

MAU34103- Lie groups and Lie algebras

Adam Keilthy

Trinity College Dublin

Contents

0	A brief introduction	2
1	Lie algebras	3
1.1	First definitions and examples	3
1.2	Derivations and Lie algebras	7
1.3	Ideals and quotient Lie algebras	9
1.4	Geometric Lie algebras	11
1.4.1	Smooth manifolds and tangent vectors	11
1.4.2	The Lie algebra of vector fields	13
2	Lie groups	14
2.1	The matrix exponential	16
2.2	Lie algebras and linear groups	19
2.3	Lie group homomorphisms and the Lie correspondence.	25
3	Representations of Lie algebras	28
4	Nilpotent and solvable Lie algebras	31
4.1	Nilpotent Lie algebras	31
4.2	Solvable Lie algebras	34
5	Semisimple Lie algebras	38
5.1	Decomposing representations	40
5.2	Decomposing Lie algebras	42
5.3	Cartan's criterion and the Killing form	45
5.4	Semisimplicity and the Killing form	47
6	Representations of semisimple Lie algebras	49

0 A brief introduction

Lie groups and Lie algebras (named after the Norwegian mathematician Sophus Lie, not deceits and falsehood) sit in the intersection of algebra and analysis. Lie groups arise naturally as *symmetry groups*. In the case of finite groups, we could interpret groups such as the dihedral group and the symmetry group naturally in terms of symmetries of polygons, or orderings respectively. Many finite groups admit such an interpretation as collection of structure preserving symmetries of some discrete set.

Lie groups come about when we ask the question of what are structure preserving symmetries of continuous sets (or more specifically, manifolds). The general linear group $GL_n(\mathbb{R})$ is a Lie group describing symmetries of $\mathbb{R}^n$ as a vector space. $\mathbb{R}^n$ as a vector space is an example of a Lie group, describing certain distance-preserving symmetries of $\mathbb{R}^n$ as a metric space. The unit circle is a Lie group describing the orientation preserving symmetries of a circle. These groups of continuous symmetries appear everywhere, and so we would like to further our understanding of their structure and their actions, analogous to our understanding of finite groups.

Unfortunately, the kind of algebra involved can get quite difficult, particularly if we cannot find a nice presentation for these groups. Here, one final property of Lie groups comes to the rescue: Lie groups are manifolds themselves. These means we can exploit ideas from analysis and geometry to help us understand Lie groups. In particular, we can use the differential structure to approximate elements of the group “to first order” and convert our group theory problem into one of almost-linear algebra via Lie algebras

Geometrically, a Lie algebra is the tangent space to a Lie group, equipped with a multiplication coming from the group structure. As the tangent space is a vector space, we can appeal to all the theory of linear algebra to help us untangle and simplify Lie algebras and their actions as “infinitesimal symmetries”. But what is even better is that, under some topological assumptions, a Lie group and its actions are entirely determined by this Lie algebra and its “infinitesimal” actions. We can essentially forget the group entirely and just work with the algebra!

Even more excitingly, Lie algebras can be discussed entirely algebraically. The analysis (both implicit and explicit) in the study of Lie groups can be forgotten about entirely. The manifold structure and need for coordinate charts do not play a role if we are discussing Lie algebras in isolation. Thus, this provides us with a great entry point into the course.

1 Lie algebras

1.1 First definitions and examples

In all that follows, we will fix a field K . There is little lost in assuming that $K = \mathbb{R}$ (and we will later take this to be the case) but all our definitions in this section hold more generally. We first recall the definition of an algebra.

Definition 1.1. *An algebra over K is a K -vector space A , equipped with a bilinear map*

$$\begin{aligned} \cdot : A \times A &\rightarrow A, \\ (a, b) &\mapsto a \cdot b. \end{aligned}$$

We call an algebra:

- Associative if $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in A$,
- Commutative if $a \cdot b = b \cdot a$ for all $a, b \in A$,
- Unital if there exists $1_A \in A$ such that $a \cdot 1_A = 1_A \cdot a = a$.

A homomorphism between algebras is a linear map $f : A \rightarrow B$ such that $f(a \cdot b) = f(a) \cdot f(b)$.

A Lie algebra is a special case of a non-associative, (usually) non-commutative, non-unital algebra, with product given by a *Lie bracket*.

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

$$\begin{aligned} [-, -] : L \times L &\rightarrow L, \\ (a, b) &\mapsto [a, b], \end{aligned}$$

called a Lie bracket, satisfying

1. Anticommutativity: $[x, x] = 0$ for all $x \in L$,
2. Jacobi identity: for all $x, y, z \in L$,

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

A linear map $f : L_1 \rightarrow L_2$ is called a Lie algebra homomorphism if it is an algebra homomorphism, i.e.

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

for all $x, y \in L_1$.

Remark 1.3. It might look a little odd to call $[x, x] = 0$ an anticommutativity condition, but

$$[x + y, x + y] = 0 \quad \Rightarrow \quad [x, y] = -[y, x]$$

so it implies a more obvious anticommutativity. If $\text{char } K \neq 2$, then the reverse implication holds

$$[x, y] = -[y, x] \quad \Rightarrow \quad [x, x] = 0$$

but this is not the case in characteristic 2.

Example 1.4. Let A be your favourite associative algebra. Then we define the commutator of two elements of A by

$$[x, y] = xy - yx.$$

Then the algebra $L(A) = (A, [-, -])$ given by taking A as the underlying vector space, and the commutator as the product is a Lie algebra.

Example 1.5. If A is the algebra of $n \times n$ matrices, $M_n(K)$, then taking the commutator as in the previous example gives an important example of a Lie algebra. We define the general linear Lie algebra

$$\mathfrak{gl}_n(K) := (M_n(K), [-, -]) = L(M_n(K))$$

More generally if V is a K -vector space, the space of its endomorphisms

$$\text{End}(V) := \{f : V \rightarrow V \mid f \text{ linear}\}$$

is an associative algebra and we define

$$\mathfrak{gl}(V) := (\text{End}(V), [-, -]) = L(\text{End}(V))$$

If V is n -dimensional, then it is not hard to show $\mathfrak{gl}(V) \cong \mathfrak{gl}_n(K)$ as Lie algebras.

Example 1.6. Not every Lie bracket has to be interesting! Let V be a vector space and define a Lie bracket

$$[x, y] = 0 \quad \text{for all } x, y \in V.$$

Then $(V, [-, -])$ is a Lie algebra. We similarly get a trivial Lie bracket on $L(A)$ if A is a commutative associative algebra.

Definition 1.7. We say $x, y \in L$ commute if $[x, y] = 0$. A Lie algebra L is called abelian if every pair of elements commute, i.e. if the Lie bracket is always 0.

Remark 1.8. In a commutator Lie algebra $L(A)$, x, y commute in $L(A)$ if and only if they commute in A . $L(A)$ is abelian if and only if A is commutative.

By taking subspace of $\mathfrak{gl}_n(K)$ closed under the Lie bracket, we obtain a whole range of further examples.

Definition 1.9. Let L be a Lie algebra over K . A Lie subalgebra is a vector subspace $Z \subset L$ that is closed under the induced Lie bracket:

$$[z, w] \in Z \quad \text{for all } z, w \in Z.$$

Note that this implies $(Z, [-, -])$ is also a Lie algebra.

Example 1.10. Consider the vector subspace of $\mathfrak{gl}_n(K)$ given by matrices with 0 trace:

$$\mathfrak{sl}_n(K) := \{x \in \mathfrak{gl}_n(K) \mid \text{tr}(x) = 0\}$$

We claim this is a Lie subalgebra. As the trace is linear and satisfies

$$\text{tr}(xy) = \text{tr}(yx)$$

we have that

$$\text{tr}([x, y]) = \text{tr}(xy - yx) = \text{tr}(xy) - \text{tr}(yx) = 0.$$

Thus $\mathfrak{sl}_n(K)$ is a Lie subalgebra of $\mathfrak{gl}_n(K)$, called the special linear Lie algebra.

Example 1.11. As the subspaces

$$\begin{aligned} \mathfrak{b}_n &:= \{x \in M_n(K) \mid x_{i,j} = 0 \text{ if } i > j\}, \\ \mathfrak{n}_n &:= \{x \in M_n(K) \mid x_{i,j} = 0 \text{ if } i \geq j\}, \\ \mathfrak{t}_n &:= \{x \in M_n(K) \mid x_{i,j} = 0 \text{ if } i \neq j\}, \end{aligned}$$

of upper triangular, strictly upper triangular, and diagonal matrices (respectively) each form subalgebras of $M_n(K)$, they are also all Lie subalgebras of $\mathfrak{gl}_n(K)$.

Example 1.12. Given a bilinear map $\sigma : V \times V \rightarrow K$, we can define a Lie algebra

$$\mathfrak{o}(V, \sigma) := \{x \in \mathfrak{gl}(V) \mid \sigma(xv, w) + \sigma(v, xw) = 0 \text{ for all } v, w \in V\}$$

called the orthogonal Lie algebra. Let us check that this is, indeed, a Lie subalgebra of $\mathfrak{gl}(V)$. It is enough to check

$$\sigma([x, y]v, w) + \sigma(v, [x, y]w) = 0.$$

We have that

$$\begin{aligned} \sigma([x, y]v, w) &= \sigma(xyv - yxv, w) \\ &= \sigma(xyv, w) - \sigma(yxv, w) \\ &= -\sigma(yv, xw) + \sigma(xv, yw) \\ &= \sigma(v, yxw) - \sigma(v, xyw) \\ &= -\sigma(v, (xy - yx)w) = -\sigma(v, [x, y]w) \end{aligned}$$

as needed.

If $V \cong K^n$, then by picking a basis, we can find a matrix $S \in M_n(K)$ such that

$$\sigma(v, w) = v^T S w$$

and so the condition defining $\mathfrak{o}(V, \sigma)$ becomes

$$x^T S + S x = 0$$

Example 1.13. If in addition we have that σ is non-degenerate (so that S is invertible), then we can rewrite this condition as

$$x + S^{-1} x^T S = 0.$$

Taking traces, as using that trace is invariant under conjugation, we find

$$0 = \text{tr}(x) + \text{tr}(S^{-1} x^T S) = \text{tr}(x) + \text{tr}(x^T) = 2 \text{tr}(x).$$

Hence, at least if K has characteristic different to 2, $\mathfrak{o}(V, \sigma)$ is (isomorphic to) a Lie subalgebra of $\mathfrak{sl}_n(K)$ when V is finite dimensional and σ is non-degenerate. Hence, when these conditions apply, we write

$$\mathfrak{so}_n(K, S) := \mathfrak{o}(K^n, \sigma)$$

and call it the special orthogonal Lie algebra. If it is possible to choose a basis such that $S = I$ (for example, if σ is symmetric and $K = \mathbb{C}$), then we usually omit S from the notation, as we can identify the Lie algebra with

$$\mathfrak{o}_n(K) = \mathfrak{so}_n(K) = \{x \in \mathfrak{gl}_n(K) \mid x + x^T = 0\}$$

Example 1.14. If V is finite dimensional, σ is symmetric and non-degenerate, and $K = \mathbb{R}$, then we can always find a basis such that S is diagonal with non-zero entries ± 1 :

$$S = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$$

for some $p + q = \dim(V)$. In this case, we denote the corresponding Lie algebra by $\mathfrak{so}(p, q)$, and refer to it as the indefinite special orthogonal Lie algebra of signature (p, q) . Such Lie algebras play a substantial role in relativity. In particular, the Lorentz Lie algebra $\mathfrak{so}(1, 3)$ can be thought of as “infinitesimal” rotations of space-time.

Example 1.15. If V is finite dimensional, σ is skew-symmetric and non-degenerate, and $K = \mathbb{R}$, then we can find a basis of V such that

$$S = \begin{pmatrix} 0 & I_p \\ -I_p & 0 \end{pmatrix}$$

In this case, we write $\mathfrak{sp}(2p) := \mathfrak{so}(2p, S)$, and call this the symplectic Lie algebra.

Remark 1.16. One can analogously define the unitary (and special unitary) Lie algebra over $\mathbb{R}$, given by the subspace of the $\mathbb{R}$ -vector space $M_n(\mathbb{C})$ consisting of matrices such that $x + \bar{x}^T = 0$ (and $\text{tr}(x) = 0$ in the special case) equipped with the commutator. These Lie algebras, denoted by $\mathfrak{u}(n)$ and $\mathfrak{su}(n)$ respectively play an important role in modern physics.

While every example we have seen so far has Lie bracket that can be written as a commutator in some larger algebra, it is not the case that every Lie algebra is of the form $L(A)$ for some associative algebra A . A simple example of this is given by

$$\mathfrak{sl}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a + d = 0 \right\}$$

which is a 3-dimensional vector space with basis

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

We can easily compute

$$[x, y] = h, \quad [h, x] = 2x, \quad [h, y] = -2y$$

and from this, every Lie bracket can be computed. Although these Lie brackets are computed in terms of commutators, we nevertheless cannot have $\mathfrak{sl}_2(\mathbb{R}) = L(A)$, as if we did, then

$$xy = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{R})$$

which it is not.

Let us finish out this list of examples with an example not defined in terms of matrices and commutators.

Example 1.17. Denote by $\mathbb{C}[z, z^{-1}]$ the space of Laurent polynomials, and $\mathbb{C}q$ a 1-dimensional complex vector space spanned by an element q . Define the Heisenberg algebra to be the Lie algebra with underlying vector space $\mathcal{H} = \mathbb{C}[z, z^{-1}] \oplus \mathbb{C}q$, and Lie bracket

$$\begin{aligned} [f, q] &= [q, f] = 0 \text{ for all } f \in \mathbb{C}[z, z^{-1}], \\ [f, g] &= \text{coeff}_{z^{-1}}(f'g)q \end{aligned}$$

Let us check that this is antisymmetric.

$$\begin{aligned} [f, g] + [g, f] &= \text{coeff}_{z^{-1}}(f'g) + \text{coeff}_{z^{-1}}(fg') \\ &= \text{coeff}_{z^{-1}}((fg)') = 0 \end{aligned}$$

as no Laurent polynomial has a derivative containing a z^{-1} . Checking the Jacobi identity follows similarly.

1.2 Derivations and Lie algebras

Another important class of Lie algebras (and our first hint that Lie algebras capture something “infinitesimal”) are derivations of algebras

Definition 1.18. Let A be an algebra. A linear map $\delta : A \rightarrow A$ is called a *derivation* if it satisfies Leibniz rule

$$\delta(ab) = \delta(a)b + a\delta(b)$$

for all $a, b \in A$.

We denote the space of derivations of A by $\text{Der}(A)$.

Lemma 1.19. Let A be a K -algebra. Then $\text{Der}(A)$, equipped with the commutator

$$[\delta, \partial] = \delta \circ \partial - \partial \circ \delta$$

is a Lie subalgebra of $\mathfrak{gl}(A)$.

Proof. It is an easy check to see that $\text{Der}(A) \subset \mathfrak{gl}(A)$ as vector spaces, so it suffices to check that $[\delta, \partial] \in \text{Der}(A)$. We check this directly:

$$\begin{aligned} [\delta, \partial](ab) &= \delta(\partial(ab)) - \partial(\delta(ab)) \\ &= \delta(\partial(a)b) + \delta(a\partial(b)) - \partial(\delta(a)b) - \partial(a\delta(b)) \\ &= \delta(\partial(a))b + \partial(a)\delta(b) + \delta(a)\partial(b) + a\delta(\partial(b)) \\ &\quad - \partial(\delta(a))b - \delta(a)\partial(b) - \partial(a)\delta(b) - a\partial(\delta(b)) \\ &= [\delta, \partial](a)b + a[\delta, \partial](b), \end{aligned}$$

as needed. □

Example 1.20. The Witt algebra is defined to be $\text{Der } \mathbb{C}[t]$. This contains an obvious derivation $\partial = \frac{d}{dt}$. What others can exist? Note that a derivation δ is entirely determined by its action on $\{t^n\}_{n \geq 0}$. We compute that

$$\begin{aligned} \delta(t^2) &= \delta(t)t + t\delta(t) = 2\delta(t)t = \delta(t)\partial(t^2) \\ \delta(t^3) &= \delta(t^2)t + t^2\delta(t) = 3\delta(t)t^2 = \delta(t)\partial(t^3) \\ &\vdots \\ \delta(t^n) &= n\delta(t)t^{n-1} = \delta(t)\partial(t^n). \end{aligned}$$

Thus, using linearity, $\delta(f(t)) = \delta(t)\partial(f(t))$ and we can conclude that $\text{Der } \mathbb{C}[t] = \mathbb{C}[t]\partial$, with Lie bracket

$$[f\partial, g\partial] = (fg' - f'g)\partial.$$

Analogously to how we can study a group G via the map

$$\begin{aligned} G &\rightarrow \text{Aut}(G) \\ g &\mapsto (h \mapsto ghg^{-1}) \end{aligned}$$

an important class of derivation Lie algebras are derivations of Lie algebras.

