}'$.

($\Longleftarrow$) Suppose $K(X,Y) = 0$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}'$. By Proposition 16.4, $\mathfrak{g}^{(2)}$ is solvable. But this implies that $\mathfrak{g}'$ is solvable, so $\mathfrak{g}$ is solvable. □

Example 16.6. Let $\mathfrak{g}$ be the non-trivial 2-dimensional solvable Lie algebra with basis $B = \{X, Y\}$ and $[X,Y] = X$. Then $[\mathrm{ad}_X]_B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $[\mathrm{ad}_Y]_B = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$, $K(X,X) = 0$, $K(X,Y) = 0$, $K(Y,Y) = 1$, so the associated matrix to the Killing form is $[K]_B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. It is a degenerate matrix. Moreover, $K(X,X) = K(Y,X) = 0$ so $K(\mathfrak{g}, \mathfrak{g}') = 0$ since $\mathfrak{g}' = \mathrm{span}_{\mathbb{C}}\{X\}$, as expected by Cartan's first criterion.

Before moving on to semisimple Lie algebras, let us first recall some basic notions in linear algebra:

- If $\{e_i\}_{i=1}^n$ is a basis of V, then $\{f_i\}_{i=1}^n$ is a basis of the dual space $V^* = \mathrm{Hom}_{\mathbb{C}}(V, \mathbb{C})$, where $f_i(e_j) = \delta_{ij}$.
- Given a subspace $W \subseteq V$ and a symmetric bilinear form $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$, the orthogonal complement of W is the vector subspace
$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \ \forall w \in W\}.$$

Definition 16.7. $\langle \cdot, \cdot \rangle$ is *non-degenerate* if $V^\perp = 0$.

Proposition 16.8. *$\langle \cdot, \cdot \rangle$ is non-degenerate iff its matrix w.r.t. a basis of V is invertible.*

Proof. ($\Longrightarrow$) Suppose $\langle\cdot,\cdot\rangle$ is non-degenerate. Then, $V^\perp = 0$. Let $\{e_i\}_{i=1}^n$ be a basis for V, and A be the matrix of $\langle\cdot,\cdot\rangle$ w.r.t. $\{e_i\}_{i=1}^n$ (i.e., $A_{ij} = \langle e_i, e_j\rangle$). Then, $A_{ij} \neq 0$ for all i, j. If Trace$(A) = 0$, this implies $\sum_{i=1}^n A_{ii} = 0$. That is, $\sum_{i=1}^n \langle e_i, e_i\rangle = 0$. Suppose this holds for every base. Then, given j,

$$\langle e_1 + e_j, e_1 + e_j\rangle + \sum_{i=2}^n \langle e_i, e_i\rangle = 0,$$

implying

$$\langle e_1, e_j\rangle + \langle e_j, e_1\rangle + \langle e_1, e_1\rangle = 0.$$

In particular, by choosing $j = 1$, we get $\langle e_i, e_i\rangle = 0$ for all i. But then by symmetry, the above implies $\langle e_1, e_j\rangle = 0$ for all j. Clearly, this means $e_1 \in V^\perp$, a contradiction. So, there must exist a basis for which the trace is non-zero (and so the matrix is invertible).

($\Longleftarrow$) Let $\{e_i\}_{i=1}^n$ be the basis of V such that the matrix A is invertible. If $V^\perp$ is not non-degenerate, then there exists non-zero $v = \sum_{i=1}^n c_i e_i$ such that $\langle e_i, v\rangle = 0$ for all $1 \le i \le n$. So, $\left\langle e_i, \sum_{j=1}^n c_j e_j\right\rangle = \sum_{j=1}^n c_j\langle e_i, e_j\rangle = 0$. Then, $-\langle e_j, e_j\rangle = \sum_{i=1, i\neq j}^n c_i\langle c_j^{-1} e_j, e_i\rangle$, and summing RHS yields 0. So, for non-zero c_j, get sum of $\langle e_j, e_j\rangle = 0$. Since trace non-zero, sum $\langle e_j, e_j\rangle \neq 0$ for zero c_j. □

Any non-degenerate bilinear form defines an isomorphism $V \to V^*$, $v \mapsto \langle v, -\rangle : V \to \mathbb{C}$.

Proposition 16.9. *Let $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$ be non-degenerate and $W \subseteq V$ a subspace. Then $\dim(W) + \dim\left(W^\perp\right) = \dim(V)$. Thus, if $W \cap W^\perp = 0$, then $V = W \oplus W^\perp$.*

Proof. We show $V = W + W^\perp$. Let $v \in V$, $\{w_i\}_{i=1}^{\dim(W)}$ be a basis for W, and $\{w_i^\perp\}_{i=1}^{\dim(W)^\perp}$ be a basis for $W^\perp$. Let $w = \sum_{i=1}^{\dim(W)} \langle v, w_i\rangle w_i$ and $w^\perp = v - w$. Then,

$$\langle w^\perp, w_k^\perp\rangle = \langle v, w_k^\perp\rangle - \sum_{i=1}^{\dim(W)} \langle v, w_i\rangle\langle w_i, w_k^\perp\rangle = \langle v, w_k^\perp\rangle.$$

So,

$$\langle w^\perp, w_k\rangle = \left\langle v, w_k - \sum_{i=1}^{\dim(W)} w_i\langle w_i, w_k\rangle\right\rangle.$$

□

Finally, given an ideal of a Lie algebra $I \subseteq \mathfrak{g}$, we may ask how the Killing form K_I and $K_\mathfrak{g}$ are related?

Lemma 16.10. *Suppose $I \subseteq \mathfrak{g}$ is an ideal. Then,*

(1) *$K_I(X, Y) = K_\mathfrak{g}(X, Y)$ for all $X, Y \in I$;*
(2) *$I^\perp \subseteq \mathfrak{g}$ is an ideal, where $I^\perp = \{X \in \mathfrak{g} \mid K(X, Y) = 0\ \forall Y \in I\}$.*

Proof. Let B a basis of I and extend it to a basis $\widetilde{B}$ of $\mathfrak{g}$. As $I \subseteq \mathfrak{g}$ is an ideal, ad_X maps $\mathfrak{g}$ to I for all $X \in I$, so

$$[\mathrm{ad}_X]_{\widetilde{B}} = \begin{bmatrix} [\mathrm{ad}_X]_B & N_X \\ 0 & 0 \end{bmatrix},$$

and

$$[\mathrm{ad}_X]_{\widetilde{B}}[\mathrm{ad}_Y]_{\widetilde{B}} = \begin{bmatrix} [\mathrm{ad}_X]_B[\mathrm{ad}_Y]_B & [\mathrm{ad}_X]_B N_Y \\ 0 & 0 \end{bmatrix}.$$

Hence,

$$K_\mathfrak{g}(X, Y) = \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y) = \mathrm{Tr}([\mathrm{ad}_X]_B[\mathrm{ad}_Y]_B) = K_I(X, Y),$$

which proves (1).

Clearly, $0 \in I^\perp$, so non-empty. Also, if $X, Y \in I^\perp$, then $K(X,Z) = K(Y,Z) = 0$ for all $Z \in I$. Hence, $K(X+Y,Z) = K(X,Z) + K(Y,Z) = 0$, by bilinearity, and also $K(\lambda X, Z) = \lambda K(X,Z) = 0$. So, $I^\perp$ is a vector subspace of $\mathfrak{g}$. If $X \in \mathfrak{g}$, $Y \in I^\perp$ and $Z \in I$, then

$$K([X,Y],Z) = K(Z,[X,Y]) = K([Z,X],Y) = K(Y,[Z,X]) = 0,$$

since $[Z,X] \in I$ (as I is an ideal). Thus, $[X,Y] \in I^\perp$, and so $I^\perp$ is an ideal (proving (2)). □

Theorem 16.11. (Cartan's second criterion). *$\mathfrak{g}$ is semi-simple iff K is non-degenerate.*

Proof. ($\Longrightarrow$) Suppose $\mathfrak{g}$ is semi-simple. Then, by previous lemma, $\mathfrak{g}^\perp \subseteq \mathfrak{g}$ is an ideal and $K(X,Y) = 0$ for all $X, Y \in \mathfrak{g}^\perp$. By Proposition 16.4, $\mathfrak{g}^\perp$ is solvable and hence $\mathfrak{g}^\perp \subseteq \mathrm{rad}(\mathfrak{g}) = 0$, so $\mathfrak{g}^\perp = 0$, i.e., K is non-degenerate.

($\Longleftarrow$) We show that if $\mathfrak{g}$ is non-semisimple, then K is degenerate, i.e., $\mathfrak{g}^\perp \neq 0$. Suppose $\mathfrak{r} := \mathrm{rad}(\mathfrak{g}) \neq 0$ and consider its derived series

$$\mathfrak{r} \supset \mathfrak{r}^{(1)} \supset \mathfrak{r}^{(2)} \supset \ldots \supset \mathfrak{r}^{(t)} = 0.$$

Then, $\mathfrak{a} = \mathfrak{r}^{(t-1)}$ is a non-zero abelian ideal of $\mathfrak{g}$. We show that $\mathfrak{a} \subseteq \mathfrak{g}^\perp$.

Claim: $\mathrm{ad}_a \circ \mathrm{ad}_x \in \mathfrak{gl}(\mathfrak{g})$ is nilpotent for all $a \in \mathfrak{a}$ and $x \in \mathfrak{g}$. Indeed, for all $y \in \mathfrak{g}$ we have

$$(\mathrm{ad}_a \circ \mathrm{ad}_x \circ \mathrm{ad}_a)(y) = [a, \underbrace{[x, \overbrace{[a,y]}^{\in \mathfrak{a}}]}_{\in \mathfrak{a}}] = 0,$$

since $\mathfrak{a}$ is an abelian ideal. So, $(\mathrm{ad}_a \circ \mathrm{ad}_x)^2 = 0$. Hence, $\mathrm{Tr}(\mathrm{ad}_a \circ \mathrm{ad}_x) = K(a,x) = 0$, which implies that $0 \neq \mathfrak{a} \subseteq \mathfrak{g}^\perp \subseteq \mathfrak{g}$. □

Lemma 16.12. *Suppose $\mathfrak{g}$ is semi-simple and $0 \neq I \subseteq \mathfrak{g}$ is an ideal. Then, we have a K-orthogonal decomposition:*

$$\mathfrak{g} = I \oplus I^\perp,$$

where I is itself semi-simple.

Proof. Since $K(X,Y) = 0$ for all $X, Y \in I \cap I^\perp$, by Cartan's first criterion $I \cap I^\perp$ is solvable. As $\mathfrak{g}$ is semi-simple, $I \cap I^\perp = 0$, and since K is non-degenerate by Cartan's second criterion, $\mathfrak{g} = I \oplus I^\perp$ by previous proposition. As $I \cap I^\perp = 0$, the restriction of K to I is non-degenerate and by Lemma 16.10 part (1), $K|_I = K_I$, so K_I is non-degenerate and I is semi-simple by Cartan's second criterion. □

Theorem 16.13. *$\mathfrak{g}$ is semi-simple iff $\mathfrak{g}$ is a direct sum of simple ideals, i.e., $\mathfrak{g} = \bigoplus_{i=1}^r \mathfrak{g}_i$ (each $\mathfrak{g}_i$ simple ideals).*

Note. This means that $\mathfrak{g} = \bigoplus_{i=1}^r \mathfrak{g}_i$ as a vector space, and $[\mathfrak{g}_i, \mathfrak{g}_j] \subseteq \mathfrak{g}_i \cap \mathfrak{g}_j = 0$ for all $i \neq j$.

Proof. ($\Longrightarrow$) Proceed by induction on $\dim(\mathfrak{g})$. Suppose $\mathfrak{g}$ is semi-simple. Base case: the statement is clear if $\mathfrak{g}$ is simple, and in particular if $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$, the lowest dimensional simple Lie algebra.

Now, let $0 \neq I \subsetneq \mathfrak{g}$ be an ideal. By the previous lemma, $\mathfrak{g} = I \oplus I^\perp$ and the ideals $I, I^\perp$ are semi-simple. By the inductive hypothesis, $I = \bigoplus_{i=1}^m L_i$ and $I^\perp = \bigoplus_{i=1}^n K_i$, where L_i, K_i are simple ideals of I, $I^\perp$, respectively. Then,

$$\begin{aligned} [L_i, \mathfrak{g}] &= [L_i, I \oplus I^\perp] = [L_i, I] \oplus [L_i, I^\perp] \\ &= [L_i, I] \subseteq L_i, \end{aligned}$$

since $[L_i, I^\perp] \subseteq [I, I^\perp] \subseteq I \cap I^\perp = 0$. Hence, all L_i's are ideals of $\mathfrak{g}$ and similarly for the K_i's. Hence,

$$\mathfrak{g} = L_1 \oplus \ldots \oplus L_m \oplus K_1 \oplus \ldots \oplus K_n$$

is a direct sum of simple ideals.

($\Longleftarrow$) Suppose $\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_r$ where $\mathfrak{g}_i$ are simple ideals. Let $I \subset \mathfrak{g}$ be a solvable ideal. Then, $[I, \mathfrak{g}_i] \subseteq I \cap \mathfrak{g}_i$, so $[I, \mathfrak{g}_i] = 0$ since $\mathfrak{g}_i$ is simple and I is solvable. Hence,

$$[I, \mathfrak{g}] = \bigoplus_{i=1}^{r} [I, \mathfrak{g}_i] = 0$$

and $I \subseteq Z(\mathfrak{g}) = \bigoplus_{i=1}^{r} Z(\mathfrak{g}_i) = 0$. □

Our goal for the remainder of the course is to classify all the complex simple Lie algebras up to isomorphism. This will require three main ingredients:

(1) representation theory $\mathfrak{sl}_2(\mathbb{C})$;
(2) prove the existence of maximal abelian subalgebras consisting of semi-simple (i.e., diagonalisable) elements, called *Cartan subalgebras*;
(3) use representation theory of abelian Lie algebras to get a *weight decomposition* of $\mathfrak{g}$.

This combined will give a bijective correspondence between simple Lie algebras and irreducible root systems. This translates the classification of simple Lie algebras to that of irreducible root systems, which is readily doable.

16.1. **Jordan decomposition in semi-simple Lie algebras.** Recall that $X \in \mathfrak{g} \subseteq \mathfrak{gl}(V)$ has a unique Jordan-Chevalley decomposition $X = X_{SS} + X_n$, $[X_{SS}, X_n] = 0$, and in general $X_{SS}, X_n \notin \mathfrak{g} \subseteq \mathfrak{gl}(V)$.

Example 16.14. $\mathfrak{g} = \text{span}_{\mathbb{C}}(X) \subseteq \mathfrak{gl}(V)$. Then $X = X_{SS} + X_n$, but $X_{SS}, X_n \notin \mathfrak{g}$ unless $X = X_{SS}$ or $X = X_n$.

However, if $\mathfrak{g}$ is semi-simple, then X_{SS}, X_n do lie in $\mathfrak{g}$. In fact, we have such a decomposition for any abstract semi-simple Lie algebra. We will not prove the following statements, but you can find proofs in sections 9.5 and 9.6 of Erdmann-Wildon.

