Lie Algebras and their Representations

Lectures by David Stewart
Notes by David Mehrle
dfm33@cam.ac.uk

Cambridge University
Mathematical Tripos Part III
Michaelmas 2015

Contents

1 Introduction 3

2 Lie Groups 8

3 Representations of $\mathfrak{sl}(2)$ 19

4 Major Results on Lie Algebras 24

5 Representations of Semisimple Lie Algebras 34

6 Classification of Complex Semisimple Lie Algebras 60

Last updated May 29, 2016.

Contents by Lecture

Lecture 1 3

Lecture 2 4

Lecture 3 7

Lecture 4 9

Lecture 5 11

Lecture 6 14

Lecture 7 17

Lecture 8 20

Lecture 9 23

Lecture 10 26

Lecture 11 29

Lecture 12 31

Lecture 13 34

Lecture 14 37

Lecture 15 39

Lecture 16 42

Lecture 17 44

Lecture 18 47

Lecture 19 49

Lecture 20 53

Lecture 21 59

Lecture 22 59

Lecture 23 63

1 Introduction

There are lecture notes online.

We'll start with a bit of history because I think it's easier to understand something when you know where it comes from. Fundamentally, mathematicians wanted to solve equations, which is a rather broad statement, but it motivates things like the Fermat's Last Theorem – solving $x^3 + y^3 = z^3$. Galois theory begins with wanting to solve equations. One of Galois's fundamental ideas was to not try to write a solution but to study the symmetries of the equations.

Sophus Lie was motivated by this to do the same with differential equations. Could you say something about the symmetries of the solutions? This technique is used a lot in physics. This led him to the study of Lie groups, and subsequently, Lie algebras.

Example 1.1. The prototypical Lie group is the circle.

A Lie group G is, fundamentally, a group with a smooth structure on it. The group has some identity $e \in G$. Multiplying e by $a \in G$ moves it the corresponding point around the manifold. Importantly, if the group isn't abelian, then $aba^{-1}b^{-1}$ is *not* the identity. We call this the **commutator** $[a, b]$. Letting a, b tend to zero, we get tangent vectors X, Y and a commutator $[X, Y]$ by letting $[a, b]$ tend to 0. These points of the tangent space are elements of the Lie algebra.

We'l make this all precise later. We'll classify the simple Lie algebras, which ends up using the fundamental objects called **root systems**. Root systems are so fundamental to math that this course might better be described as an introduction to root systems by way of Lie algebras.

Definition 1.2. Let k be a field. A Lie algebra $\mathfrak{g}$ is a vector space over k with a bilinear bracket $[-, -] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying

(i) Antisymmetry $[X, X] = 0$, for all $X \in \mathfrak{g}$;

(ii) Jacobi identity $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

The Jacobi Identity is probably easier to think of as

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

Bracketing with X satisfies the chain rule! Actually $[X, -]$ is a derivation.

Definition 1.3. A **derivation** δ of an algebra A is an endomorphism $A \to A$ that satisfies $\delta(ab) = a\delta(b) + \delta(a)b$.

Remark 1.4. From now on, our Lie algebras $\mathfrak{g}$ will always be finite dimensional. Most of the time, $k = \mathbb{C}$ (but not always!). We'll sometimes point out how things go wrong in characteristic $p > 0$.

Example 1.5.

(i) If V is any vector space, equip V with the trivial bracket $[a, b] = 0$ for all $a, b \in V$.

(ii) If A is any associative algebra, equip A with $[a, b] = ab - ba$ for all $a, b \in A$.

(iii) Let $\mathfrak{g} = M_{n\times n}(k)$, the $n \times n$ matrices over a field k. This is often written $\mathfrak{gl}_n(k)$ or $\mathfrak{gl}(n)$ when the field is understood. This is an example of an associative algebra, so define $[A, B] = AB - BA$.

There is an important basis for $\mathfrak{gl}(n)$ consisting of E_{ij} for $1 \leqslant i, j \leqslant n$, which is the matrix whose entries are all zero except in the (i, j)-entry which is 1. First observe

$$[E_{ij}, E_{rs}] = \delta_{jr}E_{is} - \delta_{is}E_{rj}.$$

This equation gives the **structure constants** for $\mathfrak{gl}(n)$.

We can calculate that

$$[E_{ii} - E_{jj}, E_{rs}] = \begin{cases} 0 & \{i, j\} \neq \{r, s\} \\ E_{rs} & i = r, j \neq s \\ -E_{rs} & j = r, i \neq s \\ 2E_{rs} & i = r, j = s \end{cases}$$

(iv) If A is any algebra over k, $\operatorname{Der}_k A \subset \operatorname{End}_k A$ is a Lie algebra, the **derivations of** A. For $\alpha, \beta \in \operatorname{Der} A$, define $[\alpha, \beta] = \alpha \circ \beta - \beta \circ \alpha$. This will be a valid Lie algebra so long as $[\alpha, \beta]$ is still a derivation.

Definition 1.6. A subspace $\mathfrak{h} \subset \mathfrak{g}$ is a **Lie subalgebra** if $\mathfrak{h}$ is closed under the Lie bracket of $\mathfrak{g}$.

Definition 1.7. Define the **derived subalgebra** $\mathcal{D}(\mathfrak{g}) = \langle [X, Y] \mid X, Y \in \mathfrak{g} \rangle$.

Example 1.8. An important subalgebra of $\mathfrak{gl}(n)$ is $\mathfrak{sl}(n)$, $\mathfrak{sl}(n) := \{X \in \mathfrak{gl}(n) \mid \operatorname{tr} X = 0\}$. This is a simple Lie algebra of type A_{n-1}. In fact, you can check that $\mathfrak{sl}(n)$ is the derived subalgebra of $\mathfrak{gl}(n)$,

$$\mathfrak{sl}(n) = [\mathfrak{gl}(n), \mathfrak{gl}(n)] = \mathcal{D}(\mathfrak{gl}(n).$$

Example 1.9. Lie subalgebras of $\mathfrak{gl}(n)$ which preserve a bilinear form. Let $Q \colon V \times V \to k$ be a bilinear form. Then we say $\mathfrak{gl}(V)$ **preserves** Q if the following is true:

$$Q(Xv, w) + Q(v, Xw) = 0$$

for all $v, w \in V$. Recall that if we pick a basis for V, we can represent Q by a matrix M. Then $Q(v, w) = v^T M w$. Then X preserves Q if and only if

$$v^T X^T M w + v^T M X w = 0,$$

if and only if

$$X^T M + MX = 0.$$

Recall that a Lie algebra $\mathfrak{g}$ is a k-vector space with a bilinear operation $[-, -] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying antisymmetry and the Jacobi identity.

We had some examples, such as $\mathfrak{g} = \mathfrak{gl}(V) = \operatorname{End}_k(V)$. If you pick a basis, this is $M_{n\times n}(V)$. Given any associative algebra, we can turn it into a Lie algebra with bracket $[X, Y] = XY - YX$.

Example 1.10. Another example is if $Q\colon V \times V \to k$ is a bilinear form, the set of $X \in \mathfrak{gl}(V)$ preserving Q is a Lie subalgebra of $\mathfrak{gl}(V)$. Taking a basis, Q is represented by a matrix M with $Q(\vec{v}, \vec{w}) = \vec{v}^T M \vec{w}$. X preserves Q if and only if $X^T M + MX = 0$.

The most important case is where Q is **non-degenerate**, i.e. $Q(v, w) = 0 \; \forall \; w \in V$ if and only if $v = 0$.

Example 1.11. Consider the bilinear form where

$$M = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

M represents an alternating form and the set of endomorphisms of $V = k^{2n}$ is the **symplectic Lie algebra** (of rank n) and denoted $\mathfrak{sp}(2n)$. If $X \in \mathfrak{gl}(V)$ is written

$$X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

Check that X preserves M if and only if

$$X = \begin{bmatrix} A & B \\ C & -A^T \end{bmatrix}$$

with B, C symmetric matrices. A basis for this consists of elements

- $H_{i,i+n} = E_{ii} - E_{i+n,i+n}$
- $E_{ij} - E_{j+n,i+n}$
- $E_{i,j+n} + E_{j,i+n}$
- $E_{i+n,j} + E_{j+n,i}$

for $1 \leqslant i, j \leqslant n$. This is a simple Lie algebra of type C_n for $\operatorname{char} k \neq 2$.

Example 1.12. There are also

- orthogonal Lie algebras of type $D_n = \mathfrak{so}(2n)$, preserving

$$M = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}.$$

- orthogonal lie algebras of type $B_n = \mathfrak{so}(2n+1)$, preserving

$$M = \begin{bmatrix} 0 & I_n & \\ I_n & 0 & \\ & & 1 \end{bmatrix}.$$

Example 1.13. $\mathfrak{b}_n$ is the **borel algebra** of upper triangular $n \times n$ matrices and $\mathfrak{n}_n$ is the **nilpotent algebra** of strictly upper triangular $n \times n$ matrices.

Definition 1.14. A linear map $f\colon \mathfrak{g} \to \mathfrak{h}$ between two Lie algebras $\mathfrak{g}, \mathfrak{h}$, is a **Lie algebra homomorphism** if $f([X, Y]) = [f(X), f(Y)]$.

Definition 1.15. We say a **subspace** $\mathfrak{j}$ is a subalgebra of $\mathfrak{g}$ if $\mathfrak{j}$ is closed under the Lie bracket. A subalgebra $\mathfrak{j}$ is an **ideal** of $\mathfrak{g}$ if $[X, Y] \in \mathfrak{j}$ for all $X \in \mathfrak{g}$, $Y \in \mathfrak{j}$.

Definition 1.16. The **center of** $\mathfrak{g}$, denoted $Z(\mathfrak{g})$ is

$$Z(\mathfrak{g}) = \{X \in \mathfrak{g} \mid [X, Y] = 0 \ \forall\ Y \in \mathfrak{g}\}.$$

Exercise 1.17. Check that $Z(\mathfrak{g})$ is an ideal using the Jacobi identity.

Proposition 1.18.

(1) If $f\colon \mathfrak{h} \to \mathfrak{g}$ is a homomorphism of Lie algebras, then $\ker f$ is an ideal;

(2) If $\mathfrak{j} \subset \mathfrak{g}$ is a linear subspace, then $\mathfrak{j}$ is an ideal if and only if the quotient bracket $[X + \mathfrak{j}, Y + \mathfrak{j}] = [X, Y] + \mathfrak{j}$ makes $\mathfrak{g}/\mathfrak{j}$ into a Lie algebra;

(3) If $\mathfrak{j}$ is an ideal of $\mathfrak{g}$ then the quotient map $\mathfrak{g} \to \mathfrak{g}/\mathfrak{j}$ is a Lie algebra homomorphism;

(4) If $\mathfrak{g}$ and $\mathfrak{h}$ are both Lie algebras, then $\mathfrak{g} \oplus \mathfrak{h}$ becomes a Lie algebra under $[(X, A), (Y, B)] = ([X, Y], [A, B])$.

Exercise 1.19. Prove Proposition 1.18.

Remark 1.20. The category of Lie algebras, **Lie**, forms a semi-abelian category. It's closed under taking kernels but not under taking cokernels. The representation theory of Lie algebras does, however, form an abelian category.

Definition 1.21. The following notions are really two ways of thinking about the same thing.

(a) A **representation** of $\mathfrak{g}$ on a vector space V is a homomorphism of Lie algebras $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$.

(b) An action of $\mathfrak{g}$ on a vector space V is a bilinear map $r\colon \mathfrak{g} \times V \to V$ satisfying $r([X, Y], v) = r(X, r(Y, v)) - r(Y, r(X, v))$. We also say that V is a $\mathfrak{g}$**-module** if this holds.

Given an action r of $\mathfrak{g}$ on V, we can make a representation of $\mathfrak{g}$ by defining $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ by $\rho(X)(v) = r(X, v)$.

Example 1.22. This is the *most important* example of a representation. For any Lie algebra $\mathfrak{g}$, one always has the **adjoint representation**, $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ defined by $\mathrm{ad}(X)(Y) = [X, Y]$. The fact that ad gives a representation follows from the Jacobi identity.

Definition 1.23. If W is a subspace of a $\mathfrak{g}$-module W, then W is a $\mathfrak{g}$-submodule if W is stable under action by $\mathfrak{g}$: $\mathfrak{g}(W) \subseteq W$.

Example 1.24.

(1) Suppose $\mathfrak{j}$ is an ideal in $\mathfrak{g}$. Then $\mathrm{ad}(X)(Y) = [X, Y] \in \mathfrak{j}$ for all $Y \in \mathfrak{j}$, so $\mathfrak{j}$ is a submodule of the adjoint representation.

(2) If $W \subseteq V$ is a submodule then V/W is another $\mathfrak{g}$-module via $X(v+W) = Xv + W$.

(3) If V is a $\mathfrak{g}$-module, then the dual space $V^* = \mathrm{Hom}_k(V, k)$ has the structure of a $\mathfrak{g}$-module via $X\phi(v) = -\phi(Xv)$.

Last time we developed a category of Lie algebras, and said what homomorphisms of Lie algebras were, as well as defining kernels and cokernels. There are a few more definitions that we should point out.

Definition 1.25. A Lie algebra is **simple** if it has no nontrivial ideals.

We also moved on and discussed representations. Recall

Definition 1.26. A **representation or $\mathfrak{g}$-module** of $\mathfrak{g}$ on V is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{gl}(V)$.

To complete the category **$\mathfrak{g}$-Mod** of $\mathfrak{g}$-modules, let's define a map of $\mathfrak{g}$-modules.

Definition 1.27. Let V, W be $\mathfrak{g}$-modules. Then a linear map $\phi\colon V \to W$ is a g**-module map** if $X\phi(v) = \phi(Xv)$ for all $X \in \mathfrak{g}$.

$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[d, "X"] & W \arrow[d, "X"] \\ V \arrow[r, "\phi"] & W \end{tikzcd}$$

Proposition 1.28. If $\phi\colon V \to W$ is a $\mathfrak{g}$-module map, then $\ker\phi$ is a submodule of V

Exercise 1.29. Prove Proposition 1.28.

Definition 1.30. A $\mathfrak{g}$-module V (resp. representation) is **simple** (resp. **irreducible**) if V has no non-trivial submodules.

We write $V = V_1 \oplus V_2$ if V_1 and V_2 are submodules with $V = V_1 \oplus V_2$ as vector spaces.

How can you build new representations from old ones? There are several ways. If V, W are $\mathfrak{g}$-modules, then so is $V \oplus W$ becomes a $\mathfrak{g}$-module via $X(v, w) = (Xv, Xw)$.

There's another way to build new representations via the tensor product. In fact, **$\mathfrak{g}$-Mod** is more than just an abelian category, it's a monoidal category via the tensor product. Given V, W representations of $\mathfrak{g}$, we can turn the tensor product into a representation, denoted $V \otimes W$, by defining the action on simple tensors as

$$X(v \otimes w) = (Xv) \otimes w + v \otimes (Xw)$$

and then extending linearly.