Definition 1.21. Let L be a Lie algebra. Given $x \in L$, we define its joint action on L via the map

$$\begin{aligned}\mathrm{ad}_x : L &\rightarrow L \\ y &\mapsto [x, y].\end{aligned}$$

We call the map

$$\begin{aligned}\mathrm{ad} : L &\rightarrow \mathfrak{gl}(L) \\ x &\mapsto \mathrm{ad}_x\end{aligned}$$

the adjoint representation of L .

Lemma 1.22. For any $x \in L$, the map ad_x is a derivation of L , called a inner derivation. Furthermore, the map

$$\mathrm{ad} : L \rightarrow \mathrm{Der}(L) \subset \mathfrak{gl}(L)$$

is a Lie algebra homomorphism.

Proof. Let us first check that $\mathrm{ad}_x \in \mathrm{Der}(L)$:

$$\begin{aligned}\mathrm{ad}_x([y, z]) &= [x, [y, z]] \\ &= -[y, [z, x]] - [z, [x, y]] \\ &= [y, [x, z]] + [[x, y], z] \\ &= [y, \mathrm{ad}_x(z)] + [\mathrm{ad}_x(y), z]\end{aligned}$$

using the Jacobi identity and anticommutativity. This holds for all $y, z \in L$. Thus ad_x is a derivation. To see that ad is a Lie algebra homomorphism, it is enough to show that

$$\mathrm{ad}_{[x, y]}(z) = [\mathrm{ad}_x, \mathrm{ad}_y](z)$$

for all $x, y, z \in L$. Indeed

$$\begin{aligned}\mathrm{ad}_{[x, y]}(z) &= [[x, y], z] \\ &= -[z, [x, y]] \\ &= [x, [y, z]] + [y, [z, x]] \\ &= [x, [y, z]] - [y, [x, z]] \\ &= \mathrm{ad}_x(\mathrm{ad}_y(z)) - \mathrm{ad}_y(\mathrm{ad}_x(z)) \\ &= [\mathrm{ad}_x, \mathrm{ad}_y](z).\end{aligned}$$

□

1.3 Ideals and quotient Lie algebras

Definition 1.23. Given a Lie algebra L , and (vector) subspaces $M, N \subset L$, we define $[M, N]$ to be the vector subspace generated by brackets $[m, n]$ with $m \in M$, $n \in N$.

A subspace $I \subset L$ is called an ideal of L if $[L, I] \subset I$, i.e. $[x, y] \in I$ if $y \in I$.

Two important examples of ideals are the centre and the derived algebra of L .

Definition 1.24. *The centre of L is the set of elements that commute with L :*

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

Definition 1.25. *The derived algebra of L is $L^1 := [L, L]$, the linear span of all brackets of elements of L .*

Example 1.26. • $Z(L)$ and L^1 are ideals of L , for any Lie algebra L .

- If $f : L_1 \rightarrow L_2$ is a homomorphism, then $\ker(f)$ is an ideal. Indeed, if $f(y) = 0$, then

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

Example 1.27. If $L = \mathfrak{sl}_2(\mathbb{R})$, then the only ideals are 0 and L itself. Indeed, we know that, as a vector space $L \cong \mathbb{R}\langle x, y, h \rangle$. We have

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

Suppose I is non-zero, containing a non-zero vector

$$v = ax + by + ch.$$

We will show that I must then contain a basis vector. Suppose, $a \neq 0$. Then

$$[y, v] = -ah + 2cx \in I$$

and so

$$[x, [y, v]] = -2ay \in I.$$

Hence $y \in I$, and so $[x, y] = h \in I$. Thus, $[h, x] = 2x \in I$, and so we must have $I = L$. The cases where $b \neq 0$ or $c \neq 0$ can be handled similarly.

Definition 1.28. Given an ideal $I \subset L$, we define the quotient space L/I to be the set of equivalence classes of L under the equivalence $x \sim y$ if $x - y \in I$.

Lemma 1.29. L/I is a Lie algebra with Lie bracket

$$[x + I, y + I] := [x, y] + I$$

Proof. That it is a well defined vector space follows from the corresponding result on vector spaces. Thus, we really just need to show that the bracket is independent of choice of representatives.

Suppose

$$x + I = x' + I \quad \text{and} \quad y + I = y' + I.$$

This implies that $x = x' + i$ and $y = y' + j$ for some $i, j \in I$. Then

$$\begin{aligned} [x, y] &= [x' + i, y' + j] \\ &= [x', y'] + [x', j] + [i, y'] + [i, j] = [x', y'] + I \end{aligned}$$

Hence, the bracket operation on L/I is well defined. $\square$

Example 1.30. If we take $L = \mathfrak{b}_2(K)$ to be the Lie algebra of upper triangular matrices, it is not hard to see that $\mathfrak{n}_2(K)$, the Lie algebra of strictly upper triangular matrices, is an ideal. Since

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} - \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \begin{pmatrix} 0 & r \\ 0 & 0 \end{pmatrix}$$

if and only if $a = A$ and $b = B$, we see that

$$\mathfrak{b}_2(K)/\mathfrak{n}_2(K) \cong \mathfrak{t}_2(K)$$

is isomorphic to the Lie algebra of diagonal matrices.

1.4 Geometric Lie algebras

Now that we have a catalogue of Lie algebras under our belt, let's consider some of the earliest examples of Lie algebra, coming from the geometry of manifolds. We will first briefly recap some of the key concepts from the theory of smooth manifolds. As this is not a course in smooth manifolds, we will sweep some subtleties under the rug, and so the definitions may differ slightly to those you are familiar with.

1.4.1 Smooth manifolds and tangent vectors

Definition 1.31. Let M be a topological space. A chart on M is a pair consisting of an open set $U \subset M$ and a continuous map $\phi : U \rightarrow \mathbb{R}^d$ such that $\phi(U)$ is open in $\mathbb{R}^d$ and

$$\phi : U \rightarrow \phi(U)$$

is a homeomorphism.

An (smooth) atlas on M is a collection of charts $(U_i, \phi_i)_{i \in I}$ such that

- $M = \cup_{i \in I} U_i$,
- The composition

$$\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \rightarrow \phi_j(U_i \cap U_j)$$

is smooth (C^∞ -differentiable).

Note that as this composition is a homeomorphism, the existence of a (smooth) atlas implies there is a singular d such that $\phi_i(U_i) \subset \mathbb{R}^d$ for all i .

A (smooth) manifold (of dimension d) is a Hausdorff topological space M with a smooth atlas (up to some equivalence). If $M \subset \mathbb{R}^n$, we say that M is embedded into $\mathbb{R}^n$.

Intuitively, a smooth manifold is a topological space that “locally” looks like a patch of $\mathbb{R}^d$. As such, many familiar topological spaces give examples of manifolds: the unit circle, the torus, , the open ball. The implicit function theorem

is a very powerful tool for identifying manifolds, guaranteeing the existence of an atlas for sets to which it applies.

Using this local identification, we can essentially “lift” all our classical multivariable calculus to calculus on a manifold. In particular, we can define a tangent vector.

Definition 1.32. Let M be a d -dimensional manifold, and let $(U_i, \phi_i)_{i \in I}$ be an atlas. Let $p \in U_i \subset M$. We call two smooth curves $\gamma_1, \gamma_2 : (-\varepsilon, \varepsilon) \rightarrow M$ such that $\gamma_1(0) = \gamma_2(0) = p$ equivalent if

$$\lim_{t \rightarrow 0} \frac{\phi_i(\gamma_1(t)) - \phi_i(\gamma_2(t))}{t} = 0.$$

We denote the equivalence class of γ by $\dot{\gamma}(0)$, and call this a tangent vector at p . The set of equivalence classes is called the tangent space to M at p and is denoted by $T_p M$.

As ϕ_i and γ are smooth, we have that

$$\phi_i(\gamma(t)) \approx \phi_i(p) + \dot{\gamma}(0)t + \text{Something small}$$

for some vector $\dot{\gamma}(0) \in \mathbb{R}^d$. An equivalence class $\dot{\gamma}(0)$ is uniquely determined by this vector, and so we can identify $T_p M \cong \mathbb{R}^d$.

Remark 1.33. If M is embedded in $\mathbb{R}^n$, then the tangent space and tangent vectors are exactly what we expect from classic multivariable calculus, with the tangent vector to a curve given by the vector

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

Under the identification $T_p M \cong \mathbb{R}^d$, it is standard to denote the standard basis by

$$e_1 = \frac{\partial}{\partial x_1}|_p, e_2 = \frac{\partial}{\partial x_2}|_p, \dots, e_d = \frac{\partial}{\partial x_d}|_p.$$

This is because, if

$$\phi_i(\gamma(t)) = (\gamma_1(t), \dots, \gamma_d(t)) \in \mathbb{R}^d$$

then we find

$$\dot{\gamma}(0) = \gamma'_1(0) \frac{\partial}{\partial x_1}|_p + \dots + \gamma'_d(0) \frac{\partial}{\partial x_d}|_p.$$

This corresponds exactly to taking a sort of directional derivative: if M is embedded in $\mathbb{R}^d$, and $f \in C^\infty(M)$ is a smooth function near p , then we can take the directional derivative via the limit

$$(\mathbf{D}_{\vec{v}} f)(p) := \lim_{t \rightarrow 0} \frac{f(\gamma(t)) - f(p)}{t}$$

if $\gamma(t) \approx p + \vec{v}t + \dots$. Computing this limit using chain rule, we find

$$(\mathbf{D}_{\vec{v}} f)(p) = (f \circ \gamma)'(0) = \sum_{i=1}^d \frac{\partial f}{\partial x_i}|_p \gamma'_i(0) = \left(\sum_{i=1}^d \gamma'_i(0) \frac{\partial}{\partial x_i}|_p \right) f$$

This holds even if M is not embedding in real space, and so we can think of tangent vectors as directional derivatives. Furthermore, via this interpretation, we find a way to see tangent vectors almost as derivations:

$$D_\gamma(fg)(p) = D_\gamma(f)(p)g(p) + f(p) D_\gamma(g)(p).$$

1.4.2 The Lie algebra of vector fields

Unfortunately, we don't have a way of making sense of a Lie algebra of derivations of functions "at" p , but by taking families of tangent vectors, we can define a Lie algebra structure on $\text{Der } C^\infty(M)$.

Definition 1.34. *The union $TM = \cup_{p \in M} T_p M$ defines a smooth $2d$ -dimensional manifold, called the tangent bundle of M . A vector field on M is a smooth map $X : M \rightarrow TM$ such that $X_p \in T_p M$ for every $p \in M$.*

Definition 1.35. *Given a vector field X , we define its Lie derivative to be the map*

$$\begin{aligned} L_X : C^\infty(M) &\rightarrow C^\infty(M) \\ f &\mapsto L_X(f) \end{aligned}$$

defined by $L_X(f)(p) = X_p(f)$.

As each X_p is a tangent vector, which acts as a directional derivative, we have the following.

Lemma 1.36. *For any vector field X , the Lie derivative is a derivation: $L_X \in \text{Der } C^\infty(M)$.*

In fact, similarly to how we handled the Witt algebra, one can show

$$\text{Der } C^\infty(M) = \{L_X \mid X \text{ is a vector field on } M\}.$$

As $\text{Der } C^\infty(M)$ is a Lie algebra, this equality implies that

$$[L_X, L_Y] = L_Z$$

for some unique vector field Z , allowing us to "lift" the Lie structure to vector fields.

Proposition 1.37. *Defining $[X, Y]$ to be the unique vector field such that $[L_X, L_Y] = L_{[X, Y]}$ equips the space of vector fields with the structure of a Lie algebra.*

Lemma 1.38. *Writing $X = \sum_{i=1}^d X_i \frac{\partial}{\partial x_i}$, $Y = \sum_{i=1}^d Y_i \frac{\partial}{\partial x_i}$ in local coordinates on a chart U ($X_i, Y_i \in C^\infty(U)$), then*

$$[X, Y] = \sum_{j=1}^d \left(\sum_{i=1}^d X_i \frac{\partial Y_j}{\partial x_i} - Y_i \frac{\partial X_j}{\partial x_i} \right) \frac{\partial}{\partial x_j}$$

on U .

Proof. It is enough to compute $[L_X, L_Y](f)(p)$ for all $p \in U$. Writing $\partial_i = \frac{\partial}{\partial x_i}$, we have

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

□

Remark 1.39. If $d = 1$, this collapses to $[f\partial, g\partial] = (fg' - f'g)\partial$, just like in the Witt algebra.

Know the Lie algebra of vector fields tells you a lot about the analytic structure of the manifold M , but requires us to already know TM , so isn't particularly helpful in isolation.

However, if we have a way of identifying $T_p M$ with $T_{p_0} M$ in some smooth way, we could define a Lie algebra on $T_{p_0} M$. Given a Lie algebra structure on a tangent space, we can use this identification to essentially determine M . This identification is exactly what we get from Lie groups.

2 Lie groups

Definition 2.1. A Lie group G is a group G equipped with the structure of a smooth manifold, such that the maps defining the group structure

$$\begin{aligned}
G \times G &\rightarrow G, & G &\rightarrow G \\
(x, y) &\mapsto xy, & x &\mapsto x^{-1}
\end{aligned}$$

are smooth maps.

Exercise 2.2. Show that it is sufficient that the map

$$\begin{aligned}
G \times G &\rightarrow G \\
(x, y) &\mapsto xy^{-1}
\end{aligned}$$

be smooth.

Many familiar infinite groups provide examples of Lie groups!

Example 2.3. • The group $(\mathbb{R}^n, +)$ is a Lie group of dimension n ,

- The group $(\mathbb{C}^n, +)$ is a Lie group of dimension $2n$,
- The general linear group

$$\mathrm{GL}_n(\mathbb{R}) = \{M \in \mathrm{M}_n(\mathbb{R}) \mid \det M \neq 0\}$$

is an open subset of $\mathbb{R}^{n^2}$ and is therefore a manifold with a single chart. Since the multiplication and inversion maps are expressed in terms of polynomials and the inverse determinant, they are smooth maps. It is therefore a Lie group of dimension n^2 .

- Similarly $\mathrm{GL}_n(\mathbb{C})$ is a Lie group of dimension $2n^2$.
- The circle

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

is a manifold of dimension 1 by the implicit function theorem. We can give it a group structure by addition of angles, which can be expressed in terms of smooth trigonometric functions. Thus S^1 is a Lie group of dimension 1.

- The special linear group

$$\mathrm{SL}_n(\mathbb{R}) = \{M \in \mathrm{M}_n(\mathbb{R}) \mid \det M = 1\}$$

is a manifold of dimension $n^2 - 1$ by the implicit function theorem, and forms a Lie group just like the general linear group.

Remark 2.4. Not every Lie group is a matrix group, though those that are not can be difficult to describe. Although maybe a bit difficult to describe without more theory, such groups still play an important role. For example, the metaplectic group, which plays a key role in theoretical physics.

As the vast majority of Lie groups we can currently describe are matrix groups, we will focus primarily on these. Specifically, we will look at a class of Lie groups called linear groups.

Definition 2.5. A linear group is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$, i.e. a subgroup that is closed as a subset in the usual topology on $\mathrm{GL}_n(\mathbb{R})$ inherited from $\mathbb{R}^{n^2}$.

It will be a while before we can formally show that every linear group is a Lie group, but in most examples we will see, it would follow from the implicit function theorem.

Example 2.6. • The general linear group $\mathrm{GL}_n(\mathbb{R})$ is clearly closed in itself.

- The special linear group $\mathrm{SL}_n(\mathbb{R})$ is closed as a subset of $\mathbb{R}^{n^2}$ and a subset of $\mathrm{GL}_n(\mathbb{R})$, hence a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$.

- The orthogonal group

$$\mathrm{O}_n(\mathbb{R}) = \{M \in \mathrm{M}_n(\mathbb{R}) \mid M^T M = I\}$$

is a linear group. More generally, given a bilinear map $\sigma : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ gives an associated linear group

$$\mathrm{O}_n(\sigma) := \{M \in \mathrm{GL}_n(\mathbb{R}) \mid \sigma(Mv, Mw) = \sigma(v, w) \text{ for all } v, w \in \mathbb{R}^n\}.$$

This contains a number of more familiar examples. For example, if σ is non-degenerate, we can find a basis such that

$$\sigma(v, w) = v^T \begin{pmatrix} I_p & 0 \\ 0 & I_q \end{pmatrix} w$$

for some $p + q = n$. In this case, we write $\mathrm{O}(p, q)$ for $\mathrm{O}_n(\sigma)$, and call it the indefinite orthogonal group of signature (p, q) .

This also covers the symplectic group: if $n = 2p$ and there is a basis for which

$$\sigma(v, w) = v^T \begin{pmatrix} 0 & I_p \\ -I_p & 0 \end{pmatrix} w$$

In this instance, we write $\mathrm{Sp}(n)$ for $\mathrm{O}_{2p}(\sigma)$.