Theorem 16.15. *Let $\mathfrak{g}$ be a semi-simple Lie algebra over $\mathbb{C}$. Then every $X \in \mathfrak{g}$ has a unique decomposition $X = X_{SS} + X_n$ such that*

(1) $X_{SS}, X_n \in \mathfrak{g}$ *and* $[X_{SS}, X_n] = 0$*;*
(2) $\text{ad}_{X_{SS}}, \text{ad}_{X_n} \in \mathfrak{gl}(\mathfrak{g})$ *are semi-simple and nilpotent respectively.*

Proposition 16.16. *If $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is semi-simple, then the abstract Jordan decomposition in the above theorem is the same as the Jordan-Chevalley decomposition of X as a linear transformation $V \to V$.*

Proposition 16.17. *Let $\mathfrak{g}$ be semi-simple and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ a representation. Suppose $X \in \mathfrak{g}$ has a Jordan decomposition $X = X_{SS} + X_n$ as in the above theorem. Then, $\rho(X) = \rho(X_{SS}) + \rho(X_n)$ in the Jordan-Chevalley decomposition of $\rho(X)$.*

17. Basic representation theory

Before looking at representations of $\mathfrak{sl}_2(\mathbb{C})$, we need to establish some basic definitions and properties of representations in general.

Recall that a representation of $\mathfrak{g}$ is a Lie homomorphism $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$, where V is a complex vector space.

Example 17.1. $\rho : \mathfrak{g} \to \mathfrak{gl}(\mathbb{C})$ defined $\rho(X) = 0$ for all $X \in \mathfrak{g}$ is called the *trivial representation.*

Example 17.2. ad : $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(\mathfrak{sl}_2(\mathbb{C}))$ and recall $B = \left\{ e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$ is the standard basis of $\mathfrak{sl}_2(\mathbb{C})$. We also recall $[e, f] = h$, $[e, h] = -2e$ and $[f, h] = 2f$. Therefore,

$$[\mathrm{ad}_e]_B = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad [\mathrm{ad}_h]_B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad [\mathrm{ad}_f]_B = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Since $Z(\mathfrak{sl}_2(\mathbb{C})) = 0$, ad is a faithful representation of $\mathfrak{sl}_2(\mathbb{C})$ on itself, as a vector space $\mathfrak{sl}_2(\mathbb{C}) \cong \mathbb{C}^3$.

Example 17.3. Let $\mathfrak{g}$ be the non-trivial 2-dimensional Lie algebra spanned by $B = \{X, Y\}$ with $[X, Y] = X$. We get $\rho : \mathfrak{g} \to \mathfrak{gl}(\mathbb{C}^2)$ determined by $\rho(X) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\rho(Y) = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$ is a faithful representation, so $\rho(\alpha X + \beta Y) = \begin{pmatrix} -\beta & \alpha + \beta \\ 0 & 0 \end{pmatrix}$ for any $\alpha X + \beta Y \in \mathfrak{g}$, $\alpha, \beta \in \mathbb{C}$.

It is often naturally convenient to drop ρ and write $X \cdot v$ instead of $\rho(X)(v)$. This defines a map $\mathfrak{g} \times V \to V$, $(X, v) \mapsto X \cdot v := \rho(X)(v)$ which satisfies

- $(X_1 + X_2) \cdot v = X_1 \cdot v + X_2 \cdot v$;
- $X \cdot (\alpha v_1 + \beta v_2) = \alpha X \cdot v_1 + \beta X \cdot v_2$; and
- $[X_1, X_2] \cdot v = X_1 \cdot (X_2 \cdot v) - X_2 \cdot (X_1 \cdot v)$.

We say that V is a $\mathfrak{g}$-*module* if there is a map $\mathfrak{g} \times V \to V$ satisfying the above conditions. $\mathfrak{g}$-modules and representations of $\mathfrak{g}$ are equivalent concepts.

Definition 17.4. Let V be a $\mathfrak{g}$-module. A subspace $W \subseteq V$ is a $\mathfrak{g}$-*submodule* if W is $\mathfrak{g}$-invariant, i.e., $\mathfrak{g} \cdot W \subseteq W$.

Note. Observe that if ρ is the trivial representation, then every subspace is a $\mathfrak{g}$-submodule.

Example 17.5. Let $\mathfrak{g}$ be a $\mathfrak{g}$-module with respect to the adjoint representation. $W \subseteq \mathfrak{g}$ is a submodule iff $\mathfrak{g} \cdot W \subseteq W$, i.e, $[\mathfrak{g}, W] \subseteq W$, which is equivalent to W being an ideal of $\mathfrak{g}$.

Definition 17.6. A $\mathfrak{g}$-module V is *irreducible* iff 0 and V are the only submodules.

Note. In some textbooks, *irreducible* modules are called *simple*.

Example 17.7. Any one-dimensional $\mathfrak{g}$-module is irreducible. For V is one-dimensional, then the only subspaces of V are itself and 0.

Example 17.8. The submodules of the adjoint $\mathfrak{g}$-module $\mathfrak{g}$ are ideals of $\mathfrak{g}$. Hence, the adjoint module $\mathfrak{g}$ is irreducible iff $\mathfrak{g}$ is a simple Lie algebra. Thus, ad : $\mathfrak{gl}_2(\mathbb{C}) \to \mathfrak{gl}_3(\mathbb{C})$ from our previous example is irreducible (where $\mathfrak{gl}_3(\mathbb{C}) \supset \mathfrak{gl}(\mathfrak{sl}_2(\mathbb{C}))$).

Definition 17.9. Let $W \subseteq V$ be a $\mathfrak{g}$-submodule. Then V/W is a *quotient* $\mathfrak{g}$-*module* with the $\mathfrak{g}$-action: $X \cdot (v + W) = X \cdot v + W$, for all $X \in \mathfrak{g}$ and $v \in V$.

Proposition 17.10. *The* $\mathfrak{g}$-*action above is well-defined (i.e., independent of representations).*

Proof. Suppose $W \subseteq V$ is a $\mathfrak{g}$-submodule, and that $v_1 + W = v_2 + W$ (where $v_1, v_2 \in W$). Suppose $X \in \mathfrak{g}$. Since W is $\mathfrak{g}$-invariant, $X \cdot (v_1 - v_2) \in W$ (note $v_1 - v_2 \in W$ because $v_1 + W = v_2 + W$). So, $X \cdot v_1 - X \cdot v_2 \in W$, which implies $X \cdot v_1 + W = X \cdot v_2 + W$. Thus, the $\mathfrak{g}$-action is indeed well-defined. □

Definition 17.11. If V is a $\mathfrak{g}$-module, then a sequence of submodules

$$0 = V_0 \subset V_1 \subset V_2 \subset \ldots \subset V_r = V$$

where V_i/V_{i-1} is irreducible for all i, is called the *composition series* of V.

Definition 17.12. Let V be a $\mathfrak{g}$-module. Suppose $V = U_1 \oplus \ldots \oplus U_r$ (i.e., $V = U_1 + \ldots + U_r$ and $U_i \cap \left(\sum_{i \neq j} U_j\right) = 0$ for all i), and each U_i is a submodule. Then we say that V is a *direct sum* of submodules $U_1, \ldots, U_r$. If each U_i is irreducible, then we call V a *completely reducible* $\mathfrak{g}$-module.

Example 17.13. Let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ and $V = \mathfrak{g} \oplus \mathbb{C}$, where $\mathfrak{g}$ is the adjoint module and $\mathbb{C}$ the trivial representation. Then V is completely reducible.

Example 17.14. Let $\mathfrak{g} = \mathfrak{b}_2(\mathbb{C})$ acting on $V = \mathbb{C}^2$. The only non-trivial (i.e., not 0 nor V) submodule of V is $U = \mathrm{span}_{\mathbb{C}}(e_1)$, so V is not irreducible (i.e., V is reducible) but $V \neq U \oplus U$, so V is not completely reducible.

Note. From these examples, we see that modules over solvable Lie algebras are not completely reducible. On the other hand, it is a deep theorem by Weyl that all finite dimensional $\mathbb{C}$-modules over semi-simple Lie algebras are completely reducible. Thus, to classify representations of semi-simple Lie algebras, it suffices to classify irreducible modules up to isomorphism.

Definition 17.15. Let V and W be $\mathfrak{g}$-modules. A *homomorphism* of $\mathfrak{g}$-modules is a linear map $\varphi : V \to W$ such that $\varphi(X \cdot v) = X \cdot \varphi(v)$ for all $X \in \mathfrak{g}$ and $v \in V$. An *isomorphism* is a bijective homomorphism, also called an *intertwiner* of $\mathfrak{g}$-modules.

In terms of representations, $\rho_V : \mathfrak{g} \to \mathfrak{gl}(V)$, $\rho_W : \mathfrak{g} \to \mathfrak{gl}(W)$ a homomorphism $\varphi : V \to W$ satisfies

$$\varphi(\rho_V(X)(v)) = \rho_W(X)(\varphi(v)),$$

i.e., $\varphi \circ \rho_V = \rho_W \circ \varphi$. If φ is an isomorphism, then $\varphi \circ \rho_V \circ \varphi^{-1} = \rho_W$. So φ is a similarity transformation. Namely, we can find bases B_V and B_W such that $[\rho_V(X)]_{B_V} = [\rho_W(X)]_{B_W}$ for all $X \in \mathfrak{g}$.

18. Representations of $\mathfrak{sl}_2(\mathbb{C})$

We will classify all the irreducible $\mathfrak{sl}_2(\mathbb{C})$-modules (up to isomorphism), showing that there is exactly one of each dimension.

Construction: Let $\mathbb{C}[X, Y]$ be the vector space of all polynomials in X, Y. Let $V_d = \mathrm{span}_{\mathbb{C}}\{X^d, X^{d-1}Y, \ldots, XY^{d-1}, Y^d\}$ be the subspace of all *homogeneous* polynomials of degree $d \in \mathbb{N}_0$. Thus, $\dim(V_0) = d + 1$.

Let e, f, h denote the standard basis: $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$. Define $\rho : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{gl}(V_d)$ by $\rho(e) = X\frac{\partial}{\partial Y}$, $\rho(f) = Y\frac{\partial}{\partial X}$ and $\rho(h) = X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}$, and extend by linearity.

Proposition 18.1. *ρ is a representation, so V_d is a $(d+1)$-dimensional $\mathfrak{sl}_2(\mathbb{C})$-module.*

Proof. We need to check ρ is a Lie homomorphism. For $b \geq 1$,

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

and

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

Hence, $\rho([h,e]) = [\rho(h), \rho(e)]$.

For $a \geq 1$,

$$\begin{aligned}[\rho(h), \rho(f)]\big(X^aY^b\big) &= \rho(h)\rho(f)\big(X^aY^b\big) - \rho(f)\rho(h)\big(X^aY^b\big) \\ &= \rho(h)\big(aX^{a-1}Y^{b+1}\big) - \rho(f)\big((a-b)X^aY^b\big) \\ &= a(a-b+2)X^{a-1}Y^{b+1} - (a-b)aX^{a-1}Y^{b+1} \\ &= 2aX^{a-1}Y^{b+1} = -2\rho(f)\big(X^aY^b\big),\end{aligned}$$

and

$$[\rho(h), \rho(f)]\big(Y^d\big) = \rho(h)(0) - \rho(f)\big(-dY^d\big) = 0 = -2\rho(f)\big(Y^d\big).$$

Hence, $\rho([h,f]) = [\rho(h), \rho(e)]$.

Finally, for $a, b \geq 1$,

$$\begin{aligned}[\rho(e), \rho(f)]\big(X^aY^b\big) &= \rho(e)\rho(f)\big(X^aY^b\big) - \rho(f)\rho(e)\big(X^aY^b\big) \\ &= \rho(e)\big(aX^{a-1}Y^{b+1}\big) - \rho(f)\big(bX^{a+1}Y^{b-1}\big) \\ &= a(b+1)X^aY^b - b(a+1)X^aY^b \\ &= (a-b)X^aY^b = \rho(h)X^aY^b,\end{aligned}$$

and we also have

$$[\rho(e), \rho(f)]X^d = \rho(e)\big(YdX^{d-1}\big) - \rho(f)(0) = dX^d = \rho(h)\big(X^d\big)$$

and

$$[\rho(e), \rho(f)]Y^d = \rho(e)(0) - \rho(f)\big(XdY^{d-1}\big) = -dY^d = \rho(h)\big(Y^d\big).$$

Thus, $\rho([e,f]) = [\rho(e), \rho(f)]$, and we are done. □

Proposition 18.2. *Let $B = \{X^d, X^{d-1}Y, \ldots, XY^{d-1}, Y^d\}$ be the basis of V_d. Then,*

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

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

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

Proof. □

Example 18.3.
- $V_0 = \mathbb{C}$ is the trivial representation.
- $V_1 = \mathbb{C}^2$ is the fundamental representation where $\mathfrak{sl}_2(\mathbb{C})$ acts by 2×2 matrices.

- $V_2 = \mathfrak{sl}_2(\mathbb{C})$ is isomorphic to the adjoint representation, where we recall

$$[\mathrm{ad}_e]_B = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad [\mathrm{ad}_h]_B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad [\mathrm{ad}_f]_B = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

We must find the change of basis matrix, that transforms these to

$$[\rho(e)]_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad [\rho(h)]_B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad [\rho(f)]_B = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Next, we show that V_d is irreducible for all $d \in \mathbb{N}_0$ and these are all the irreducible $\mathfrak{sl}_2(\mathbb{C})$-modules.

By Weyl's complete reducibility theorem, any other $\mathfrak{sl}_2(\mathbb{C})$-module decomposes into a direct sum of the V_d's.

Proposition 18.4. *V_d is irreducible.*

Proof. Suppose $U \subset V_d$ is a non-zero submodule, so U is preserved by $\rho(e)$, $\rho(f)$, $\rho(h)$. Since $\rho(h)$ has $d+1$ distinct eigenvalues, the eigenvalues of the restriction $\rho(h)\mid_U\colon U \to U$ are also distinct and U contains an eigenvector for $\rho(h)$. The eigenspaces of $\rho(h)$ are 1-dimensional spanned by the basis vectors in $B = \{X^{d-j}Y^j\}_{j=0,\dots,d}$. Hence, $X^aY^b \in U$ for some a, b. Applying $\rho(e)$ to this successively gives

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

and applying $\rho(f)$ gives

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

Hence, $U = V_d$, so V_d is irreducible. □

Lemma 18.5. *Let V be a finite dimensional $\mathfrak{sl}_2(\mathbb{C})$-module.*

- *If $v \in V$ with $h \cdot v = \lambda v$, then*

$$h \cdot (e \cdot v) = (\lambda + 2)e \cdot v$$
$$h \cdot (f \cdot v) = (\lambda - 2)f \cdot v.$$

- *V contains an eigenvector $w \neq 0$ for h such that $e \cdot w = 0$.*

Proof. We first prove (1). Observe that

$$\begin{aligned} h \cdot (e \cdot v) &= e \cdot (h \cdot v) + [h, e] \cdot v \\ &= \lambda e \cdot v + 2e \cdot v = (\lambda + 2)e \cdot v \\ h \cdot (f \cdot v) &= f \cdot (h \cdot v) + [h, f] \cdot v \\ &= \lambda f \cdot v - 2f \cdot v = (\lambda - 2)f \cdot v. \end{aligned}$$

We now prove (2). Note that the linear map $v \mapsto h \cdot v$ has an eigenvector (since the field is $\mathbb{C}$), say $h \cdot v = \lambda v$. Consider $v, e \cdot v, e^2 \cdot v, \dots$. If all of them non-zero, by (1), these are eigenvectors for h with distinct eigenvalues, hence they are linearly independent. As V is finite dimensional, there exists k such that $e^k \cdot v \neq 0$ but $e^{k+1} \cdot v = 0$. Then $w = e^k \cdot v$. □

Theorem 18.6. *If V is a finite dimensional irreducible $\mathfrak{sl}_2(\mathbb{C})$-module, then $V \cong V_d$ for some $d \in \mathbb{N}_0$.*

Proof. By Lemma 18.5 (2), there is a $0 \neq w \in V$ such that $e \cdot w = 0$, $h \cdot w = \lambda w$. Moreover, there exists a $d \in \mathbb{N}$ such that $f^d \cdot w \neq 0$, $f^{d+1} \cdot w = 0$.