We can iterate on multiple copies of V, say, to get **tensor powers**

$$V^{\otimes r} = \underbrace{V \otimes V \otimes \cdots \otimes V}_{r \text{ times}}$$

Definition 1.31. The r**-th symmetric power** of V, with basis $e_1, \dots, e_n$ is the vector space with basis $e_{i_1} \cdots e_{i_r}$ for $i_1 \leqslant i_2 \leqslant, \dots, \leqslant i_r$. This is denoted $S^r(V)$. The action of $\mathfrak{g}$ on $S^r(V)$ is

$$X(e_{i_1} \cdots e_{i_r}) = X(e_{i_1})e_{i_2} \cdots e_{i_r} + e_{i_1}X(e_{i_2}) \cdots e_{i_r} + \dots + e_{i_1}e_{i_2} \cdots e_{i_{r-1}}X(e_{i_r}).$$

Definition 1.32. The r**-th alternating power** of a $\mathfrak{g}$-module V, denoted $\bigwedge^r(V)$, is the vector space with basis $\{e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_r} \mid i_1 < i_2 < \dots < i_r\}$, if V has basis $e_1, \dots, e_n$. The action is functionally the same as on the symmetric power:

$$X(e_{i_1} \wedge \cdots \wedge e_{i_r}) = X(e_{i_1}) \wedge e_{i_2} \wedge \cdots \wedge e_{i_r} + \dots + e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_{r-1}} \wedge X(e_{i_r}).$$

We also have the rule that

$$e_{i_1} \wedge \dots \wedge e_{i_j} \wedge \dots \wedge e_{i_k} \wedge \dots \wedge e_{i_r} = -e_{i_1} \wedge \dots \wedge e_{i_k} \wedge \dots \wedge e_{i_j} \wedge \dots \wedge e_{i_r}.$$

Exercise 1.33. What is the dimension of the symmetric powers / alternating powers?

Example 1.34. Let

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

and let $V = k^2$ with basis $\{e_1, e_2\}$. Let $\mathfrak{g} = kX \subset \mathfrak{gl}(V)$. Then $V \otimes V$ has basis $e_1 \otimes e_1$, $e_1 \otimes e_2$, $e_2 \otimes e_1$, and $e_2 \otimes e_2$.

Observe $Xe_1 = 0$, $Xe_2 = e_1$. Therefore,

$$\begin{aligned} X(e_1 \otimes e_1) &= 0 \\ X(e_1 \otimes e_2) &= e_1 \otimes e_1 \\ X(e_2 \otimes e_1) &= e_1 \otimes e_1 \\ X(e_2 \otimes e_1) &= e_1 \otimes e_2 + e_2 \otimes e_1 \end{aligned}$$

As a linear transformation $V \otimes V \to V \otimes V$, X is represented by the matrix

$$X = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A basis for $\bigwedge^2 V$ is $\{e_1 \wedge e_2\}$, and here $X(e_1 \wedge e_2) = Xe_1 \wedge e_2 + e_1 \wedge Xe_2 = 0 \wedge e_2 + e_1 \wedge e_1 = 0$. So X is the zero map on the alternating square.

Exercise 1.35. Work out the preceding example for the symmetric square, and the tensor cube.

2 Lie Groups

Lots of stuff in this section requires differential geometry and some analysis.

Definition 2.1. A Hausdorff, second countable topological space X is called a **manifold** if each point has an open neighborhood (nbhd) homeomorphic to an open subset U of $\mathbb{R}^d$ by a homeomorphism $\phi\colon U \xrightarrow{\sim} \phi(U) \subset \mathbb{R}^N$.

The pair (U, ϕ) of a homeomorphism and open subset of M is called a **chart**: given open subsets U and V of X with $U \cap V \neq \varnothing$, and charts (U, ϕ_U) and (V, ϕ_V), we have a diffeomorphism $\phi_V \circ \phi_U^{-1}\colon \phi_U(U \cap V) \to \phi_V(U \cap V)$ of open subsets of $\mathbb{R}^N$.

We think of a manifold as a space which looks locally like $\mathbb{R}^N$ for some N.

Example 2.2.

(a) $\mathbb{R}^1, S^1$ are one-dimensional manifolds;

(b) S^2 or $S^1 \times S^1$ are two-dimensional manifolds. The torus carries a group structure; S^2 does not.

Definition 2.3. A function $f\colon M \to N$ is called **smooth** if composition with the appropriate charts is smooth. That is, for $U \subseteq M$, $V \subseteq N$, and charts $\phi\colon U \to \mathbb{R}^M$, $\psi\colon V \to \mathbb{R}^N$, then $\psi \circ f \circ \phi^{-1}\colon \mathbb{R}^M \to \mathbb{R}^N$ is smooth where defined.

Definition 2.4. A **Lie group** is a manifold G together with the structure of a group such that the multiplication $\mu\colon G \times G \to G$ and inversion $i\colon G \to G$ maps are smooth.

Exercise 2.5. Actually, the fact that the inverse is smooth follows from the fact that multiplication is smooth and by looking in a neighborhood of the identity. Prove it!

To avoid some subtleties of differential geometry, we will assume that M is embedded in $\mathbb{R}^N$ for some (possibly large) N. This is possible under certain tame hypotheses by Nash's Theorem.

Example 2.6.

(1) $\mathrm{GL}(n) := \{n \times n \text{ matrices over } \mathbb{R} \text{ with non-zero determinant}\}$. There is only a single chart: embed it into $\mathbb{R}^{n^2}$.

(2) $\mathrm{SL}(n) := \{g \in \mathrm{GL}(n) \mid \det g = 1\}$.

(3) If $Q\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear form, then

$$G(Q) = \{g \in M(n) \mid Q(v, w) = Q(gv, gw)\}$$

for all $v, w \in \mathbb{R}^n$.

Recall that a Lie group is a manifold with a group structure such that the group operations are smooth. For example, $\mathrm{SL}(n)$.

Definition 2.7. Let G and H be two Lie groups. Then a map $f\colon G \to H$ is a **Lie group homomorphism** if f is a group homomorphism and a smooth map of manifolds.

Let G be a Lie group and let G° be the connected component of G containing the identity.

Proposition 2.8. For any Lie group G, the set G° is an open normal subgroup of G. Moreover, if U is any open neighborhood of the identity in G°, then $G^\circ = \langle U \rangle$.

Proof. The first thing we need to show is that G° is a subgroup. Since G is a manifold, its connected components are path connected. Suppose $a, b \in G^\circ$. Then we can find paths $\gamma, \delta \colon [0,1] \to G^\circ$ with $\gamma(0) = e = \delta(0)$, $\gamma(1) = a$ and $\delta(1) = b$. Then taking the path $\mu(\gamma(t), \delta(t))$ gives a path from the identity to ab. Hence, G° is closed under multiplication. Similarly, $i(\gamma(t))$ gives a path from e to a^{-1}, and G° is closed under inverse.

Why is G° normal? Well, the map $g \mapsto aga^{-1}$ gives a diffeomorphism of G° with that fixes e and therefore also G°.

By replacing U with $U \cap U^{-1}$, we can arrange that U contains the inverse of every element in U. Now let $U^n = U \cdot U \cdots U = \{u_1 u_2 \cdots u_n \mid u_i \in U\}$ is open as it is the union of open cosets $u_1 u_2 \cdots u_{n-1} U$ over all $(u_1, \ldots, u_{n-1})$. Then set $H = \bigcup_{n \geqslant 0} U^r$ This is an open subgroup of G^0 containing $\langle U \rangle$. It is also closed since the set of cosets $\bigcup_{a \notin H} aH = G^\circ \backslash H$ is open as the union of diffeomorphic translates of an open set, so H is the compliment of an open set.

The connected component G° is a minimal set that is both open and closed, so $H = G$. $\square$

Definition 2.9. Any open neighborhood of the identity is called a **germ** or **nucleus**.

Corollary 2.10. If f and g are two homomorphisms from G to H with G connected, then $f = g$ if and only if $f|_U = g|_U$ for any germ of g.

Definition 2.11. Let $M \subset \mathbb{R}^N$ be a manifold. The **tangent space** of $p \in M$ is

$$T_p(M) = \left\{ v \in \mathbb{R}^N \;\middle|\; \text{there is a curve } \phi \colon (-\varepsilon, \varepsilon) \to M \text{ with } \phi(0) = p, \phi'(0) = v \right\}$$

One can show that this is a vector space. Scalar multiplication is easy: take the curve $\phi(\lambda t) = \lambda v$. Addition follows by looking at addition in charts.

Let's single out a very important tangent space when we replace M with a Lie group G.

Definition 2.12. If G is a Lie group, then we denote $T_e(G)$ by $\mathfrak{g}$ and call it the **Lie algebra** of G.

We don't *a priori* know that this is actually a Lie algebra as we defined it previously, but we can at least see what kind of vectors live in a Lie algebra by looking at the Lie group.

Example 2.13. Let's calculate $\mathfrak{sl}_n = T_{I_n}(\mathrm{SL}_n)$. If $v = {}^{d}\!/_{dt}\big|_{t=0} g(t) = g'(0)$. By the condition on membership in SL_n, we have $\det g(t) = 1 = \det(g_{ij}(t))$. Write out the determinant explicitly.

$$1 = \det(g_{ij}(t)) = \sum_{\sigma \in S_n} (-1)^{\mathrm{sgn}(\sigma)} \prod_{i=1}^{n} g_{i\sigma(i)}(t)$$

Now differentiate both sides with respect to t and evaluate at $t = 0$:

$$0 = \sum_{\sigma \in S_n} (-1)^{\mathrm{sgn}(\sigma)} \sum_{j=1}^{n} g'_{j\sigma(j)}(0) \prod_{i \neq j} g_{i\sigma(i)}(0)$$

Observe that g is a path through the identity, so $g(0) = I_n$. Thus, $g_{ij}(0) = \delta_{ij}$. Therefore, we are left with

$$0 = \sum_{j=1}^{n} \left.\frac{d}{dt}\right|_{t=0} g_{jj}(t) = \operatorname{tr} v$$

Hence, $\mathfrak{sl}(n)$ is the traceless $n \times n$ matrices.

Because manifolds and tangent spaces have the same dimension, this tells us that $\dim \mathrm{SL}_n = \dim \mathfrak{sl}_n = n^2 - 1$.

Example 2.14. Now consider $G = G(Q) = \{g \in \mathrm{GL}(n) \mid Q(v,v) = Q(gv, gv)\}$. If Q is represented by a matrix M, we have $Q(v,w) = v^T M w$. We have that $g \in G(Q) \iff g^T M g = M$.

Now let $g(t)$ be a path through the identity of $G(Q)$. We see that

$$0 = \left.\frac{d}{dt}\right|_{t=0} \Big(g(t)^T M g(t)\Big) = g'(0)^T M + M g'(0),$$

so the Lie algebra here is those matrices g such that $X^T M + MX = 0$.

Definition 2.15. Let $f\colon M \to N$ be a map of manifolds. Then given $p \in M$, we define $df_p\colon T_pM \to T_{f(p)}N$ as follows. Given $v \in T_pM$ with $v = \phi'(0)$ for some path ϕ, define $df_p(v) = w$ where $w = (f \circ \phi)'(0) = {}^{d}\!/_{dt}\big|_{t=0}(f \circ \phi)$.

We need to check that df_p is well-defined, that is, given another path ϕ_1 through p, we have

$$\left.\frac{d}{dt}\right|_{t=0} (f \circ \phi)(t) = \left.\frac{d}{dt}\right|_{t=0} (f \circ \phi_1)(t),$$

but this is true by the multivariable chain rule.

2.1 The Exponential Map

If we have a Lie group G and a Lie algebra T_eG, we want to find a way to map elements of the algebra back onto the group. This is the exponential map. It somehow produces a natural way of molding $T_e(G)$ onto G such that a exp is a homomomorphism on restriction to any line in $T_e(G)$, that is, $\exp(aX)\exp(bX) = \exp((a+b)X)$.

Remark 2.16. Some preliminaries for the exponential map.

- Left multiplication by $g \in G$ gives a smooth map $L_g \colon G \to G$, and differentiating gives a map
$$(dL_g)_h \colon T_h(G) \to T_{gh}(G).$$
- Recall the theorem alternatively called the **Picard-Lindelöf Theorem** or the **Cauchy-Lipschitz Theorem**: for any first order ODE of the form $y' = F(x, y)$, there is a unique continuously differentiable local solution containing the point (x_0, y_0).

The following definition is not rigorous, but it will suffice for what we're trying to do. To define it (marginally more) rigorously, we need to talk about vector bundles on a manifold.

Definition 2.17. A **vector field** on M is a choice of tangent vector for each $m \in M$, varying smoothly over M, which we write as a map $M \to TM$. (More precisely, a vector field is a smooth section of the tangent bundle $\pi \colon TM \to M$, that is, an $X \colon M \to TM$ such that $\pi \circ X = \mathrm{id}_M$).

Definition 2.18. If G is a Lie group, then a vector field v is **left-invariant** if $dL_g(v(h)) = v(gh)$.

Given $X \in T_h(G)$, we construct a left-invariant vector field v_X by $v_X(g) = (dL_{gh^{-1}})_h(X)$. It's clear that all left-invariant vector fields arise in this way. As soon as you know any left-invariant vector field at a point, then you know it anywhere.

Proposition 2.19 ("The flow of v_X is complete")**.** Let $g \in G$ and $X \in T_g(G)$ with v_X the associated left-invariant vector field. Then there exists a unique curve $\gamma_g \colon \mathbb{R} \to G$ such that $\gamma_g(0) = g$ and $\gamma_g'(t) = v_X(\gamma(t))$.

Proof. First we will reduce to the case when $g = e$. By defining $\gamma_g(t) = g\gamma_e(t)$, we have that
$$\gamma_g(0) = g\gamma_e(0) = ge = g,$$
and moreover,
$$\gamma_g'(t) = (dL_g)_e(\gamma_e'(t)) = (dL_g)_e(v_X(\gamma_e(t))),$$
now apply left invariance to see that
$$(dL_g)_e(v_X(\gamma_e(t))) = v_X(g\gamma_e(t)) = v_X(\gamma_g(t)).$$
Therefore, if we define $\gamma_g(t) = g\gamma_e(t)$, we have that $\gamma_g'(t) = v_X(\gamma_g(t))$ and $\gamma_g(0) = g\gamma_e(0) =$ So we have reduced it to the case where $g = e$.

Now to establish the existence of $\gamma_e(t)$, we will solve the equation $v_X(\gamma_e(t)) = \gamma_e'(t)$ with initial condition $\gamma_e(0) = X$ in a small neighborhood of zero, and then push the solutions along to everything.

Using the existence part of the Cauchy-Lipschitz theorem for ODE's, there is some $\varepsilon > 0$ such that γ_ε can be defined on an open interval $(-\varepsilon, \varepsilon)$. We show that there is no maximum such ε. For each $s \in (-\varepsilon, \varepsilon)$, define a new curve $\alpha_s \colon (-\varepsilon + |s|, \varepsilon - |s|) \to G$ via $\alpha_s(t) = \gamma_e(s+t)$.