- The special orthogonal group $\mathrm{SO}_n(\mathbb{R}) = \mathrm{SL}_n(\mathbb{R}) \cap \mathrm{O}_n(\mathbb{R})$ is a linear group
- By embedding $\mathrm{GL}_n(\mathbb{C}) \hookrightarrow \mathrm{GL}_{2n}(\mathbb{R})$, we can view $\mathrm{GL}_n(\mathbb{C})$ as a linear group. Similarly, the unitary group

$$\mathrm{U}(n) = \{M \in \mathrm{GL}_n(\mathbb{C}) \mid \overline{M}^T M = I\}$$

- The groups of upper triangular, strictly upper triangular, and diagonal matrices are linear groups
- The additive group $(\mathbb{R}, +)$ is a linear group via the isomorphism:

$$(\mathbb{R}, +) \cong \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mid x \in \mathbb{R} \right\}$$

2.1 The matrix exponential

Our goal for the next while will be to establish a relationship between Lie algebras and Lie groups in the special case of linear groups. Essentially every result we show holds in greater generality, though interpreting what the statements mean requires more theory.

We start by considering the matrix exponential and associated properties.

Definition 2.7. Define the exponential map

$$\exp : \mathfrak{gl}_n(\mathbb{R}) \rightarrow \mathrm{GL}_n(\mathbb{R})$$

by the exponential series

$$\exp(x) = \sum_{n \geq 0} \frac{x^n}{n!}, \quad (x^0 := I)$$

We will often write $e^x = \exp(x)$, though it is important to remember that this is not the real exponential.

We can show that this series converges by introducing a norm, such as

$$\|x\|_2 := \left(\sum_{i,j} x_{i,j}^2 \right)^{\frac{1}{2}}$$

which satisfies

1. $\|x\| = 0$ if and only if $x = 0$,
2. $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{R}$, $x \in \mathfrak{gl}_n(\mathbb{R})$,
3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathfrak{gl}_n(\mathbb{R})$,
4. $\|xy\| \leq \|x\| \|y\|$ for all $x, y \in \mathfrak{gl}_n(\mathbb{R})$.

The convergence of the series in this norm follows then from the convergence of $\exp(\|x\|)$.

Exercise 2.8. Prove the series converges rigorously. See the corresponding exercise for Analysis in Several Real Variables if you need a hint.

To see the $\exp(x) \in \mathrm{GL}_n(\mathbb{R})$ as claimed, we need the following lemma.

Lemma 2.9. If $xy = yx$, then $e^x e^y = e^{x+y} = e^y e^x$

Proof. It is enough to show $e^{x+y} = e^x e^y$. Using that $xy = yx$ (and absolute convergence), we have that

$$\begin{aligned} e^{x+y} &= \sum_{n \geq 0} \frac{(x+y)^n}{n!} \\ &= \sum_{n \geq 0} \sum_{i=0}^n \binom{n}{i} \frac{x^i y^{n-i}}{n!} \\ &= \sum_{i,j \geq 0} \binom{i+j}{i} \frac{x^i y^j}{(i+j)!} \\ &= \sum_{i,j \geq 0} \frac{x^i}{i!} \frac{y^j}{j!} = e^x e^y \end{aligned}$$

□

Corollary 2.10. *The exponential of a matrix is invertible: $\exp(x) \in \mathrm{GL}_n(\mathbb{R})$.*

Proof. The matrices x and $-x$ commute, so

$$e^x e^{-x} = e^{x-x} = e^0 = I$$

Hence e^x is invertible, with inverse e^{-x} . □

We can also prove this more directly as follows.

Lemma 2.11. *For $g \in \mathrm{GL}_n(\mathbb{R})$ and $x \in \mathfrak{gl}_n(\mathbb{R})$, $e^{gxg^{-1}} = ge^xg^{-1}$.*

Proof. As $(gxg^{-1})^n = gx^n g^{-1}$, we have

$$e^{gxg^{-1}} = \sum_{n \geq 0} \frac{(gxg^{-1})^n}{n!} = g \left(\sum_{n \geq 0} \frac{x^n}{n!} \right) g^{-1} = ge^xg^{-1}.$$

□

Corollary 2.12. *For any $x \in \mathfrak{gl}_n(\mathbb{R})$, $\det(e^x) = e^{\mathrm{tr} x}$. In particular $\det e^x > 0$.*

Proof. There exists $g \in \mathrm{GL}_n(\mathbb{R})$ such that gxg^{-1} is upper triangular with diagonal given by the eigenvalues $\lambda_1, \dots, \lambda_n$. As

$$\begin{pmatrix} \lambda_1 & \cdots & \cdot \\ 0 & \ddots & \vdots \\ 0 & 0 & \lambda_n \end{pmatrix}^m = \begin{pmatrix} \lambda_1^m & \cdots & \cdot \\ 0 & \ddots & \vdots \\ 0 & 0 & \lambda_n^m \end{pmatrix}$$

for any upper triangular matrix, we have that $\exp(gxg^{-1})$ is upper triangular, with diagonal given by $e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n}$. Hence

$$\det(e^x) = \det(ge^xg^{-1}) = \det(e^{gxg^{-1}}) = e^{\lambda_1} e^{\lambda_2} \cdots e^{\lambda_n} = e^{\mathrm{tr}(x)}.$$

□

By direct computation, one can show that $\exp$ is a smooth map, with invertible derivative at 0, so one could use the inverse function theorem to show that $\exp$ defines a local homeomorphism near 0. It will be more useful to us to prove this more explicitly by constructing a local inverse.

Lemma 2.13. *The series*

$$\log(I + M) = - \sum_{m \geq 1} \frac{(-M)^m}{m}$$

is absolutely convergent for $M \in \mathfrak{gl}_n(\mathbb{R})$ with $\|M\| < 1$. Whenever the expressions involved make sense, then

$$\begin{aligned} e^{\log(I+M)} &= I + M \quad (\|M\| < 1) \\ \log(e^M) &= M \quad (\|M\| < \log 2). \end{aligned}$$

Proof. Convergence follows from comparison to the series for the real logarithm. The identities hold at the level of formal power series, and so hold whenever everything is well defined. So all that needs to be checked is that $\|e^M - I\| < 1$ if $\|M\| \log 2$. By the triangle inequality

$$\|e^M - I\| \leq \sum_{m \geq 1} \frac{\|M\|^m}{m!} = e^{\|M\|} - 1$$

and so $\|e^M - I\| < 1$ if $e^{\|M\|} < 2$, as claimed. $\square$

Remark 2.14. *It is not relevant to the content of this course, so the claim about formal power series in the preceding proof will be handled in an appendix.*

Theorem 2.15. *The map $\exp : \mathfrak{gl}_n(\mathbb{R}) \rightarrow \mathrm{GL}_n(\mathbb{R})$ is a local homeomorphism at 0: there is an open neighbourhood of 0 $U \subset \mathfrak{gl}_n(\mathbb{R})$ such that $V = \exp(U)$ is open and*

$$\exp : U \rightarrow V$$

is a homeomorphism.

Proof. We define

$$W = \{B \in \mathrm{GL}_n(\mathbb{R}) \mid \|B - I\| < 1\}$$

$$U_\varepsilon = \{A \in \mathfrak{gl}_n(\mathbb{R}) \mid \|A\| < \varepsilon\}$$

where $\varepsilon > 0$ is chosen such that $\exp(U) \subset W$. We then have continuous maps

$$\begin{aligned} \exp : \mathfrak{gl}_n(\mathbb{R}) &\rightarrow \mathrm{GL}_n(\mathbb{R}) \\ \log : W &\rightarrow \mathfrak{gl}_n(\mathbb{R}) \end{aligned}$$

such that

$$\begin{aligned} \log(e^A) &= A \text{ for } A \in U, \\ e^{\log B} &= B \text{ for } B \in W. \end{aligned}$$

Let $V = \exp(U) \subset W$. If $B \in V$, then $B = e^A$ for some $A \in U$ and so $\log B = A$. Hence $\exp : U \rightarrow V$ and $\log : V \rightarrow U$ are inverse functions, giving a homeomorphism. Finally to see that V is open in $\mathrm{GL}_n(\mathbb{R})$, we can note that $V = \log^{-1}(U)$ as a subset of W , and so V is open in W , which is open in $\mathrm{GL}_n(\mathbb{R})$. The claim then follows. $\square$

2.2 Lie algebras and linear groups

Why introduce the matrix exponential? It will give us a way to show linear groups are Lie groups and to relate Lie groups and Lie algebras. We will proceed as follows.

Definition 2.16. Given a linear group G , define its Lie algebra to be the tangent space to G at I :

$$\begin{aligned}\mathfrak{lg} &:= T_I G \subset \mathfrak{gl}_n(\mathbb{R}), \\ &= \{\dot{A}(0) \mid A : (-\varepsilon, \varepsilon) \rightarrow G \text{ smooth with } A(0) = I\}.\end{aligned}$$

Remark 2.17. Strictly speaking, until we know that G is a smooth manifold, it doesn't make sense to define something as the tangent space, but since linear groups embed into $\mathrm{GL}_n(\mathbb{R}) \subset \mathbb{R}^{n^2}$, we can still define tangent vectors as the derivative of smooth curved at I . Thus we obtain a subspace of $\mathbb{R}^{n^2} \cong \mathfrak{gl}_n(\mathbb{R})$.

In order to reduce the number of limits we will need to write down, we introduce *little o-notation*. Namely, we say that

$$f(t) = g(t) + o(t)$$

near 0 if

$$\lim_{t \rightarrow 0} \frac{f(t) - g(t)}{t} = 0.$$

We should think of this as saying that $f(t)$ and $g(t)$ “agree to first order”.

Proposition 2.18. The set $\mathfrak{lg}$ is a Lie algebra when equipped with the commutator.

Proof. Let us first check that it is a vector space. Clearly $0 \in \mathfrak{lg}$, as this is the derivative of the constant map $A(t) = I$.

If $x \in \mathfrak{lg}$, then $x = \dot{A}(0)$ for some $A : (-\varepsilon, \varepsilon) \rightarrow G$, i.e.

$$A(t) = I + tx + o(t).$$

For non-zero $\lambda \in \mathbb{R}$, define $A_\lambda : (-\varepsilon/\lambda, \varepsilon/\lambda) \rightarrow G$ by $A_\lambda(t) := A(\lambda t)$. Via chain rule

$$\dot{A}_\lambda(0) = \lambda \dot{A}(0) = \lambda x$$

and so $\lambda x \in \mathfrak{lg}$.

If $x, y \in \mathfrak{lg}$ are tangent vectors to $A(t)$ and $B(t)$ respectively, then we can consider the smooth curve $C(t) = A(t)B(t)$. By product rule

$$\begin{aligned}\dot{C}(0) &= \dot{A}(0)B(0) + A(0)\dot{B}(0) \\ &= xI + Iy = x + y\end{aligned}$$

and so $x + y \in \mathfrak{lg}$. Thus $\mathfrak{lg}$ is a vector space. To see that it is a Lie algebra, we consider $x, y \in \mathfrak{lg}$ tangent to $A(t), B(t)$ respectively. Consider

$$C_s(t) = A(s)B(t)A^{-1}(s).$$

We have that $C_s(0) = I$, so $\dot{C}_s(0) \in \mathfrak{lg}$ for all small s . As such, for all small, non-zero s

$$\frac{1}{s}A(s)yA^{-1}(s) \in \mathfrak{lg}.$$

Since we have shown $\mathfrak{lg}$ to be a vector subspace of $\mathfrak{gl}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$, it is a closed subset. Thus

$$\lim_{s \rightarrow 0} \frac{A(s)yA^{-1}(s)}{s} \in \mathfrak{lg}.$$

This is just the derivative at $s = 0$, so we compute

$$A'(0)yA^{-1}(0) + A(0)y(A^{-1})'(0) = xy - yx \in \mathfrak{lg}.$$

Thus, we get a Lie algebra. $\square$

We can give a very explicit description of this Lie algebra in term of the exponential. Define

$$L(G) = \{x \in \mathfrak{gl}_n(\mathbb{R}) \mid e^{tx} \in G \text{ for all } t \in \mathbb{R}\}$$

Proposition 2.19. *For any linear group G with Lie algebra $\mathfrak{lg}$, we have $L(G) = \mathfrak{lg}$.*

Proof. Clearly $A(t) = e^{tx}$ defines a smooth map to G with $A(0) = I$. As

$$A(t) = I + tx + o(t)$$

we have $\dot{A}(0) = x$. Hence $L(G) \subset \mathfrak{lg}$. To see that the reverse inclusion holds, suppose we have a smooth path A such that $A(0) = I$, $\dot{A}(0) = x$.

Define a sequence

$$\begin{aligned} x_n &= \frac{1}{n} \log \left(A \left(\frac{1}{n} \right) \right) \\ &= \frac{A \left(\frac{1}{n} \right) - I}{n} \left(I - \frac{A \left(\frac{1}{n} \right) - I}{2} + \frac{(A \left(\frac{1}{n} \right) - I)^2}{3} - \dots \right) \end{aligned}$$

As $n \rightarrow \infty$, $A \left(\frac{1}{n} \right) \rightarrow I$. Since the infinite series is bounded by a geometric series, we can take limits of each term separately:

$$\begin{aligned} \lim_{n \rightarrow \infty} x_n &= \lim_{n \rightarrow \infty} \left(\frac{A \left(\frac{1}{n} \right) - I}{n} \right) \lim_{n \rightarrow \infty} \left(I - \frac{A \left(\frac{1}{n} \right) - I}{2} + \frac{(A \left(\frac{1}{n} \right) - I)^2}{3} - \dots \right) \\ &= \lim_{n \rightarrow \infty} \left(\frac{A \left(\frac{1}{n} \right) - I}{n} \right) I \\ &= \lim_{n \rightarrow \infty} \left(\frac{I + \frac{x}{n} + o \left(\frac{1}{n} \right) - I}{n} \right) \\ &= x. \end{aligned}$$

As $\exp$ is continuous,

$$\lim_{n \rightarrow \infty} e^{x_n} = e^{\lim_{n \rightarrow \infty} x_n} = e^x.$$

Since $e^{x_n} \in G$, by definition, and G is closed, we have $e^x \in G$. This holds for all $x \in \mathfrak{lg}$ and in particular for tx for all $t \in \mathbb{R}$. Thus $x \in L(G)$. $\square$

This gives an effective way to determine the Lie algebras of many linear groups.

Example 2.20. As e^x is invertible for all $x \in \mathfrak{gl}_n(\mathbb{R})$, $L(\mathrm{GL}_n(\mathbb{R})) = \mathfrak{gl}_n(\mathbb{R})$.

Example 2.21. We claim $L(\mathrm{SL}_n(\mathbb{R})) = \mathfrak{sl}_n(\mathbb{R})$. Suppose that $e^{tx} \in \mathrm{SL}_n(\mathbb{R})$ for all $t \in \mathbb{R}$. This implies

$$\det e^{tx} = 1 \text{ for all } t \in \mathbb{R}$$

and hence

$$e^{t \operatorname{tr} x} = 1 \text{ for all } t \in \mathbb{R}.$$

This holds if and only if $\operatorname{tr} x = 0 \Leftrightarrow x \in \mathfrak{sl}_n(\mathbb{R})$.

Example 2.22. Let's determine $L(\mathrm{O}_n(\mathbb{R}))$. Suppose $x \in L(\mathrm{O}_n(\mathbb{R}))$. This implies

$$(e^{tx})^T e^{tx} = I$$

for all $t \in \mathbb{R}$. As $(e^{tx})^T = e^{tx^T}$, we can take the derivative of this equality to find that

$$x^T e^{tx^T} e^{tx} + e^{tx^T} e^{tx} x = 0$$

for all $t \in \mathbb{R}$. In particular, if $t = 0$, we find $x^T + x = 0$. Thus $L(\mathrm{O}_n(\mathbb{R})) \subset \mathfrak{o}_n(\mathbb{R})$.

To see that these are equal, note that if $x^T + x = 0$, then x^T and x commute, Hence

$$(e^{tx})^T e^{tx} = e^{tx^T} e^{tx} = e^{tx^T + tx} = e^0 = I$$

and so $x \in L(\mathrm{O}_n(\mathbb{R}))$. Thus $L(\mathrm{O}_n(\mathbb{R})) = \mathfrak{o}_n(\mathbb{R})$.

But $\mathfrak{o}_n(\mathbb{R}) = \mathfrak{so}_n(\mathbb{R})$. So what is $L(\mathrm{SO}_n(\mathbb{R}))$? As $\mathrm{SO}_n(\mathbb{R}) \subset \mathrm{O}_n(\mathbb{R})$, we must have

$$L(\mathrm{SO}_n(\mathbb{R})) \subset L(\mathrm{O}_n(\mathbb{R}))$$

But note also that, since $\det(\exp(x)) > 0$ for all $x \in \mathfrak{gl}_n(\mathbb{R})$, we have that

$$\det(\exp(tx)) = 1 \text{ for all } x \in \mathfrak{o}_n(\mathbb{R})$$

(since $\det M = \pm 1$ for all $M \in \mathrm{O}_n(\mathbb{R})$). Hence $e^{tx} \in \mathrm{SO}_n(\mathbb{R})$ for all $t \in \mathbb{R}$ and $x \in \mathfrak{o}_n(\mathbb{R})$. Thus

$$\mathfrak{so}_n(\mathbb{R}) = L(\mathrm{SO}_n(\mathbb{R})) = L(\mathrm{O}_n(\mathbb{R})) = \mathfrak{o}_n(\mathbb{R}).$$

Lie algebras detect a lot about their groups, but not everything!