We claim $B = \{w, f \cdot w, \dots, f^d \cdot w\}$ is a basis of V, consisting of h-eigenvectors with distinct eigenvalues $\lambda, \lambda-2, \dots, \lambda-2d$. The latter claim follows by Lemma 18.5

(1). To show that B spans V, let $U = \operatorname{span}_{\mathbb{C}}(B)$. Then clearly $f \cdot U \subseteq U$ and $h \cdot U \subseteq U$. Moreover, we show by induction on k that $e \cdot f^k \cdot w \in \operatorname{span}_{\mathbb{C}}\{w, f \cdot w, \ldots, f^{k-1} \cdot w\}$ for all $k \leq d$. This is clearly true for $k = 0$, as $e \cdot w = 0$. Assume this is true for $k - 1$. Then

$$\begin{aligned} e \cdot f^k \cdot w &= e \cdot f \cdot \left(f^{k-1} \cdot w\right) \\ &= (f \cdot e + [e, f])\left(f^{k-1} \cdot w\right) \\ &= (f \cdot e + h)\left(f^{k-1} \cdot w\right). \end{aligned}$$

By the inductive hypothesis, $e \cdot f^{k-1} \cdot w \in \operatorname{span}_{\mathbb{C}}\{w, f \cdot w, \ldots, f^{k-2} \cdot w\}$, so $f \cdot e \cdot f^{k-1} \cdot w \in \operatorname{span}_{\mathbb{C}}\{w, f \cdot w, \ldots, f^{k-1} \cdot w\}$ and so is $h \cdot f^{k-1} \cdot w$. This completes the induction and shows that $e \cdot U \subseteq U$. Hence we have shown that U is an $\mathfrak{sl}_2(\mathbb{C})$-submodule of V, and as V is irreducible, we have $U = V$.

In the basis B, we have

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

Moreover, $h = [e, f] \in \mathfrak{g}'$, so $\operatorname{Tr}[h]_B = \operatorname{Tr}[e, f] = 0$. Thus,

$$\lambda + (\lambda - 2) + \ldots + (\lambda - 2d) = 0,$$

implying $(d + 1)\lambda = d(d + 1)$ and hence $\lambda = d$.

Now, we can show that $V \cong V_d$. The $\mathfrak{sl}_2(\mathbb{C})$-module V has a basis $B = \{w, f \cdot w, \ldots, f^d \cdot w\}$ and V_d has a basis $\widetilde{B} = \{X^d, fX^d, \ldots, f^d X^d\}$. Both bases consist of h-eigenvectors with eigenvalues d, $d - 2$, $\ldots$, $-d$. Define $\varphi : V \to V_d$ by $\varphi\left(f^k \cdot w\right) = f^k X^d$, $0 \leq k \leq d$.

Clearly, φ is a bijection, so we need to show that is an isomorphism of $\mathfrak{sl}_2(\mathbb{C})$-modules, i.e. (by linearity) $\varphi(x \cdot \varphi(v))$ for all $v \in B$ and $x = e, f, h$. We have

$$\begin{aligned} x = f \qquad f \cdot \varphi\left(f^k \cdot w\right) &= f \cdot \left(f^k \cdot X^d\right) = f^{k+1} \cdot X^d \\ &= \varphi\left(f \cdot \left(f^k \cdot w\right)\right) \\ x = h \qquad h \cdot \varphi\left(f^k \cdot w\right) &= h \cdot f^k \cdot X^d \\ &= (d - 2k) f^k X^d \\ &= \varphi\left(h \cdot f^k \cdot w\right). \end{aligned}$$

For $x = e$, we show that

$$e \cdot \varphi\left(f^k \cdot w\right) = \varphi\left(e \cdot f^k \cdot w\right)$$

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

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

as $e \cdot w = 0$. Now assume the equality holds for $k - 1$. Then

$$\begin{aligned} \varphi\left(e \cdot f^k \cdot w\right) &= \varphi\left((f \cdot e + h) \cdot f^{k-1} \cdot w\right) \\ &= f \cdot \varphi\left(e \cdot f^{k-1} \cdot w\right) + h \cdot \varphi\left(f^{k-1} \cdot w\right) \\ \text{IH} \qquad &= f \cdot e \cdot \varphi\left(f^{k-1} \cdot w\right) + h \cdot \varphi\left(f^{k-1} \cdot w\right) \\ &= e \cdot f \cdot \varphi\left(f^{k-1} \cdot w\right) \\ &= e \cdot \varphi\left(f^k \cdot w\right), \end{aligned}$$

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

19. Cartan subalgebras

While a simple Lie algebra $\mathfrak{g}$ has no non-trivial ideals (other than $\mathfrak{g}$ itself), it can have non-trivial subalgebras. Consider the adjoint representation,

$$\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$$

which is faithful for simple Lie algebras (since $Z(\mathfrak{g}) = 0$). If we can find a non-trivial abelian subalgebra $\mathfrak{h} \subset \mathfrak{g}$, then by the representation theory of abelian Lie algebras, the restriction

$$\mathrm{ad}\,|_{\mathfrak{h}} : \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g})$$

will give a decomposition of $\mathfrak{g}$ into generalised weight spaces

$$\mathfrak{g} = \bigoplus_{\lambda \in \sigma(\mathrm{ad}|_{\mathfrak{h}})} \mathfrak{g}_\lambda.$$

Now, if we can find an abelian subalgebra $\mathfrak{h}$ consisting of *semi-simple* elements, i.e., ad_X is diagonalisable for all $X \in \mathfrak{h}$, then the decomposition will be into actual *weight spaces*:

$$\mathfrak{g} = \bigoplus_{\alpha \in \sigma(\mathrm{ad}|_{\mathfrak{h}})} \mathfrak{g}_\alpha, \;\; \mathfrak{g}_\alpha = \{x \in \mathfrak{g} \mid \mathrm{ad}_h(x) = \alpha(h)x \;\forall h \in \mathfrak{h}\} \neq 0.$$

Note. $[\mathrm{ad}_{h_1}, \mathrm{ad}_{h_2}] = \mathrm{ad}_{[h_1,h_2]} = 0$.

The weights for the adjoint representations are called *roots.* If $\Phi := \sigma(\mathrm{ad}\,|_{\mathfrak{h}}) \setminus \{0\}$ denotes the non-zero roots, then we can write the *root decomposition* as

$$\mathfrak{g} = \mathfrak{g}_0 \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha.$$

This idea turns out to be the key to obtaining a complete classification of simple Lie algebras. Hence, the strategy is to:

(1) Find an abelian subalgebra $\mathfrak{h} \subset \mathfrak{g}$ consisting of semi-simple elements.
(2) Decompose $\mathfrak{g}$ into a direct sum of root (i.e., weight) spaces of the adjoint representation $\mathrm{ad}\,|_{\mathfrak{h}}$.
(3) Use the root decomposition to pin down structure constants.

Example 19.1. $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C}) = \mathrm{span}\{e, f, h\}$ where $\mathfrak{h} = \mathrm{span}\{\mathfrak{h}\}$ is an abelian subalgebra and ad_h is diagonalisable.

Example 19.2. $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{C})$, where $\dim(\mathfrak{g}) = 3^2 - 1 = 8$. Let $\mathfrak{h}$ be the subalgebra consisting of diagonal traceless matrices in $\mathfrak{g}$. Then, $h \in \mathfrak{h}$ iff $h = \begin{pmatrix} a_1 & & 0 \\ & a_2 & \\ 0 & & a_3 \end{pmatrix}$, where $a_1 + a_2 + a_3 = 0$. So, $\dim(\mathfrak{h}) = 2$. Now, if $e_{ij} \in \mathfrak{sl}_3(\mathbb{C})$, $i \neq j$, then $\mathrm{ad}_h(e_{ij}) = (a_i - a_j)e_{ij}$, so ad_h is diagonalisable and we get

$$\mathfrak{sl}_3(\mathbb{C}) = \mathfrak{h} \bigoplus_{i \neq j = 1}^{3} \mathrm{span}(e_{ij})$$

is the root space decomposition. The roots of ad_h corresponding to each of these root (or weight) spaces are

root space	root
$\mathfrak{h}$	0
$\mathrm{span}(e_{ij})$	$\varepsilon_i - \varepsilon_j$

where $\varepsilon_i : \mathfrak{g} \to \mathbb{C}$ is defined by $h = \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix} \mapsto a_i$.

Next, we prove the existence and characterise maximal abelian subalgebras $\mathfrak{h}$, called *Cartan subalgebras*, for any simple Lie algebra $\mathfrak{g}$.

First, let us record some properties of the root decomposition $\mathfrak{g} = \mathfrak{g}_0 \oplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$. Note that $\mathfrak{h} \subseteq \mathfrak{g}_0 = \{x \in \mathfrak{g} \mid \mathrm{ad}_h(x) = 0\ \forall h \in \mathfrak{h}\}$ and recall that Φ are non-zero weights $\alpha \in \mathfrak{h}^* = \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$ for which $\mathfrak{g}_\alpha \neq 0$.

Proposition 19.3. *Let $\alpha, \beta \in \mathfrak{h}^* = \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$. Then,*

(1) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$*;*
(2) *if $\alpha + \beta \neq 0$, then $K(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$, i.e., $\mathfrak{g}_\alpha \perp \mathfrak{g}_\beta$ w.r.t. the Killing form K; and*
(3) *if $\mathfrak{g}$ is semi-simple, then $K\mid_{\mathfrak{g}_0}$ is non-degenerate (i.e., $\mathfrak{g}_0 \cap \mathfrak{g}_0^\perp = 0$).*

Proof. Let $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$. For all $h \in \mathfrak{h}$,

$$\begin{aligned}\mathrm{ad}_h([X,Y]) &= [\mathrm{ad}_h(X), Y] + [X, \mathrm{ad}_h(Y)]\\ &= \alpha(h)[X,Y] + \beta(h)[X,Y]\\ &= (\alpha+\beta)(h)[X,Y].\end{aligned}$$

Hence, $[X,Y] \in \mathfrak{g}_{\alpha+\beta}$, proving (1).

Let $h \in \mathfrak{h}$ with $(\alpha+\beta)(h) \neq 0$. For any $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$, we have

$$\begin{aligned}\alpha(h)K(X,Y) = K([h,X],Y) &= -K([X,h],Y)\\ &= -K(X,[h,Y]) = -\beta(h)K(X,Y).\end{aligned}$$

Hence, $(\alpha+\beta)(h)K(X,Y) = 0$, so $K(X,Y) = 0$ (proving (2)).

Let $Y \in \mathfrak{g}_0 \cap \mathfrak{g}_0^\perp$. Then $K(\mathfrak{g}_0, Y) = 0$. Now, any $X \in \mathfrak{g}$ can be written

$$X = X_0 + \sum_{\alpha\in\Phi} X_\alpha,$$

where $X_\alpha \in \mathfrak{g}_\alpha$. By (2), $K(\mathfrak{g}_0, \mathfrak{g}_\alpha) = 0$ if $\alpha \neq 0$, so $K(X_\alpha, Y) = 0$ for all $X \in \Phi$. Also, $K(X_0, Y) = 0$ by assumption. Hence, $K(X,Y) = 0$ for all $X \in \mathfrak{g}$, so $Y \in \mathfrak{g}^\perp$. As $\mathfrak{g}$ is semi-simple, K is non-degenerate, so $\mathfrak{g}^\perp = 0$. Hence, $Y = 0$, so $\mathfrak{g}_0 \cap \mathfrak{g}_0^\perp$. □

Corollary 19.4. *If $X \in \mathfrak{g}_\alpha$, $\alpha \neq 0$, then ad_X is nilpotent.*

Proof. For any root $\alpha \in \Phi \cup \{0\}$,

$$\begin{aligned}\mathrm{ad}_X(\mathfrak{g}_\beta) &\subseteq \mathfrak{g}_{\alpha+\beta}\\ \mathrm{ad}_X^2(\mathfrak{g}_\beta) &\subseteq \mathfrak{g}_{2\alpha+\beta}\\ &\vdots\\ \mathrm{ad}_X^r(\mathfrak{g}_\beta) &\subseteq \mathfrak{g}_{r\alpha+\beta},\end{aligned}$$

by previous proposition part (1). However, Φ is finite, so for some $r \in \mathbb{N}$ we have $r\alpha+\beta \notin \Phi$. This means that $(\mathrm{ad}_X)^r(\mathfrak{g}_\beta) = 0$. Again, since Φ is finite, and for each $\beta \in \Phi \cup \{0\}$ we can find r such that $(\mathrm{ad}_X)^r(\mathfrak{g}_\beta) = 0$, we can take the maximum of these r to get $(\mathrm{ad}_X)^r(\mathfrak{g}_\beta) = 0$ for all $\beta \in \Phi \cup \{0\}$. Since $\mathfrak{g} = \bigoplus_{\beta\in\Phi\cup\{0\}} \mathfrak{g}_\beta$, follows $(\mathrm{ad}_X)^r = 0$ on $\mathfrak{g}$. □

Now, each choice of abelian subalgebra $\mathfrak{h} \subset \mathfrak{g}$ consisting of semi-simple elements gives rise to a direct sum decomposition $\mathfrak{g} = \mathfrak{g}_0 \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$. Hence, it makes sense to seek the largest such subalgebra.

Definition 19.5. A subalgebra $\mathfrak{h}$ of a semi-simple Lie algebra $\mathfrak{g}$ is called a *Cartan subalgebra* if $\mathfrak{h}$ is abelian, every element $h \in \mathfrak{h}$ is semi-simple (i.e., ad_h is diagonalisable) and $\mathfrak{h}$ is maximal among subalgebras satisfying these two properties.

Note. A Cartan subalgebra is normally defined as a nilpotent self-normalising subalgebra $\mathfrak{h}$ (i.e., $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$) of $\mathfrak{g}$. Over an algebraically closed field (e.g., $\mathbb{C}$) and when $\mathfrak{g}$ is semi-simple, this definition is equivalent to the one above.