Then $\alpha_s(0) = \gamma_e(s)$ and

$$\alpha_s'(t) = \gamma_e'(s+t) = v_X(\gamma_e(s+t)) = v_X(\alpha_s(t)). \tag{1}$$

By the uniqueness part of Cauchy-Lipschitz theorem, we must have a unique solution to (1) for $|s| + |t| < \varepsilon$. But notice that $\gamma(s)\gamma(t)$ is another solution to (1), because

$$\begin{aligned}(\gamma_e(s)\gamma_e(t))' &= dL_{\gamma_e(s)}\gamma_e'(t)\\ &= dL_{\gamma_e(s)}v_X(\gamma_e(t))\\ &= v_X(L_{\gamma_e(s)}\gamma_e(t))\\ &= v_X(\gamma_e(s)\gamma_e(t)).\end{aligned}$$

Therefore, we have that

$$\gamma_e(s+t) = \gamma_e(t)\gamma_e(s).$$

We can use this equation to extend the range of γ_e to $(-3\varepsilon/2, 3\varepsilon/2)$, via

$$\gamma_e(t) = \begin{cases}\gamma_e(-\varepsilon/2)\gamma_e(t+\varepsilon/2) & t \in (-3\varepsilon/2, \varepsilon/2)\\ \gamma_e(\varepsilon/2)\gamma_e(t-\varepsilon/2) & t \in (-\varepsilon/2, 3\varepsilon/2)\end{cases}$$

Repeating this infinitely defines a curve $\gamma_e \colon \mathbb{R} \to G$ with the required properties. □

Definition 2.20. The curves γ_g guaranteed by the previous proposition are called **integral curves** of v_X.

Definition 2.21. The **exponential** map of G is the map $\exp\colon \mathfrak{g} = T_eG \to G$ given by $\exp(X) = \gamma_e(1)$, where $\gamma_e \colon \mathbb{R} \to G$ is the integral curve associated to the left-invariant vector field v_X. (We choose $\gamma_e(1)$ because we want exp to be it's own derivative, like e^X.)

Proposition 2.22. Every Lie group homomorphism $\phi\colon \mathbb{R} \to G$ is of the form $\phi(t) = \exp(tX)$ for $X = \phi'(0) \in \mathfrak{g}$.

Proof. Let $X = \phi'(0) \in \mathfrak{g}$ and v_X the associated left-invariant vector field. We have $\phi(0) = e$, and by the uniqueness part of the Cauchy-Lipschitz Theorem, we need only show that ϕ is an integral curve of v_X. Note that, as in the proof of Proposition 2.19,

$$\phi(s+t) = \phi(s)\phi(t),$$

which implies that $\phi'(s) = (dL_{\phi(s)})_e(\phi'(0)) = v_X(\phi(s))$.

Conversely, we must show that $\phi(t) = \exp(tX)$ is a Lie group homomorphism $\mathbb{R} \to G$, that is, $\exp((t+s)X) = \exp(tX)\exp(sX)$. To do this, we will use

ODE uniqueness. Let γ be the integral curve associated to X, which is a solution to the equation $v_X(\gamma(t)) = \gamma'(t)$ with initial condition $\gamma(0) = e$. Let θ be the integral curve associated to aX for $a \in \mathbb{R}$, solving the equation

$$\theta'(t) = v_{aX}(\theta(t)) \tag{2}$$

with initial condition $\theta(0) = aX$. We will show that $\gamma(at)$ is a solution to (2), and therefore $\gamma(at) = \theta(t)$ by ODE uniqueness.

To that end,

$$\begin{aligned} v_{aX}(\gamma(at)) &= dL_{\gamma(at)}(aX) \\ &= adL_{\gamma(at)}X \\ &= av_X(\gamma(at)) \\ &= a\gamma'(at) = \frac{d}{dt}(\gamma(at)). \end{aligned}$$

Therefore, $\theta(t) = \gamma(at)$. But as in the proof of Proposition 2.19, notice that $\theta(s+t)$ and $\theta(s)\theta(t)$ are both solutions to $\theta'(s+t) = v_X(\theta(s+t))$, hence

$$\gamma(t(a+b)) = \theta(a+b) = \theta(a)\theta(b) = \gamma(at)\gamma(bt).$$

In particular, setting $t = 1$, we have that $\gamma(a+b) = \gamma(a)\gamma(b)$. But $\exp(aX) = \theta(1) = \gamma(a)$ by definition, so

$$\exp((a+b)X) = \gamma(a+b) = \gamma(a)\gamma(b) = \exp(aX)\exp(bX). \qquad \square$$

Exercise 2.23. Let $\delta\colon \mathbb{R} \to \mathfrak{g}$ be the curve tX for some $X \in \mathfrak{g}$. Show that

$$\left.\frac{d}{dt}\right|_{t=0} \exp(\delta(t)) = X.$$

This is really just saying that $(d\exp)_0\colon \mathfrak{g} \to \mathfrak{g}$ is the identity map.

Example 2.24. For $G = \mathrm{GL}(V)$, the exponential map is just the map

$$X \mapsto 1 + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \ldots = \sum_{n=0}^{\infty} \frac{X^n}{n!}.$$

One can check that ${}^d/_{dt}|_{t=0}\exp(tX) = X$, and that it is a homomorphism when restricted to any line through zero.

This definition is not something that's not used much at all, but it's nice to have the terminology.

Definition 2.25. We refer to the images of $\mathbb{R} \subseteq G$ under the exponential map as **1-parameter subgroups.**

Let's summarize what we've done so far:

(1) We defined left invariant vector fields v such that $(dL_g)_h(v(h)) = v(gh)$. (In the notes the subscript $(\ )_h$ is often dropped, and each such is v_X for $X \in \mathfrak{g}$).

(2) We used analysis to prove the existence of integral curves ϕ corresponding to v_X and defined $\exp(X) = \phi(1)$.

(3) exp restricts to a homomorphism of every line of $\mathfrak{g}$ through 0.

(4) The image of any such is called a **1-parameter subgroup**.

Example 2.26. For SL(2), if $H = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in \mathfrak{sl}(2)$, then

$$\exp(tH) = \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix},$$

which is part of the **split torus** $\left\{\left(\begin{smallmatrix} t & 0 \\ 0 & t \end{smallmatrix}\right) \mid t \in \mathbb{R}\right\}$. If we look at

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \exp(tX) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

$$Z = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \exp(tZ) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$$

The image of $t \mapsto \exp(tZ)$ is isomorphic to S^1 and called a **non-split torus**.

2.2 The Lie Bracket

We need a way to multiply exponentials. That is, a formula of the form

$$\exp(X)\exp(Y) = \exp(X + Y + C) \tag{3}$$

where C encodes the non-commutativity of X and Y.

To develop such a formula, let's look to GL(n) for inspiration. The left hand side of Equation 3 becomes

$$\begin{aligned} \text{LHS of Equation 3} &= \left(1 + X + \frac{X^2}{2} + \dots\right)\left(1 + Y + \frac{Y^2}{2} + \dots\right) \\ &= 1 + X + Y + \frac{X^2}{2} + XY + \frac{Y^2}{2} + \dots \end{aligned}$$

So what do we need for C on the right hand side? Up to quadratic terms, what we want is equal to

$$\exp\left(X + Y + \tfrac{1}{2}[X, Y] + \dots\right).$$

This is where the Lie bracket becomes important.

Observe that since $(d\exp)_0 \colon \mathfrak{g} \to \mathfrak{g}$ is the identity mapping, it follows from the inverse/implicit function theorem that exp is a diffeomorphism of restriction to some open neighborhood U of 0. Call the inverse log.

Definition 2.27. For g near the identity in GL(n),

$$\log(g) = (g - 1) - \frac{(g-1)^2}{2} + \frac{(g-1)^3}{3} - \dots$$

Exercise 2.28. Check that $\log \circ \exp$ and $\exp \circ \log$ are the identity on $\mathfrak{g}$ and G where defined.

Moreover, there must be a possibly smaller neighborhood V of U such that multiplication $\mu\colon \exp V \times \exp V \to G$ has image in $\exp U$. It follows that there is a unique smooth mapping $\nu\colon V \times V \to U$ such that

$$\exp(X)\exp(Y) = \mu(\exp X, \exp Y) = \exp(\nu(X,Y)).$$

Notice that $\exp(0)\exp(0) = \exp(0+0) = \exp(0)$, so it must be that $\nu(0,0) = 0$. Now Taylor expand ν around $(0,0)$ to see that

$$\nu(X,Y) = \nu_1(X,Y) + \frac{1}{2}\nu_2(X,Y) + \big(\text{higher order terms}\big),$$

where ν_1 is the linear terms, ν_2 is quadratic terms, etc.

Let's try to figure out what ν_1 and ν_2 are.

Since exp is a homomorphism on lines in G, we have that $\nu(aX, bX) = (a+b)X$. In particular, $\nu(X,0) = X$ and $\nu(0,Y) = Y$. Since ν_1 is the linear terms, $\nu_1(X,Y)$ is linear in both X and Y, so we see that

$$(a+b)X = \nu(aX,bX) = \nu_1(aX,bX) + \frac{1}{2}\nu_2(aX,bX) + \ldots$$

But the terms higher than linear vanish by comparing the left hand side to the right hand side. Then set $a = 0, b = 1$ to see that $\nu_1(0,X) = X$ and likewise, $b = 0, a = 1$ to see that $\nu_1(X,0) = X$. Therefore, $\nu_1(X,Y) = \nu_1(X,0) + \nu_1(0,Y) = X + Y$.

So we have that

$$\nu(X,Y) = (X+Y) + \frac{1}{2}\nu_2(X,Y) + \big(\text{higher order terms}\big),$$

To figure out what ν_2 is, consider

$$X = \nu(X,0) = X + 0 + \frac{1}{2}\nu_2(X,0) + \ldots$$

Therefore, $\nu_2(X,0) = 0$. Similarly, $\nu_2(0,Y) = Y$. So the quadratic term ν_2 contains neither X^2 nor Y^2. Similarly, $2X = \nu(X,X) = \nu_1(X,X)$, so $\nu_2(X,X) = 0$. Therefore, if $\nu_2(X,Y)$ must be antisymmetric in the variables.

Definition 2.29. The antisymmetric, bilinear form $[-,-]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ defined by $[X,Y] = \nu_2(X,Y)$ is called the **Lie bracket** on $\mathfrak{g}$.

So this justifies calling $\mathfrak{g}$ an algebra, if not a Lie algebra. To know that $\mathfrak{g}$ is a Lie algebra, we need to know that $[-,-]$ obeys the Jacobi identity.

Proposition 2.30. If $F\colon G \to H$ is a Lie group homomorphism and $X \in \mathfrak{g} = T_e(G)$, then $\exp(dF_e(X)) = F(\exp(X))$, that is, the following diagram commutes

$$\begin{tikzcd} G \arrow[r, "F"] & H \\ \mathfrak{g} \arrow[u, "\exp"] \arrow[r, "dF"] & \mathfrak{h} \arrow[u, "\exp"'] \end{tikzcd}$$

Proof. Let $\gamma(t) = F(\exp(tX))$. This gives us a line through the identity in G, so we get a Lie group homomorphism $\mathbb{R} \to H$. Take the derivative

$$\gamma'(0) = dF_e(X)$$

by the chain rule. Now from Proposition 2.22, any Lie group homomorphism $\phi\colon \mathbb{R} \to G$ is of the form $\exp(tY)$ for $Y = \phi'(0)$, so $\gamma(t) = \exp(t\, dF_e(X))$. Plug in $t = 1$ to get the proposition. □

Proposition 2.31. If G is a connected Lie group and $f, g\colon G \to H$ are homomorphisms then $f = g$ if and only if $df = dg$.

Proof. If $f = g$, then it's clear that $df = dg$.

Conversely, assume that $df = dg$. Then there is an open neighborhood U of e in G such that exp is invertible with inverse log. Then for $a \in U$, by Proposition 2.30, we have

$$\begin{aligned} f(a) &= f(\exp(\log(a))) \\ &= \exp(df_e(\log a)) \\ &= \exp(dg_e(\log(a))) \\ &= g(\exp(\log(a))) = g(a). \end{aligned}$$

So now by Corollary 2.10, it must be that f and g agree everywhere. □

Proposition 2.32. If $f\colon G \to H$ is a Lie group homomormphism then df is a homomorphism of Lie algebras. That is, $df([X, Y]) = [df X, df Y]$.

Proof. Take $X, Y \in \mathfrak{g}$ sufficiently close to zero. Then

$$f(\exp(X)\exp(Y)) = f(\exp(X)) f(\exp(Y)).$$

But also, if we expand the left hand side,

$$\begin{aligned} f(\exp(X)\exp(Y)) &= f(\exp(X + Y + \tfrac{1}{2}[X, Y] + \ldots)) \\ &= \exp(df_e(X + Y + \tfrac{1}{2}[X, Y] + \ldots))). \end{aligned}$$

On the right hand side,

$$\begin{aligned} f(\exp(X)) f(\exp(Y)) &= \exp(df_e(X)) \exp(df_e(Y)) \\ &= \exp\left(df_e X + df_e Y + \tfrac{1}{2}[df_e X, df_e Y] + \ldots\right) \end{aligned}$$

using Proposition 2.30 to pull the df inside on the left and the right. Therefore, we have that

$$\exp(df_e X + df_e Y + \tfrac{1}{2} df_e[X, Y] + \ldots)) = \exp\left(df_e X + df_e Y + \tfrac{1}{2}[df_e X, df_e Y] + \ldots\right)$$

Taking logs and comparing quadratic terms gives the result. □

Given $f\colon G \to H$, Proposition 2.30 tells us that $f \exp X = \exp(df_e(X))$, and Proposition 2.32 tells us that $df_e([x, y]) = [df_e X, df_e Y]$. We're going to use these to prove the Jacobi identity.

2.3 The Lie bracket revisited

Define a map $\psi_\bullet : G \to \mathrm{Aut}(G)$ by $g \mapsto \psi_g$, where ψ_g is the conjugation by g: $\psi_g(h) = ghg^{-1}$. We can easily check that ψ_g is a homomorphism, and that $gh \mapsto \psi_g \psi_h$, so the map $\psi_\bullet$ is a homomorphism as well.

Notice that $\psi_g(e) = e$ so that $d\psi_g : T_eG \to T_eG$. By the chain rule,

$$d\psi_{gh} = d\psi_g d\psi_h.$$

So the maps $d\psi_\bullet$ can be thought of as homomorphisms of groups $G \to \mathrm{GL}(T_eG) = \mathrm{GL}(\mathfrak{g})$. We call these $\mathrm{Ad}: G \to \mathrm{GL}(\mathfrak{g})$, and define them on $X = {}^{d}\!/_{dt}\big|_{t=0} h(t) \in \mathfrak{g}$ by

$$(\mathrm{Ad}\, g)X = \left.\frac{d}{dt}\right|_{t=0} g h(t) g^{-1} = gXg^{-1}.$$

In particular, notice that $\mathrm{Ad}\, e$ is the identity in $\mathrm{GL}(\mathfrak{g})$. So we can differentiate again at the identity to get a map $\mathrm{ad} = d\,\mathrm{Ad}: \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$.

Proposition 2.33. We have $(\mathrm{ad}\, X)(Y) = [X, Y]$. In particular, we have the Jacobi identity

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

Proof. By definition, $\mathrm{Ad}\, g = d\psi_g$. In order to compute $d\psi_g(Y)$ for $Y \in \mathfrak{g}$, we need to compute $\gamma'(0)$ for $\gamma(t) = g(\exp tY)g^{-1}$. Moreover, since $(d\exp)_0$ is the identity mapping on $\mathfrak{g}$, we may as well compute $\beta'(0)$ where $\beta = \exp^{-1} \circ \gamma$.