Example 2.23. Consider the linear group $(\mathbb{R}, +)$, viewed as a linear group in $\mathrm{GL}_2(\mathbb{R})$. If $e^{tx} = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$, then x must have eigenvalues 0. Hence x is conjugate to $\begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix}$. It is then not difficult to convince yourself that x must be equal to a matrix of this form, and hence $L(\mathbb{R}) = \mathbb{R}$.

We could also have concluded this geometrically. The tangent space is 1-dimensional, and $\mathbb{R}$ is the only 1-dimensional Lie algebra. Similarly, we can conclude that $L(S^1) = \mathbb{R}$, as it must be 1-dimensional.

These instances of linear groups with the same Lie algebra relates to the fact that $\exp$ only gives us a local homeomorphism rather than a global one. Nevertheless, we still see that linear groups with the same Lie algebra have some relationship.

Let us now show that linear groups are Lie groups. It is clear that the maps will be smooth, as they are essentially polynomials, so it is enough to show that every linear group is a smooth manifold. For this, we need the following lemma.

Lemma 2.24. *Let G be a linear group, $A_n \in G$, and $\alpha_n \in \mathbb{R}$ sequences such that*

$$\lim_{n \rightarrow \infty} A_n = I, \quad \lim_{n \rightarrow \infty} \alpha_n(A_n - I) = x.$$

Then $x \in \mathfrak{g}$.

Proof. Similarly to the proof of Proposition 2.19, we have that

$$\lim_{n \rightarrow \infty} t\alpha_n \log(A_n) = \lim_{n \rightarrow \infty} t\alpha_n(A_n - I) = tx.$$

Thus, if $t\alpha_n \in \mathbb{Z}$ for every n , we could say that

$$e^{tx} = \lim_{n \rightarrow \infty} (A_n)^{t\alpha_n}.$$

If $t\alpha_n \in \mathbb{Z}$, then the right hand side is an element of G for every n , and so by the closed-ness of G , e^{tx} would be in G and the claim would follow.

We claim we can replace $t\alpha_n$ with a sequence of integers with the limiting properties. Let $a_n = \lfloor t\alpha_n \rfloor$. We claim

$$\lim_{n \rightarrow \infty} a_n(A_n - I) = \lim_{n \rightarrow \infty} t\alpha_n(A_n - I)$$

as then the claim would follow.

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n(A_n - I) &= \lim_{n \rightarrow \infty} (a_n - t\alpha_n + t\alpha_n)(A_n - I) \\ &= \lim_{n \rightarrow \infty} (a_n - t\alpha_n) + \lim_{n \rightarrow \infty} (t\alpha_n)(A_n - I) \\ &= 0 + tx = tx \end{aligned}$$

as $|a_n - t\alpha_n| < 1$. □

Proposition 2.25. *The map $\exp : \mathfrak{g} = L(G) \rightarrow G$ is a local homeomorphism.*

Proof. From the proof of Theorem 2.15, for all small enough $\varepsilon > 0$, $\exp$ is a local homeomorphism from

$$U_\varepsilon = \{x \in \mathfrak{gl}_n(\mathbb{R}) \mid \|x\| < \varepsilon\}$$

to $V_\varepsilon = \exp(U_\varepsilon)$, with inverse given by the $\log$. Hence, we get a continuous map

$$\exp : U_\varepsilon \cap L(G) \rightarrow V_\varepsilon \cap G$$

which almost has an inverse

$$\log : V_\varepsilon \cap G \rightarrow U_\varepsilon.$$

It would therefore be enough to show that

$$\log(V_\varepsilon \cap G) \subset L(G)$$

for some $\varepsilon > 0$. Suppose no such ε exists. Then we can find

$$g_n \in V_{1/n} \cap G$$

such that $\log(g_n) \notin L(G)$ for all sufficiently large n .

Results from linear algebra tell us that we can choose a splitting

$$\mathfrak{gl}_n(\mathbb{R}) = L(G) \oplus Q$$

and hence write

$$\log(g_n) = x_n + y_n, \quad x_n \in L(G), \quad y_n \in Q,$$

with both $x_n \rightarrow 0$ and $y_n \rightarrow 0$. Since $\log(g_n) \notin L(G)$ for all (sufficiently large) n , we must have $y_n \neq 0$. Thus, we obtain a well defined sequence of norm 1 $\left\{ \frac{y_n}{\|y_n\|} \right\} \subset Q$. By Bolzano-Weierstrass, this has a convergent subsequence, converging to some $y \in Q$ of norm 1. We will show such a y cannot exist, and hence conclude that $y_n = 0$ for some n , giving the desired result.

Let $h_n = e^{-x_n} e^{x_n + y_n} \in G$. Expanding the second factor in terms of series, we find

$$\begin{aligned} h_n &= e^{-x_n} \left(I + x_n + y_n + \frac{(x_n + y_n)^2}{2} + \dots \right) \\ &= e^{-x_n} \left(I + x_n + y_n + \frac{x_n^2}{2} + \frac{x_n y_n + y_n x_n + y_n^2}{2} + \dots \right) \\ &= e^{-x_n} (e^{x_n} + y_n + \dots) \\ &= I + e^{-x_n} y_n + e^{-x_n} S \end{aligned}$$

where S is an absolutely convergent series, whose terms are bounded above in norm by $\|y_n\|$ multiplied by something tending to 0. Hence

$$\lim_{n \rightarrow 0} \frac{h_n - I}{\|y_n\|} = e^{-x_n} \frac{y_n}{\|y_n\|} + e^{-x_n} \frac{S}{\|y_n\|} = y$$

But, by the previous lemma, we must then have $y \in L(G)$. But $L(G) \cap Q = 0$, so $y = 0$, providing the needed contradiction. $\square$

Corollary 2.26. *Every linear group G is a Lie group.*

Proof. As remarked earlier, it is enough to show that G is a smooth manifold. From Proposition 2.25, we have an open set V around I that is homeomorphic to some $U \subset L(G) \subset \mathbb{R}^{n^2}$. Both $\exp$ and $\log$ can be shown to be smooth maps, and so $(V, \log)$ provides a chart around I .

Note also that, for any $g \in G$, multiplication by g is a smooth map, and a homeomorphism (with inverse given by multiplication by g^{-1}). Hence

$$V_g = \{gh \mid h \in V\}$$

is an open set around g , and the map

$$\begin{aligned} \log_g : V_g &\rightarrow U \\ h &\mapsto \log(g^{-1}h) \end{aligned}$$

is a smooth map. Thus, we obtain a chart around every point. To see that this defines an atlas (up to some finiteness conditions), it is enough to check that

$$\log_{g_2} \circ \log_{g_1}^{-1} : U \rightarrow U$$

is a smooth map. This follows as

$$\log_{g_2} \circ \log_{g_1}^{-1}(x) = \log(g_2^{-1}g_1 \exp(x))$$

is a composition of smooth maps. $\square$

Remark 2.27. This also gives a sleek way to show $\mathfrak{g} = L(G)$. The map $\exp : L(G) \rightarrow G$ is a smooth local homeomorphism, and hence induces an isomorphism of tangent spaces. As $L(G)$ is a vector space,

$$L(G) = T_0 L(G) = T_0 U \cong T_I V = T_I G.$$

2.3 Lie group homomorphisms and the Lie correspondence.

Definition 2.28. A map of Lie groups $f : G \rightarrow H$ is called a Lie group homomorphism if it is both a group homomorphism and a smooth map of manifolds.

Example 2.29. The adjoint representation of G

$$\begin{aligned} \text{Ad} : G &\rightarrow \text{GL}(\mathfrak{g}), \\ g &\mapsto \text{Ad}_g \end{aligned}$$

defined by $\text{Ad}_g(x) = gxg^{-1}$ is a Lie group homomorphism. To see that it is well defined, note that if $x \in \mathfrak{g}$, then

$$e^{tgxg^{-1}} = ge^{tx}g^{-1} \in G \text{ for all } t \in \mathbb{R}.$$

Thus $\text{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}$. It is clearly linear and invertible, so $\text{Ad}_g \in \text{GL}(\mathfrak{g})$. We have that

$$\text{Ad}_{gh}(x) = ghxh^{-1}g^{-1} = \text{Ad}_g \text{Ad}_h(x)$$

for all $x \in \mathfrak{g}$, so it is a homomorphism. Smoothness follows easily.

Theorem 2.30. *Let G, H be linear groups. Then, given any Lie group homomorphism $f : G \rightarrow H$, there exists a unique Lie algebra homomorphism*

$$df : \mathfrak{g} \rightarrow \mathfrak{h}$$

called the differential, such that

$$e^{df(x)} = f(e^x).$$

Proof. To see that this map is unique, suppose we have a linear map $F : \mathfrak{g} \rightarrow \mathfrak{h}$ such that

$$e^{F(x)} = f(e^x)$$

for all $x \in \mathfrak{g}$. Differentiating $e^{tF(x)} = f(e^{tx})$, and evaluating at $t = 0$, we see that

$$F(x) = \frac{d}{dt} (f(e^{tx}))|_{t=0} = J(f)_I x$$

and hence $F(x)$ is uniquely determined if it exists. Thus, we just need to show that

$$df(x) := \frac{d}{dt} (f(e^{tx}))|_{t=0} = J(f)_I x$$

is a Lie algebra homomorphism satisfying the given property.

We first note that, by chain rule, we can compute $df(x)$ as

$$df(x) := \frac{d}{dt} (f(A(t)))|_{t=0}$$

for any smooth path through I such that $\dot{A}(0) = x$.

To see that $df(x) \in \mathfrak{h}$, note that $f(e^{tx})$ defines a smooth path in H , and so $\frac{d}{dt} (f(e^{tx}))|_{t=0} \in \mathfrak{h}$.

To see that the map is linear, consider

$$\begin{aligned} df(x+y) &= \frac{d}{dt} (f(e^{tx} e^{ty}))|_{t=0} \\ &= \frac{d}{dt} (f(e^{tx}) f(e^{ty}))|_{t=0} \\ &= \frac{d}{dt} (f(e^{tx}))|_{t=0} f(I) + f(I) \frac{d}{dt} (f(e^{ty}))|_{t=0} \\ &= df(x) + df(y). \end{aligned}$$

We handle scalar multiplication similarly.

To see that df is a Lie algebra homomorphism, we note that

$$e^{tgxg^{-1}} = ge^{tx}g^{-1} \quad \Rightarrow \quad df(gxg^{-1}) = f(g)df(x)f(g)^{-1}.$$

Next, we note that

$$e^{tx}ye^{-tx} = y + t(xy - yx) + o(t)$$

so

$$[x, y] = \frac{d}{dt} (e^{tx} y e^{-tx})|_{t=0}.$$

As differentiation commutes with linear maps

$$df([x, y]) = \frac{d}{dt} df(e^{tx} y e^{-tx})|_{t=0} = \frac{d}{dt} (f(e^{tx}) df(y) f(e^{-tx}))|_{t=0}.$$

Computing the derivative, we find

$$df([x, y]) = [df(x), df(y)].$$

Finally, to see that df satisfies its defining property, let $y = df(x)$. The paths

$$P(t) = e^{ty}, \quad \text{and} \quad Q(t) = f(e^{tx})$$

in H have the same derivative at 0. We claim they are equal.

As $\exp((s+t)u) = \exp(su)\exp(tu)$, the paths satisfy

$$P(s+t) = P(s)P(t), \quad Q(s+t) = Q(s)Q(t)$$

Differentiating with respect to s and evaluating at $s = 0$, we find

$$\dot{P}(t) = yP(t), \quad \dot{Q}(t) = yQ(t).$$

This gives a differential equation with a unique solution given an initial condition. Since $P(0) = Q(0) = I$, we must therefore have $P(t) = Q(t)$. Taking $t = 1$ gives the desired result. $\square$

Remark 2.31. *The argument in the final part of the proof is a statement about the uniqueness of a 1-parameter subgroup, which is the image $\mathbb{R}$ under a Lie group homomorphism. The notion of a 1-parameter subgroup gives us a way to extend the notion of exponentials in most of these arguments to general Lie groups.*

Example 2.32. *The Lie algebra homomorphism associated to $\text{Ad} : G \rightarrow \text{GL}(\mathfrak{g})$ is $\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$. We compute this as follows. For $x, y \in \mathfrak{g}$,*

$$\begin{aligned} df(x)(y) &= \frac{d}{dt} \Big|_t = 0 (\text{Ad}_{e^{tx}}(y)) \\ &= \frac{d}{dt} \Big|_t = 0 (e^{tx} y e^{-tx}) \\ &= [x, y] = \text{ad}_x(y). \end{aligned}$$

Remark 2.33. *A fun corollary of this is that, since $\text{Ad}_{e^x} = e^{\text{ad}_x}$, we have*

$$e^x y e^{-x} = y + [x, y] + \frac{1}{2}[x, [x, y]] + \cdots$$

This generalises further to the Baker-Campbell-Hausdorff formula, which says that

$$\log(e^x e^y) = x + y + \frac{1}{2}[x, y] + \cdots$$

is given by a series in brackets of x, y , modulo some convergence conditions.

Corollary 2.34. *Every close Lie subgroup $H \subset G$ induces a Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$.*

Proof. It is easy to check that if f is injective, then df is injective. Indeed, if $df(x) = 0$, then

$$f(e^{tx}) = e^{t df(x)} = e^0 = I$$

If f is injective, then $e^{tx} = I$ for all $t \in \mathbb{R}$. Taking the derivative at $t = 0$ implies $x = 0$.

Applying Theorem 2.30 to the inclusion $H \hookrightarrow G$ gives the claim. $\square$

Corollary 2.35. *If $G \cong H$ as Lie groups, then $\mathfrak{g} \cong \mathfrak{h}$ as Lie algebras*

Proof. By chain rule $d(f \circ g) = df \circ dg$. Hence, if $f \circ g = id$, $df \circ dg = id$. $\square$

There is a partial converse to this theorem, called the Lie group - Lie algebra correspondence. It consists of three theorems, including Lie's third theorem, which we summarise into a single result here.

Theorem 2.36. • *Let G, H be Lie groups with corresponding Lie algebras $\mathfrak{g}, \mathfrak{h}$. If G is connected and simply connected, there is a bijection between Lie group homomorphisms $G \rightarrow H$ and Lie algebra homomorphisms $\mathfrak{g} \rightarrow \mathfrak{h}$.*

- *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then there is a bijection between Lie subalgebras of $\mathfrak{g}$ and (not necessarily closed) connected Lie subgroups of G .*
- *There is a bijection between isomorphism classes of connected, simply connected Lie groups and finite dimensional Lie algebras.*

In particular, this means that for topologically “nice” Lie groups, we can understand their action on a vector space V via an action of the Lie algebra

$$\rho : G \rightarrow \mathrm{GL}(V) \rightarrow d\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V).$$

The latter is generally much easier to understand, as we can often break the Lie algebra down into simpler parts.

3 Representations of Lie algebras

Definition 3.1. *Let L be a Lie algebra. An L -representation is a pair (V, ρ) consisting of a (usually finite dimensional) K -vector space V and a Lie algebra homomorphism $\rho : L \rightarrow \mathfrak{gl}(V)$.*

Remark 3.2. *We can also define an L -module, to be a vector space V equipped with an action of L , given by a bilinear map*

$$\begin{aligned} L \times V &\rightarrow V \\ (x, v) &\mapsto x \cdot v \end{aligned}$$

such that

$$[x, y]v = x \cdot (y \cdot v) - y \cdot (x \cdot v).$$

This is equivalent to the structure of a representation of L on V , and can be notationally more convenient. As such, we will be lazy in referring to V as an L -representation, but using module notation, etc.

Example 3.3. • A vector space V comes with a trivial L -action: $x \cdot v := 0$ for all $x \in L$ and $v \in V$.

- V is a representation of $\mathfrak{gl}(V)$ in the obvious way, called the standard representation of $\mathfrak{gl}(V)$.
- L is an L -representation via the adjoint action $x \cdot y = [x, y]$.

Many properties of vector spaces hold for representations as well, with most results holding under some additional assumptions. In particular, we can define a notion of a quotient representation.

Definition 3.4. Let (V, ρ) be an L -representation. A subspace $U \subset V$ is called a subrepresentation if $(U, \rho|_U)$ is an L -representation:

$$\rho(x)u \in U \text{ for all } x \in L, u \in U.$$

The quotient vector space V/U of an L -representation V by a subrepresentation U is an L -representation in the obvious way

$$x \cdot (v + U) = x \cdot v + U.$$

Example 3.5. Given L -representations V, W , the vector space $\text{Hom}_K(V, W)$ of linear maps is an L -representation via

$$(x \cdot f)(v) = x \cdot (f(v)) - f(x \cdot v).$$

In particular, the dual vector space

$$V^* = \text{Hom}_K(V, K)$$

is an L -representation with $(x \cdot f)(v) = -f(x \cdot v)$.