A subalgebra consisting of semi-simple elements is called *toral*. Any toral subalgebra must be abelian (if $[h_1, h_2] \neq 0$, then $\mathrm{ad}_{[h_1,h_2]} = [\mathrm{ad}_{h_1}, \mathrm{ad}_{h_2}] \neq 0$ is diagonalisable; however, the trace is clearly 0, contradiction). Hence, a Cartan subalgebra is the same as a maximal toral subalgebra if $\mathfrak{g}$ is semi-simple.

Proposition 19.6. *Let $\mathfrak{g}$ be a finite dimensional semi-simple Lie algebra over $\mathbb{C}$. Then $\mathfrak{g}$ has a Cartan subalgebra.*

Proof. Let $X = X_{SS} + X_n \in \mathfrak{g}$ be the Jordan decomposition. Suppose $X_{SS} = 0$ for all $X \in \mathfrak{g}$. Then $\mathrm{ad}_X \in \mathrm{ad}(\mathfrak{g}) \subseteq \mathfrak{gl}(\mathfrak{g})$ is nilpotent for all $X \in \mathfrak{g}$. By Engel's theorem, $\mathrm{ad}(\mathfrak{g})$ and hence $\mathfrak{g}$ is a nilpotent Lie algebra, which contradicts semi-simplicity of $\mathfrak{g}$. Hence, there is $X \in \mathfrak{g}$ such that $X_{SS} \neq 0$. $\mathrm{span}_{\mathbb{C}}\{X_{SS}\} \subset \mathfrak{g}$ is an abelian subalgebra, so $\mathfrak{g}$ has a non-trivial abelian subalgebra consisting of semi-simple elements. Such a subalgebra of maximal dimension is a Cartan subalgebra. □

Definition 19.7. The *centraliser* of a subset $S \subseteq \mathfrak{g}$, is the subalgebra $C_{\mathfrak{g}}(S) = \{X \in \mathfrak{g} \mid [X, Y] = 0 \ \forall Y \in S\}$.

Lemma 19.8. *The centralizer $C_{\mathfrak{g}}(I)$ of an ideal $I \subseteq \mathfrak{g}$ is an ideal of a Lie algebra $\mathfrak{g}$.*

Proof. By Question 1 Part (*b*) Assignment 3, we know $C_{\mathfrak{g}}(I)$ is a subalgebra of $\mathfrak{g}$. Therefore, it suffices to show $[\mathfrak{g}, C_{\mathfrak{g}}(I)] \subseteq C_{\mathfrak{g}}(I)$. To this end, suppose $X \in \mathfrak{g}$, $Y \in C_{\mathfrak{g}}(I)$ and $Z \in I$. By Jacobi we have

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0,$$

and observe $[Y, Z] = 0$ (since $Y \in C_{\mathfrak{g}}(I)$ and $Z \in I$) and that $[Z, X] \in I$ (since $Z \in I$, $X \in \mathfrak{g}$ and I is an ideal of $\mathfrak{g}$), implying $[Y, [Z, X]] = 0$. So, $[Z, [X, Y]] = 0$, and hence $[X, Y] \in C_{\mathfrak{g}}(I)$. Thus, it clearly follows $C_{\mathfrak{g}}(I)$ is an ideal of $\mathfrak{g}$. □

Lemma 19.9. *Suppose $\mathfrak{h} \subseteq \mathfrak{g}$ is a subalgebra consisting of semi-simple elements and $C_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$. Then $\mathfrak{h}$ is a Cartan subalgebra in $\mathfrak{g}$.*

Proof. Note that $\mathfrak{h} \subseteq C_{\mathfrak{g}}(\mathfrak{h})$ iff $\mathfrak{h}$ is abelian. Hence, $\mathfrak{h}$ is abelian. We claim that if $\mathfrak{h} = C_{\mathfrak{g}}(\mathfrak{h})$ must be of maximal dimension. Otherwise, there exists a subalgebra $\widetilde{\mathfrak{h}}$ such that $\mathfrak{h} \subset \widetilde{\mathfrak{h}}$. Then $\widetilde{\mathfrak{h}} \subseteq C_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$, which gives a contradiction. □

Note. This states that a self-centralising subalgebra $\mathfrak{h} \subset \mathfrak{g}$ consisting of semi-simple elements is a Cartan subalgebra of $\mathfrak{g}$. If $\mathfrak{g}$ is a semi-simple Lie algebra, then the converse is true.

Example 19.10. We show that the abelian subalgebra $\mathfrak{h}$ consisting of traceless diagonal matrices in $\mathfrak{sl}_n(\mathbb{C})$ is a Cartan subalgebra by showing that $C_{\mathfrak{sl}_n(\mathbb{C})}(\mathfrak{h}) = \mathfrak{h}$. Clearly, $\mathfrak{h}$ is contained in $C_{\mathfrak{sl}_n(\mathbb{C})}$. It is easy to see that if a matrix commutes with $\mathfrak{h}$, then it must be a diagonal, and so we are done.

Theorem 19.11. *Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$ a Cartan subalgebra. Then $\mathfrak{h} = C_{\mathfrak{g}}(\mathfrak{h})$.*

Note. In the root space decomposition $\mathfrak{g} = \mathfrak{g}_0 \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$ we have

$$\mathfrak{g}_0 = \{X \in \mathfrak{g} \mid [X, h] = 0 \ \forall h \in \mathfrak{h}\} = C_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}.$$

Hence, $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$.

Proof. We proceed in three steps. Step 1: Choose $h \in \mathfrak{h}$ such that $\dim(C_{\mathfrak{g}}(h))$ is minimal. We will show that $C_{\mathfrak{g}}(h) = C_{\mathfrak{g}}(\mathfrak{h})$. Suppose $C_{\mathfrak{g}}(h) \neq C_{\mathfrak{g}}(\mathfrak{h})$, i.e., there exists $s \in \mathfrak{h}$ such that $C_{\mathfrak{g}}(h) \not\subseteq C_{\mathfrak{g}}(s)$, so

$$C_{\mathfrak{g}}(h) \cap C_{\mathfrak{g}}(s) \subsetneq C_{\mathfrak{g}}(h).$$

We show that there is a linear combination $h + \lambda s$, $\lambda \in \mathbb{C}$, such that

$$C_{\mathfrak{g}}(h + \lambda s) = C_{\mathfrak{g}}(h) \cap C_{\mathfrak{g}}(s) \subsetneq C_{\mathfrak{g}}(h),$$

contradicting the minimal dimension of $C_{\mathfrak{g}}(h)$. Choose a basis $c_1, \dots, c_n$ of $C_{\mathfrak{g}}(h) \cap C_{\mathfrak{g}}(s)$. Since $s \in \mathfrak{h} \subseteq C_{\mathfrak{g}}(h)$, and s is semi-simple, we can extend to a basis $c_1, \dots, c_n$, $x_1, \dots, x_p$ of $C_{\mathfrak{g}}(h)$ consisting of eigenvectors for ad_s. Also, we can extend to a basis $c_1, \dots, c_n$, $y_1, \dots, y_q$ of $C_{\mathfrak{g}}(s)$ of eigenvectors for ad_h. So, $c_1, \dots, c_n$, $x_1, \dots, x_p$, $y_1, \dots, y_q$ is a basis of $C_{\mathfrak{g}}(h) + C_{\mathfrak{g}}(s)$. Moreover,

$$[\mathrm{ad}_s, \mathrm{ad}_h] = \mathrm{ad}_{[s,h]} = 0,$$

so we can extend to a basis of $\mathfrak{g}$:

$$c_1, \dots, c_n, x_1, \dots, x_p, y_1, \dots, y_q, z_1, \dots, z_r,$$

consisting of simultaneous eigenvectors of ad_h and ad_s. Note that $[s, x_i] \neq 0$, since $x_i \notin C_{\mathfrak{g}}(s)$, and similarly $[h, y_i] \neq 0$. Let $[h, z_i] = \alpha_i z_i$, $[s, z_i] = \beta_i z_i$, $\alpha_i, \beta_i \neq 0$. We have the following table of eigenvalues ($\lambda \neq 0$):

	c_i	x_i	y_i	z_i
ad_s	0	$\neq 0$	0	β_i
ad_h	0	0	$\neq 0$	α_i
$\mathrm{ad}_s + \lambda \mathrm{ad}_h$	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, $C_{\mathfrak{g}}(s + \lambda h) = C_{\mathfrak{g}}(s) \cap C_{\mathfrak{g}}(h)$ and $\dim C_{\mathfrak{g}}(s + \lambda h) < \dim(C_{\mathfrak{g}}(h))$.

Step 2: Next, we prove that $C_{\mathfrak{g}}(h)$ is nilpotent. Let $X \in C_{\mathfrak{g}}(h)$, with Jordan decomposition $X = X_{SS} + X_n$. Then,

$$[X, h] = 0 \Longrightarrow [X_{SS}, h] = [X_n, h] = 0,$$

so $X_{SS}, X_n \in C_{\mathfrak{g}}(h)$. Also, $X_{SS} \in \mathfrak{h}$, since by step 1, $[X_{SS}, \mathfrak{h}] = 0$, so $\mathfrak{h} + \mathrm{span}\{X_{SS}\}$ is abelian and consists of semi-simple elements, so $X_{SS} \in \mathfrak{h}$ as $\mathfrak{h}$ is a Cartan subalgebra (maximality conidition).

Now, the restriction $\mathrm{ad}_{X_{SS}} : C_{\mathfrak{g}}(h) \to C_{\mathfrak{g}}(h)$ is the zero map as $X_{SS} \in \mathfrak{h}$, so the restriction $\mathrm{ad}_X : C_{\mathfrak{g}}(h) \to C_{\mathfrak{g}}(h)$ is the restriction of ad_{X_n} to $C_{\mathfrak{g}}(h)$, so it is nilpotent. Hence, $\mathrm{ad}(C_{\mathfrak{g}}(h))$ consists of nilpotent operators and by Engel's theorem $C_{\mathfrak{g}}(h)$ is a nilpotent Lie algebra.

Step 3: We now show that $C_{\mathfrak{g}}(h) \subseteq \mathfrak{h}$, so combined with step 1, we have $C_{\mathfrak{g}}(\mathfrak{h}) \subseteq \mathfrak{h}$ and thus $C_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$. By step 2, $C_{\mathfrak{g}}(h)$ is nilpotent, hence solvable, so by Lie's theorem, there exists a basis B of $C_{\mathfrak{g}}(h)$ such that $\{[\mathrm{ad}_X]_B \mid X \in C_{\mathfrak{g}}(h)\} \subseteq \mathfrak{b}_n(\mathbb{C})$ where $n = \dim(C_{\mathfrak{g}}(\mathfrak{h}))$. Let $X \in C_{\mathfrak{g}}(h)$ with Jordan decomposition $X = X_{SS} + X_n$, where by the proof of step 2, $X_{SS} \in \mathfrak{h}$, and $X_n \in C_{\mathfrak{g}}(h)$ with ad_{X_n} nilpotent. Then $[\mathrm{ad}_{X_n}]_B \in \mathfrak{n}_n(\mathbb{C})$.

Let K be the Killing form of $\mathfrak{g}$. For all $Y \in C_{\mathfrak{g}}(h)$,

$$K(X_n, Y) = \mathrm{Tr}(\mathrm{ad}_{X_n} \circ \mathrm{ad}_Y) = 0.$$

By Proposition 19.3 part (3), the restriction of K to $\mathfrak{g}_0 = C_{\mathfrak{g}}(\mathfrak{h})$ is non-degenerate. Hence, $X_n = 0$, so $X = X_{SS} \in \mathfrak{h}$. Thus, $C_{\mathfrak{g}}(h) = \mathfrak{h}$. □

Returning to our strategy, we have now established the existence of maximal abelian subalgebras $\mathfrak{h}$ consisting of semi-simple elements (i.e., Cartan subalgebras) for any semi-simple Lie algebra and the weight space decomposition w.r.t. $\mathrm{ad}\,|_{\mathfrak{h}}\colon \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g})$ is called the root space decomposition ($\mathfrak{g}_0 = C_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$)

$$\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$$

where $\Phi \subseteq \mathfrak{h}^* \setminus \{0\}$, the set of roots $\mathfrak{g}$ with respect to $\mathfrak{h}$, and

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \mid [h, X] = \alpha(h)X \ \forall h \in \mathfrak{h}\}.$$

So far we know the following structure constraints:

- $[h_1, h_2] = 0$ for all $h_1, h_2 \in \mathfrak{h}$
- $[h, x] = \alpha(h)x$ for all $h \in \mathfrak{h}$ and $x \in \mathfrak{g}_\alpha$
- $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$ for all $\alpha, \beta \in \Phi$

There is still a long way to go to pin down all possible structure constraints (i.e., Φ) for simple Lie algebras, but remarkably it is possible to do by studying the root space decomposition more closely. This is the task that we will undertake next.

20. $\mathfrak{sl}_2(\mathbb{C})$-subalgebras

We observed that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$, so in particular $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{h}$ so long as $-\alpha \in \Phi$. Not only is it always true that $\alpha \in \Phi \iff -\alpha \in \Phi$, but we will show that for any $\alpha \in \Phi$ there exists a triple $(e_\alpha, e_{-\alpha}, h_\alpha) \in \mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha} \times \mathfrak{h}$, forming an $\mathfrak{sl}_2(\mathbb{C})$-subalgebra of $\mathfrak{g}$.

Proposition 20.1. *Let $\alpha \in \Phi$. Then,*

(1) *$-\alpha \in \Phi$; and*
(2) *if $0 \neq e_\alpha \in \mathfrak{g}_\alpha$, then there exists $e_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $\mathrm{span}\{e_\alpha, e_{-\alpha}, [e_\alpha, e_{-\alpha}]\}$ is a subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.*

Proof. Let $0 \neq X \in \mathfrak{g}_\alpha$. As K is non-degenerate, there exists $Y \in \mathfrak{g}$ such that $K(X, Y) \neq 0$. Write $Y = Y_0 + \sum_{\beta \in \Phi} Y_\beta$, where $Y_0 \in \mathfrak{h}$ and $Y_\beta \in \mathfrak{g}_\beta$. Recall $K(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$, unless $\alpha + \beta = 0$. Therefore, $Y_{-\alpha} \neq 0$ and so $-\alpha \in \Phi$ (which proves (1)).

Let $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$ satisfying $K(X, Y) \neq 0$. We claim that $[X, Y] \in \mathfrak{h}$ and $[X, Y] \neq 0$. We have $[X, Y] \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{g}_0 = \mathfrak{h}$ and as $\alpha \neq 0$, there exists $u \in \mathfrak{h}$ such that $\alpha(u) \neq 0$. Then,

$$K(u, [X, Y]) = K([u, X], Y) = \alpha(u)K(X, Y) \neq 0.$$

Therefore, $[X, Y] \neq 0$. Now, let $S = \mathrm{span}\{X, Y, [X, Y]\}$. Since $[X, Y] \in \mathfrak{h}$, we have

$$[[X, Y], X] = \alpha([X, Y])X \in S$$
$$[[X, Y], Y] = -\alpha([X, Y])Y \in S.$$

Hence $S \subseteq \mathfrak{g}$ is a subalgebra. We show that $S \cong \mathfrak{sl}_2(\mathbb{C})$ by proving that $S' = S$, so the result follows by the following lemma below.