Now letting $g = \exp X$,

$$\begin{aligned}
\beta(t) &= \exp^{-1}(\exp X \exp(tY)) \exp(-x)) \\
&= \exp^{-1}(\exp(X + tY + \tfrac{1}{2}[X, tY] + \ldots) \exp(-X)) \\
&= \exp^{-1}(\exp(tY + \tfrac{1}{2}[X, tY] - \tfrac{1}{2}[tY, X] + \ldots)) \\
&= tY + [X, tY] + \big(\text{higher order terms}\big)
\end{aligned}$$

Thus, $\mathrm{Ad}(\exp X)(Y) = \beta'(0) = Y + [X, Y] +$ (higher order terms). By Proposition 2.30 we have

$$\mathrm{Ad}(\exp X) = \exp(\mathrm{ad}\, X) = 1 + \mathrm{ad}\, X + \tfrac{1}{2}(\mathrm{ad}\, X)^2 + \text{ (higher order terms)}$$

Comparing the two sides here after application to Y,

$$Y + [X, Y] + \text{ (higher order terms)} = Y + (\mathrm{ad}\, X)(Y) + \text{ (higher order terms)}$$

and therefore $(\mathrm{ad}\, X)(Y) = [X, Y]$ as required. Finally $\mathrm{ad} = d\,\mathrm{Ad}$ so ad is a Lie algebra homomorphism by Proposition 2.32. Hence, we get the Jacobi identity. □

Finally, let's see that the bracket on $\mathfrak{gl}_n$ was correct. Let $g(t)$ be a curve in G with $g'(0) = X$, and note that

$$0 = \frac{d}{dt} g(t) g(t)^{-1} = X + \left.\frac{d}{dt}\right|_{t=0} g(t)^{-1} \implies \left.\frac{d}{dt}\right|_{t=0} g(t)^{-1} = -X$$

Then,

$$\begin{aligned}(\operatorname{ad} X)(Y) &= (d\operatorname{Ad})(X)(Y)\\ &= \left.\frac{d}{dt}\right|_{t=0}(\operatorname{Ad} g(t))Y\\ &= \left.\frac{d}{dt}\right|_{t=0} g(t)Yg(t)^{-1}\\ &= XY + \left.\frac{d}{dt}\right|_{t=0} Yg(t)^{-1}\\ &= XY - YX\end{aligned}$$

3 Representations of $\mathfrak{sl}(2)$

One of the themes of Lie theory is to understand the representation theory of $\mathfrak{sl}(2)$, which can then be used to understand the representations of larger Lie algebras, which are in some sense built from a bunch of copies of $\mathfrak{sl}(2)$ put together. This is also a good flavor for other things we'll do later.

From now on, in this section, we'll work over $\mathbb{C}$. Recall

$$\mathfrak{sl}(2) = \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \;\middle|\; a, b, c \in \mathbb{C} \right\}$$

with the Lie bracket $[X, Y] = XY - YX$. There's an important basis for $\mathfrak{sl}(2)$, given by

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

These basis elements have relations

$$[H, X] = 2X, \qquad [H, Y] = -2Y, \qquad [X, Y] = H.$$

Example 3.1. What are some representations of $\mathfrak{sl}(2)$?

(1) The **trivial representation**, $\mathfrak{sl}(2) \to \mathfrak{gl}(1)$ given by $X, Y, H \mapsto 0$.

(2) The **natural/defining/standard representation** that comes from including $\mathfrak{sl}(2) \longhookrightarrow \mathfrak{gl}(2)$, wherein $\mathfrak{sl}(2)$ acts on $\mathbb{C}^2$ by the 2×2 matrices.

(3) For any Lie algebra $\mathfrak{g}$, the **adjoint representation** $\operatorname{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. For $\mathfrak{sl}(2)$, this is a map $\mathfrak{sl}(2) \to \mathfrak{gl}(3)$. Let's work out how this representation works on the basis.

	X	H	Y
$\operatorname{ad} X$	0	$-2X$	H
$\operatorname{ad} H$	$2X$	0	$-2Y$
$\operatorname{ad} Y$	$-H$	$2Y$	0

Therefore, the matrices of $\operatorname{ad} X, \operatorname{ad} Y$, and $\operatorname{ad} H$ in this representation are

$$\operatorname{ad} X = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \operatorname{ad} H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} \qquad \operatorname{ad} Y = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

(4) The map $\rho\colon \mathfrak{sl}(2) \to \mathfrak{gl}(\mathbb{C}[x,y])$ given by $X \mapsto x^{\partial}/_{\partial y}$ and $Y \mapsto y^{\partial}/_{\partial x}$, and $H \mapsto x^{\partial}/_{\partial x} - y^{\partial}/_{\partial y}$. Under ρ, the span of monomials of a given degree are stable.

- monomials of degree zero are constant functions, so this is just the trivial module.
- monomials of degree one $\lambda x + \mu y$ give the standard representation if we set $x = \binom{1}{0}$ and $y = \binom{0}{1}$.
- monomials of degree two give the adjoint representation.
- for monomials of degree k, denote the corresponding representation by Γ_k.

(5) $\Gamma_3 = \mathbb{C}\langle x^3, x^2y, xy^2, y^3\rangle$ and X, Y, H act on it as in the previous example. The matrices of the basis elements are

$$X = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad H = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \qquad Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

It turns out that all of the finite-dimensional irreducible representations of $\mathfrak{sl}(2)$ appear as the monomials of a fixed degree inside the representation on $\mathbb{C}[x,y]$. Notice that for Γ_3, the matrix of H is diagonal. It turns out that this is always the case: for any finite dimensional complex representation of $\mathfrak{sl}(2)$, say V, any diagonalizable element will map to another.

We can decompose V into eigenspaces for H, $V = \bigoplus_\lambda V_\lambda$, where λ is an eigenvalue of H which we will call a **weight** of the representation V.

Exercise 3.2. Check that the representation Example 3.1(4) is indeed a representation of $\mathfrak{sl}(2)$. That is, $\rho([A,B]) = \rho(A)\rho(B) - \rho(B)\rho(A)$ and apply to $f \in \mathbb{C}[x,y]$.

$\rho(H)$ is a diagonalizable element under any representation $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ with V finite dimensional over $\mathbb{C}$. Let $V = \bigoplus_\lambda V_\lambda$, be an eigenspace decomposition for V, where λ are the eigenvalues for the action of H. If $v \in V_\lambda$, then $Hv = \lambda v$.

We'll classify all the finite-dimensional complex irreducible representations of $\mathfrak{sl}(2)$. Let's start with an easy proposition.

Proposition 3.3. Let V be any $\mathfrak{sl}(2)$-module and let $v \in V_\alpha$. Then

(1) $Xv \in V_{\alpha+2}$;

(2) $Hv \in V_\alpha$;

(3) $Yv \in V_{\alpha-2}$.

Proof. (1) $HXv = XHv + [H,X]v = \alpha Xv + 2Xv = (\alpha+2)Xv$;

(2) $HHv = H\alpha v = \alpha Hv$;

(3) $HYv = YHv + [H, Y]v = \alpha Yv + 2Yv = (\alpha + 2)Yv$.

□

So that was easy. It won't get much harder.

Proposition 3.4. Let V be a finite dimensional representation of $\mathfrak{sl}(2)$, and let $v \in V_\alpha$ with $v \neq 0$. Then

(1) $Y^{n+1}v = 0$ for some $n \in \mathbb{N}$; and,

(2) if V is irreducible and $Xv = 0$ and n is minimal among such integers, then $V = \langle Y^n v, Y^{n-1}v, \ldots, v\rangle$.

(3) Further, with conditions as in (2), we have that as an $\langle H\rangle$-module,

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

Proof. First let's prove (1). Look at the set of $\{v, Yv, Y^2v, \ldots\}$. Because V is finite-dimensional, then we can choose $n \in \mathbb{N}$ minimal such that $v, Yv, \ldots, Y^{n+1}v$ are linearly dependant. Then, we can write

$$Y^{n+1}v = \sum_{i=0}^{n} a_i Y^i v.$$

Now apply $H - (\alpha - 2n - 2)I$ to this vector $Y^{n+1}v$. Proposition 3.3 says that $Y^{n+1}v$ is in the weight space $V_{\alpha-2n-2}$, and $H - (\alpha - 2n - 2)I$ should act as the zero operator on this weight space because every element in $V_{\alpha-2n-2}$ has H-eigenvalue $\alpha - 2n - 2$. Therefore,

$$0 = \sum_{i=0}^{n} a_i \left((\alpha - 2i) - (\alpha - 2n - 2)\right) Y^i v = \sum_{i=0}^{n} a_i (2(n - i + 1)) Y^i v.$$

Since no term $2(n - i + 1)$ is zero for $i = 0, \ldots, n$, we must have that $a_i = 0$ for all i, since the $\{Y^i v \mid 0 \leqslant i \leqslant n\}$ are linearly independent.

So $Y^{n+1}v = 0$. To establish the second claim, we use the following lemma.

□

Lemma 3.5. Let $v \in V_\alpha$, and assume $Xv = 0$. Then, $XY^m v = m(\alpha - m + 1)Y^{m-1}v$.

Proof. By induction on m. For base case $m = 1$,

$$XYv = YXv + [X, Y]v = 0 + Hv = \alpha v$$

For the inductive step,

$$\begin{aligned} XY^m V &= YXY^{m-1}v + [X, Y]Y^{m-1}v \\ &= Y(m - 1)(\alpha - m + 2)Y^{m-2}v + Y^{m-1}(\alpha - 2m + 2)v \\ &= ((m - 1)(\alpha - m + 2) + (\alpha - 2m + 2))Y^{m-1}v \\ &= m(\alpha - m + 1)Y^{m-1}v \end{aligned}$$

□

Proof of Proposition 3.4 continued. Now, given this lemma, we can prove (2). Observe that by the lemma, for $W = \langle v, Yv, \dots, Y^n v\rangle$, we have $XW \subset W$. Also by the previous result $HW \subset W$, and clearly $YW \subset W$. Therefore, W is an irreducible subspace of V and because V is irreducible, then $W = V$.

Finally, let's prove (3). Putting $m = n+1$ into Lemma 3.5, we get that $0 = (n+1)(\alpha - n)Y^n v$, so $\alpha = n$. ☐

The nice thing about $\mathfrak{sl}(2)$ representations is that they're parameterized by integers, so we can draw them! We go about it like this. For $V = \bigoplus_{i=0}^{n} V_{-n+2i}$ a decomposition into weight spaces, the picture is:

$$\begin{tikzcd}
V_{-n} \arrow[loop left, "H"] \arrow[r, shift left, "X"] & V_{-n+2} \arrow[loop above, "H"] \arrow[l, shift left, "Y"] \arrow[r, shift left, "X"] & \cdots \arrow[l, shift left, "Y"] \arrow[r, shift left, "X"] & V_{n-2} \arrow[loop above, "H"] \arrow[l, shift left, "Y"] \arrow[r, shift left, "X"] & V_n \arrow[l, shift left, "Y"] \arrow[loop right, "H"]
\end{tikzcd}$$

Corollary 3.6. To summarize what we've seen so far,

(1) each finite-dimensional irreducible representation has an eigenvector for H with a maximal possible integral value, which we call the **highest weight**.

(2) Any two irreducible modules of the same highest weight r are isomorphic. We call such a module Γ_r.

(3) Γ_r has dimension $r+1$.

(4) The eigenvalues of H are equally spaced around 0.

(5) All eigenspaces are 1-dimensional.

(6) All even steps between n and $-n$ are filled.

(7) We can reconstruct V by starting with a highest weight vector $v \neq 0$ such that $X_v = 0$. Then $V = \langle v, Yv, \dots, Y^n v\rangle$.

It follows from the theory of associative algebras that any $\mathfrak{g}$-module has a Jordan-Hölder series. That is, given any finite-dimensional representation W of $\mathfrak{sl}(2)$, we can explicitly decompose into composition factors (irreducible subquotients) by the following algorithm:

(1) identify a highest weight vector v of weight r, say;

(2) generate a submodule $\langle v\rangle$ from this vector;

(3) write down a composition factor Γ_r;

(4) repeat after replacing W by W/Γ_r.

This gives us a decomposition of W into irreducible factors.

Example 3.7.

(1) Consider the standard representation $\mathfrak{sl}(2) \hookrightarrow \mathfrak{gl}(2)$. This decomposes as $V = V_{-1} \oplus V_1$.

(2) The adjoint representation decomposes with weight spaces $\{-2, 0, 2\}$, as $V = V_{-2} \oplus V_0 \oplus V_2$.

(3) The representation W on degree 3 polynomials in $\mathbb{C}[x, y]$ has weight spaces $V_{-3} \oplus V_{-1} \oplus V_1 \oplus V_3$.

(4) Consider the $\mathfrak{sl}(2)$-submodule V of $\mathbb{C}[x, y]$ generated by all monomials of degree at most three. H acts as $x^{\partial}/_{\partial x} - y^{\partial}/_{\partial y}$, and sends $x^i y^j$ to $(i - j)x^i y^j$, and so we can calculate the weights on the basis $\{x^3, x^2y, xy^2, y^3, x^2, xy, y^2, x, y, 1\}$. The weights are (with multiplicity) $3, 2, 1, 1, 0, 0, -1, -1, -2, -3$, and

$$V = V_3 \oplus V_2 \oplus V_1 \oplus V_0 \oplus V_{-1} \oplus V_{-2} \oplus V_{-3}.$$

There is a factor of Γ_3 as $V_3 \oplus V_1 \oplus V_{-1} \oplus V_{-3}$, and from there we can decompose further.

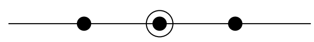
This is remarkable! For most finite simple groups, we can't classify their representations over $\mathbb{C}$. Not even for finite groups of Lie type. So this is really a simple representation theory, and remarkably it's complete.

Example 3.8. Recall that Γ_1 is the standard representation $\langle x \rangle \oplus \langle y \rangle$, where y has eigenvalue -1 and x has eigenvalue $+1$.

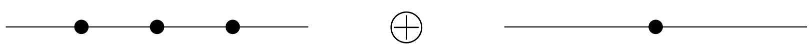
A basis for the tensor product $\Gamma_1 \otimes \Gamma_1$ is $x \otimes x, x \otimes y, y \otimes x, y \otimes y$. The action of H on this module is $Hx = x$ and $Hy = -y$. Then,

$$\begin{aligned} H(x \otimes x) &= Hx \otimes x + x \otimes Hx = 2x \otimes x \\ H(x \otimes y) &= 0 \\ H(y \otimes x) &= 0 \\ H(y \otimes y) &= -2y \otimes y \end{aligned}$$

The weight diagram is



This decomposes as $\Gamma_2 \oplus \Gamma_0$.



So $\Gamma_1 \otimes \Gamma_1 = \Gamma_2 \oplus \Gamma_0$. We can even find a basis for Γ_2 and Γ_0 in this manner: for Γ_2, the basis is $\{x \otimes x, x \otimes y + y \otimes x, y \otimes y\}$ and Γ_0 has basis $\{x \otimes y - y \otimes x\}$.