Example 3.6. Given two L -representations V, W , their direct sum is an L -representation:

$$x \cdot (v, w) := (x \cdot v, x \cdot w).$$

Example 3.7. Given two L -representations V, W , we define their tensor product $V \otimes W$ to be the vector space spanned by symbols

$$\{v \otimes w \mid v \in V, w \in W\}$$

subject to the relations

$$\begin{aligned}(v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w \\ v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2 \\ (\lambda v) \otimes w &= v \otimes (\lambda w) = \lambda(v \otimes w).\end{aligned}$$

This is an L -representation

$$x \cdot (v \otimes w) = (x \cdot v) \otimes w + v \otimes (x \cdot w).$$

Definition 3.8. A linear map $f : V \rightarrow W$ between L -representations is called a homomorphism of L -representations if

$$f(x \cdot v) = x \cdot f(v)$$

for all $x \in L, v \in V$. We call f an isomorphism if it is invertible.

Lemma 3.9. A linear map $f : V \rightarrow W$ is a homomorphism if and only if $x \cdot f = 0$ for all $x \in L$.

Proof. We have that

$$\begin{aligned}(x \cdot f)(v) &= 0 \Leftrightarrow x \cdot f(v) - f(x \cdot v) = 0 \\ &\Leftrightarrow x \cdot f(v) = f(x \cdot v).\end{aligned}$$

Thus $x \cdot f = 0$ for all $x \in L$ if and only if $x \cdot f(v) = f(x \cdot v)$ for all $x \in L$ and $v \in V$. $\square$

Such f are called invariant under the action of L . More generally we give the following definition.

Definition 3.10. Let (V, ρ) be an L -representation. We call $v \in V$ L -invariant if $x \cdot v = 0$ for all $x \in L$.

This definition may be counter-intuitive, but comes naturally from viewing L -actions as “infinitesimal” actions of a Lie group. Namely, if $L = L(G)$ for a Lie group G acting on V , then we would call $v \in V$ G -invariant if $g \cdot v = v$ for all $g \in G$. Hence

$$e^{tx} \cdot v = v$$

for all $t \in \mathbb{R}$ and $x \in L(G)$. Differentiating this, we find exactly $x \cdot v = 0$.

Our goal for much of the rest of the course is to describe a classification of “nice” Lie algebras via their representations, and further classifying the representations of such Lie algebras.

4 Nilpotent and solvable Lie algebras

4.1 Nilpotent Lie algebras

Definition 4.1. Let L be a finite dimensional Lie algebra. We define the central series to be the sequence of ideals

$$L = L^0 \supset L^1 \supset L^2 \subset \dots$$

given by

$$L^{i+1} = [L, L^i].$$

We call L nilpotent if $L^n = 0$ for some $n \geq 0$.

Remark 4.2. This is equivalent to saying that, if we take enough Lie brackets, we must obtain 0:

$$[x_1, [x_2, [\dots, [x_{n-1}, x_n]] \dots]] = 0 \text{ for all } x_1, \dots, x_n \in L.$$

This relates to the usual notion of a linear operator being nilpotent if $M^n = 0$ for some n , as this vanishing implies that the $(n-1)$ -fold composition

$$\text{ad}_{x_1} \circ \text{ad}_{x_2} \circ \dots \circ \text{ad}_{x_{n-1}} = 0$$

Example 4.3. The Lie algebra of strictly upper triangular matrices is nilpotent. Suppose $x, y \in \mathfrak{n}_n$ and that $y_{i,j} = 0$ for all $j - i \leq k$. (Note that this is true with $k = 0$ for all elements of $\mathfrak{n}_n$.)

Then we claim that

$$[x, y]_{i,j} = 0 \text{ for all } j - i \leq k + 1.$$

Indeed, we have

$$\begin{aligned} (xy)_{i,j} - (yx)_{i,j} &= \sum_{q=1}^n x_{i,q} y_{q,j} - \sum_{q=1}^n y_{i,q} x_{q,j} \\ &= \sum_{q=i+1}^{j-k} x_{i,q} y_{q,j} - \sum_{q=i+k}^{j-1} y_{i,q} x_{q,j}. \end{aligned}$$

If $j - i < k + 1$, then $i + 1 > j - k$ and $j - 1 < i + k$, so both of these sums are empty. Thus $[x, y]_{i,j} = 0$ for $j - i < k + 1$. As such, an iterated bracket satisfies

$$[x_1, [x_2, [\dots, [x_{k-1}, x_k] \dots]]_{i,j} = 0$$

for $j - i \leq k - 1$. In particular, if $k = n$, then every iterated bracket vanishes. Hence $\mathfrak{n}_n^n = 0$.

The Lie algebra of strictly upper triangular matrices is the key example of a nilpotent Lie algebra, in a way made precise in the following.

Theorem 4.4 (Engel). *Let L be a finite dimensional Lie algebra, and (V, ρ) be a finite dimensional non-zero L -representation such that $\rho(x)$ is a nilpotent operator for all $x \in L$. Then:*

1. *There is a non-zero $v \in V$ such that $\rho(x)v = 0$ for all $x \in L$, i.e. the set of L -invariants is non-empty;*
2. *There is a basis of V such that all $\rho(x)$ are strictly upper triangular in this basis;*
3. *There exists a sequence of subrepresentations*

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

such that $\rho(L) \cdot V_k \subset V_{k-1}$.

In particular, this holds for all finite dimensional non-zero representations of a nilpotent Lie algebra.

Remark 4.5. *In all the follows, we will assume we are considering finite dimensional Lie algebras and representations, but will not explicitly state this for sake of brevity.*

We first need a lemma.

Lemma 4.6. *If $x \in \mathfrak{gl}(V)$ is a nilpotent linear operator, then $\text{ad}_x : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(V)$ is nilpotent.*

Proof. We introduce linear operators $L_x, R_x : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(V)$ defined by

$$L_x(y) := xy, \quad \text{and} \quad R_x(y) := yx$$

and note that

$$\text{ad}_x = L_x - R_x, \quad \text{and} \quad L_x \circ R_x = R_x \circ L_x.$$

Hence

$$\text{ad}_x^n = (L_x - R_x)^n = \sum_{k=0}^n \binom{n}{k} L_x^k \circ R_x^{n-k}$$

for all $n \geq 0$. Suppose that $x^N = 0$. Then $L_x^N = R_x^N = 0$. As every summand in ad_x^{2N} contains at least one of L_x^N or R_x^N , we therefore have $\text{ad}_x^{2N} = 0$. $\square$

We can now prove Theorem 4.4.

Proof. 1. We will induct on $\dim \rho(L)$. Since the theorem only concerns the image of L , we may assume that $L \subset \mathfrak{gl}(V)$, suppress all reference to ρ , and assume x is nilpotent for all $x \in L$.

If $\dim L = 1$, then $L = K \cdot x$ with $x^n = 0$ for some $n > 0$. Taking n to be minimal, and $v_0 \neq 0$, then $v = x^{n-1}v_0$ provides the desired v .

Now suppose $\dim L \geq 2$ and let I be a maximal proper Lie subalgebra. We want to emulate an argument from elementary linear algebra and consider L/I , but we need I to be an ideal in order for the argument to work.

In order to show that I is an ideal, consider the normaliser of I

$$N = \{x \in L \mid [x, I] \subset I\}$$

Clearly $I \subset N \subset L$. It is easy to see that N is a vector subspace. The Jacobi identity implies that

$$[[x, y], i] = [x, [y, i]] - [y, [x, i]] \in N$$

for all $x, y \in N$, so N is a Lie subalgebra. Hence $N = I$ or $N = L$. By Lemma 4.6, I acts nilpotently on L via the adjoint action. As I is a Lie subalgebra, this action descends to the (vector space) quotient L/I . By induction, there is therefore a non-zero $x + I \in L/I$ such that

$$\text{ad}_i(x + I) = 0 + I$$

for all $i \in I$. This gives an $x \notin I$ such that $\text{ad}_i(x) \in I$ for all $i \in I$ and hence $[x, I] \subset I$. Therefore $I \not\ni x \in N$ and so $N = L$. We conclude I is an ideal.

Knowing that I is an ideal, we have that L/I is a Lie algebra. In order to apply induction, we need a representation of L/I , which V is not. Consider instead

$$V^I = \{v \in V \mid Iv = 0\}.$$

By the induction hypothesis, this is non-zero and defines an L -subrepresentation: if $v \in V^I$ and $x \in L$, then

$$\begin{aligned} i \cdot (x \cdot v) &= x \cdot (i \cdot v) - [x, i] \cdot v \\ &= 0 - 0 = 0. \end{aligned}$$

Thus, as $IV^I = 0$, we have that L/I acts on V^I . But $\dim(L/I) < \dim L$, and so there exists $v \neq 0$, $v \in V^I$ such that $(L/I)v = 0$.

As $Iv = 0$, we therefore have a $v \in V^I \subset V$ such that $Lv = 0$ as needed.

2. Take v_1 to be the vector obtained from part (1). Kv_1 is a subrepresentation as $Lv_1 = 0$. Hence V/Kv_1 is an L -representation of lower dimension. Inducting on $\dim V$, we obtain a basis $\tilde{v}_2, \dots, \tilde{v}_n$ of V/Kv_1 such that all the matrices of all $x \in L$ are simultaneously strictly upper triangular. Pickling lifts $v_2, \dots, v_n \in V$ of this basis, we obtain a basis $v_1, \dots, v_n$ of V in which all $x \in L$ are strictly upper triangular.
3. Let $v_1, \dots, v_n$ be the basis from part (2). Taking $V_1 = Kv_1$, $V_2 = Kv_1 \oplus Kv_2$, ..., we obtain the desired sequence.

□

While every representation of a nilpotent Lie algebra will be nilpotent and satisfy the above theorem, it is not usually enough to find a nilpotent representation of a Lie algebra L to say anything about the nilpotence of L . We do, however, have a partial converse.

Theorem 4.7 (Engel). *A Lie algebra L is nilpotent if and only if $\text{ad}_x \in \mathfrak{gl}(L)$ is nilpotent for all $x \in L$.*

Proof. If L is nilpotent, $\text{ad}_x \in \mathfrak{gl}(L)$ is nilpotent by direct observation. To see the converse, suppose ad_x is nilpotent for all $x \in L$. By Theorem 4.4 applied to the adjoint representation $\rho : L \rightarrow \mathfrak{gl}(V)$, we can find

$$0 = V_0 \subset V_1 \subset \cdots \subset V_n = L$$

such that $\text{ad}_x(V_{k+1}) = V_k$ for all $1 \leq k < n$ and all $x \in L$. Hence

$$\begin{aligned} L^1 &= [L, L] \subset [L, V_n] \subset V_{n-1} \\ L^2 &= [L, L^1] \subset [L, V_{n-1}] \subset V_{n-2} \\ &\vdots \\ L^n &= [L, L^{n-1}] \subset [L, V_1] \subset V_0 = 0. \end{aligned}$$

Therefore $L^n = 0$ and L is nilpotent. $\square$

Remark 4.8. *Note that Engel's theorem really do rely on the nilpotence of $\rho(L)$, not just of a generating set. Indeed, $\mathfrak{sl}_2(\mathbb{R})$ is generated by nilpotent matrices $x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, but not all elements are nilpotent and it is impossible to simultaneously upper-triangularise x and y .*

4.2 Solvable Lie algebras

Nilpotent Lie algebras are particularly nice, as we essentially only have to deal with “small” Lie brackets. The next step up in generality are solvable Lie algebras.

Definition 4.9. *Let L be a finite dimensional Lie algebra, and define its derived series to be the sequence of ideals*

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

given by

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

We call L solvable if $L^{(n)} = 0$ for some $n \geq 0$. If $I \subset L$ is an ideal that is solvable as a Lie algebra, we call it a solvable ideal.

Remark 4.10. *Similarly to how Galois theory relates solvable finite groups to solvability of polynomials by radicals, solvable Lie algebras can be related to the solvability of PDEs in explicit terms via a similar framework.*

Lemma 4.11. *If L is nilpotent, then L is solvable.*

Proof. We have that $L^{(0)} = L^0$. Suppose that for some $k \geq 0$, we have $L^{(k)} \subset L^k$. Then

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

Thus, if $L^n = 0$, $L^{(n)} = 0$. □

Example 4.12. *The Lie algebra $L = \mathfrak{b}_n$ of upper triangular matrices is solvable, as $L^{(1)}$ consists of strictly upper triangular matrices, which we know to be nilpotent. Via the same argument as in the above lemma, $L^{(n+1)} = 0$. This is the key example of a solvable Lie algebra.*

Solvability is a property that is somehow “additive” in a sense made precise by the following lemma.

Lemma 4.13. *Let L be a finite dimensional Lie algebra. Then*

1. *If L is solvable, then I and L/I are solvable for any ideal $I \subset L$,*
2. *If $I \subset L$ is a solvable ideal and the quotient L/I is solvable, then L is solvable,*
3. *If $I, J \subset L$ are solvable ideals, then so is the ideal $I + J$.*

Proof. 1. We have that $I^{(k)} \subset L^{(k)}$ for all $k \geq 0$, so if $L^{(n)} = 0$ then $I^{(n)} = 0$. Thus I is solvable. For the quotient, let $\pi : L \rightarrow L/I$ be the natural projection map. Note that $\pi(L^{(0)}) = (L/I)^{(0)}$. We claim $\pi(L^{(k)}) = (L/I)^{(k)}$. Suppose this holds for some $k \geq 0$. Then

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

Thus if $L^{(n)} = 0$, $(L/I)^{(n)} = 0$.

2. Since I and L/I are solvable, there exist $m, n \geq 0$ such that $I^{(m)} = 0$ and $(L/I)^{(n)} = 0$. As before $\pi(L^{(n)}) = (L/I)^{(n)} = 0$, and so $L^{(n)} \subset I$. Hence

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

Iterating this, we find

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

and so L is solvable.

3. This can be shown directly, but it is easiest to invoke the 2nd isomorphism theorem:

$$(I + J)/I \cong J/(I \cap J).$$

As J is solvable, $J/(I \cap J)$ is solvable. Thus $(I + J)/I$ is solvable. As I is solvable, $I + J$ is therefore solvable.

□

Corollary 4.14. *The radical of L is solvable. In fact, it is the maximal solvable ideal of L .*

$$\text{rad } L := \sum_{\substack{I \subset L \\ I \text{ solvable}}} I$$

Exercise 4.15. *Let $f : L \rightarrow M$ be a Lie homomorphism. Show that $f(\text{rad}(L)) \subset \text{rad}(M)$. Show also that $\text{rad}(L_1 \oplus L_2) = \text{rad}(L_1) \oplus \text{rad}(L_2)$.*

As with nilpotent Lie algebra, the representations of a solvable Lie algebra have a particularly nice description, at least if K is algebraically closed of characteristic 0.

Lemma 4.16. *Let L be a Lie algebra over a field of characteristic 0, $I \subset L$ an ideal, and (V, ρ) an L -representation. Then, for any linear map $\lambda : L \rightarrow K$, the space*

$$V_\lambda = \{v \in V \mid \rho(y)v = \lambda_y v \text{ for all } y \in I\}$$

is a (possible 0) L -subrepresentation of V .

Proof. If $V_\lambda = 0$, there is nothing to prove, so we suppose otherwise. Suppressing ρ we suppose we have $x \in L$ and $w \in V_\lambda$. We want to show that

$$y \cdot (x \cdot w) = \lambda_y x \cdot w.$$

Note that

$$\begin{aligned} y \cdot (x \cdot w) &= (yx) \cdot w \\ &= x \cdot (y \cdot w) - [x, y] \cdot w \\ &= x \cdot (\lambda_y w) - \lambda_{[x, y]} w \\ &= \lambda_y x \cdot w - \lambda_{[x, y]} w. \end{aligned}$$

It would therefore suffice to show that $\lambda_{[x, y]} = 0$ for all $x \in L$ and $y \in I$.

Our approach to this is to construct a subspace $U \subset V$ such that elements of I act on U by matrices of the form

$$y|_U = \begin{pmatrix} \lambda_y & \bullet & \cdots & \bullet \\ 0 & \lambda_y & \cdots & \bullet \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_y \end{pmatrix}$$

so that we can compute λ_y via a trace.