Let $h = [X, Y] \in \mathfrak{h} \setminus \{0\}$ and suppose for a contradiction that $\alpha(h) = 0$, so $[h, X] = [h, Y] = 0$. Then $\dim(S') = 1$, so S is solvable. By Lie's theorem, there exists a basis B of $\mathfrak{g}$ such that

$$\{[\mathrm{ad}_X]_B \mid X \in S\} \subseteq \mathfrak{b}_3(\mathbb{C}).$$

Then, $[\mathrm{ad}_h]_B = [\mathrm{ad}_{[X,Y]}]_B \in \mathfrak{n}_3(\mathbb{C})$, and so ad_h is nilpotent. As h is semi-simple, ad_h is diagonalisable. Hence, $h = 0$, a contradiction. Consequently, $\alpha(h) \neq 0$, so $[h, X] \neq 0$, $[h, Y] \neq 0$ and S' is 3-dimensional, so $S \cong \mathfrak{sl}_3(\mathbb{C})$. Rescaling X and Y, we get $S := \mathfrak{sl}(\alpha) = \mathrm{span}\{e_\alpha, e_{-\alpha}, h_\alpha\}$, where $e_\alpha \in \mathfrak{g}_\alpha$, $e_{-\alpha} \in \mathfrak{g}_{-\alpha}$, $h_\alpha := [e_\alpha, e_{-\alpha}] \in \mathfrak{h}$ and $[h_\alpha, e_\alpha] = 2e_\alpha$, $[h_\alpha, e_{-\alpha}] = -2e_{-\alpha}$, i.e., $\alpha(h_\alpha) = 2$. □

Lemma 20.2. *Let $\mathfrak{g}$ be a complex 3-dimensional Lie algebra with $\mathfrak{g} = \mathfrak{g}'$. Then, $\mathfrak{g} \cong \mathfrak{sl}_2(\mathbb{C})$.*

Proof. For $0 \neq X \in \mathfrak{g}$, $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ has rank 2. Namely, if $\mathfrak{g} = \mathrm{span}\{X, Y, Z\}$, then $\mathfrak{g}' = \mathrm{span}\{[X,Y],[X,Z],[Y,Z]\}$, so $[X,Y]$ and $[X,Z]$ are linearly independent. There exists $h \in \mathfrak{g}$ such that ad_h has a non-zero eigenvalue. Namely, let $0 \neq X \in \mathfrak{g}$ and assume ad_X has only eigenvalue 0. Then there is a basis $B = \{X, Y, Z\}$ of $\mathfrak{g}$, such that

$$[\mathrm{ad}_X]_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which is Jordan canonical form of ad_X, so $[X,Y] = X$, $[X,Z] = Y$. Then $\mathrm{ad}_Y(X) = -X$, so ad_Y has eigenvalue -1.

We have shown that there exist $h, X \in \mathfrak{g}$ such that $\mathrm{ad}_h(X) = \alpha X$ for $\alpha \neq 0$. Since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, the trace of any representation $\phi : \mathfrak{g} \to \mathfrak{gl}(V)$ vanishes:

$$\mathrm{Tr}(\phi(X)) = \mathrm{Tr}(\phi([A,B])) = \mathrm{Tr}([\phi(A), \phi(B)]) = 0,$$

i.e., $\phi : \mathfrak{g} \to \mathfrak{sl}(V) \subset \mathfrak{gl}(V)$, so in particular, $\mathrm{Tr}(\mathrm{ad}_h) = 0$ and the eigenvalues of ad_h are $0, \alpha, -\alpha$. So there exists a basis X, Y, h with $[h, X] = \alpha X$, $[h, Y] = -\alpha Y$. Moreover,

$$[h, [X,Y]] = [[h,X],Y] + [X,[h,Y]] = \alpha[X,Y] - \alpha[X,Y] = 0.$$

Thus, $[X,Y] \in \ker(\mathrm{ad}_X) = \mathrm{span}\{h\}$, so $[X,Y] = \lambda h$ for $\lambda \neq 0$ (as $\mathfrak{g} = \mathfrak{g}'$). Rescaling $h \mapsto \frac{2}{\alpha} h$ to take $\alpha = 2$ and rescaling $X \mapsto \frac{1}{\lambda} X$ to take $\lambda = 1$, gives

$$[h, X] = 2X, [h, Y] = -2y, [X, Y] = h.$$

□

Example 20.3. Let $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{C})$. Recall that

$$\mathfrak{g} = \mathfrak{h} \bigoplus_{i \neq j} \mathrm{span}\{e_{ij}\}$$

with roots $\varepsilon_i - \varepsilon_j$, where

$$\epsilon_i : \begin{pmatrix} a_1 & & 0 \\ & a_2 & \\ 0 & & a_3 \end{pmatrix} \mapsto a_i.$$

For $\alpha = \varepsilon_i - \varepsilon_j$, we have

$$\mathfrak{sl}(\alpha) = \mathrm{span}\{e_{ij}, e_{ji}, h_{ij}\},$$

where $h_{ij} = e_{ii} - e_{jj}$.

Recall that we obtained the root space decomposition $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$ by decomposing $\mathfrak{g}$ into weight spaces for $\mathrm{ad}\,|_{\mathfrak{h}} \colon \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g})$. Since any simple Lie algebra $\mathfrak{g}$ contains an $\mathfrak{sl}_2(\mathbb{C})$ subalgebra $\mathfrak{sl}(\alpha)$, for each $\alpha \in \Phi$, we can gain further information by studying the $\mathfrak{sl}_2(\mathbb{C})$-representation;

$$\mathrm{ad}\,|_{\mathfrak{sl}(\alpha)} \colon \mathfrak{sl}(\alpha) \to \mathfrak{gl}(\mathfrak{g})$$

and its submodules in $\mathfrak{g}$.

Proposition 20.4. *Let $\alpha \in \phi$ and $\beta \in \phi \cup \{0\}$. Then*

$$M = \bigoplus_{\substack{c \in \mathbb{Z} \\ \beta + c\alpha \in \Phi}} \mathfrak{g}_{\beta + c\alpha}$$

is an $\mathfrak{sl}(\alpha)$-submodule of $\mathfrak{g}$.

Proof. Observe

$$[\mathfrak{g}_{\beta+c\alpha}, \mathfrak{g}_{\pm\alpha}] \subseteq \mathfrak{g}_{\beta\pm(c\pm1)\alpha}$$

and

$$[\mathfrak{g}_{\beta+c\alpha}, \mathfrak{h}] \subseteq \mathfrak{g}_{\beta+c\alpha}.$$

□

Definition 20.5. We call the set of roots $\{\beta + c\alpha \mid c \in \mathbb{Z}\}$ the *α-string through β.*

Example 20.6. Let $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{C})$, $\alpha = \varepsilon_2 - \varepsilon_3$ and $\beta = \varepsilon_1 - \varepsilon_2$. The α-string through β is β, $\beta + \alpha$. The corresponding submodule is

$$M = \operatorname{span}\{e_{12}, e_{13}\} = \left\{ \begin{pmatrix} 0 & \bullet & \bullet \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\},$$

and

$$\mathfrak{sl}(\alpha) = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & \bullet & \bullet \\ 0 & \bullet & \bullet \end{pmatrix} \right\}.$$

Proposition 20.7. *Let $\alpha \in \Phi$. Then*

1. $\dim(\mathfrak{g}_\alpha) = 1$*; and*
2. *if $n\alpha \in \Phi$, $n \in \mathbb{Z} \setminus \{0\}$, then $n = \pm 1$.*

Proof. Define $W = \operatorname{span}\{e_{-\alpha}, \mathfrak{h}, \mathfrak{g}_{n\alpha} \mid n \in \mathbb{N}\}$. Then W is invariant under $\operatorname{ad}_{e_\alpha}$, $\operatorname{ad}_{e_{-\alpha}}$ and ad_h, $h \in \mathfrak{h}$. So, W is an $\mathfrak{sl}(\alpha)$-submodule. Consider the restriction $\operatorname{ad}_{e_\alpha}|_W$ to W. Then

$$\operatorname{Tr}\big(\operatorname{ad}_{[e_\alpha, e_{-\alpha}]}|_W\big) = \operatorname{Tr}(\operatorname{ad}_{h_\alpha}|_W) = \operatorname{Tr}\big([\operatorname{ad}_{e_\alpha} \mid W, \operatorname{ad}_{e_{-\alpha}} \mid W]\big) = 0.$$

Now,

$$\operatorname{ad}_{h_\alpha}|_W = \begin{pmatrix} -\alpha(h) & & & & & & \\ & 0 & & & & & \\ & & \ddots & & & & \\ & & & 0 & & & \\ & & & & n\alpha(h) & & \\ & & & & & \ddots & \\ & & & & & & n\alpha(h) \end{pmatrix},$$

where $-\alpha(h)$ comes from $e_{-\alpha}$, the zeroes come from $\mathfrak{h}$ and $n\alpha(h)$ comes from $\mathfrak{g}_{n\alpha}$. We write $h = h_\alpha$ so that $\alpha(h) = 2$.

Taking the trace, we obtain

$$0 = -\alpha(h) + \sum_{n\geq 1} n\alpha(h) \dim(\mathfrak{g}_{n\alpha}).$$

Thus, $\sum_{n\geq1} n \dim(\mathfrak{g}_{n\alpha}) = 1$, so $\dim(\mathfrak{g}_\alpha) = 1$ and $\mathfrak{g}_{n\alpha} = 0$ for all $n \geq 2$. Similarly, $\dim(\mathfrak{g}_{-\alpha}) = 1$ and $\mathfrak{g}_{-n\alpha} = 0$ for all $n \geq 2$. □

Proposition 20.8. *Let V be a finite dimensional $\mathfrak{sl}_2(\mathbb{C})$-module. Then every eigenvalue of the linear map $v \mapsto h \cdot v$, $v \in V$, is an integer.*

Note. $\mathfrak{sl}_2(\mathbb{C}) = \operatorname{span}\{e, f, h\}$.

Proof. Recall that the irreducible $\mathfrak{sl}_2(\mathbb{C})$-modules are V_d, $\dim(V_d) = d+1$, $d \in \mathbb{N}_0$, and $\operatorname{spec}\{h|_{V_d}\} = \{d, d-2, \ldots, -d\} \subset \mathbb{Z}$. Now, take a composition series for V:

$$V = W_0 \supset W_1 \supset \ldots \supset W_r = 0$$

where W_i are submodules and each quotient W_i/W_{i+1} is irreducible, so $W_i/W_{i+1} \cong V_{d_i}$ for some $d_i \in \mathbb{N}_0$. Then with respect to a suitable basis of V, h acts as

$$h = \begin{pmatrix} d_0 & & & & & & \\ & \ddots & & & & 0 & \\ & & -d_0 & & & & \\ & & & d_1 & & & \\ & & & & \ddots & & \\ & \bullet & & & & -d_1 & \\ & & & & & & \ddots \end{pmatrix},$$

hence the eigenvalues are integers. □

Using the above proposition, we can obtain further restrictions on the set of roots Φ.

Proposition 20.9. *Let $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$. Then*

(1) $\beta(h_\alpha) \in \mathbb{Z}$.
(2) *The α-string through β is*

$$\beta - r\alpha, \beta - (r-1)\alpha, \dots, \beta, \beta + \alpha, \dots, \beta + qx$$

where $r, q \geq 0$ and $\beta(h_\alpha) = r - q$.
(3) *If $\alpha + \beta \in \Phi$, then $[e_\alpha, e_\beta] = \lambda e_{\alpha+\beta}$, $\lambda \neq 0$.*
(4) $\beta - \beta(h_\alpha)\alpha \in \Phi$.

Proof. Let $M = \bigoplus_{\substack{c \in \mathbb{Z} \\ \beta + c\alpha \in \Phi}} \mathfrak{g}_{\beta + c\alpha} \subseteq \mathfrak{g}$ be the $\mathfrak{sl}_2(\alpha)$-submodule.

(1) is clear, since $\beta(h_\alpha)$ is an h_α-eigenvalue on $\mathfrak{g}_\beta$, so $\beta(h_\alpha) \in \mathbb{Z}$ by previous proposition.

For (2): Recall $\dim(\mathfrak{g}_{\beta+c\alpha}) = 0$ or 1 for all $c \in \mathbb{Z}$. So, the eigenspaces of ad_{h_α} on M are all one-dimensional. The eigenvalues are $(\beta + c\alpha)(h_\alpha) = \beta(h_\alpha) + 2c$ so they are 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).

Thus, $M \cong V_d$ for some $d \in \mathbb{N}_0$, and the eigenvalues are

$$\{\beta(h_\alpha) + 2c \mid \beta + c\alpha \in \Phi\} = \{d, d-2, \dots, -d\},$$

so (2) follows by taking $d = \beta(h_\alpha) + 2q$ and $-d = \beta(h_\alpha) - 2r$, so in particular, we obtain $\beta(h_\alpha) = r - q$ by taking the sum of these equations.

For (3): 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 $q \geq 1$, which gives a contradiction. Therefore, $[e_\alpha, e_\beta] \neq 0$, showing (3).

(4) follows from (2), since

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

and

$$-r \leq -r + q \leq q.$$

□

Recall that $[\mathfrak{h}, \mathfrak{h}] = 0$ and $[\mathfrak{h}, \mathfrak{g}]$ is determined by the roots Φ. The above proposition (3) shows that Φ also determines the brackets $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta]$ for $\alpha \neq \pm\beta$, up to scalar factor, and by construction $[e_\alpha, e_{-\alpha}] \in \mathfrak{h}$. To advance further, we will need to unravel more of the geometric properties of root system Φ.

21. Cartan subalgebras as inner product spaces

Proposition 21.1. *$K(h_1, h_2) = \sum_{\alpha\in\Phi} \alpha(h_1)\alpha(h_2)$ for all $h_1, h_2 \in \mathfrak{h}$.*

Proof. Let B be a basis of $\mathfrak{g}$; $B = \{\text{basis of } \mathfrak{g}\} \cup \{e_\alpha \mid \alpha \in \Phi\}$. For $h \in \mathfrak{h}$,

$$[\mathrm{ad}_h]_B = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & \alpha(h) & \\ & & & \ddots \end{pmatrix},$$

where 0 corresponds to $\mathfrak{h}$, and $\alpha(h)I$ corresponds to $\{e_\alpha \mid \alpha \in \Phi\}$. Hence,

$$K(h_1, h_2) = \mathrm{Tr}(\mathrm{ad}_{h_1}, \mathrm{ad}_{h_2}) = \sum_{\alpha\in\Phi} \alpha(h_1)\alpha(h_2).$$

□

Proposition 21.2. $\mathrm{span}(\Phi) = \mathfrak{h}^*$.