Observe more generally that in tensor product one simply "adds weights." Let V and W be $\mathfrak{sl}(2)$-modules. If $\{v_1, \ldots, v_n\}$ is a basis of H-eigenvectors of V and $\{w_1, \ldots, w_m\}$ is a basis of H-eigenvectors of W, then $\{v_i \otimes w_j\}$ is a bsis of H-eigenvectors of $V \otimes W$. If $Hv_i = \lambda_i v_i$ and $Hw_j = \mu_j w_j$, then $H(v_i \otimes w_j) = (\lambda_i + \mu_j) v_i \otimes w_j$.

4 Major Results on Lie Algebras

Definition 4.1. Let $\mathfrak{g}$ be a Lie algebra. Then the **derived subalgebra** $\mathcal{D}(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ is the span of the commutators of $\mathfrak{g}$. Inductively, we define the **lower central series** $\mathcal{D}_0(\mathfrak{g}) = \mathfrak{g}$, $\mathcal{D}_1(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$, and $\mathcal{D}_k(\mathfrak{g}) = [\mathfrak{g}, D_{k-1}\mathfrak{g}]$.

Similarly, the **upper central series** $\mathcal{D}^i(\mathfrak{g})$ is given by $\mathcal{D}^0(\mathfrak{g}) = \mathfrak{g}$, $\mathcal{D}^k(\mathfrak{g}) = [\mathcal{D}^{k-1}(\mathfrak{g}), \mathcal{D}^{k-1}(\mathfrak{g})]$.

It will be important for us that $D^i(\mathfrak{g})$ and $D_i(\mathfrak{g})$ are **characteristic ideals**, meaning that they are stable under all derivations of $\mathfrak{g}$.

Proposition 4.2.

(1) $\mathcal{D}_k(\mathfrak{g}) \subseteq \mathcal{D}_{k-1}(\mathfrak{g})$;

(2) $\mathcal{D}_k(\mathfrak{g})$ is a characteristic ideal;

(3) ${}^{\mathcal{D}_k(\mathfrak{g})}/_{\mathcal{D}_{k+1}(\mathfrak{g})}$ is abelian;

(4) ${}^{\mathcal{D}_k(\mathfrak{g})}/_{\mathcal{D}_{k+1}(\mathfrak{g})}$ is central in ${}^{\mathfrak{g}}/_{\mathcal{D}_{k+1}(\mathfrak{g})}$.

Proof.

(1) By induction. For $k = 1$, we have $\mathcal{D}(\mathfrak{g}) \subseteq \mathfrak{g}$ clearly. Given $\mathcal{D}_k(\mathfrak{g}) \subseteq \mathcal{D}_{k-1}(\mathfrak{g})$, we can take brackets on both sides to see that
$$\mathcal{D}_{k+1}(\mathfrak{g}) = [\mathfrak{g}, \mathcal{D}_k(\mathfrak{g})] \subseteq [\mathfrak{g}, \mathcal{D}_{k-1}(\mathfrak{g})] = \mathcal{D}_k(\mathfrak{g}).$$

(2) To see that it's an ideal, for each k we have that
$$[X, \mathcal{D}_k(\mathfrak{g})] \subseteq \mathcal{D}_{k+1}(\mathfrak{g}) \subseteq \mathcal{D}_k(\mathfrak{g})$$
for all $X \in \mathfrak{g}$. To see that this ideal is characteristic, let $\alpha \in \mathrm{Der}(\mathfrak{g})$. Then
$$\alpha(\mathcal{D}_{k+1}(\mathfrak{g})) = \alpha\left([\mathfrak{g}, \mathcal{D}_k(\mathfrak{g})]\right) = [\alpha(\mathfrak{g}), \mathcal{D}_k(\mathfrak{g})] + [\mathfrak{g}, \alpha(\mathcal{D}_k(\mathfrak{g}))]$$
But $\alpha(\mathfrak{g}) \subseteq \mathfrak{g}$, so $[\alpha(\mathfrak{g}), \mathcal{D}_k(\mathfrak{g})] \subseteq \mathcal{D}_{k+1}(\mathfrak{g})$. By induction,
$$[\mathfrak{g}, \alpha(\mathcal{D}_k(\mathfrak{g}))] \subseteq \mathcal{D}_k(\mathfrak{g}) \subseteq D_{k+1}(\mathfrak{g}).$$
Hence, $\alpha(\mathcal{D}_{k+1}(\mathfrak{g})) \subseteq \mathcal{D}_{k+1}(\mathfrak{g})$.

(3) If $X + \mathcal{D}_{k+1}(\mathfrak{g})$, $Y + \mathcal{D}_{k+1}(\mathfrak{g})$ are elements of ${}^{\mathcal{D}_k(\mathfrak{g})}/_{\mathcal{D}_{k+1}(\mathfrak{g})}$, then
$$[X + \mathcal{D}_{k+1}(\mathfrak{g}), Y + \mathcal{D}_{k+1}(\mathfrak{g})] = [X, Y] + \mathcal{D}_{k+1}(\mathfrak{g}) = 0 + \mathcal{D}_{k+1}(\mathfrak{g})$$
as required, because $X \in \mathfrak{g}$ and $Y \in \mathcal{D}_k(\mathfrak{g})$, so $[X, Y] \in \mathcal{D}_{k+1}(\mathfrak{g})$.

(4) Let $X \in \mathfrak{g}$, and $Y \in \mathcal{D}_k(\mathfrak{g})$. Then
$$[X + \mathcal{D}_{k+1}(\mathfrak{g}), Y + \mathcal{D}_{k+1}(\mathfrak{g})] = [X, Y] + \mathcal{D}_{k+1}(\mathfrak{g}).$$
Yet $[X, Y] \in [\mathfrak{g}, \mathcal{D}_k(\mathfrak{g})] = \mathcal{D}_{k+1}(\mathfrak{g})$. Hence,
$$[X + \mathcal{D}_{k+1}(\mathfrak{g}), Y + \mathcal{D}_{k+1}(\mathfrak{g})] = 0. \qquad \square$$

Proposition 4.3.

(1) $\mathcal{D}^k(\mathfrak{g}) \subseteq \mathcal{D}^{k-1}(\mathfrak{g})$;

(2) $\mathcal{D}^k(\mathfrak{g})$ is a characteristic ideal;

(3) ${}^{\mathcal{D}^k(\mathfrak{g})}\!/\!_{\mathcal{D}^{k+1}(\mathfrak{g})}$ is abelian;

(4) $\mathcal{D}^k(\mathfrak{g}) \subseteq \mathcal{D}_k(\mathfrak{g})$.

Exercise 4.4. Prove Proposition 4.3.

Definition 4.5. If $\mathcal{D}_k(\mathfrak{g}) = 0$ for some k, then we say that $\mathfrak{g}$ is **nilpotent.** If on the other hand $\mathcal{D}^k(\mathfrak{g}) = 0$ for some k, then we say that $\mathfrak{g}$ is **solvable.** If $\mathfrak{g}$ has no solvable ideals then we say that $\mathfrak{g}$ is **semisimple**.

Remark 4.6. Everyone seems to use the term "solvable" nowadays, which is an unfortunate Americanism. If you say "soluble," you will be understood.

Note that a nilpotent Lie algebra is necessarily solvable by Proposition 4.3(4), but a solvable Lie algebra need not be nilpotent.

Remark 4.7. For much of this chapter, we will have $\mathfrak{g} \subset \mathfrak{gl}(V)$. There is a theorem due to Ado which guarantees that there is a faithful representation $\mathfrak{g} \to \mathfrak{gl}(V)$. As such, we have the notion of an element $X \in \mathfrak{g}$ being V-nilpotent if it is a nilpotent endomorphism of V. This is distinct from $\mathfrak{g}$ being nilpotent as a Lie algebra.

Theorem 4.8 (Engel's Theorem)**.** Let k be an arbitrary field, and let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie algebra such that every element of $\mathfrak{g}$ is nilpotent (for every $X \in \mathfrak{g}$, there is N such that $X^N = 0$). Then there is some nonzero $v \in V$ such that $Xv = 0$ for all $X \in \mathfrak{g}$.

To prove this, we'll need a lemma.

Lemma 4.9. If $X \in \mathfrak{gl}(V)$ is nilpotent, then $\operatorname{ad} X \in \operatorname{End}(\mathfrak{g})$ is nilpotent.

Proof. Suppose $X^N = 0$. Then by induction, one can show that

$$(\operatorname{ad} X)^M Y = \sum_{i=0}^{M} (-1)^i \binom{M}{i} X^{M-i} Y X^i$$

Now take $M = 2N - 1$. Then either $M - i$ or i is bigger than N, so the right hand side vanishes identically for all Y. Hence, $(\operatorname{ad} X)^{2N-1} = 0$. □

Now that we're given this, the proof of Engel's theorem is a clever application of linear algebra by an induction argument. This is not the way it was first proved, but the proof has been cleaned up over the years to be much more elegant.

Proof of Theorem 4.8. Induction on the dimension of $\mathfrak{g}$. Let $\dim \mathfrak{g} = n$.

For $n = 1$, if $\mathfrak{g} = \langle X \rangle$, then suppose $X^N = 0$ but $X^{N-1} \neq 0$. Then there is some nonzero $v \in V$ such that $X^{N-1}v \neq 0$. And so $X^{N-1}v$ is our desired vector.

For $n > 1$, now assume that we have the result for all Lie algebras $\mathfrak{h}$ with $\dim \mathfrak{h} < n$. Claim that $\mathfrak{g}$ has a codimension 1 ideal.

To prove this claim, let $\mathfrak{h}$ be any maximal proper subalgebra. Since the subalgebra generated by one element is a proper subalgebra when $\dim \mathfrak{g} > 1$, and $\mathfrak{h}$ is maximal, it cannot be that $\mathfrak{h} = 0$.

Let $\mathfrak{h}$ act on $\mathfrak{g}$ by the adjoint action. By Lemma 4.9, $\operatorname{ad} \mathfrak{h} \subseteq \mathfrak{gl}(\mathfrak{g})$ consists of nilpotent endomorphisms of $\mathfrak{g}$, and indeed also of $\mathfrak{g}/\mathfrak{h}$. Note that $\dim {}^{\mathfrak{g}}/_{\mathfrak{h}} < n$ since $\dim \mathfrak{h} > 1$, so by induction there is $Y \in \mathfrak{g} \backslash \mathfrak{h}$ such that $Y + \mathfrak{h}$ is killed by $\mathfrak{h}$. In particular, $(\operatorname{ad} h)Y \subseteq \mathfrak{h}$, so $\mathfrak{h} \oplus \langle Y \rangle$ is a subalgebra of $\mathfrak{g}$. But $\mathfrak{h}$ is maximal among proper subalgebras, so $\mathfrak{g} \cong \mathfrak{h} \oplus Y$ as a vector space. Thus, $\mathfrak{h}$ is a codimension 1 ideal.

Let $W = \{v \in V \mid \mathfrak{h}v = 0\}$. This is nontrivial by the inductive hypothesis and $\dim h < \dim \mathfrak{g}$. But now $YW \subseteq W$ because for any $X \in \mathfrak{h}$, $w \in W$, then $XYw = YXw + [X, Y]w$. $YXw = 0$ because $X \in \mathfrak{h}$ and $w \in W$, and $[X, Y]w = 0$ because $[X, Y] = (\operatorname{ad} X)Y = 0$.

Now Y is nilpotent on W, so $Y^N = 0$ but $Y^{N-1} \neq 0$ for some N. Thus, there is $w \in W$ such that $Y^{N-1}w \neq 0$, but $Y(Y^{N-1}w) = 0$. Therefore, $\mathfrak{g}(Y^{N-1}w) = 0$. □

Remark 4.10. This is basically the only theorem we'll talk about that works over fields of arbitrary characteristic. The rest of the theorems we'll talk about will fail in general, or at least for positive characteristic.

Last time we proved Engel's theorem. Before we move on, let's point out a corollary to this.

Corollary 4.11. Under the same hypotheses of Engel's theorem, Theorem 4.8, then there is a basis of V with respect to which all elements of $\mathfrak{g}$ can be represented by strictly upper triangular matrices.

Proof. Theorem 4.8 guarantees a nonzero $v \in V$ such that $Xv = 0$ for all $X \in \mathfrak{g}$. Now by induction on dimension, there is a basis $v_2 + \langle v \rangle, \dots, v_n + \langle v \rangle$ of the quotient module of $V/\langle v \rangle$ satisfying the conclusion. That is, $Xv_i + \langle v \rangle \in \langle v_{i+1}, \dots, v_n \rangle + \langle v \rangle$, so $\{v, v_2, \dots, v_m\}$ is the desired basis. □

Now we'll do the other major theorem of Lie algebras that allows the theory of complex semisimple Lie algebras to go so far with so little work.

Theorem 4.12 (Lie's Theorem)**.** Let $k = \mathbb{C}$. Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra. Suppose $\mathfrak{g}$ is solvable. Then there is a common eigenvector for all of the elements of $\mathfrak{g}$.

Proof. By induction on the dimension of $\mathfrak{g}$. Let $n = \dim \mathfrak{g}$.

If $n = 1$, this is trivial because for any nonzero $X \in \mathfrak{g}$, the fact that $\mathbb{C}$ is algebraically closed guarantees an eigenvector. Any other nonzero element of $\mathfrak{g}$ is a multiple of X, and therefore shares this eigenvector.

Now assume the result for all $\mathfrak{h}$ with $\dim \mathfrak{h} < n$. We first find a codimension 1 ideal of $\mathfrak{g}$. For this, observe that $\mathcal{D}(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ is strictly contained in $\mathfrak{g}$ by solvability (if not, then it's never the case that $\mathcal{D}^k(\mathfrak{g})$ is zero). Also observe that the quotient $\mathfrak{g}/\mathcal{D}(\mathfrak{g})$ is abelian. Now by the first isomorphism theorem, each ideal of $\mathfrak{g}/\mathcal{D}(\mathfrak{g})$ corresponds to an ideal of $\mathfrak{g}$ containing $\mathcal{D}(\mathfrak{g})$.

Any subspace of $\mathfrak{g}/\mathcal{D}(\mathfrak{g})$ is an ideal since it's abelian. So let $\mathfrak{h}$ be the lift of any codimension 1 subspace of $\mathfrak{g}/\mathcal{D}(\mathfrak{g})$, and this is the required codimension 1 ideal of $\mathfrak{g}$. Observe that $\mathfrak{h}$ is also solvable since $\mathcal{D}^k(\mathfrak{h}) \subseteq \mathcal{D}^k(\mathfrak{g})$. So we can apply the inductive hypothesis to obtain a nonzero $v \in V$ such that for all $X \in \mathfrak{h}$, $Xv = \lambda(X)v$ for some $\lambda \in \mathfrak{h}^* = \mathrm{Hom}_{\mathbb{C}}(\mathfrak{h}, \mathbb{C})$.

Take $Y \in \mathfrak{g}\backslash\mathfrak{h}$ such that $\mathfrak{g} = kY \oplus \mathfrak{h}$ as a vector space.

Let $W = \{w \in V \mid Xw = \lambda(X)w \ \forall X \in \mathfrak{h}\}$. We know that W is nonempty because we have found one such vector for which $Xv = \lambda(X)v$ by applying our inductive hypothesis.