Take some non-zero $w_1 \in V_\lambda$, and let $w_{k+1} = xw_k$. We claim $U = \text{Span}(w_1, w_2, \dots)$ is such a subspace. We have $y \cdot w_1 = \lambda_y w_1$, $y \cdot w_2 = \lambda_y w_2 - \lambda_{[x,y]} w_1$, and more generally suppose

$$y \cdot w_k - \lambda_y w_k + \sum_{i=1}^{k-1} a_i w_i.$$

for all $y \in I$. Then

$$\begin{aligned} y \cdot w_{k+1} &= x \cdot (\lambda_y w_k) + \sum_{i=1}^{k-1} a_i w_i - [x, y] \cdot w_k \\ &= \lambda_y w_{k+1} + \sum_{i=1}^k b_i w_k \end{aligned}$$

for some b_i . Thus, I acts as described on U . As $U \subset V$, it is a finite dimensional space, and we must have

$$(\dim U) \lambda_{[x,y]} = \text{tr}(\rho([x, y])|_U) = \text{tr}(\rho(x)|_U \rho(y)|_U) - \text{tr}(\rho(y)|_U \rho(x)|_U) = 0.$$

Thus $\lambda_{[x,y]} = 0$ as needed. $\square$

Theorem 4.17 (Lie). *Let L be a solvable, non-zero Lie algebra over an algebraically closed field K of characteristic 0, and let (V, ρ) is a finite dimensional, non-zero L -representation. Then*

1. *There exists a non-zero $v \in V$ such that v is a common eigenvector for all $\rho(x)$, $x \in L$. That is, $\rho(L)v \subset Kv$.*
2. *There exists a basis of V such that all $\rho(x)$ are simulatenously upper triangular in this basis.*
3. *There is a chain of subrepresentations*

$$0 = V_0 \subset V_1 \subset \dots \subset V_n = V$$

$$\text{such that } \dim V_k = 0 \Leftrightarrow \dim(V_k/V_{k-1}) = 1.$$

Proof. 1. As L is solvable and non-zero, we must have

$$L = L^{(0)} \supsetneq L^{(1)} \supsetneq \dots \supsetneq L^{(n)} = 0$$

with strict inclusions. As $[L, L] = L^{(1)} \subsetneq L$, and L is finite dimensional, we can find vectors $x_1, \dots, x_m \notin [L, L]$ such that

$$L = Kx_1 \oplus Kx_2 \oplus \dots \oplus Kx_m \oplus [L, L].$$

Let $I = Kx_2 \oplus \dots \oplus Kx_m \oplus [L, L]$. This is an ideal by construction. Indeed, $[L, L]$ is an ideal, and $[x_i, x_j] \in [L, L]$. By induction on $\dim L$, we can find nonzero $v \in V$ such that $\rho(x)v = \lambda_x v$ for some $\lambda_x \in K$ for all $x \in I$.

By construction, we must have that $x \mapsto \lambda_x$ is a linear map. By Lemma 4.16, V_λ is a non-zero subrepresentation of L and so $\rho(x_1)V_\lambda \subset V_\lambda$.

As K is algebraically closed, $\rho(x_1)$ has an eigenvector $v_0 \in V_\lambda$. But v_0 is a common eigenvector for I , and hence v_0 is a common eigenvector for all of L .

2. Let v_1 be a common eigenvector and consider the L -representation V/Kv_1 . By induction, we can find a basis of V/Kv_1 such that L acts via upper triangular matrices. Lifting this basis to V and adding v_1 , we obtain a basis of V such that $\rho(x)$ is upper triangular for all $x \in L$.
3. Let $v_1, \dots, v_n$ be a basis such that $\rho(x)$ is upper triangular for all $x \in L$. Then taking $V_k = Kv_1 \oplus \dots \oplus Kv_k$ gives such a chain of subrepresentations. $\square$

We do not have a direct analogue to the partial converse in the nilpotent setting, but something similar is given by the following corollary.

Corollary 4.18. *Let L be a Lie algebra over a field K of characteristic 0. Then L is solvable if and only if $[L, L]$ is nilpotent.*

Proof. It is clear that if $[L, L]$ is nilpotent, then L is solvable. Conversely, if L is solvable, then we can find a basis of L , possibly over an algebraic closure of K , such that

$$\text{ad} : L \rightarrow \mathfrak{gl}(L)$$

is given by upper triangular matrices. Thus

$$\text{ad} : [L, L] \rightarrow \mathfrak{gl}([L, L]) \subset \mathfrak{gl}(L)$$

is given by strictly upper triangular matrices, and hence is nilpotent. Therefore, $[L, L]$ is nilpotent. $\square$

Solvable Lie algebras play a substantial role in the classification of Lie algebras, due to the *Levi decomposition*. This is a result which, under mild conditions, tells us that every finite dimensional Lie algebra is a semidirect product of a solvable Lie algebra, and what is called a semisimple Lie algebra. While we will not be able to prove this result, we will see a similar result at the level of representations and will study semisimple Lie algebras in some detail.

5 Semisimple Lie algebras

Definition 5.1. *Let L be a (finite dimensional) Lie algebra. We call L*

- Simple if L is not abelian and L has no non-zero proper ideals,
- Semisimple if L contains no non-zero abelian ideals,
- Reductive if $Z(L) = \text{rad}(L)$, i.e. the maximal abelian ideal is equal to the maximal solvable ideal.

Note that $Z(L)$ is abelian, and hence solvable, so $Z(L) \subset \text{rad}(L)$ for all L .

Lemma 5.2. *The following are equivalent:*

- i) L is semisimple,
- ii) L contains no non-zero solvable ideals,
- iii) $\text{rad}(L) = 0$.

Proof. The equivalence of (ii) and (iii) is clear. To see that (i) and (ii) are equivalent, note that any abelian ideal is solvable, and any solvable ideal contains an abelian one. Indeed, if $I^{(n)} = 0$ with n minimal, then $[I^{(n-1)}, I^{(n-1)}] = 0$, so $I^{(n-1)}$ is non-zero and abelian. $\square$

Corollary 5.3. *Every simple Lie algebra is semisimple. Every semisimple Lie algebra is reductive.*

Example 5.4. • We have seen that $\mathfrak{sl}_2(\mathbb{R})$ is simple!

- If we define $L = \mathfrak{sl}_2(\mathbb{R}) \oplus \mathfrak{sl}_2(\mathbb{R})$ with Lie bracket determined by

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

it is not too hard to show that the only ideals are 0 , $\mathfrak{sl}_2(\mathbb{R}) \oplus 0$, $0 \oplus \mathfrak{sl}_2(\mathbb{R})$ and L . Hence. L is not simple. But only 0 is abelian, so L is semisimple.

- $\mathfrak{gl}_2(K)$ is reductive. We know that

$$Z(\mathfrak{gl}_2(K)) = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \mid a \in K \right\}$$

As every matrix can be written as

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + M$$

with $\text{tr}(M) = 0$. Thus, we have

$$\mathfrak{gl}_2(K) = Z(\mathfrak{gl}_2(K)) \oplus \mathfrak{sl}_2(K)$$

as vector spaces. Since $[Z(\mathfrak{gl}_2(K)), \mathfrak{sl}_2(K)] = 0$, this is a decomposition into Lie subalgebras, and in fact ideals.

Hence, any ideal of $\mathfrak{gl}_2(K)$ can be uniquely written as a sum of an ideal of $Z(\mathfrak{gl}_2(K))$ and of $\mathfrak{sl}_2(K)$. Thus

$$\begin{aligned} \text{rad}(\mathfrak{gl}_2(K)) &= \text{rad}(Z(\mathfrak{gl}_2(K))) \oplus \text{rad}(\mathfrak{sl}_2(K)) \\ &= Z(\mathfrak{gl}_2(K)). \end{aligned}$$

Exercise 5.5. • IF I is a solvable ideal, show that $\text{rad}(L)/I \cong \text{rad}(L/I)$. Hence conclude that $\mathfrak{gl}_2(K)$ is reductive.

- Show that if L is reductive, then $L = Z(L) \oplus \tilde{L}$ for a semisimple Lie algebra $\tilde{L}$.

5.1 Decomposing representations

Definition 5.6. Let L be a finite dimensional Lie algebra and (V, ρ) an L -representation. Then we call (V, ρ) :

- i) A faithful representation if ρ is injective,
- ii) An irreducible representation if it has no non-zero proper subrepresentations. We also call V a simple L -module in this case,
- iii) A completely reducible representation if it is a sum of irreducible representations. We also call V a semisimple L -module in this case.

Remark 5.7. For sake of brevity, it is common to refer to an irreducible representation as an irrep.

Example 5.8. Let I be an ideal of L . By definition, I is a subrepresentation of (L, ad) . Hence, (L, ad) is irreducible if and only if L is a simple Lie algebra. We will later see that L is semisimple as a Lie algebra if and only if (L, ad) is completely reducible.

Lemma 5.9. Let L be a Lie algebra and (V, ρ) be an L -representation. Then the following are equivalent:

- i) V is a sum of irreducible representations,
- ii) V is a direct sum of irreducible representations,
- iii) Any subrepresentation $W \subset V$ has a complement $U \subset V$ such that $V = U \oplus W$.

Proof. Let us show (i) implies (ii) first. Let $V = \sum_{i=1}^n U_i$ with U_i irreducible. Assume that n is minimal. We want to show that $U_i \cap U_j = 0$ for $i \neq j$. We can induct on n to assume that

$$W = \sum_{i=1}^{n-1} U_i = U_1 \oplus U_2 \oplus \cdots \oplus U_{n-1}$$

is a direct sum. As U_n is irreducible, and $W \cap U_n$ is a subrepresentation of U_n , it must be either 0 and we are done, or $W \cap U_n = U_n$, in which case

$$V = W = U_1 \oplus \cdots \oplus U_{n-1}$$

contradicting the minimality of n . Hence $V = U_1 \oplus \cdots \oplus U_n$.

Now suppose (ii) holds, and let $V = U_1 \oplus \cdots \oplus U_n$. Suppose $W \subset V$ is a subrepresentation. Consider the projection

$$\pi : V \rightarrow \tilde{V} := U_1 \oplus \cdots \oplus U_{n-1}.$$

Inducting on n , we must have that $\pi(W)$ has a complement in $\tilde{V}$ and so

$$\tilde{V} = \pi(W) \oplus \tilde{N}.$$

As before, $W \cap U_n$ must either be 0 or U_n . If $W \cap U_n = 0$, then $W \cong \pi(W)$, and so

$$V \cong W \oplus \tilde{N} \oplus \ker \pi = W \oplus (\tilde{N} \oplus \ker \pi).$$

If $W \cap U_n = U_n$, then $\ker \pi = U_n \subset W$ and so

$$V \cong \pi(W) \oplus \tilde{N} \oplus \ker \pi \cong W \oplus \tilde{N}.$$

Now suppose (iii) holds. Either V is irreducible and we are done, or V has a subrepresentation W with complement U . Inducting on the dimension of our representations, we can assume W and U are sums of irreducibles and hence V is a sum of irreducibles. $\square$

Corollary 5.10. *Subrepresentations and quotients of completely reducible representations are completely reducible.*

Example 5.11. *Consider the Lie subalgebra $L = \mathbb{R} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \subset \mathfrak{gl}_2(\mathbb{R})$. Then $\mathbb{R}^2$ is a representation of L , with a subrepresentation $\mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. However, this subrepresentation cannot have a complement, as $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has no other eigenvectors.*

In contrast, if we consider the 1-dimensional Lie subalgebra spanned by a symmetric matrix, then $\mathbb{R}^2$ is completely reducible, as we can diagonalise a real symmetric matrix.

Lemma 5.12 (Schur). *Let V be an irreducible representation of L over an algebraically closed field K . Then, if $f : V \rightarrow V$ is a homomorphism of L -representations, $f(v) = \lambda v$ for all $v \in V$. That is, $\text{End}_L(V) = K \text{id}$.*

More generally, if V, W are irreducible representations of L over any field K and $f : V \rightarrow W$ is a homomorphism, then f is either an isomorphism or the zero map.

Proof. Let λ be an eigenvalue of f . Then

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

is a subrepresentation:

$$f(x \cdot v) = x \cdot f(v) = x \cdot (\lambda v) = \lambda(x \cdot v).$$

It is non-empty, as f must have an eigenvector, and hence $V_\lambda = V$.

For a map $f : V \rightarrow W$, note that $\text{im } f \subset W$ is a subrepresentation and $\ker f \subset V$ is a subrepresentation. Hence either $\text{im } f = 0$, $\ker f = V$ and f is the zero map; or $\ker f = 0$, $\text{im } f = W$ and f is an isomorphism. $\square$

Remark 5.13. *Schur's Lemma implies that all maps between completely reducible representations are essentially described by the decomposition into irreducible representations. Indeed, every map can be decomposed into a combination of rescalings (if K is algebraically closed) and a permutation of copies of each irreducible.*

5.2 Decomposing Lie algebras

A powerful tool for decomposing vector spaces into subspaces in a systematic way was the notion of an inner product. A similar decomposition can be obtained in the theory of Lie algebras, but with some additional constraints.

Definition 5.14. Let $\langle -, - \rangle : L \times L \rightarrow K$ be a bilinear form. We call the map

- i) Symmetric if $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in L$,
- ii) Non-degenerate if $\langle x, L \rangle = 0$ if and only if $x = 0$,
- iii) L -invariant if

$$\langle [x, y], z \rangle = \langle x, [y, z] \rangle$$

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

Remark 5.15. A bilinear form is L -invariant if and only if

$$\langle -, - \rangle : L \otimes L \rightarrow K$$

is a map of L -representations, where $L \otimes L$ is equipped with the tensor product of adjoint representations and K is the trivial representation.

An L -invariant bilinear form is precisely the structure we need to decompose a Lie algebra in a sensible way and to define an orthogonal complement.

Lemma 5.16. Let $\langle -, - \rangle : L \times L \rightarrow K$ be an L -invariant bilinear form. Then for every ideal $I \subset L$, the orthogonal complement

$$I^\perp := \{x \in L \mid \langle x, I \rangle = 0\}$$

is again an ideal.

Proof. Standard linear algebra ensures that $I^\perp$ is a vector subspace, so we just need to check that $[y, x] \in I^\perp$ for all $x \in I^\perp$, $y \in L$. Namely, we need to show

$$\langle [y, x], z \rangle = 0 \text{ for all } z \in I.$$

We have

$$\begin{aligned} \langle [y, x], z \rangle &= -\langle [x, y], z \rangle \\ &= -\langle x, [y, z] \rangle = 0 \end{aligned}$$

as $x \in I^\perp$ and $[y, z] \in I$. □

An easy source of symmetric, L -invariant (but possibly degenerate) bilinear forms comes from the representations of L .

Lemma 5.17. Let (V, ρ) be an L -representation. Then the bilinear form

$$\langle x, y \rangle_\rho := \text{tr}_V(\rho(x), \rho(y))$$

is symmetric and L -invariant.

Proof. Symmetry follows from the symmetry of trace. To see L -invariance, note that

$$\begin{aligned}
\langle [x, y], z \rangle_\rho &= \text{tr}_V([x, y]z) \\
&= \text{tr}_V(xyz - yxz) \\
&= \text{tr}_V(xyz) - \text{tr}_V(yxz) \\
&= \text{tr}_V(xyz) - \text{tr}_V(xzy) \\
&= \text{tr}_V(x[y, z]) = \langle x, [y, z] \rangle_\rho
\end{aligned}$$

where we have suppressed ρ in our computations. $\square$

However, this is usually degenerate, and so there exist elements orthogonal to themselves. In fact, non-degeneracy of such a bilinear form implies L is reductive!

Theorem 5.18. *Let L be a Lie algebra over a field of characteristic 0, and let (V, ρ) be an L -representation. If $\langle -, - \rangle_\rho$ is non-degenerate, then L is reductive.*

Proof. We will again suppress ρ from our computations, treating V as an L -module. We need to show that $\text{rad}(L) \subset Z(L)$. Equivalently, we want to show $[\text{rad}(L), L] = 0$.

Under the assumption of non-degeneracy, $[\text{rad}(L), L] = 0$ if and only if

$$\langle [I, L], L \rangle_\rho = 0.$$

Since $\text{tr}_V = \text{tr}_U + \text{tr}_{V/U}$ for any subrepresentation U , it would be enough to show

$$\langle [I, L], L \rangle_{(U, \rho)} = \langle [I, L], L \rangle_{(V/U, \rho)} = 0$$

By iterating this, it would be enough to show

$$\langle [I, L], L \rangle_{(W, \rho_W)}$$

for all W such that W has no non-zero proper subrepresentation, i.e. an irreducible representation. As such, we will assume (V, ρ) is irreducible.

Considering V as a representation of $\text{rad}(L)$ over the algebraic closure of our field, Theorem 4.17, there exists $v_0 \in V$ such that

$$x \cdot v_0 = \lambda_x v_0$$

for all $x \in \text{rad}(L)$, with $\lambda : \text{rad}(L) \rightarrow K$ a linear map. We know that V_λ is a subrepresentation of V and is non-zero, so $V = V_\lambda$. Thus, for all $v \in V$, $x \cdot v = \lambda_x v$ for all $x \in \text{rad}(L)$.