Proof. Suppose for a contradiction that $V := \mathrm{span}(\Phi) \subsetneq \mathfrak{h}^*$. Then

$$\mathrm{Ann}_{\mathfrak{h}}(V) = \{h \in \mathfrak{h} \mid f(h) = 0 \ \forall f \in V\}$$

is non-zero, as it has dimension

$$\dim(\mathfrak{h}^*) - \dim(V) \neq 0.$$

Hence, there exists $0 \neq h \in \mathfrak{h}$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$. Then $K(h, \mathfrak{h}) = 0$ by previous proposition, and $K(h, \mathfrak{g}_\alpha) = 0$ for all $\alpha \in \Phi$ by Proposition 19.3 (2). Hence, $h \in \mathfrak{h}^\perp = 0$, a contradiction. □

By Proposition 19.3 (3), the restriction $K\mid_{\mathfrak{h}}$ of the Killing form to $\mathfrak{h}$ is non-degenerate. Define

$$\theta : \mathfrak{h} \to \mathfrak{h}^*, h \mapsto \theta_h = K(h, -).$$

$K\mid_{\mathfrak{h}}$ being non-degenerate implies that θ is injective. Since $\dim(\mathfrak{h}) = \dim(\mathfrak{h}^*)$, $\theta : \mathfrak{h} \to \mathfrak{h}^*$ is an isomorphism.

Proposition 21.3. *For each $\alpha \in \Phi$, there exists a unique $t_\alpha \in \mathfrak{h}$ such that*

$$\alpha(X) = K(t_\alpha, X), \ \forall X \in \mathfrak{h}.$$

Proof. Follows by the isomorphism θ. □

Proposition 21.4. *For all $\alpha \in \Phi$, $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$, we have $[X, Y] = K(X, Y)t_\alpha$.*

Proof. For any $h \in \mathfrak{h}$,

$$\begin{aligned} K(h, [X, Y]) &= K([h, X], Y) = \alpha(h)K(X, Y) \\ &= K(t_\alpha, h)K(X, Y) \\ &= K(h, K(X, Y)t_\alpha). \end{aligned}$$

Since $K\mid_{\mathfrak{h}}$ is non-degenerate,

$$[X, Y] = K(X, Y)t_\alpha.$$

□

Proposition 21.5. (1) $t_\alpha = \frac{h_\alpha}{K(e_\alpha, e_{-\alpha})}$,

(2) $h_\alpha = \frac{2t_\alpha}{K(t_\alpha,t_\alpha)}$,
(3) $K(t_\alpha, t_\alpha)K(h_\alpha, h_\alpha) = 4$.

Proof. We have $h_\alpha = [e_\alpha, e_{-\alpha}]$, so by previous proposition,

$$h_\alpha = [e_\alpha, e_{-\alpha}] = K(e_\alpha, e_{-\alpha})t_\alpha,$$

which proves (1).

$\alpha(h_\alpha) = 2$, so using (1), we get

$$2 = K(t_\alpha, h_\alpha) = K(t_\alpha, t_\alpha)K(e_\alpha, e_{-\alpha}),$$

so

$$h_\alpha = t_\alpha K(e_\alpha, e_{-\alpha}) = \frac{2t_\alpha}{K(t_\alpha, t_\alpha)},$$

which proves (2).

(3) follows from (1), since

$$K(t_\alpha, t_\alpha)K(h_\alpha, h_\alpha) = K(t_\alpha, t_\alpha)\frac{K(2t_\alpha, 2t_\alpha)}{K(t_\alpha, t_\alpha)^2} = 4.$$

□

Proposition 21.6. *If $\alpha, \beta \in \Phi$, then*
(1) *$K(h_\alpha, h_\beta) \in \mathbb{Z}$; and*
(2) *$K(t_\alpha, t_\beta) \in \mathbb{Q}$.*

Proof. For (1), observe

$$K(h_\alpha, h_\beta) = \sum_{\gamma\in\Phi} \gamma(h_\alpha)\gamma(h_\beta) \in \mathbb{Z}.$$

For (2), by previous proposition and part (1),

$$K(t_\alpha, t_\beta) = K(h_\alpha, h_\beta)\frac{K(t_\alpha, t_\alpha)}{2}\frac{K(t_\beta, t_\beta)}{2} \in \mathbb{Q}.$$

□

The isomorphism $\theta : \mathfrak{h} \to \mathfrak{h}^*$, $h \mapsto K(h, -)$ naturally induces a non-degenerate inner product on $\mathfrak{h}^* = \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$. For every $f_1, f_2 \in \mathfrak{h}^*$:

$$\langle f_1, f_2\rangle := K\big(\theta^{-1}(f_1), \theta^{-1}(f_2)\big).$$

For $\alpha, \beta \in \Phi \subset \mathfrak{h}^*$,

$$\langle \alpha, \beta\rangle = K(t_\alpha, t_\beta).$$

Recall $\mathrm{span}(\Phi) = \mathfrak{h}^*$, so there exists a basis $\alpha_1, \dots, \alpha_k$ of $\mathfrak{h}^*$ with all $\alpha_i \in \Phi$.

Proposition 21.7. *If $\beta \in \Phi$, then $\beta = \sum_{i=1}^k r_i\alpha_i$, $r_i \in \mathbb{Q}$.*

Proof. Let $\beta = \sum_{i=1}^k r_i\alpha_i$, $r_i \in \mathbb{C}$. Then $\langle \beta, \alpha_j\rangle = \sum_{i=1}^k r_i\langle \alpha_i, \alpha_j\rangle$. As $K\mid_{\mathfrak{h}}$ is non-degenerate, the matrix $A_{ij} = \langle \alpha_i, \alpha_j\rangle$ is invertible and all entries of A are in $\mathbb{Q}$ by previous proposition, and also $\langle \beta, \alpha_j\rangle \in \mathbb{Q}$, so $r_i \in \mathbb{Q}$ for all i. □

Definition 21.8. Let E be the real span of $\alpha_1, \dots, \alpha_k \in \mathfrak{h}^*$, i.e., $E = \left\{\sum_{i=1}^k r_i\alpha_i \;\middle|\; r_i \in \mathbb{R}\right\}$. By previous proposition,
(1) E is independent of the choice of basis,
(2) $\Phi \subseteq E$,
(3) $E = \mathrm{span}_{\mathbb{R}}\{\Phi\}$.

Proposition 21.9. *The bilinear form $\langle\cdot,\cdot\rangle$ is a real-valued inner product on E.*

Proof. For $\alpha, \beta \in \Phi$, $\langle \alpha, \beta \rangle = K(t_\alpha, t_\beta) \in \mathbb{R}$. Hence, $\langle \cdot, \cdot \rangle$ is real-valued on E. Let $f \in E$, $f = \theta_h = K(h, -)$ for some $h \in \mathfrak{h}$. Then

$$\begin{aligned}\langle f, f \rangle = \langle \theta_h, \theta_h \rangle = K(h,h) &= \sum_{\gamma \in \Phi} \gamma(h)^2 \\ &= \sum_{\gamma \in \Phi} K(t_\gamma, h)^2 \\ &= \sum_{\gamma \in \Phi} \langle \gamma, \theta_h \rangle^2.\end{aligned}$$

As $\langle \gamma, \theta_h \rangle \in \mathbb{R}$, this sows that $\langle \theta_h, \theta_h \rangle \geq 0$. If $\langle \theta_h, \theta_h \rangle = 0$, then $\gamma(h) = 0$ for all $\gamma \in \Phi$. Hence, $h = 0$, so $\theta_h = 0$. □

22. Root systems

Let E be a finite dimensional real vector space with an inner product $(\cdot, \cdot)$. For $0 \neq v \in E$, the *reflection* $S_v : E \to E$ is defined by

$$S_v(x) = x - 2\frac{(v,x)}{(v,v)}v,$$

for all $x \in E$. Note that S_v sends $v \mapsto -v$ and fixes every vector in $\text{span}_{\mathbb{R}}(v)^\perp$.

Proposition 22.1. *S_v preserves the inner product, i.e.,*

$$(S_v(x), S_v(y)) = (x, y)$$

for all $x, y \in E$. So, $S_v \in O(E, (\cdot, \cdot))$, the orthogonal group of E.

Note. In terms of notation, write $\langle x, v \rangle := \frac{2(x,v)}{(v,v)}$; all previous occurrences of $\langle \cdot, \cdot \rangle$ is to be thought of as $(\cdot, \cdot)$ from now on. $\langle \cdot, \cdot \rangle$ is linear in x, but not in v.

Proof. Observe

$$\begin{aligned}(S_v(x), S_v(y)) &= (x - \langle x, v \rangle v, y - \langle y, v \rangle v) \\ &= (x, y - \langle y, v \rangle v) - \langle x, v \rangle (v, y - \langle y, v \rangle v) \\ &= (x,y) - \langle y, v \rangle (x, v) - \langle x, v \rangle (v, y) + \langle y, v \rangle \langle x, v \rangle (v, v) \\ &= (x,y) - 2\frac{(y,v)(x,v)}{(v,v)} - 2\frac{(x,v)(y,v)}{(v,v)} + 4\frac{(x,v)(y,v)}{(v,v)} \\ &= (x,y).\end{aligned}$$

□

Definition 22.2. A subset $R \subset E$ is a *root system* if

(1) R is finite, $0 \notin R$, and $E = \text{span}_{\mathbb{R}}(R)$.
(2) For $\alpha \in \mathbb{R}$, 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 $\alpha, \beta \in R$, $\langle \beta, \alpha \rangle = \frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.

Elements of R are called *roots*, and $\dim(E)$ is called the *rank* of R.

Example 22.3. The only root system of rank 1 is

$$-\alpha \longleftarrow\!\bullet\!\longrightarrow \alpha$$

Proposition 22.4. *Let $\mathfrak{g}$ be a semi-simple Lie algebra over $\mathbb{C}$, $\mathfrak{h} \subset \mathfrak{g}$ a Cartan subalgebra, and root decomposition:*

$$\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha.$$

Let $E = \text{span}_{\mathbb{R}}(\Phi) \subseteq \mathfrak{h}^*$ with inner product $(\cdot, \cdot)$. Then $\Phi \subset E$ is a root system.

Proof. Axiom (1) is clear, axiom (2) we proved in previous proposition. For axiom (3), let $\alpha, \beta \in \Phi$. Then $S_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha$. Claim: $\langle \beta, \alpha \rangle = \beta(h_\alpha)$. Indeed,

$$\begin{aligned} \beta(h_\alpha) &= K(t_\beta, h_\alpha) \\ &= K\left(t_\beta, \frac{2t_\alpha}{K(t_\alpha, t_\alpha)}\right) \\ &= \frac{2(\beta, \alpha)}{(\alpha, \alpha)} \\ &= \langle \beta, \alpha \rangle. \end{aligned}$$

Hence, $S_\alpha(\beta) = \beta - \beta(h_\alpha)\alpha \in \Phi$ by a previous proposition part (4). Finally, axiom (4) follows from $\langle \beta, \alpha \rangle = \beta(h_\alpha) \in \mathbb{Z}$ by that same previous proposition part (1). □

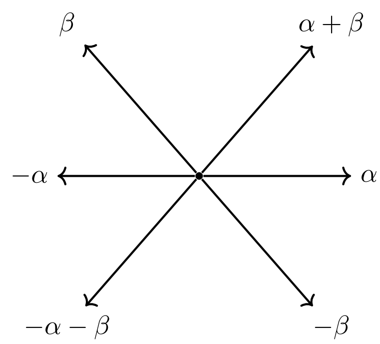
Example 22.5. $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$. A Cartan subalgebra is $\mathfrak{h} = \{\text{diagonal traceless matrices}\}$. The corresponding root spaces are $\mathfrak{g}_{\alpha_{ij}} = textspan_{\mathbb{C}}\{e_{ij}\}$ with root $\alpha_{ij} \in \Phi$, $\alpha_{ij} = \varepsilon_i - \varepsilon_j$, $i \neq j$,

$$\varepsilon_i : \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \mapsto a_i.$$

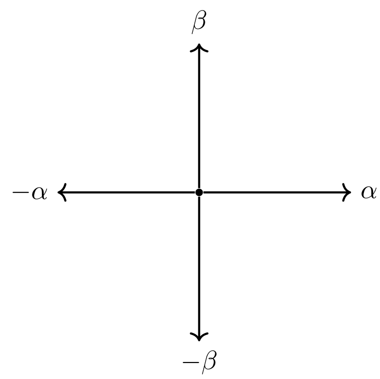
$E = \text{span}_{\mathbb{R}}(\Phi) = \{\sum_{i=1}^n \lambda_i \varepsilon_i \mid \sum_{i=1}^n \lambda_i = 0\}$, and $\text{rank}(\mathfrak{g}) := \dim(E) = n - 1$. Inner product:

$$\left(\sum_{i=1}^n \lambda_i \varepsilon_i, \sum_{j=1}^n \mu_j \varepsilon_j\right) = \sum_{i=1}^n \lambda_i \mu_i.$$

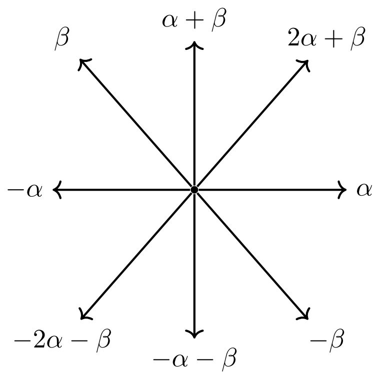
Example 22.6. Some rank 2 root systems: $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{C})$, $\alpha = \epsilon_1 - \epsilon_2$, $\beta = \epsilon_2 - \epsilon_3$ and $\Phi = \{\alpha, \beta, \alpha + \beta, -\alpha, -\beta, -\alpha - \beta\}$. We have $(\alpha, \beta) = -1$, $(\alpha, \alpha) = 2$, so the angle between α and β is $\cos^{-1}\left(\frac{1}{2}\right) = \frac{2\pi}{3} = 120°$. We have root system (a):



We also have root system (b):



Finally, we have root system (c):



Definition 22.7. We say that root systems $R \subseteq E$, $R' \subseteq E'$ are *isomorphic* if there exists a vector space isomorphism $\varphi : E \to E'$ such that

(1) $\varphi(R) = R'$;
(2) $(\varphi(\alpha), \varphi(\beta)) = (\alpha, \beta)$ for all $\alpha, \beta \in R$.

Definition 22.8. A root system $R \subseteq E$ is *reducible* if $R = R_1 \cup R_2$, where $R_i \neq \varnothing$ and $(\alpha, \beta) = 0$ for any $\alpha \in R_1$, $\beta \in R_2$. Otherwise, R is *irreducible.*

Note. In the previous examples, (a) and (c) are irreducible, but (b) is reducible since it is the union of the unique root of rank 1,

$$-\alpha \longleftarrow\!\!\bullet\!\!\longrightarrow \alpha$$

with itself, corresponding to the semi-simple Lie algebra $\mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C})$.