We want to see that $YW \subseteq W$. If we can do this, then considering Y as a linear transformation on W, Y has an eigenvector in W. Since $\mathfrak{g} = \langle Y \rangle \oplus \mathfrak{h}$ as vector spaces, $\mathfrak{g}\langle w \rangle = \langle w \rangle$. The fact that $YW \subseteq W$ will follow from Lemma 4.13 (which is more general). □

Lemma 4.13. Let $\mathfrak{h}$ be an ideal in $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ with $\lambda \colon \mathfrak{h} \to \mathbb{C}$ a linear functional. Set $W = \{v \in V \mid Xv = \lambda(X)v \ \forall X \in \mathfrak{h}\}$. Then $YW \subseteq W$ for any $Y \in \mathfrak{g}$.

Proof. Apply $X \in \mathfrak{h}$ to Yv for some $v \in W$. We have that

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

We want to show now that $\lambda([X, Y]) = 0$. This is a bit of work. Take $w \in W$ and consider $U = \langle w, Yw, Y^2w, \ldots \rangle$. Clearly, $YU \subseteq U$. We claim that $XU \subseteq U$ for all $X \in \mathfrak{h}$, and according to a basis $\{w, Yw, Y^2w, \ldots, Y^iw\}$ for U, X is represented by an upper triangular matrix with $\lambda(X)$ on the diagonal.

We prove this claim by induction on i. For $i = 0$, $Xw = \lambda(X)w \in U$.

Now for $k \leqslant i$,

$$XY^kw = YXY^{k-1}w + [X, Y]Y^{k-1}w.$$

Note that $[X, Y] \in \mathfrak{h}$ and $Y^{k-1}w$ is a previous basis vector, so by induction we may express $[X, Y]Y^{k-1}w$ as a linear combination of $w, Yw, Y^2w, \ldots, Y^{k-1}w$.

$$[X, Y]Y^{k-1}w = \sum_{i=0}^{k-1} a_i Y^i w. \tag{4}$$

And

$$YXY^{k-1}w = Y\lambda(X)Y^{k-1}w + \big(\text{linear combination of } w, Yw, Y^2w, \ldots Y^{k-1}\,\big) \tag{5}$$

So by (4) and (5), we have that

$$XY^kw = \lambda(X)Y^kw + \big(\text{linear combination of } w, Yw, Y^2w, \ldots Y^{k-1}\,\big).$$

Therefore, according to the basis $\{w, Yw, \dots, Y^i w\}$, X looks like

$$\begin{bmatrix} \lambda(X) & * & * & * \\ & \lambda(X) & * & * \\ & & \ddots & * \\ & & & \lambda(X) \end{bmatrix}$$

Therefore, for any $X \in \mathfrak{h}$, $\operatorname{tr} X|_U = (\dim U)\lambda(X)$. This holds in particular for $[X, Y]$, so

$$\operatorname{tr}([X, Y]|_U) = (\dim U) \cdot \lambda([X, Y])$$

but the trace of a commutator is zero. So we get that $\lambda([X, Y]) = 0$, as required. $\square$

Corollary 4.14. Let V be a $\mathfrak{g}$-module where $\mathfrak{g}$ is solvable. Then there is a basis of V with respect to which $\mathfrak{g}$ acts by upper triangular matrices.

The proof of this theorem is very similar to the proof of Corollary 4.11, so we won't repeat it here. It's in the notes.

Proposition 4.15 (Jordan Decomposition)**.** Let $X \in \mathfrak{gl}(V)$. Then there exist polynomials $P_s(t)$, $P_n(t)$ such that

(1) $X_s = P_s(X)$ is diagonalizable, $X_n = P_n(X)$ is nilpotent, and $X = X_s + X_n$.

(2) $[X_s, X_n] = [X, X_n] = [X, X_s] = 0$

(3) If $A \in \mathfrak{gl}(V)$ for which $[X, A] - 0$, then $[X_n, A] = [X_s, A] = 0$.

(4) If $XW \subseteq W$ for any $W \subseteq V$, then $X_n W \subseteq W$ and $X_s W \subseteq W$

(5) If D and N are such that $[D, N] = 0$ and $X = D + N$ with D diagonalizable, N nilpotent, then $D = X_s$ and $N = X_n$.

Proof. The hard part of this theorem is constructing the polynomials. Everything else follows from that. Let $\chi_X(t)$ be the characteristic polynomial of X. We can factor this as

$$\chi_X(t) = (t - \lambda_1)^{e_1}(t - \lambda_2)^{e_2} \cdots (t - \lambda_r)^{e_r},$$

where the λ_i are the distinct eigenvalues of X. Note that $(t - \lambda_i)$ is coprime to $(t - \lambda_j)$ for all $i \neq j$.

Then by the Chinese Remainder Theorem, we can find a polynomial $P_s(t)$ such that

$$P_s(t) \equiv \lambda_i \pmod{(t - \lambda_i)^{e_i}}$$

for all i. Define further $P_n(t) = t - P_s(t)$. Let $X_s = P_s(X)$ and $X_n = P_n(X)$.

(1) Clearly we have that $X = X_s + X_n$, since $t = P_s(t) + P_n(t)$. Let $V_i = \ker(X - \lambda_i I)^{e_i}$ be the generalized eigenspaces of X, and note that $V = \bigoplus_i V_i$. Since $X_s = P_s(X) = \lambda_i + (X - \lambda_i I)^{e_i} q_i(X)$ for some $q_i(X)$, we have that X_s acts diagonalizably on V_i with eigenvalue λ_i. By definition $X_n = X - X_s$ so $X_n|_{V_i} = X - \lambda_i I$. So X_n is nilpotent as required.

(2) Since X_s and X_n are polynomials in X, then $[X_s, X_n] = 0$.

(3) If $A \in \mathfrak{gl}(V)$ commutes with X, then it also commutes with X_s and X_n because they are polynomial in X.

(4) Likewise, if $W \subseteq V$ is stable under X, it is also stable under X_n and X_s.

(5) We have $X_s - D = N - X_n$ with everything in sight commuting. So $X_s - D$ is both diagonalizable and nilpotent, and is thus zero. $\square$

Observe that if X is a matrix in Jordan block form

$$X = \begin{bmatrix} \lambda_1 & 1 & & & & & & \\ & \lambda_1 & \ddots & & & & & \\ & & \ddots & 1 & & & & \\ & & & \lambda_1 & 0 & & & \\ & & & & \lambda_2 & 1 & & \\ & & & & & \ddots & \ddots & \\ & & & & & & \lambda_n & 1 \end{bmatrix}$$

with the λ_i not necessarily distinct, then X_s has λ_i on the diagonal.

$$X_s \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & & \\ & & & \lambda_2 & & \\ & & & & \ddots & \\ & & & & & \lambda_n \end{bmatrix}$$

and X_n is the matrix with 1's immediately above the diagonal.

Corollary 4.16. For $X \in \mathfrak{gl}(V)$, we have $(\operatorname{ad} X)_n = \operatorname{ad} X_n$ and $(\operatorname{ad} X)_s = \operatorname{ad} X_s$.

Proof. According to some basis $\{v_1, \ldots, v_n\}$ of V, X_s is diagonal with entries $d_1, \ldots, d_n$. Relative to this basis, let E_{ij} be the standard basis of $\mathfrak{gl}(V)$. Then calculate

$$(\operatorname{ad} X_s)(E_{ij}) = [X_s, E_{ij}] = (d_i - d_j)E_{ij}$$

So the E_{ij} are a basis of eigenvectors for $\operatorname{ad} X_s$. So $\operatorname{ad} X_s$ is diagonalizable.

Furthermore, $\operatorname{ad} X_n$ is nilpotent by Lemma 4.9. Also,

$$\operatorname{ad} X = \operatorname{ad}(X_s + X_n) = \operatorname{ad} X_s + \operatorname{ad} X_n$$

and as $[X_n, X_s] = 0$, then $[\operatorname{ad} X_n, \operatorname{ad} X_s] = \operatorname{ad}[X_n, X_s] = 0$. $\square$

We've seen that taking traces can be a useful tool. This continues to be the case, and is formalized in the following definition.

Definition 4.17. Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of a Lie algebra $\mathfrak{g}$. Then the **Killing form with respect to ρ (or V)** is the symmetric bilinear form given by

$$B_V(X,Y) = \operatorname{tr}_V \rho(X)\rho(Y)$$

for $X, Y \in \mathfrak{g}$. When $\rho = \operatorname{ad}$ and $V = \mathfrak{g}$, we have **'the' Killing form**

$$B(X,Y) = \operatorname{tr} \operatorname{ad} X \operatorname{ad} Y$$

Remark 4.18 (Historical interlude)**.** Killing invented Lie algebras independently of Lie when he was thinking about the infinitesimal transformations of a space, whereas Lie wanted to study differential equations. Killing is more-or-less responsible for the program of the classification of Lie algebras, but it was completed by Élie Cartan.

Note that the Killing form is a symmetric bilinear form; this isn't too hard to see because the trace is linear and the definition is symmetric. The Killing form has the nice property that it's invariant under the adjoint action of $\mathfrak{g}$.

Proposition 4.19. $B_\rho([X,Y],Z) = B_\rho(X,[Y,Z])$.

Proof. Use the cyclic invariance of trace.

$$\begin{aligned}
\operatorname{tr}(\rho[X,Y]\rho Z) &= \operatorname{tr}\left((\rho(X)\rho(Y) - \rho(Y)\rho(X))\rho(Z)\right) \\
&= \operatorname{tr}\left(\rho(X)\rho(Y)\rho(Z)\right) - \operatorname{tr}(\rho(Y)\rho(X)\rho(Z)) \\
&= \operatorname{tr}\left(\rho(X)\rho(Y)\rho(Z)\right) - \operatorname{tr}\left(\rho(X)\rho(Z)\rho(Y)\right) \\
&= \operatorname{tr}\left(\rho(X)(\rho(Y)\rho(Z) - \rho(Z)\rho(Y))\right) \\
&= \operatorname{tr}\left(\rho(X)\rho[Y,Z]\right)
\end{aligned}$$ ∎

Theorem 4.20 (Cartan's Criterion)**.** Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra. If $B_V(X,Y) = \operatorname{tr}(XY)$ is identically zero on $\mathfrak{g} \times \mathfrak{g}$, then $\mathfrak{g}$ is solvable.

Proof. It suffices to show that every element of $\mathcal{D}(\mathfrak{g})$ is nilpotent, since then by Corollary 4.11, we have that some basis of V with respect to which all elements of $\mathfrak{g}$ can be represented by strictly upper triangular matrices, and repeated commutators of strictly upper triangular matrices eventually vanish. More precisely, if every element of $\mathcal{D}(\mathfrak{g})$ is nilpotent, then $\mathcal{D}(\mathfrak{g})$ is a nilpotent ideal, so $\mathcal{D}_k(\mathcal{D}(\mathfrak{g})) = 0$ for some k. Now by induction $\mathcal{D}^i(\mathfrak{g}) \subseteq \mathcal{D}_i(\mathfrak{g})$, so $\mathcal{D}^{k+1}(\mathfrak{g}) = 0$.

So take $X \in \mathcal{D}(\mathfrak{g})$ and write $X = D + N$ for D diagonalizable and N nilpotent. Work with a basis such that D is diagonal, say with entries $\lambda_1, \ldots, \lambda_n$. We will show that

$$\operatorname{tr} \overline{D} X = \sum_i \lambda_i \overline{\lambda_i} = 0.$$

where $\overline{D}$ is complex-conjugate matrix of D. It suffices to show this because $\lambda_i \overline{\lambda}_i$ is always nonnegative, and a sum of nonnegative things is only zero when each is zero individually.

Since X is a sum of commutators, $[Y_i, Z_i]$ say, it will suffice to show that

$$\operatorname{tr}(\overline{D}[Y,Z]) = 0$$

for $Y, Z \in \mathfrak{g}$. But

$$\operatorname{tr}(\overline{D}[Y,Z]) = \operatorname{tr}([\overline{D}, Y]Z)$$

by Proposition 4.19. By hypothesis, we will be done if we can show that $\operatorname{ad} \overline{D}$ takes $\mathfrak{g}$ to itself, in which case we say that $\overline{D}$ **normalizes** $\mathfrak{g}$.

Since $\operatorname{ad} D = \operatorname{ad} X_s = (\operatorname{ad} X)_s$ is a polynomial in $\operatorname{ad} X$ by Corollary 4.16, we have that $\operatorname{ad} D$ normalizes $\mathfrak{g}$. Taking a basis of $\mathfrak{gl}(V)$ relative to which $\operatorname{ad} D$ is diagonal, $\operatorname{ad} \overline{D}$ is also diagonal with eigenvalues the complex conjugates of the eigenvalues of $\operatorname{ad} D$, and moreover they stabilize the same subspaces. In particular, they stabilize $\mathfrak{g}$. □

Remark 4.21 (Very very tangential aside)**.** This proof is kind of cheating. We proved it for specifically the complex numbers. The statement is true for any algebraically closed field of characteristic zero, but we can use the **Lefschetz principle** that says that any statement in any first order model theory that holds for any algebraically closed field of characteristic zero is true for all such fields. So really we should check that we can express this statement in some first order model theory. But this remark is safe to ignore for our purposes.

Corollary 4.22. A Lie algebra $\mathfrak{g}$ is solvable if and only if $B(\mathfrak{g}, \mathcal{D}(\mathfrak{g})) = 0$.

Proof. Assume first that $\mathfrak{g}$ is solvable and consider the adjoint representation of $\mathfrak{g}$. By the corollary to Lie's theorem, Corollary 4.14, there is a basis of $\mathfrak{g}$ relative to which each endomorphism $\operatorname{ad} X$ is upper triangular. But now $\operatorname{ad}[X, Y] = [\operatorname{ad} X, \operatorname{ad} Y]$ and the commutator of any two upper triangular matrices is strictly upper triangular.

So for $X \in \mathfrak{g}$, $Y \in \mathcal{D}(\mathfrak{g})$, the previous paragraph shows that Y is strictly upper triangular, as the sum of commutators, so $\operatorname{ad} Y$ is as well. And by our choice of basis $\operatorname{ad} X$ is upper triangular. The product of an upper-triangular matrix and strictly upper-triangular matrix is strictly upper triangular, so

$$B(X, Y) = \operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y) = 0$$

Conversely, assume $B(\mathfrak{g}, \mathcal{D}(\mathfrak{g}))$ is identically zero. Then $B(\mathcal{D}(\mathfrak{g}), \mathcal{D}(\mathfrak{g})) = 0$ and so by Cartan's Criterion (Theorem 4.20), we have that $\operatorname{ad} \mathcal{D}(\mathfrak{g})$ is solvable. Then, $\mathcal{D}^k(\operatorname{ad} \mathcal{D}(\mathfrak{g})) = 0$ for some k. But ad is a Lie algebra homomorphism, so $\operatorname{ad} \mathcal{D}^{k+1}(\mathfrak{g}) = 0$ as well.