Hence, for all $x \in \text{rad}(L)$, $y, z \in L$ and $v \in V$, we have

$$\begin{aligned}
[x, y]zv &= xyzv - yxzv \\
&= \lambda_x yzv - y \cdot (\lambda_x zv) \\
&= 0.
\end{aligned}$$

Hence $\langle [x, y], z \rangle_\rho = \text{tr}_V([x, y]z) = 0$ for all $x \in \text{rad}(L)$, $y, z \in L$, from which the claim follows. $\square$

Remark 5.19. If $\rho(x) = 0$, then $\langle x, y \rangle_\rho = 0$ for all $y \in L$. Thus, if $\langle -, - \rangle_\rho$ is non-degenerate, $\rho(x) = 0$ if and only if $x = 0$. If we get a non-degenerate bilinear form from a representation, it must therefore be a faithful representation.

Corollary 5.20. For K a field of characteristic 0, the Lie algebras $\mathfrak{gl}_n(K)$ and $\mathfrak{u}(n)$ are reductive.

Proof. Consider first the standard representation of $\mathfrak{gl}_n(K)$ on $V = K^n$. Then

$$\langle x, y \rangle_\rho = \text{tr}(xy) = \sum_{i,j=1}^n x_{i,j} y_{j,i}.$$

By taking $y = E_{j_0, i_0}$, we see that

$$\langle x, y \rangle_\rho = 0 \quad \Leftrightarrow \quad x_{i_0, j_0} = 0.$$

and so $\langle x, y \rangle_\rho = 0$ for all y implies $x = 0$. Thus, we have a non-degenerate bilinear form and so $\mathfrak{gl}_n(K)$ is reductive.

To handle the unitary Lie algebra, we again consider the standard representation of $\mathfrak{u}(n)$ on $V = \mathbb{C}^n$, viewed as a $2n$ -dimensional $\mathbb{R}$ -vector space. By taking an appropriate basis, the complex matrix $A + iB$ corresponds to the real matrix

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

and so

$$\text{tr}_V(A + iB) = 2 \text{tr}_{\mathbb{R}^n}(A).$$

We have that

$$\text{tr}_V((x + iy)(z + iw)) = \text{tr}_V((xz - yw) + i(yz + xw)) = 2 \text{tr}_{\mathbb{R}^n}(xz - yw).$$

By carefully choosing z and w , as for $\mathfrak{gl}_n(K)$, we can again show that the bilinear form is non-degenerate. Hence $\mathfrak{u}(n)$ is reductive. $\square$

As remarked earlier, representations detecting reductivity are necessarily faithful. We have a partial converse, giving a way to detect both reductivity and semisimplicity of a Lie algebra.

Theorem 5.21. Let L be a Lie algebra over an algebraically closed field K of characteristic 0, with a faithful irreducible representation (V, ρ) . Then L is reductive and $\dim Z(L) \leq 1$. If $\rho(L) \subset \mathfrak{sl}(V)$, then L is semisimple.

Proof. As ρ is injective, we can treat L as a subalgebra of $\mathfrak{gl}(V)$. As before, we can find a linear map

$$\lambda : \text{rad}(L) \rightarrow K$$

such that

$$V_\lambda = \{v \mid xv = \lambda_x v \text{ for all } x \in \text{rad}(L)\}$$

is equal to V (viewed as a representation over K). Thus, $x \in \text{rad}(L)$ acts on V by $\lambda_x \text{id}_V$ and so

$$\text{rad}(L) \subset K \text{id}$$

As every such matrix commutes with everything in $\mathfrak{gl}(V)$, we have $\text{rad}(L) \subset Z(L)$ and hence $\text{rad}(L) = Z(L)$. We conclude L is reductive, and

$$\dim(Z(L)) \leq \dim K \text{id} = 1.$$

If $L \subset \mathfrak{sl}(V)$, and $\text{rad}(L) \neq 0$, we must have

$$\text{id} \in \text{rad}(L) \subset L \subset \mathfrak{sl}(V).$$

But $\text{tr}(\text{id}) \neq 0$, so this is impossible. Hence, if $L \subset \mathfrak{sl}(V)$, we must have $\text{rad}(L) = 0$ and L is semisimple. $\square$

Corollary 5.22. *Over an algebraically closed field of characteristic 0, $\dim Z(\mathfrak{gl}_n(K)) \leq 1$.*

Corollary 5.23. *The Lie algebras $\mathfrak{sl}_n(K)$, $\mathfrak{so}_n(K)$, $\mathfrak{sp}_{2n}$, $\mathfrak{su}(n)$ are semisimple.*

Proof. These all admit faithful representations whose image lies in $\mathfrak{sl}(V)$, just by taking the standard representations. The claim then follows if these are irreducible. We will show this for $\mathfrak{sl}(K)$ and leave the remainder as an exercise.

Suppose

$$v = c_1 e_1 + \cdots c_n e_n \in W \subset K^n$$

is an element of a subrepresentation of K^n , and suppose $c_m \neq 0$. Then

$$\frac{1}{c_m} E_{i,m} v = e_i \quad \text{and} \quad \frac{1}{c_m} E_{i,m} \in \mathfrak{sl}_n(K)$$

if $i \neq m$. Thus, $e_i \in W$ for some i , and hence $e_k = E_{k,i} e_i \in W$ for each $k \neq i$. Hence $W = K^n$ as it contains every basis vector. $\square$

Remark 5.24. *We can also show this by writing $\mathfrak{gl}_n(K) = \text{rad}(\mathfrak{gl}_n(K)) \oplus \mathfrak{sl}_n(K)$. Hence*

$$\text{rad}(\mathfrak{sl}_n(K)) \cong \text{rad}(\mathfrak{gl}_n(K)) / \text{rad}(\text{rad}(\mathfrak{gl}_n(K))) \cong 0.$$

By restricting to appropriate subspaces, this can also be used to show the other cases.

5.3 Cartan's criterion and the Killing form

We can also use bilinear forms to detect solvability of a Lie algebra, assuming we have a nice representation, via a result due to Cartan. As in the previous section, we will assume our base field K is algebraically closed of characteristic 0 throughout.

Theorem 5.25 (Cartan). *Let L be a finite dimensional Lie algebra and (V, ρ) a faithful representation. Then L is solvable if and only if*

$$\langle [x, y], z \rangle_\rho = 0 \text{ for all } x, y, z \in L.$$

In order to prove this, we need a lemma. We will not prove this lemma currently, returning to it much later, but will need it to give a criterion for nilpotence.

Lemma 5.26. *Let $P \subsetneq Q \subset \mathfrak{gl}(V)$ be vector subspaces, and let*

$$M = \{x \in \mathfrak{gl}(V) \mid [x, Q] \subset P\}$$

Then, if $x \in M$ satisfies $\text{tr}(xy) = 0$ for all $y \in M$, x is a nilpotent operator.

We can now prove Cartan's theorem.

Proof. As ρ is injective, we identify L with its image in $\mathfrak{gl}(V)$. We first suppose that L is solvable. We can then find a basis in which L is upper triangular and so $[x, y]$ is strictly upper triangular for all $x, y \in L$.

This implies that $[x, y]z$ is strictly upper triangular for all $x, y, z \in L$, and so

$$\langle [x, y], z \rangle_\rho = \text{tr}([x, y]z) = 0$$

as claimed.

Conversely, suppose $\langle [L, L], L \rangle = 0$. To show that L is solvable, it is enough to show that $[L, L]$ is nilpotent. By Theorem 4.7, it is enough to show that ad_x is nilpotent for $x \in [L, L]$, which in turn would follow from x being nilpotent as an operator.

We apply Lemma 5.26 with $P = [L, L]$, $Q = L$, and

$$M = \{x \in \mathfrak{gl}(V) \mid [x, L] \subset [L, L]\}.$$

As $P \subset M$, to show $x \in [L, L]$ is nilpotent, it is enough to show $\text{tr}(xy) = 0$ for all $y \in M$. Writing

$$x = \sum [x_i, \tilde{x}_i]$$

as a sum of commutators, the L -invariance of our bilinear form implies

$$\text{tr}([x_i, \tilde{x}_i]y) = \text{tr}(x_i[\tilde{x}_i, y]).$$

for all $y \in M$. But $x_i \in L$ is upper triangular, and $[\tilde{x}_i, y] \in [L, L]$ is strictly upper triangular, so this trace is 0. Hence $\text{tr}(xy) = 0$ for all $y \in M$, and x is nilpotent. $\square$

Unfortunately, it can be difficult to find a faithful representation. Fortunately, much like the converse to Engel's theorem, the adjoint representation is good enough to act as a stand-in

Definition 5.27. Define the Killing form of L to be the bilinear L -invariant form associated to the adjoint representation

$$\kappa_L(x, y) = \langle x, y \rangle_{\text{ad}} = \text{tr}(\text{ad}_x \circ \text{ad}_y)$$

Lemma 5.28. Let $I \subset L$ be an ideal. Then the Killing form is the restriction of the Killing form on L : $\kappa_I = \kappa_L|_I$.

Proof. Note that if $x \in I$, then $\text{ad}_x(L) \subset I$. Hence, if we choose a basis $x_1, \dots, x_n$ of L such that $x_1, \dots, x_k$ is a basis of I , the matrix of ad_x in this basis is 0 below the k^{th} row. Hence, we must have

$$\text{tr}_L(\text{ad}_x) = \text{tr}_I(\text{ad}_x|_I)$$

Similarly, as $\text{ad}_x \circ \text{ad}_y : L \rightarrow I$, this is given by a matrix of the same form and hence the claim follows. $\square$

Theorem 5.29 (Cartan's criterion). A Lie algebra L is solvable if and only if $\kappa_L([L, L], L) = 0$.

Proof. If L is solvable, then $\text{ad}(L) \subset \mathfrak{gl}(V)$ is solvable. Hence, applying Theorem 5.25 to the faithful representation V of $\text{ad}(L)$, we have

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

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

Conversely, if $\kappa_L([L, L], L) = 0$, then

$$\text{tr}([\text{ad}_x, \text{ad}_y] \text{ad}_z) = 0$$

for all $x, y, z \in L$. Considering V as a representation of $L/\ker(\text{ad})$, we have a faithful representation of $L/\ker(\text{ad})$ such that the associated bilinear form vanishes as in Theorem 5.25. Thus $L/\ker(\text{ad})$ is solvable.

As $\ker(\text{ad}) = Z(L)$ is abelian, it is also solvable, and therefore L is solvable. $\square$

5.4 Semisimplicity and the Killing form

In addition to detecting solvability, the Killing form can also be used to identify semisimple Lie algebras, and plays a key role in describing their structure.

Theorem 5.30. Over an algebraically closed field of characteristic 0, the following are equivalent:

- i) L is semisimple,
- ii) The Killing form κ_L is non-degenerate,
- iii) L is a direct sum of ideals that are simple as Lie algebras,
- iv) L is a direct sum of simple Lie algebras.

Proof. Clearly (iii) implies (iv). Let us first show (i) $\Rightarrow$ (iii). If L is simple, we are done. Otherwise, suppose $I \subset L$ is a non-trivial ideal. By Lemma 5.16,

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

is also an ideal. As $I \cap I^\perp$ is therefore an ideal, we have

$$\kappa_{I \cap I^\perp} = \kappa_L|_{I \cap I^\perp} = 0$$

by definition of $I^\perp$. By Theorem 5.29, $I \cap I^\perp$ is solvable, and hence 0 if L is semisimple. Thus $L = I \oplus I^\perp$. Iterating the argument with each component gives (iii).

Now suppose (iv) holds, and write $L = L_1 \oplus L_2 \oplus \cdots \oplus L_n$ as a sum of simple Lie algebras. If $I \subset L$ is an abelian ideal, then $I \cap L_i$ is an abelian ideal of L_i . But L_i is simple, so $I \cap L_i = 0$. Hence $I = 0$ and L contains no non-trivial abelian ideals. Thus L is semisimple: (i).

It remains to show that (i) $\Leftrightarrow$ (ii). If L is semisimple, then we consider the ideal

$$L^\perp = \{x \in L \mid \kappa_L(x, L) = 0\}.$$

By definition, $\kappa_{L^\perp} = \kappa_L|_{L^\perp} = 0$, and so $L^\perp$ is solvable. As L is semisimple, this means $L^\perp = 0$, and so κ_L is non-degenerate.

If κ_L is non-degenerate, then Theorem 5.18 implies L is reductive. Hence $\text{rad}(L) = Z(L)$. But if $x \in \text{rad}(L) = Z(L)$, then $\text{ad}_x(L) = 0$ and so

$$\kappa_L(x, y) = \text{tr}(\text{ad}_x \circ \text{ad}_y) = \text{tr}(0) = 0$$

for all $y \in L$. By nondegeneracy, $x = 0$, and so $\text{rad}(L) = 0$. Thus L is semisimple. $\square$

Between this, Engel's theorem, and Cartan's criterion, we see that the Killing form and the adjoint representation capture much of the structure of L . As such, the following is not surprising.

Theorem 5.31. *Let L be a finite dimensional semisimple Lie algebra over an algebraically closed field of characteristic 0. Then*

$$\text{ad} : L \rightarrow \text{Der}(L) \subset \mathfrak{gl}(L)$$

is an isomorphism.

Proof. As L is semisimple

$$\ker(\text{ad}) = Z(L) = \text{rad}(L) = 0$$

and so ad is injective. We just need to show it is surjective onto derivations.

Note that if $\partial \in \text{Der}(L)$, then

$$\begin{aligned} [\text{ad}_x, \partial](y) &= \text{ad}_x(\partial(y)) - \partial(\text{ad}_x(y)) \\ &= [x, \partial(y)] - \partial([x, y]) \\ &= [x, \partial(y)] - [\partial(x), y] - [x, \partial(y)] \\ &= [\partial(x), y] = -\text{ad}_{\partial(x)}(y) \end{aligned}$$

for all $y \in L$. Thus $[\text{ad}_x, \partial] = -\text{ad}_{\partial(x)}$ for all $\partial \in \text{Der}(L)$ and $\text{ad}(L)$ is an ideal of $\text{Der}(L)$. Taking the orthogonal complement with respect to the Killing form, $\text{ad}(L)^\perp$ is also an ideal of $\text{Der}(L)$.

As $\text{ad}(L) \cong L$ is semisimple,

$$\kappa(\text{ad}(L)) = \kappa_{\text{Der}(L)}|_{\text{ad}(L)}$$

is non-degenerate. In particular

$$\text{ad}(L) \cap \text{ad}(L)^\perp = 0 \quad \Rightarrow \quad \text{Der}(L) = \text{ad}(L) \oplus \text{ad}(L)^\perp.$$

We claim $\text{ad}(L)^\perp = 0$. Suppose $\partial \in \text{ad}(L)^\perp$. Then

$$-\text{ad}_{\partial(x)} = -[\text{ad}_x, \partial] \in \text{ad}(L) \cap \text{ad}(L)^\perp = 0.$$

But ad is injective, so $\partial(x) = 0$, for all $x \in L$. Thus $\partial = 0$. Hence

$$\text{Der}(L) = \text{ad}(L) \oplus \text{ad}(L)^\perp = \text{ad}(L) \oplus 0 = \text{ad}(L).$$

□

In particular, this means that any semisimple Lie algebra has a faithful representation and can be written in terms of matrices!

6 Representations of semisimple Lie algebras

The representation theory of semisimple Lie algebras is particularly nice, as we obtain complete reducibility in all cases. Again, we will assume our field K is algebraically closed of characteristic 0.

Theorem 6.1 (Weyl). *Every representation of a semisimple Lie algebra is completely reducible.*

This implies that, to understand the representations of a semisimple Lie algebra, it is enough to understand the irreducible representations. This is a much more manageable task, and we can often give a complete description of these. This is *not* the case for solvable Lie algebras, where we cannot reliably break a representation down into simpler components.

However, in order to prove Weyl's theorem, we need a number of auxiliary results and ideas.