Proposition 22.9. *Let $\mathfrak{g}$ be a semi-simple Lie algebra with root system Φ, with respect to a choice of Cartan subalgebra $\mathfrak{h}$. If Φ is irreducible, then $\mathfrak{g}$ is simple.*

Proof. We have $\mathfrak{g} = \mathfrak{h} \oplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$. Suppose $\mathfrak{g}$ is not simple, so it has an ideal $I \neq 0$. Then $[\mathfrak{h}, I] \subseteq I$ and $\mathrm{ad}_\mathfrak{h}$ is simultaneously diagonalisable on $\mathfrak{g}$, hence also on I. Thus, I has a basis of common eigenvectors for $\mathrm{ad}_\mathfrak{h}$, so $I = \mathfrak{h}_1 \oplus_{\alpha\in\Phi_1} \mathfrak{g}_\alpha$, where $\mathfrak{h}_1 \subseteq \mathfrak{h}$, $\Phi_1 \subseteq \Phi$. Similarly, $I^\perp = \mathfrak{h}_2 \oplus_{\alpha\in\Phi_2} \mathfrak{g}_\alpha$ and as $\mathfrak{g} = I \oplus I^\perp$, $\mathfrak{h} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$, $\Phi = \Phi_1 \cup \Phi_2$, $\Phi_1 \cap \Phi_2 = \varnothing$. If $\Phi_2 = \varnothing$, then $\Phi_1 = \Phi$ so I contains all $\mathfrak{g}_\alpha$, $\alpha \in \Phi$, and therefore all $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$, which span $\mathfrak{h}$. Hence, $I = \mathfrak{g}$, a contradiction. Thus, $\Phi_i \neq \varnothing$ 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 previous proposition. Hence, $(\alpha, \beta) = 0$ showing that Φ is reducible. □

23. Classification of semi-simple Lie algebras

The classification has many contributors (Killing, Cartan, Weyl, ...). Recall that the root system Φ depends on the choice of Cartan subalgebra $\mathfrak{h}$. However, we have the following theorem.

Theorem 23.1. *Let $\mathfrak{g}$ be a semi-simple Lie algebra over $\mathbb{C}$, with Cartan subalgebras $\mathfrak{h}_1$, $\mathfrak{h}_2$, and corresponding root systems Φ_1 and Φ_2. Then, $\Phi_1 \cong \Phi_2$.*

The proof (omitted in this course) relies on the fact that all Cartan subalgebras are conjugate, i.e., there exists

$$g \in \mathrm{Aut}(\mathfrak{g}) = \{X \in \mathrm{GL}(\mathfrak{g}) \mid g([X, Y]) = [g(X), g(Y)], \forall X, Y \in \mathfrak{g}\}$$

such that $g(\mathfrak{h}_1) = \mathfrak{h}_2$. Hence, every semi-simple Lie algebra over $\mathbb{C}$ has a unique root system. The converse also holds.

Theorem 23.2. *For any root system Φ, there exists a unique (up to isomorphism) semi-simple Lie algebra $\mathfrak{g}$ over $\mathbb{C}$ with root system Φ.*

A detailed proof is omitted, but existence follows by constructing a simple Lie algebra for any given irreducible root system (these are classified). These are

$$\mathfrak{sl}_n, \mathfrak{so}_n, \mathfrak{sp}_n, \mathfrak{g}_2, \mathfrak{f}_4, \mathfrak{e}_6, \mathfrak{e}_7, \mathfrak{e}_8.$$

Alternative approach to existence is to use Serre's theorem (which constructs the Lie algebras using generators and relations).

Uniqueness can be proved abstractly, or by proving the existence of a Chevalley basis over $\mathbb{Z}$ with respect to which the structure constants are uniquely determined up to sign by the root data. Recall that for $\alpha, \beta \in \Phi$, the α-string through β is

$$\beta - r\alpha, \ldots, \beta + q\alpha.$$

Then the Chevalley basis is $\{h_\alpha, e_\alpha\}_{\alpha\in\Phi}$ and the structure constants are

$$\begin{aligned} [h_\alpha, h_\beta] &= 0 \\ [h_\alpha, e_\beta] &= \beta(h_\alpha)e_\beta \text{ and } \beta(h_\alpha) = r - q \\ [e_\alpha, e_{-\alpha}] &= h_\alpha \\ [e_\alpha, e_\beta] &= \begin{cases} 0 & \text{if } \alpha + \beta \notin \Phi \cup \{0\}; \\ \pm(q+1)e_{\alpha+\beta} & \text{if } \alpha + \beta \in \Phi. \end{cases} \end{aligned}$$

Conclusion: Classification of simple Lie algebras is equivalent to classification of irreducible root systems. In the remainder of these notes, we will look closer at irreducible root systems and prove that they are up to isomorphism classified by connected *Dynkin diagrams*. We will conclude listing all such diagrams.

24. Irreducible root systems

Let $R \subseteq E$ be a root system. Recall that $\langle \beta, \alpha \rangle = 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}$, where $\alpha, \beta \in R$. Also,

$$(\alpha, \beta) = \|\alpha\|\|\beta\| \cos(\theta),$$

where θ is the angle between α and β.

Proposition 24.1. *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, then $\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = 4\cos^2(\theta) \leq 4$. Moreover, it is not equal to 4, otherwise $\cos^2(\theta) = 1$ and $\theta = n\pi$ so $\beta = \pm\alpha$. □

Proposition 24.2. *Let $\alpha, \beta \in R$, $\beta \neq \pm\alpha$, and assume that $(\beta, \beta) \geq (\alpha, \alpha)$. Then the possible values for $\langle \alpha, \beta \rangle$, $\langle \beta, \alpha \rangle$ and 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	$\pi/2$	undetermined
1	1	$1/2$	$\pi/3$	1
-1	-1	$-1/2$	$2\pi/3$	1
1	2	$1/\sqrt{2}$	$\pi/4$	2
-1	-2	$-1/\sqrt{2}$	$3\pi/4$	2
1	3	$\sqrt{3}/2$	$\pi/6$	3
-1	-3	$-\sqrt{3}/2$	$5\pi/6$	3

Proposition 24.3. *Let θ be the angle between $\alpha, \beta \in R$, and $(\beta, \beta) \geq (\alpha, \alpha)$.*

(1) *If* $\theta > \frac{\pi}{2}$*, then* $\alpha + \beta \in R$.
(2) *If* $\theta < \frac{\pi}{2}$*, then* $\alpha - \beta \in R$.

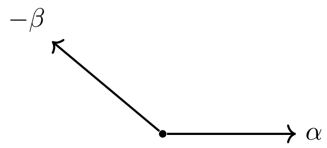
Proof. By Axiom (3), for root systems, $S_\beta(\alpha) = \alpha - \langle \alpha, \beta \rangle \beta \in R$ and from previous proposition, $\theta > \frac{\pi}{2}$ implies that $\langle \alpha, \beta \rangle = -1$ and $\theta < \frac{\pi}{2}$ implies that $\langle \alpha, \beta \rangle = 1$. □

Using Proposition 24.2, we can classify all root systems of rank 2 "by hand".

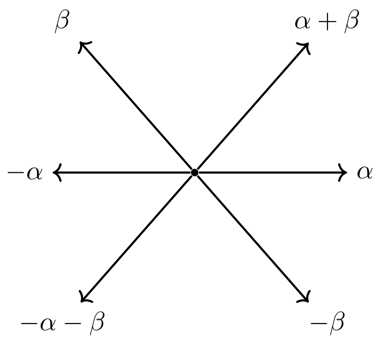
24.1. **Classification of root systems of rank 2.** Let $R \subseteq \mathbb{R}^2$ be a rank 2 root system. Pick $\alpha, \beta \in R$ with $\alpha \neq \pm\beta$, and the angle θ between α and β as large as possible ($\theta \geq \frac{\pi}{2}$). By Proposition 24.2, the possibilities are

$$\theta = \frac{2\pi}{3}, \frac{3\pi}{4}, \frac{5\pi}{6} \text{ or } \frac{\pi}{2}.$$

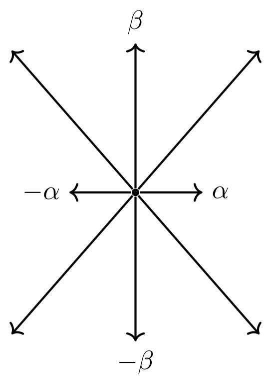
If $\theta = \frac{2\pi}{3}$, then α, β have the same length, so we have



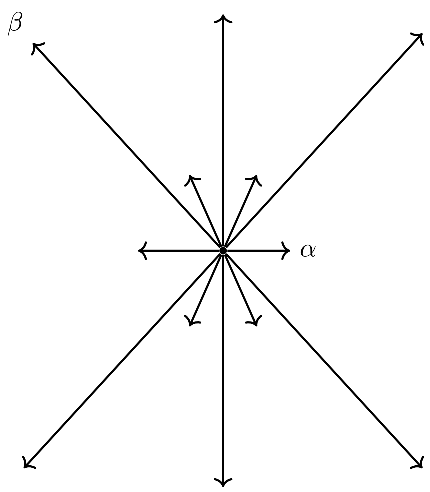
Applying reflections, we get A_2:



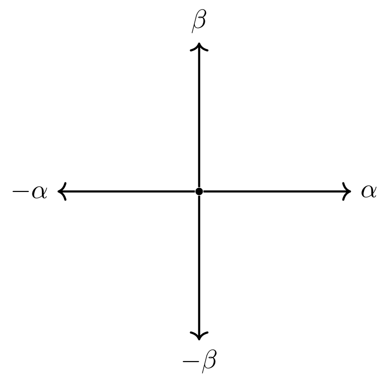
If $\theta = \frac{3\pi}{4}$, then $(\beta, \beta) = 2(\alpha, \alpha)$, so we get B_2:



If $\theta = \frac{5\pi}{6}$, then $(\beta, \beta) = 3(\alpha, \alpha)$, so we get G_2:



If $\theta = \frac{\pi}{2}$ we get $A_1 + A_1$ (which is reducible):



Definition 24.4. The *Weyl group* $W(R)$ of a root system $R \subseteq E$ is the group generated by the reflections

$$S_\alpha(x) = x - \frac{2(\alpha, x)}{(\alpha, \alpha)}\alpha$$

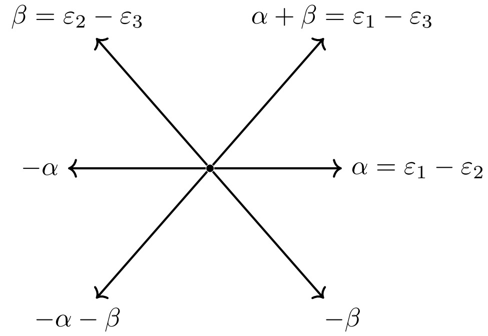
for all $\alpha \in E$, i.e.,

$$W(R) = \langle S_\alpha \mid \alpha \in R \rangle \subset \mathrm{GL}(E).$$

Proposition 24.5. *$W(R)$ is a finite group.*

Proof. By axiom (3) of root systems, each reflection S_α gives a permutation of the finite set R, so we have a homomorphism $\varphi : W(R) \to S_{|R|}$ where S_n is the permutation group on n elements (as usual). If $g \in \ker(\varphi)$, then $g(\alpha) = \alpha$ for all $\alpha \in R$, hence as $E = \mathrm{span}_{\mathbb{R}}(R)$, $g = 1$. Therefore, $\ker(\varphi) = \{1\}$, so φ is injective and $W(R) \cong \mathrm{im}(\varphi) \subseteq S_{|R|}$ and $\left|S_{|R|}\right| = |R|! < \infty$. □

Example 24.6. For $R = A_2$, the we have the roots:



$W(R) = \langle S_\alpha, S_\beta, S_{\alpha+\beta} \rangle$. The action on the standard basis $\epsilon_1, \epsilon_2, \epsilon_3$ of $\mathbb{R}^3$ is:

$$S_\alpha : \varepsilon_1 \leftrightarrow \varepsilon_2, \quad \varepsilon_3 \mapsto \varepsilon_3$$
$$S_\beta : \varepsilon_2 \leftrightarrow \varepsilon_3, \quad \varepsilon_1 \mapsto \varepsilon_1$$
$$S_{\alpha+\beta} : \varepsilon_1 \leftrightarrow \varepsilon_3, \quad \varepsilon_2 \mapsto \varepsilon_2$$

Hence, $W(A_2) \cong S_3$.

For $R = A_{n-1}$, the roots are $\epsilon_i - \epsilon_j$ where $i \neq j \in [n]$. Reflection $S_{\epsilon_i - \epsilon_j}$ sends $\epsilon_i \leftrightarrow \epsilon_j$ and fixes the other basis vectors. Hence, $W(A_{n-1}) \cong S_n$.

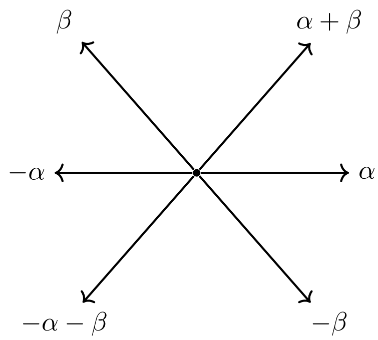
Definition 24.7. A subset $B \subseteq R$ is a *base* of the root system R if

(1) $E = \mathrm{span}_{\mathbb{R}}(B)$
(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 in B.

We say that $\beta \in R$ is a *positive root* with respect to base B if all $n_\alpha \geq 0$, and similarly define *negative roots*. Let R_+ and R_- denote the subsets of positive and negative roots, respectively. Then $R = R_+ \cup R_-$, a disjoint union. The set $B \subset R_+$ are called *simple roots* and S_α, $\alpha \in B$, are called *simple reflections*.

Note. The terms "positive" and "negative" roots are always taken with respect to a choice of a fixed basis B.

Example 24.8. For $R = A_2$:



A base is $B = \{\alpha, \beta\}$, so $R_+ = \{\alpha, \beta, \alpha + \beta\}$. Another base is $\widetilde{B} = \{\alpha, -\alpha - \beta\}$ and $\widetilde{R_+} = \{\alpha, -\alpha - \beta, -\beta\}$.

We find a Weyl group element $w \in W(A_2) \cong S_3$, such that $w(B) = \widetilde{B}$. Recall $\alpha = \varepsilon_1 - \varepsilon_2$, $\beta = \varepsilon_2 - \varepsilon_3$ and $-\alpha - \beta = \varepsilon_3 - \varepsilon_1$. Hence, $\varepsilon_2 \mapsto \varepsilon_1$, $\varepsilon_3 \mapsto \varepsilon_2$, $\varepsilon_1 \mapsto \varepsilon_3$ yields the desired result, which corresponds to $(1, 3, 2) = (1, 3)(1, 2)$ in S_3, and $w = S_{\alpha+\beta} S_\alpha$ in $W(R)$.

Theorem 24.9. (1) *Every root system has a base.*
(2) *If $B, \widetilde{B}$ are bases of R, then there exists a unique $w \in W(R)$ such that $w(B) = \widetilde{B}$.*

Proof. Omitted, see Erdmann-Wildon. □

Note. It follows that R has precisely $|W(R)|$ different bases.

Example 24.10. A_2 has 6 distinct bases, all of the form $w(\{\alpha, \beta\})$ for $w \in W(R) \cong S_3$.