Therefore, $\mathcal{D}^k(\mathfrak{g}) \subseteq \ker \operatorname{ad}$, and $\ker \operatorname{ad}$ is abelian. So $\mathcal{D}^{k+1}(\mathfrak{g})$ is abelian. Hence $\mathcal{D}^{k+2}(\mathfrak{g}) = 0$. □

"The Killing Form", which sounds kind of like a television detective drama. Previously on the Killing Form, we saw Cartan's Criterion: if $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ and B_V is identically zero on $\mathfrak{g}$, then $\mathfrak{g}$ is solvable. We also showed that $\mathfrak{g}$ is solvable if and only if $B(\mathfrak{g}, \mathcal{D}(\mathfrak{g})) = 0$.

There are a bunch of easy-ish consequences of Cartan's Criterion.

Definition 4.23. For an alternating or symmetric bilinear form $F \colon V \times V \to k$, the **radical of** F is

$$\operatorname{rad}(F) = \{v \in V \mid F(v, w) = 0 \text{ for all } w \in V\}.$$

and if $W \subseteq V$ is a subspace, define

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

Note that $\operatorname{rad} F = V^{\perp}$. If $\operatorname{rad} F = 0$ we say that F is **non-degenerate.**

Corollary 4.24 (Corollary of Theorem 4.20)**.** The Lie algebra $\mathfrak{g}$ is semisimple if and only if B is non-degenerate.

Proof. Assume $\mathfrak{g}$ is semisimple. Consider $\operatorname{rad} B$. If $Y, Z \in \mathfrak{g}$, and $X \in \operatorname{rad} B$, then

$$0 = B(X, [Y, Z]) = B([X, Y], Z)$$

Since Z was arbitrary, this tells us that $[X, Y] \in \operatorname{rad} B$, Hence, $\operatorname{rad} B$ is an ideal. But B vanishes identically on $\operatorname{rad} B$, so Cartan's Criterion (Theorem 4.20) shows us that $\operatorname{rad} B$ is a solvable ideal. But $\mathfrak{g}$ is semisimple, so $\operatorname{rad} B = 0$, which implies B is nondegenerate.

Conversely, assume $\mathfrak{g}$ is not semisimple. Take $\mathfrak{b}$ a non-trivial solvable ideal. Then for some k, we have that $\mathcal{D}^{k+1}(\mathfrak{b}) = 0$ but $\mathcal{D}^{k}(\mathfrak{b}) \neq 0$. Now take some nonzero $X \in \mathcal{D}^{k}(\mathfrak{b})$. For any $Y \in \mathfrak{g}$, consider $(\operatorname{ad} X \operatorname{ad} Y)^2$.

Since $\mathcal{D}^{i}(\mathfrak{b})$ are characteristic ideals, they are stable under $\operatorname{ad} Y$. Now apply $(\operatorname{ad} X \operatorname{ad} Y)^2$ to $\mathfrak{g}$.

$$\mathfrak{g} \xrightarrow{\operatorname{ad} Y} \underbrace{\mathfrak{g} \xrightarrow{\operatorname{ad} X} \mathcal{D}^{k}(\mathfrak{b}) \xrightarrow{\operatorname{ad} Y} \mathcal{D}^{k}(\mathfrak{b})}_{\mathcal{D}^{k}(\mathfrak{b}) \text{ is a characteristic ideal in } \mathfrak{b}} \xrightarrow{\operatorname{ad} X} \mathcal{D}^{k+1}(\mathfrak{b}) = 0$$

So $\operatorname{ad} X \operatorname{ad} Y$ is a nilpotent endomorphism, and therefore has trace zero. Therefore, $X \in \operatorname{rad} B$ but $X \neq 0$. □

Corollary 4.25. If $\mathfrak{g}$ is a semisimple Lie algebra and I is an ideal, then $I^{\perp}$ is an ideal and $\mathfrak{g} = I \oplus I^{\perp}$. Moreover B is nondegenerate on I.

Proof. Recall that $I^{\perp} = \{X \in \mathfrak{g} \mid B(X, Y) = 0 \forall Y \in I\}$. This is an ideal, because given any $X \in I^{\perp}$ and $Y \in \mathfrak{g}$, let $Z \in I$. Then

$$B([X, Y], Z) = B(X, [Y, Z]) = 0,$$

because $[Y, Z] \in I$ as I is an ideal. Hence, $[X, Y] \in I^{\perp}$ since Z was arbitrary.

By general considerations of vector spaces, $\mathfrak{g} = I + I^{\perp}$.

Now consider $I \cap I^{\perp}$. This is an ideal of $\mathfrak{g}$ on which B is identically zero. Therefore, by Cartan's criterion, $\operatorname{ad}(I \cap I^{\perp})$ is solvable. So there is some k such that $\mathcal{D}^{k}(\operatorname{ad}(I \cap I^{\perp})) = 0$.

Since ad is a Lie algebra homomorphism, then $\operatorname{ad}(D^{k}(I \cap I^{\perp})) = 0$ as well. Hence $\mathcal{D}^{k}(I \cap I^{\perp}) \subseteq \ker \operatorname{ad}$, but $\ker \operatorname{ad}$ is abelian, so

$$\mathcal{D}^{k+1}(I \cap I^{\perp}) = [\mathcal{D}^{k}(I \cap I^{\perp}), \mathcal{D}^{k}(I \cap I^{\perp})] = 0.$$

Hence, $I \cap I^{\perp}$ is a solvable ideal of $\mathfrak{g}$, which means that $I \cap I^{\perp} = 0$ since $\mathfrak{g}$ is semisimple.

Finally, since B is nondegenerate on $\mathfrak{g}$, then for any $X \in I$ there is $Y \in \mathfrak{g}$ such that $B(X, Y) \neq 0$. We have that $Y \neq 0$, and if $Y \in I^{\perp}$, then it must be that $B(X, Y) = 0$. So $Y \in I$ since $\mathfrak{g} = I \oplus I^{\perp}$. □

Corollary 4.26. If $\mathfrak{g}$ is semisimple, then $\mathfrak{g} = \mathcal{D}(\mathfrak{g})$. (Terminology: $\mathfrak{g}$ is **perfect**).

Proof. Let $\mathfrak{h} = \mathcal{D}(\mathfrak{g})^{\perp}$. Claim that $\mathfrak{h}$ is an ideal. If $X \in \mathfrak{h}$, $Y \in \mathfrak{g}$ and $Z \in \mathcal{D}(\mathfrak{g})$, then $[Y, Z] \in \mathcal{D}(\mathfrak{g})$, and

$$B([X, Y], Z) = B(X, [Y, Z]) = 0$$

Therefore, $[X, Y] \in \mathfrak{h}$, so $\mathfrak{h}$ is an ideal.

By Corollary 4.25 $\mathfrak{g} \cong \mathfrak{h} \oplus \mathcal{D}(\mathfrak{g})$ as vector spaces, but $h \cong {}^{\mathfrak{g}}/_{\mathcal{D}(\mathfrak{g})}$ is abelian. So $\mathfrak{h}$ is a solvable ideal, and hence zero. □

We can get lots of mileage out of Cartan's Criterion.

Corollary 4.27. Let $\mathfrak{g}$ be semisimple and $\rho \colon \mathfrak{g} \to \mathfrak{gl}(V)$ with $\dim V = 1$. Then ρ is the trivial representation.

Proof. $\rho(\mathfrak{g})$ is abelian, and a quotient of $\mathfrak{g}$ by $\ker \rho$. It therefore factors through the largest abelian quotient, $\mathfrak{g}/\mathcal{D}(\mathfrak{g})$, so $\mathcal{D}(\mathfrak{g}) \subseteq \ker \rho$. But by Corollary 4.26, $\mathfrak{g} = \mathcal{D}(\mathfrak{g}) \subseteq \ker \rho$. □

So why is "semisimple" a good word for "has no solvable ideals?"

Corollary 4.28. Let $\mathfrak{g}$ be semisimple. Then $\mathfrak{g} \cong \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_n$, where each $\mathfrak{g}_i$ is a simple ideal of $\mathfrak{g}$.

Proof. If $\mathfrak{g}$ is not simple, then let $\mathfrak{h}$ be any nontrivial ideal. Then as before, $\mathfrak{h}^{\perp}$ is an ideal and the Killing form vanishes identically on $\mathfrak{h} \cap \mathfrak{h}^{\perp}$, so $\mathfrak{h} \cap \mathfrak{h}^{\perp} = 0$ and therefore $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}$. Repeat with $\mathfrak{h}^{\perp}$ and $\mathfrak{h}$ until they are themselves simple. This terminates because $\dim \mathfrak{h} \leqslant \dim \mathfrak{g}$ and $\dim \mathfrak{h}^{\perp} \leqslant \dim \mathfrak{g}$. □

Corollary 4.29. Let $\rho \colon \mathfrak{g} \to \mathfrak{h}$ be any homomorphism with $\mathfrak{g}$ semisimple. Then $\rho(\mathfrak{g})$ is also semisimple.

Proof. $\ker \rho$ is an ideal, so as before $\mathfrak{g} \cong \ker \rho \oplus (\ker \rho)^{\perp}$, with $B|_{\ker \rho}$ and $B|_{(\ker \rho)^{\perp}}$ nondegenerate. So $\rho(\mathfrak{g}) \cong \mathfrak{g}/\ker \rho \cong (\ker \rho)^{\perp}$ is semisimple. □

Corollary 4.30. Let $\mathfrak{g}$ be semisimple. If $\rho \colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation of $\mathfrak{g}$ then B_V is non-degenerate on $\rho(\mathfrak{g})$

Proof. We know that the image of $\mathfrak{g}$ under ρ is also semisimple. Let $\mathfrak{s} = \{Y \in \rho(\mathfrak{g}) \mid B_V(X, Y) = 0 \ \forall \ X \in \rho(\mathfrak{g})\}$. Then as usual $\mathfrak{s}$ is an ideal of $\rho(\mathfrak{g})$ on which B_V is identically zero, and thus is zero by Cartan's Criterion. □

Remark 4.31. There is a huge (infinite-dimensional) associative algebra $U(\mathfrak{g})$ called the **universal enveloping algebra** such that $\mathfrak{g} \hookrightarrow U(\mathfrak{g})$ as a Lie algebra homomorphism. The representation theory of $U(\mathfrak{g})$ is the same as that of $\mathfrak{g}$. It's constructed as the quotient of the tensor algebra

$$T(\mathfrak{g}) = \bigoplus_{i=0}^{\infty} \mathfrak{g}^{\otimes n}$$

by the ideal $I = \langle (X \otimes Y - Y \otimes X) - [X, Y] \mid X, Y \in \mathfrak{g} \rangle$. $U(\mathfrak{g}) = {}^{T(\mathfrak{g})}/_{I}$.

Remark 4.32. Many infinite-dimensional Lie algebras start by considering the **Loop algebra** $\mathfrak{g} \otimes \mathbb{C}[t, t^{-1}]$ with $\mathfrak{g}$ some finite-dimensional complex semisimple Lie algebra. This is not the direct sum of simple Lie algebras, but it does not have solvable ideals. To get **Kac-Moody algebras**, one takes a **central extension** $\widehat{\mathfrak{g}}$ such that $\widehat{\mathfrak{g}}$ sits in the exact sequence

$$0 \to \mathbb{C}c \to \widehat{\mathfrak{g}} \to \mathfrak{g} \otimes \mathbb{C}[t, t^{-1}] \to 0$$

with $\mathbb{C}c \subseteq Z(\mathfrak{g})$. Then $\widetilde{\mathfrak{g}}$ is $\widehat{\mathfrak{g}} + \mathbb{C}d$ with d acting as $t^{\partial}/_{\partial t}$ on $\widehat{\mathfrak{g}}$ and as 0 on t.

5 Representations of Semisimple Lie Algebras

In this section, we will explore the representation theory of semisimple Lie algebras. The first result is Weyl's theorem on complete reducibility of representations. To that end, we first define the Casimir operator, which is a distinguished (up to scalar multiples) of $Z(U(\mathfrak{g}))$.

Definition 5.1. Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$, and let B_V be the Kiling form relative to V. If $\mathfrak{g}$ is semisimple, then B_V is non-degenerate on $\mathfrak{g}$. Take a basis for $\mathfrak{g}$, say $U_1, U_2, \ldots, U_{\dim \mathfrak{g}}$ and let $\{U_i' \mid 1 \leqslant i \leqslant \dim \mathfrak{g}\}$ be the dual basis under B_V, that is, $B(U_i, U_i') = \delta_{ij}$. Then **"the" Casimir operator with respect to** V is

$$C_V = \sum_{i=1}^{\dim \mathfrak{g}} U_i U_i'.$$

Exercise 5.2. The word "the" is in quotes above because it's not obvious the definition doesn't depend on the choice of basis. Check that C_V doesn't depend on the choice of basis for $\mathfrak{g}$.

"The Casimir operator" sounds like the name of a spy thriller. Let's see an example.

Example 5.3. Let $\mathfrak{g} = \mathfrak{sl}(2) \subseteq \mathfrak{gl}(2)$. Then as before,

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Then,

$$\begin{aligned} B_V(X, Y) &= 1 \\ B_V(Y, Y) &= B_V(X, X) = B_V(X, H) = B_V(Y, H) = 0 \\ B_V(H, H) &= 2 \end{aligned}$$

So if $\{U_i\} = \{X, Y, H\}$ is a basis for $\mathfrak{g}$ and $\{U_i'\} = \{Y, X, \frac{1}{2}H\}$ is in the dual basis for $\mathfrak{g}$ under B_V, then

$$C_V = XY + YX + \frac{1}{2}H^2$$

As an element of $\mathfrak{gl}(2)$,

$$C_V = \begin{bmatrix} 3/2 & 0 \\ 0 & 3/2 \end{bmatrix}$$

Proposition 5.4. Let C_V be the Casimir operator for $\mathfrak{g}$ with respect to a representation $\mathfrak{g} \to \mathfrak{gl}(V)$. Then

(1) $\operatorname{tr} C_V = \dim \mathfrak{g}$;

(2) if $W \subseteq V$ is a $\mathfrak{g}$-submodule, then $C_V W \subseteq W$;

(3) for $X \in \mathfrak{g}$, $[X, C_V] = 0$.

Proof.

(1) $\operatorname{tr}\left(\sum_i U_i U_i'\right) = \sum_i \operatorname{tr}(U_i U_i') = \sum_i B_V(U_i, U_i') = \sum_{i=1}^{\dim g} 1 = \dim g$.

(2) Follows from $U_i \in \mathfrak{g}$ and $U_i' \in \mathfrak{g}$.

(3) Define coefficients a_{ij} by $[X, U_i] = \sum_j a_{ij} U_j$. We have that

$$a_{ij} = B_V([X, U_i], U_j') = -B_V(U_i, [X, U_j'])$$

and therefore, $[X, U_j'] = -\sum_k a_{kj} U_k'$. So

$$\begin{aligned}
[X, C_V] &= \sum_i [X, U_i U_i'] \\
&= \sum_i U_i [X, U_i'] + \sum_i [X, U_i] U_i' \\
&= \sum_{i,j} U_i a_{ij} U_j' + \sum_{ij} a_{ij} U_j U_i' \\
&= \sum_{i,j} -U_i a_{ji} U_j' + \sum_{i,j} a_{ji} U_i U_j' \\
&= 0
\end{aligned}$$

□

Lemma 5.5 (Schur's Lemma)**.** Let $\mathfrak{g}$ be a Lie algebra over an algebraically closed field. Let V be an irreducible finite-dimensional representation of $\mathfrak{g}$. Then $\dim \operatorname{Hom}_{\mathfrak{g}}(V, V) = 1$

Proof. Let θ be a non-zero map in $\operatorname{Hom}_{\mathfrak{g}}(V, V)$. Because we work over an algebraically closed field, θ has a non-trivial eigenvector with eigenvalue λ say. Then $\theta - \lambda I$ is clearly a $\mathfrak{g}$-module map, with a nonzero kernel. But $\ker(\theta - \lambda I)$ is a $\mathfrak{g}$-submodule of V, and V is irreducible, so $\ker(\theta - \lambda I) = V$. Hence, $\theta = \lambda I$. □

The next theorem says that representations of semisimple Lie algebras are completely reducible into a direct sum of irreducible representations, much like representations of finite groups.