Definition 6.2. *Let L be a Lie algebra equipped with a non-degenerate bilinear form, with a basis $x_1, \dots, x_n$. Let $x^1, \dots, x^n$ be the dual basis. Then, given an L -module V , we define the Casimir operator*

$$C : V \rightarrow V$$

$$v \mapsto \sum_{i=1}^n x_i x^i \cdot v$$

Lemma 6.3. *The Casimir operator C is independent of choice of basis, and defines a homomorphism of L -representations.*

Proof. Let $y_1, \dots, y_n$ be another basis with dual basis $y^1, \dots, y^n$. Then, there exists invertible matrices A and B such that

$$x_i = \sum_{j=1}^n A_{i,j} y_j, \quad \text{and} \quad x^i = \sum_{j=1}^n B_{i,j} y^j$$

We must have that $A^{-1} = B^T$, as

$$\delta_{k,\ell} = \langle x_k, x^\ell \rangle = \sum A_{k,i} B_{\ell,j} \langle y_i, y^j \rangle = \sum A_{k,i} B_{\ell,j} \delta_{i,j} = \sum A_{k,i} B_{\ell,i} = (AB^T)_{k,\ell}$$

where $\delta_{k,\ell} = 1$ if $k = \ell$ and 0 otherwise. Then

$$\begin{aligned} \sum_i x_i x^i &= \sum_{i,j,k} A_{i,j} y_j B_{i,k} y^k \\ &= \sum_{j,k} \delta_{j,k} y_j y^k = \sum_j y_j y^j. \end{aligned}$$

To see that we get a homomorphism, let $x \in L$ and

$$[x_k, x] = \sum c_{k,i} x_i, \quad \text{and} \quad [x, x^\ell] = \sum d_{i,\ell} x^i.$$

Note that

$$c_{k,\ell} = \langle [x_k, x], x^\ell \rangle = \langle x_k, [x, x^\ell] \rangle = d_{k,\ell}.$$

Consider, for $x \in L$, $v \in V$:

$$\begin{aligned} C(xv) - xC(v) &= \sum x_i x^i x v - x x_i x^i v \\ &= \sum (x_i x^i x - x x_i x^i) v \\ &= \sum (x_i [x^i, x] + x_i x x^i - [x, x_i] x^i - x_i x x^i) v \\ &= \sum (c_{i,j} x_i x^j - d_{i,j} x_i x^j) v = 0. \end{aligned}$$

Hence $C(xv) = xC(v)$. □

Lemma 6.4. *If L is semisimple, and (V, ρ) is a faithful representation, then $\langle -, - \rangle_\rho$ is non-degenerate, and $\text{tr } C = \dim L$.*

Proof. Let $L^\perp$ be the orthogonal complement of L with respect to the bilinear form. If $L^\perp = 0$, then the form is non-degenerate. We have

$$\langle [L^\perp, L^\perp], L^\perp \rangle_\rho \subset \langle L^\perp, L^\perp \rangle_\rho = 0$$

and hence $L^\perp$ is solvable. As L is semisimple $L^\perp = 0$.

Finally, note that

$$\text{tr}(C) = \text{tr}\left(\sum x_i x^i\right) = \sum \text{tr}(x_i x^i) = \sum \langle x_i x^i \rangle_\rho = \sum 1 = \dim L.$$

□

Remark 6.5. Often the Casimir operator refers specifically to the Killing form for semisimple L , and if you see reference to a Casimir operator without a bilinear form, you should assume the Killing form is in use.

We next need a way to canonically split a representation into (possibly 0) subrepresentations.

Definition 6.6. Given an L -module V , define submodules

$$\begin{aligned} V^L &= \{v \in V \mid Lv = 0\} \\ LV &= \text{Span}\{xv \mid x \in L, v \in V\} \end{aligned}$$

Lemma 6.7. Let L be a semisimple Lie algebra. Then, for representations V, W ,

i) $V = V^L \oplus LV$,

ii) If $f : V \rightarrow W$ is a surjective homomorphism, then $f(V^L) = W^L$.

Proof. i) We identify L with its image in $\mathfrak{gl}(V)$, as we are only concerned with its action on V . If $L = 0$, then $V = V^L$ and $LV = 0$, so the claim is trivial. Otherwise, L is non-zero and semisimple. Via our identification, we can assume the map $\rho : L \rightarrow \mathfrak{gl}(V)$ is injective and so $\langle -, - \rangle_\rho$ is non-degenerate. Let $C : V \rightarrow V$ be the corresponding Casimir operator.

We can choose a basis of V such that C is in Jordan normal form, giving a splitting

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

into generalised eigenspaces. As $C(xv) = xC(v)$, each V_i will be a subrepresentation.

If $k > 1$, then we can induct on $\dim V$ to assume $V_i = V_i^L \oplus LV_i$, giving the desired splitting

$$\begin{aligned} V &= (V_1^L \oplus \cdots \oplus V_k^L) \oplus (LV_1 \oplus \cdots \oplus LV_k) \\ &= V^L \oplus LV. \end{aligned}$$

If $k = 1$, then C has only a single eigenvalue λ , and

$$\dim L = \text{tr } C = \lambda \dim V.$$

In particular, $\lambda \neq 0$, and so C is invertible. Since $CV^L = 0$, we therefore have $V^L = 0$. We also have $V = CV \subset LV$, and so $V = LV = V^L \oplus LV$.

ii) Note that, for all $v \in V^L$, we have

$$xf(v) = f(xv) = f(0) = 0$$

so $f(V^L) \subset W^L$. We also have $f(xv) = xf(v)$, so $f(LV) \subset LW$. As

$$V = V^L \oplus LV, \quad \text{and} \quad W = W^L \oplus LW$$

surjectivity of f implies that $f(V^L) = W^L$ (and $f(LV) = LW$).

□

We can finally prove Weyl's theorem

Theorem (Weyl). *Every representation of a semisimple Lie algebra is completely reducible.*

Proof. It is enough to show that any subrepresentation $W \subset V$ has a complement. For any subrepresentation, we have a surjective linear map

$$\begin{aligned} \text{Hom}(V, W) &\rightarrow \text{Hom}(W, W) \\ f &\mapsto f|_W. \end{aligned}$$

We can check that this is a map of representations:

$$(x \cdot f)|_W(w) = xf(w) - f(xw) = x \circ f|_W(w) - f|_W \circ x(w) = x \cdot f|_W(w).$$

Hence, we have a surjective map

$$\text{Hom}(V, W)^L \rightarrow \text{Hom}(W, W)^L.$$

The first space consists of maps $f : V \rightarrow W$ such that

$$xf(v) - f(xv) = 0 \quad \Leftrightarrow \quad xf(v) = f(xv).$$

Thus $\text{Hom}(V, W)^L$ consists of homomorphisms of representations. The second space similarly consists of homomorphisms $W \rightarrow W$.

By surjectivity, we therefore have a homomorphism $f : V \rightarrow W$ such that $f|_W = \text{id}$. Then $\ker f \subset V$ is a subrepresentation, which we claim is a complement to W . To see that every $v \in V$ is a sum of an element of $\ker f$ and an element of W , write

$$v = (v - f(v)) + f(v).$$

By definition, $f(v) \in W$, and

$$f(v - f(v)) = f(v) - f(f(v)) = f(v) - f(v) = 0$$

as $f|_W = \text{id}$. Thus $v - f(v) \in \ker f$. Hence $v = W + \ker f$.

Now suppose $v \in W \cap \ker f$. Then

$$v = \text{id}(v) = f|_W(v) = f(v) = 0.$$

Therefore $V = W \oplus \ker f$, and so V is completely reducible. □

Corollary 6.8. *The following are equivalent:*

- i) L is reductive,
- ii) The adjoint representation is a direct sum of irreducibles,
- iii) L is a direct sum of $Z(L)$ and a semisimple Lie algebra (as a Lie algebra).

Proof. Assume (i) holds, and consider the adjoint representation $\text{ad} : L \rightarrow \mathfrak{gl}(L)$. As $\ker(\text{ad}) = Z(L) = \text{rad}(L)$, this induces a representation of $L/\text{rad}(L)$, which is semisimple by construction. Thus L splits as a direct sum of irreducible representations of $L/\text{rad}(L)$. As $[\text{rad}(L), L] = 0$ for reductive L , this gives a splitting as a direct sum of irreducible representations of L .

Next assume (ii) holds, and let $L = I_1 \oplus \cdots \oplus I_m$ be the sum of irreducibles. As there are L -representations, they are ideals of L and so $[L, I_k] \subset I_k$. In particular, we must have

$$[I_j, I_k] \subset I_j \cap I_k = 0$$

if $j \neq k$.

If $J \subset I_k$ is an ideal of I_k as a Lie algebra, then

$$[L, J] = [I_k, J] \subset J$$

and so J is a subrepresentation of L . Hence $J = 0$ or $J = I_k$. This implies I_k is either abelian or simple.

If I_k is abelian, then $[L, I_k] = [I_k, I_k] = 0$, and so $I_k \subset Z(L)$. Hence we can write

$$\begin{aligned} L &= I_1 \oplus I_2 \oplus \cdots \oplus I_m \\ &= \left(\bigoplus_{I_k \subset Z(L)} I_k \right) \oplus \left(\bigoplus_{I_k \text{ simple}} I_k \right) \\ &= I_{ab} \oplus I_{ss} \end{aligned}$$

as a direct sum of an abelian Lie algebra and a semisimple one. As $L/I_{ab} \cong I_{ss}$ is semisimple, I_{ab} was the maximal solvable ideal and hence

$$Z(L) \supset I_{ab} = \text{rad}(L) \subset Z(L).$$

Thus $L = Z(L) \oplus I_{ss}$.

Finally assume (iii) holds. Since $L/Z(L)$ is semisimple, $Z(L)$ must be the maximal solvable ideal, and so $Z(L) = \text{rad}(L)$. $\square$

Appendix A: Formal power series

Definition 6.9. A formal power series (in $\mathbb{R}$) is a sequence $\{a_n\}_{n \geq 0}$ of real numbers, represented as an infinite formal sum

$$f(x) = \sum_{n \geq 0} a_n x^n$$

considered without any reference to convergence. The set of formal power series is denoted $\mathbb{R}[[x]]$.

Remark 6.10. The field $\mathbb{R}$ can easily be replaced by $\mathbb{C}$ without any change to the following.

Proposition 6.11. The set of formal power series is a ring with

$$\{a_n\} + \{b_n\} = \{a_n + b_n\}, \quad \{a_n\}\{b_n\} = \left\{ \sum_{i+j=n} a_i b_j \right\}$$

Equivalently

$$\sum_{n \geq 0} a_n x^n + \sum_{n \geq 0} b_n x^n = \sum_{n \geq 0} (a_n + b_n) x^n$$

and

$$\left(\sum_{n \geq 0} a_n x^n \right) \left(\sum_{n \geq 0} b_n x^n \right) = \sum_{n \geq 0} \left(\sum_{i=0}^n a_i b_{n-i} \right) x^n$$

Note that these are precisely the operations obtained by formally adding or multiplying the relevant series. As such, we often treat formal power series as “functions” and use the corresponding notation and terminology. For example, we will refer to a_0 as the constant term of the corresponding series.

Lemma 6.12. Let $f(x) \in \mathbb{R}[[x]]$ be a formal power series with non-zero constant term. Then $f(x)$ has a multiplicative inverse in $\mathbb{R}[[x]]$

Proof. Consider the equality of formal power series

$$f(x)g(x) = 1$$

with $f(x) = \sum_{n \geq 0} a_n x^n$, $g(x) = \sum_{n \geq 0} b_n x^n$. This implies

$$a_0 b_0 = 1, \quad a_0 b_1 + a_1 b_0 = 0, \quad \dots, \quad a_0 b_n + a_1 b_{n-1} + \dots + b_0 a_n = 0.$$

The first of these equations implies that $b_0 = a_0^{-1}$. As $a_0 \neq 0$, we can then solve the second equation for b_1 . More generally, as $a_0 \neq 0$, once we have determined $b_0, b_1, \dots, b_{n-1}$, there is a unique $b_n \in \mathbb{R}$ such that the coefficient of x^n in $f(x)g(x)$ is zero. As such, we can iteratively determine a multiplicative inverse g . $\square$

Most relevant to our purposes is the notation of composition of formal power series.

Definition 6.13. Let $f(x), g(x) \in \mathbb{R}[[x]]$ and suppose $g(x)$ has constant term 0. Then we define their composition to be the formal series obtained by expanding

$$f(g(x)) := \sum_{n \geq 0} a_n g(x)^n$$

Lemma 6.14. The composition $f(g(x))$ is well defined.

Proof. In the context of formal power series, “well defined” just means that the coefficient of each x^n is finite. As $g(x) = b_1x + b_2x^2 + \dots$, we have that

$$g(x)^n = x^n(b_1^n + nb_1^{n-1}b_2x + \dots)$$

involves only terms of order x^n or higher. As such, the coefficient of x^n in $f(g(x))$ involves only contributions from the coefficient of x^n in

$$a_0, a_1g(x), \dots, a_ng(x)^n.$$

Hence, it is a finite sum and the composition is well defined. To low order, it is given by

$$f(g(x)) = a_0 + a_1b_1x + (a_1b_2 + a_2b_1^2)x^2 + (a_1b_3 + 2a_2b_1b_2 + a_3b_1^3)x^3 + \dots$$

□

While there do exist formulae for determining the coefficients of the composition explicitly, we often instead make use of the uniqueness of Taylor series for analytic functions to compute the composition instead.

Proposition 6.15. Let $f(x), g(x) \in \mathbb{R}[[x]]$, and suppose $g(x)$ has zero constant term. Suppose further that $f(x)$ and $g(x)$ are absolutely convergent series on $(-r, r)$ and $(-R, R)$ respectively, and hence define real analytic functions.

Assume further that that $|g(x)| < r$ on $(-R, R)$, shrinking R if necessary. Then the composition $(f \circ g)(x)$ of functions is a real analytic function whose Taylor series is given by the composition $f(g(x))$ of formal power series.

Proof. By uniqueness of Taylor series, if we can show that the function $(f \circ g)(x)$ is equal to the corresponding series for $|x|$ small enough, we simultaneously show $(f \circ g)(x)$ is real analytic and that the series is given by the formal expression.

Let $x \in (-R, R)$, so that

$$g(x) = \sum_{n \geq 1} b_n x^n \in \mathbb{R}$$

By assumption, R is small enough such that $|g(x)| < r$, and so we have that

$$(f \circ g)(x) = \sum_{n \geq 0} a_n g(x)^n \in \mathbb{R}$$

Letting $b_m(n)$ be the coefficient of x^m in the series expansion of $g(x)^n$, we have that

$$(f \circ g)(x) = \sum_{n \geq 0} \sum_{m \geq 1} a_n b_m(n) x^m$$

Note that, if we were to swap the order of summation here, we would obtain the series

$$\sum_{m \geq 0} \left(\sum_{n \geq 0} a_n b_m(n) \right) x^m$$

which is precisely the formal series $f(g(x))$. Thus, if we can justify this reordering of the sum, we are done. To see that this is possible, it is enough to do so, we need to show that this double sum is absolutely convergent.

First note that

$$|b_m(n)| = \left| \sum_{m_1 + \dots + m_n = m} b_{m_1} b_{m_2} \dots b_{m_n} \right| \leq \sum_{m_1 + \dots + m_n = m} |b_{m_1}| |b_{m_2}| \dots |b_{m_n}| =: B_m(n)$$

and that this right hand side is equal to the coefficient of x^m in

$$\left(\sum_{k \geq 1} |b_k| x^k \right)^n.$$

As $g(x)$ is defined by an absolutely convergent series, $\sum_{k \geq 1} |b_k| x^k$ is absolutely convergent and hence defines a real analytic function $G(x)$. By absolute convergence, we have that $B_m(n)$ is the coefficient of x^m in $G(x)^n$.

Thus, we have that

$$\begin{aligned} \left| \sum_{n \geq 0} \sum_{m \geq 1} a_n b_m(n) x^m \right| &\leq \sum_{n \geq 0} |a_n| \sum_{m \geq 1} |b_m(n)| |x|^m \\ &\leq \sum_{n \geq 0} |a_n| \sum_{m \geq 1} B_m(n) |x|^m \\ &\leq \sum_{n \geq 0} |a_n| (G(|x|))^n \end{aligned}$$

which is finite by absolute convergence of $\sum_{n \geq 0} a_n x^n$, assuming $G(|x|) < r$. But $G(0) = 0$ and $G(x)$ is continuous, so we can ensure this, shrinking R if necessary.

Hence, the double sum is absolutely convergent and we can swap the order of summation to obtain

$$(f \circ g)(x) = \sum_{m \geq 0} \left(\sum_{n \geq 0} a_n b_m(n) \right) x^m$$

for small enough x . Thus $(f \circ g)(x)$ is real analytic near 0 with series $f(g(x))$. $\square$

Remark 6.16. *If we were working over $\mathbb{C}$, then the proof that $f \circ g$ is analytic would follow immediately from chain rule and standard results about holomorphic functions. However, we would miss out on computing the associated series, so we instead give the proof for real coefficients.*

As the Taylor series of a real analytic function is unique, we obtain the following corollary.

Corollary 6.17. *Suppose $f(x), g(x) \in \mathbb{R}[[x]]$ are as in Proposition 6.15, and suppose that $f(g(x)) = x$ as a real analytic function. Then $f(g(x)) = x$ as a formal power series.*

Corollary 6.18. *As formal power series*

$$\exp(x) = \sum_{n \geq 0} \frac{x^n}{n!}, \quad \log(1+x) = - \sum_{n \geq 1} \frac{(-x)^n}{n}$$

satisfy

$$\exp(\log(1+x)) = 1+x, \quad \log(\exp(x)) = x.$$

Proof. The series

$$\exp(x) - 1 = \sum_{n \geq 1} \frac{x^n}{n!}, \quad \log(1+x) = - \sum_{n \geq 1} \frac{(-x)^n}{n}$$

define real analytic functions for $|x| < \varepsilon$ small enough, and these real analytic functions satisfy that

$$\exp(\log(1+x)) - 1 = x, \quad \log(1 + (\exp(x) - 1)) = x.$$

Hence, by Corollary 6.17, the formal power series are compositional inverses, from which the claim follows. $\square$

Remark 6.19. *It is possible to prove this final corollary by direction computing the coefficients of the composition of formal power series and some creative combinatorial arguments, but this analytic approach is much more generalisable.*