Let $B \subset \mathbb{R}$ be a basis and fix an ordinary of the elements:

$$B = \{\alpha_1, \alpha_2, \dots, \alpha_\ell\}, \ell = \operatorname{rank}(R).$$

The *Cartan matrix* C of R is the matrix with integral entries $C_{ij} := \langle \alpha_i, \alpha_j \rangle \in \mathbb{Z}$, $\alpha_i \in B$. Since $\langle S_\beta(\alpha_i), S_\beta(\alpha_j) \rangle = \langle \alpha_i, \alpha_j \rangle$ for any root $\beta \in R$, $W(R)$ is generated by the reflections S_β, $\beta \in R$, and any two bases of R are related by a unique $w \in W(R)$, it follows that the $\ell \times \ell$-matrix C is uniquely specified by the root system R, up to reordering of the basis elements.

We will show that the Cartan matrix determines a root system R up to isomorphism. Another way to encode the information in C is in terms of a graph, called *Dynkin diagram.* The Dynkin diagram $\Delta(R)$ of a root system R with a base B is constructed as follows:

vertices : elements of B
edges : join $\alpha, \beta \in B$ 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 that there are at most two different root lengths. Given a Dynkin diagram, one can read off the numbers $\langle \alpha_i, \alpha_j \rangle$, and obtain the Cartan matrix, and vice versa.

Example 24.11. A_{n-1} with roots $\epsilon_i - \epsilon_j$, $i \neq j \in [n]$. Base:

$$B = \left\{ \overbrace{\epsilon_j - \epsilon_{j+1}}^{\alpha_j} \;\middle|\; j \in [n-1] \right\}.$$

This is a base as for $i < j$:

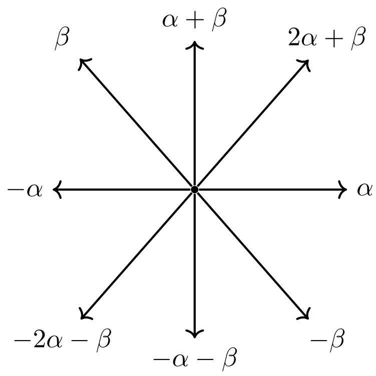
$$\epsilon_i - \epsilon_j = \alpha_i + \alpha_{i+1} + \ldots + \alpha_{j-1}.$$

Thus,

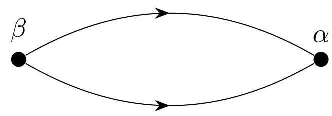
$$R_+ = \{\epsilon_i - \epsilon_j \mid i < j\}$$
$$R_- = \{\epsilon_i - \epsilon_j \mid i > j\}$$

$d_{\alpha_i \alpha_{i+1}} = \langle \alpha_i, \alpha_{i+1} \rangle = 1$ and $d_{\alpha_i \alpha_j} = 0 = 0$ if $j \neq i \pm 1$. The Dynkin diagram is:

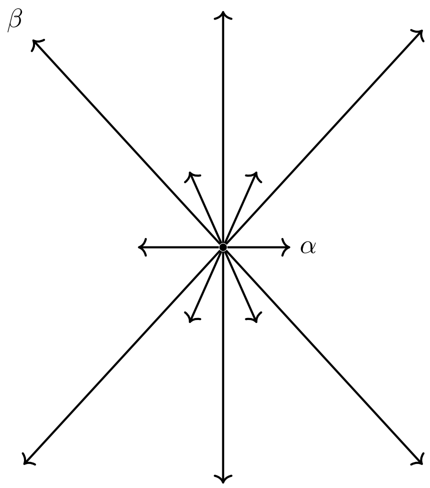
Example 24.12. For B_2 recall the roots are:



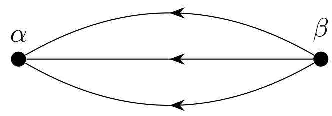
We have base $B = \{\alpha, \beta\}$ where we have α shorter than β, and $\langle \beta, \alpha \rangle = 2$. So, the Dynkin diagram is



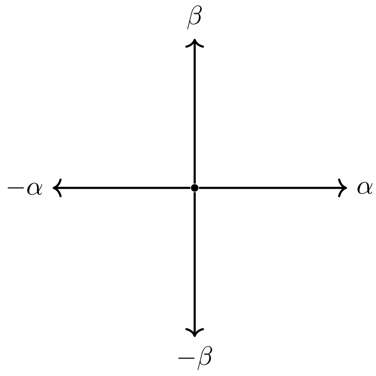
Example 24.13. For G_2 recall the roots are:



We have base $B = \{\alpha, \beta\}$ where we have β longer than α, and $\langle \beta, \alpha \rangle = 3$. So the Dynkin diagram is:



Example 24.14. For A_2, recall the roots are:



We have base $B = \{\alpha, \beta\}$ with same length, and $\langle \beta, \alpha \rangle = 0$, so no edges drawn yielding two isolated vertices as the Dynkin diagram:

Proposition 24.15. *A root system is irreducible iff its Dynkin diagram is connected.*

Proof. ($\Longrightarrow$) Suppose $R \subseteq E$ is an irreducible root system with base B. Then there does not exist non-empty R_1, R_2 with $(\alpha, \beta) = 0$ for each $\alpha \in R_1$ and $\beta \in R_2$ such that $R = R_1 \cup R_1$. Suppose $\Delta(R)$ is disconnected (with components C_1 and C_2). Then, there exists $\alpha \in V(C_1)$ such that for each $\beta \in V(C_2)$, $\langle \beta, \alpha \rangle = 0$, which implies $(\beta, \alpha) = 0$ and similarly $(\alpha, \beta) = 0$. One takes $R_1 = \operatorname{span}(V(C_1))$ and $R_2 = \operatorname{span}(V(C_2))$, which clearly is a reducible pair of R, contradiction.

($\Longleftarrow$) Suppose $\Delta(R)$ is connected. To derive a contradiction, suppose $R = R_1 \cup R_2$ with R_1, R_2 non-empty and $(\alpha, \beta) = 0$ for each $\alpha \in R_1$ and $\beta \in R_2$. One need only take $B \cap R_1$ and $B \cap R_2$ as vertices to see that $\Delta(R)$ is disconnected, a contradiction (note that say $B \cap R_1$ could be empty, which would imply $B \subseteq R_2$, so we need only preform a change of basis; necessarily, the change of basis preserves the Dynkin diagram structure, i.e., will be disconnected under new basis). $\square$

Proposition 24.16. *Let $B \subset R$ be a base for a root system and define*

$$W_0 = \langle S_\alpha \mid \alpha \in B \rangle \subset W(R).$$

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 $\beta \in R_+$, so

$$\beta = \sum_{\gamma \in B} k_\gamma \gamma, \quad k_\gamma \in \mathbb{N}_0.$$

Define $ht(\beta) = \sum_{\gamma \in B} k_\gamma$. We proceed by induction on $ht(\beta)$. If $ht(\beta) = 1$, then $\beta \in B$ and take $\alpha = \beta$, $w = 1$. Now assume $ht(\beta) \geq 2$. By axiom (2) of a root system, at least two k_γ's are non-zero.

We 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 $\beta \in R$ (axiom (1)).

We claim that $S_{\gamma_0}(\beta) \in R_+$. Indeed, $S_{\gamma_0}(\beta) = \beta - \langle \beta, \gamma_0 \rangle \gamma_0$, so $S_{\gamma_0}(\beta)$ has at least one coefficient $k_\gamma > 0$, hence $S_{\gamma_0}(\beta) \in R_+$.

By the previous two claims,

$$ht(S_{\gamma_0}(\beta)) = ht(\beta) - \langle \beta, \gamma_0 \rangle < ht(\beta).$$

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. □

Now we can prove that the Cartan matrix, or equivalently, the Dynkin diagram determines the root system up to isomorphism.

Theorem 24.17. *Let $R \subseteq E$ and $\widetilde{R} \subseteq \widetilde{E}$ be root systems with the same Dynkin diagram. Then $R \cong \widetilde{R}$.*

Proof. We have $R \subseteq E$, $\widetilde{R} \subseteq \widetilde{E}$ with bases $B = \{\alpha_1, \dots, \alpha_n\}$ and $\widetilde{B} = \{\widetilde{\alpha}_1, \dots, \widetilde{\alpha}_n\}$, respectively, such that $\langle \alpha_i, \alpha_j \rangle = \langle \widetilde{\alpha}_i, \widetilde{\alpha}_j \rangle$ for all i, j. Define a linear map $\varphi : E \to \widetilde{E}$ by $\varphi(\alpha_i) = \widetilde{\alpha}_i$ for all $1 \le i \le n$. We need to show that $\varphi(R) = \widetilde{R}$, since this implies that $R \cong \widetilde{R}$. We use the previous proposition to obtain

$$\{w_0(\alpha) \mid \alpha \in B, w_0 \in W_0\} = R.$$

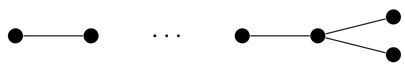
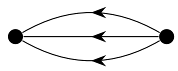
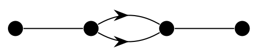
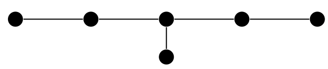
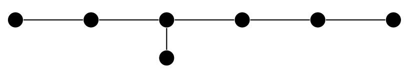
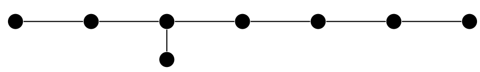
Now,

$$\begin{aligned} \varphi(S_{\alpha_i}(\alpha_j)) &= \varphi(\alpha_j - \langle \alpha_j, \alpha_i \rangle \alpha_i) \\ &= \widetilde{\alpha}_j - \langle \widetilde{\alpha}_j, \widetilde{\alpha}_i \rangle \widetilde{\alpha}_i \\ &= S_{\widetilde{\alpha}_i}(\widetilde{\alpha}_j) \in \widetilde{R}, \end{aligned}$$

by axiom (3). Hence, for $w_0 \in W_0$, $\varphi(w_0(\alpha)) \in \widetilde{R}$, so $\varphi(R) \subseteq \widetilde{R}$. The same argument for φ^{-1} gives $\varphi^{-1}\left(\widetilde{R}\right) \subseteq R$. Hence, $\varphi(R) = \widetilde{R}$. □

It follows that semi-simple Lie algebras are classified by their Dynkin diagrams and to classify simple Lie algebras, we need to classify all connected Dynkin diagrams.

Theorem 24.18. *Classification of connected Dynkin diagrams (equivalently, irreducible root systems, and also simple complex Lie algebras):*

Dynkin diagram	Notation	Simple Lie algebra
	$A_n, n \ge 1$	$\mathfrak{sl}_{n+1}(\mathbb{C})$
	$B_n, n \ge 2$	$\mathfrak{so}_{2n+1}(\mathbb{C})$
	$C_n, n \ge 2$	$\mathfrak{sp}_{2n}(\mathbb{C})$
	$D_n, n \ge 3$	$\mathfrak{so}_{2n}(\mathbb{C})$
	G_2	$\mathfrak{g}_2$
	F_4	$\mathfrak{f}_4$
	E_6	$\mathfrak{e}_6$
	E_7	$\mathfrak{e}_7$
	E_8	$\mathfrak{e}_8$

Note. The exceptional Lie algebras are not easy to construct. The classical Lie algebras, however, have been constructed as subalgebras of $\mathfrak{sl}_n(\mathbb{C})$, and one can prove that they are simple directly. Namely, the following proposition allows us to show that they are semi-simple and by verifying that their Dynkin diagrams are connected, it follows by Proposition 24.15 and Proposition 22.9 that they are simple.

Proposition 24.19. *Let* $\mathfrak{g}$ *be a finite dimensional Lie algebra over* $\mathbb{C}$ *with* $Z(\mathfrak{g}) = 0$. *Assume* $\mathfrak{h}$ *is a Cartan subalgebra and* $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$, *where* $\Phi \subseteq \mathfrak{g}^* \setminus \{0\}$. *Suppose*

1. $\dim(\mathfrak{g}_\alpha) = 1$, *for all* $\alpha \in \Phi$;
2. *if* $\alpha \in \Phi$, *then* $-\alpha \in \Phi$; *and*
3. $[[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}], \mathfrak{g}_\alpha] \neq 0$, *for all* $\alpha \in \Phi$.

Then $\mathfrak{g}$ *is semi-simple.*

Proof. Suppose for a contradiction that $\mathfrak{g}$ is not semi-simple. Then $\mathfrak{g}$ has a solvable ideal, so it has an abelian ideal $I \neq 0$. Now, $[\mathfrak{h}, I] \subseteq I$ and $\mathrm{ad}_\mathfrak{h}$ is diagonalisable on $\mathfrak{g}$, hence also on I. I is therefore a sum of eigenspaces of $\mathrm{ad}_\mathfrak{h}$, so

$$I = I \cap \mathfrak{h} \bigoplus_{\alpha \in \Phi} (I \cap \mathfrak{g}_\alpha).$$

If $I \cap \mathfrak{g}_\alpha \neq 0$, for some α, then $\mathfrak{g}_\alpha \subseteq I$ by (1), so

$$[[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}], \mathfrak{g}_\alpha] \subseteq [[I, \mathfrak{g}_{-\alpha}], I] \subseteq [I, I] = 0,$$

a contradiction with (3). Therefore, $I = I \cap \mathfrak{h}$, i.e., $I \subset \mathfrak{h}$. As $Z(\mathfrak{g}) = 0$, there exists $\alpha \in \Phi$ such that $[I, \mathfrak{g}_\alpha] \neq 0$. But

$$[I, \mathfrak{g}_\alpha] \subseteq I \cap [\mathfrak{h}, \mathfrak{g}_\alpha] \subseteq I \cap \mathfrak{g}_\alpha = 0,$$

a contradiction. □

We demonstrate this for $\mathfrak{sl}_n(\mathbb{C})$:

Theorem 24.20. $\mathfrak{sl}_n(\mathbb{C})$, $n \geq 2$, *is simple.*

Proof. We have $Z(\mathfrak{sl}_n(\mathbb{C})) = 0$. Let $\mathfrak{h} = \{\text{diagonal traceless matrices}\} \subset \mathfrak{sl}_n(\mathbb{C})$ is a Cartan subalgebra, and

$$\mathfrak{sl}_n(\mathbb{C}) = \mathfrak{h} \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha.$$

We check that $(1) - (3)$ of previous proposition holds:

1. the root spaces $\mathfrak{g}_\alpha = \mathrm{span}_\mathbb{C}\{e_{ij}\}$ are one-dimensional;
2. $\epsilon_i - \epsilon_j \in \Phi$ implies $\epsilon_j - \epsilon_i \in \Phi$; and
3. if $\mathfrak{g}_\alpha = \mathrm{span}_\mathbb{C}\{e_{ij}\}$, then $\mathfrak{g}_{-\alpha} = \mathrm{span}_\mathbb{C}\{e_{ji}\}$ and $[[e_{ij}, e_{ji}], e_{ij}] = 2e_{ij} \neq 0$.

Hence, $\mathfrak{sl}_n(\mathbb{C})$ is semi-simple by previous proposition. Finally, the Dynkin diagram of $\mathfrak{sl}_n(\mathbb{C})$

is connected. □

Similarly, one shows that $\mathfrak{sp}_{2n}(\mathbb{C})$ and $\mathfrak{so}_n(\mathbb{C})$ are simple for $n \geq 2$, apart from $\mathfrak{so}_2(\mathbb{C})$ and $\mathfrak{so}_4(\mathbb{C})$.