Theorem 5.6 (Weyl's Theorem)**.** Let $\mathfrak{g}$ be a semisimple complex Lie algebra, and let V be a representation of $\mathfrak{g}$ with $W \subseteq V$ a $\mathfrak{g}$-submodule. Then there is a $\mathfrak{g}$-stable compliment to W in V, that is, a $\mathfrak{g}$-submodule W' such that $V \cong W \oplus W'$.

Proof. This is yet another incredibly clever piece of linear algebra. There are several cases, which we prove in order of increasing generality.

Case 1: Assume first that W is codimension 1 in V and irreducible.

Proof of Case 1. First observe that ${}^{V}\!/_{W}$ is a 1-dimensional $\mathfrak{g}$-module which (by Corollary 4.27) is trivial for $\mathfrak{g}$. That is, $\mathfrak{g}V \subseteq W$. This implies that $C_V V \subseteq W$. Because $[X, C_V] = 0$ by Proposition 5.4, we have that $X(C_V(v)) = C_V(X(v))$ for all $v \in V$. Therefore, C_V is a $\mathfrak{g}$-module map. So $C_V|_W = \lambda I_W$ by Lemma 5.5 (using W irreducible). Now ${}^{1}\!/_{\lambda} C_V$ is a projection homomorphism from V to W. Dividing by λ is okay, since $\operatorname{tr} C_V = \dim \mathfrak{g} \Rightarrow \lambda \neq 0$. Thus, $V \cong W \oplus \ker({}^{1}\!/_{\lambda} C_V)$. Hence, V is reducible. ☐

Case 2: Assume that W is codimension 1.

Proof of Case 2. By induction on $\dim V$. If $\dim V = 1$ there's nothing to check.

In general, we can assume that W has a nontrivial $\mathfrak{g}$-submodule Z (or else W is irreducible and we refer to case 1). Consider ${}^{V}\!/_{Z} \supset {}^{W}\!/_{Z}$. By an isomorphism theorem, we have that

$$V/W \cong \frac{V/Z}{W/Z}.$$

So ${}^{W}\!/_{Z}$ is codimension 1 in ${}^{V}\!/_{Z}$ and so by induction we can assume there is W' a $\mathfrak{g}$-submodule of V such that

$$V/Z \cong W'/Z \bigoplus W/Z.$$

But Z is codimension 1 in W', so by induction again there is $U \subset W'$ a $\mathfrak{g}$-submodule such that $W' \cong U \oplus Z$. So $V \cong W \oplus U$ by the following chain of isomorphisms

$$V/W \cong \frac{W/Z \oplus W'/Z}{W/Z} \cong W'/Z \cong (Z \oplus U)/Z \cong U$$

☐

Case 3: Assume that W is irreducible.

Proof of Case 3. Consider $\operatorname{Hom}_{\mathbb{C}}(V, W)$. We know that this is a $\mathfrak{g}$-module via $(X\alpha)(v) = -\alpha(Xv) + X(\alpha(v))$. Similarly, there is an action of $\mathfrak{g}$ on $\operatorname{Hom}_{\mathbb{C}}(W, W)$.

Consider a restriction map $R\colon \operatorname{Hom}_{\mathbb{C}}(V, W) \to \operatorname{Hom}_{\mathbb{C}}(W, W)$. This is a $\mathfrak{g}$-module homomorphism, because for $w \in W$,

$$\begin{aligned} X(R(\alpha))(w) &= X(R(\alpha)(w)) - R(\alpha)(X(w)) \\ &= X(\alpha(w)) - \alpha(X(w)) \\ &= (X(\alpha))(w) \\ &= R(X(\alpha))(w) \end{aligned}$$

Now note that $X(\alpha) = 0$ for all $X \in \mathfrak{g}$ precisely if α is a $\mathfrak{g}$-module map. By this observation, $\mathrm{Hom}_{\mathbb{C}}(W, W)$ contains a $\mathfrak{g}$-submodule $\mathrm{Hom}_{\mathfrak{g}}(W, W)$, which is trivial and 1-dimensional by Lemma 5.5 since W is irreducible. The module $M := R^{-1}(\mathrm{Hom}_{\mathfrak{g}}(W, W))$ is a submodule of $\mathrm{Hom}_{\mathbb{C}}(V, W)$.

Now $\ker(R|_M)$ has codimension 1, as it's image $\mathrm{Hom}_{\mathfrak{g}}(W, W)$ has dimension 1. So by Case 2 we have that $M \cong \ker(R|_M) \oplus \mathbb{C}\psi$ for some $\psi \in \mathrm{Hom}_{\mathbb{C}}(V, W)$. But $\mathbb{C}\psi$ is 1-dimensional, so $\mathfrak{g}$ acts trivially on this space. Therefore, ψ is a $\mathfrak{g}$-module map. Moreover, ψ is nonzero because otherwise $R(M) = 0$. Again, by scaling, we can arrange that ψ is a projection $V \to W$, so V has a compliment W' to W, that is, $V \cong W \oplus W'$. □

Case 4: The whole theorem.

Proof of Case 4. Proof by induction on $\dim V$. If $\dim V = 1$, then we are done by Corollary 4.27.

If $\dim V > 1$, let $W \subseteq V$. If W is irreducible, then this was done in Case 3. Otherwise, W has a nontrivial submodule Z; pick Z maximal among nontrivial submodules of W. Then W/Z is irreducible, because submodules of ${}^{W}\!/_{Z}$ are submodules of W containing Z, and Z is maximal so there are none. Since Z is nontrivial, $\dim\left({}^{V}\!/_{Z}\right) < \dim V$, so by induction ${}^{W}\!/_{Z}$ has a compliment in ${}^{V}\!/_{Z}$, of the form ${}^{W'}\!/_{Z}$. So

$$ {}^{V}\!/_{Z} \cong {}^{W}\!/_{Z} \oplus {}^{W'}\!/_{Z}. $$

Then because Z is nontrivial, $\dim {}^{W'}\!/_{Z} < \dim V$. So by induction, Z has a compliment U in W', with $W' \cong U \oplus Z$. Then

$$ {}^{V}\!/_{W} \cong \frac{{}^{W}\!/_{Z} \oplus {}^{W'}\!/_{Z}}{{}^{W}\!/_{Z}} \cong {}^{W'}\!/_{Z} \cong {}^{(Z \oplus U)}\!/_{Z} \cong U. $$

Hence, $V \cong W \oplus U$ and U, W are $\mathfrak{g}$-invariant. □

This concludes the proof of Theorem 5.6. □

Exercise 5.7. Show that if $\pi \colon V \to V$ satisfies $\pi^2 = \pi$, then $V = \operatorname{im} \pi \oplus \ker \pi$.

Remark 5.8 (Important Examinable Material)**.** Previously, on *Complete Reducibility*, Rick Moranis invents a machine to reduce things to their component parts. By a cruel twist of fate, he is the victim of his own invention, and thereby his consciousness gets trapped in a single glucose molecule. This is the story of that glucose molecule's fight to reunite itself with the rest of its parts, and thereby reform Rick Moranis.

By Weyl's Theorem, we see that if $\mathfrak{g}$ is complex semisimple finite-dimensional Lie algebra, and V is a finite-dimensional representation, then any submodule W has a complement W' and $V \cong W \oplus W'$.

Corollary 5.9. A simple induction yields that under these hypotheses, $V \cong W_1 \oplus W_2 \oplus \ldots \oplus W_n$, where W_i are simple modules.

Theorem 5.10. Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be semisimple. For any $X \in \mathfrak{g}$, let $X = X_s + X_n$ be the Jordan decomposition of X into semisimple X_s and nilpotent X_n parts. Then $X_s, X_n \in \mathfrak{g}$.

Proof. The idea here is to write $\mathfrak{g}$ as the intersection of some subalgebras for which the result is obvious. Let W be a simple submodule of V, and define $\mathfrak{s}_W$ to be the component of the stabilizer of W in $\mathfrak{gl}(V)$ that is traceless, that is,

$$\mathfrak{s}_W = \{X \in \mathfrak{gl}(V) \mid XW \subseteq W \text{ and } \operatorname{tr} X|_W = 0\}.$$

Claim that $\mathfrak{g} \subseteq \mathfrak{s}_W$. First, W is a submodule so it is stabilized by $\mathfrak{g}$, and also the image of $\mathfrak{g}$ in $\mathfrak{gl}(W)$, $\overline{\mathfrak{g}}$, is by Corollary 4.29 also semisimple, so has $\mathcal{D}(\overline{\mathfrak{g}}) = \overline{\mathfrak{g}}$. Therefore, every element of $\overline{\mathfrak{g}}$ is a sum of commutators, all of whose trace must be zero. So we conclude that $\mathfrak{g} \subseteq \mathfrak{s}_W$. This tells us that $\operatorname{tr} X|_W = 0$.

Note that X_s and X_n, being polynomials in X, stabilize everything that X does, and also the trace of $X_n|_W$ is zero since $X_n|_W$ is nilpotent, and

$$\operatorname{tr}(X_s|_W) = \operatorname{tr}(X|_W - X_n|_W) = \operatorname{tr}(X|_W) - \operatorname{tr}(X_n|_W) = 0.$$

Therefore, $X_s, X_n \in \mathfrak{s}_W$ for each W.

Now let $\mathfrak{n}$ be the normalizer of $\mathfrak{g}$ in $\mathfrak{gl}(V)$,

$$\mathfrak{n} = \mathfrak{n}_{\mathfrak{gl}(V)}(\mathfrak{g}) = \{X \in \mathfrak{gl}(V) \mid [X, \mathfrak{g}] \subseteq \mathfrak{g}\}.$$

Clearly $\mathfrak{g} \subseteq \mathfrak{n}$, and also $X_s, X_n \in \mathfrak{n}$, being polynomials in X.

To finish the proof, claim that $\mathfrak{g}$ is precisely the intersection

$$\mathfrak{g}' = \mathfrak{n} \cap \left(\bigcap_{\substack{W \subseteq V \\ W \text{ irred.}}} \mathfrak{s}_W \right)$$

Since $\mathfrak{g}' \subset \mathfrak{n}$, $\mathfrak{g}$ is an ideal of $\mathfrak{g}'$. Then $\mathfrak{g}$ is a submodule of $\mathfrak{g}'$ under the adjoint action of $\mathfrak{g}$. So by Weyl's Theorem, $\mathfrak{g}' \cong \mathfrak{g} \oplus U$ as $\mathfrak{g}$-modules for some $\mathfrak{g}$-submodule U.

So we want to show that $U = 0$. Take $Y \in U$. We have $[Y, \mathfrak{g}] \subseteq \mathfrak{g}$ as $\mathfrak{g}$ is an ideal of $\mathfrak{g}'$. But also $\operatorname{ad} \mathfrak{g} U \subseteq U$, so $[Y, \mathfrak{g}] \subseteq U$. Therefore, $[Y, \mathfrak{g}] \subseteq U \cap \mathfrak{g} = 0$.

Thus, Y commutes with every element of $\mathfrak{g}$. Hence, Y is a $\mathfrak{g}$-module map from V to V. So Y stabilizes every irreducible submodule W, so by Schur's lemma $Y|_W = \lambda \mathrm{id}_W$ for some scalar λ.

Now $\operatorname{tr} Y|_W = 0$ for all irreducible $W \subseteq V$ because $Y \in \mathfrak{s}_W$ for each W. Therefore, $\operatorname{tr} \lambda \mathrm{id}_W = 0 \implies \lambda = 0$, so $Y|_W = 0$ for all irreducible W. But $V \cong \bigoplus_i W_i$ for W_i irreducible. So $Y = 0$. □

If $\mathfrak{g}$ is as in the theorem, we can define an abstract Jordan decomposition by $\operatorname{ad} X = (\operatorname{ad} X)_s + (\operatorname{ad} X)_n$. And because ad is faithful for $\mathfrak{g}$ semisimple, we have that $\mathfrak{g} \cong \operatorname{ad} \mathfrak{g} \subseteq \mathfrak{gl}(\mathfrak{g})$.

But by the theorem $(\operatorname{ad} X)_s$ and $(\operatorname{ad} X)_n$ are elements of $\operatorname{ad} \mathfrak{g}$, and hence are of the form $\operatorname{ad} X_s$ and $\operatorname{ad} X_n$ for some elements X_s and X_n of $\mathfrak{g}$. Therefore, $\operatorname{ad} X = \operatorname{ad}(X_s + X_n)$ and the faithfulness of ad implies $X = X_s + X_n$.

Suppose $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ for some V, then write $X = X_s + X_n$ relative to V. By Corollary 4.16, $\operatorname{ad} X_s = (\operatorname{ad} X)_s$ and $\operatorname{ad} X_n = (\operatorname{ad} X)_n$. So the abstract Jordan decomposition as just defined agrees with the usual notion.

Moreover, this is true under any representation.

Corollary 5.11 (Preservation of Jordan Decomposition)**.** Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be any representation of a semisimple Lie algebra $\mathfrak{g}$. Let $X \in \mathfrak{g}$ with abstract Jordan decomposition $X = X_s + X_n$. Then $\rho(X)_s = \rho(X_s)$ and $\rho(X)_n = \rho(X_n)$ is a Jordan decomposition.

We really rely on the semisimplicity of $\mathfrak{g}$ here. This fails spectacularly in positive characteristic.

Proof. The idea is that we want to compare $\rho(\operatorname{ad} X)$ with $\operatorname{ad}\rho(X)$ in some sense (because these are not quite well-defined).

By Corollary 4.29, $\rho(\mathfrak{g})$ is semisimple. Therefore, by Corollary 5.11, $\rho(X)_s, \rho(X)_n \in \rho(\mathfrak{g})$. So it remains to check that $\rho(X_s)$ is semisimple and $\rho(X_n)$ is nilpotent, and then we may apply Proposition 4.15(5) to claim that $\rho(X_s) = \rho(X)_s$ and $\rho(X_n) = \rho(X)_n$.

Let Z_i be a basis of eigenvectors of $\operatorname{ad} X_s$ in $\mathfrak{g}$. That is,

$$\operatorname{ad}(X_s)Z_i = \lambda_i Z_i$$

for some λ_i. Then $\rho(Z_i)$ span $\rho(\mathfrak{g})$ and

$$\operatorname{ad}(\rho(X_s))\rho(Z_i) = [\rho(X_s), \rho(Z_i)] = \rho([X_s, Z_i]) = \lambda_i \rho(Z_i)$$

so that $\operatorname{ad}(\rho(X_s))$ has a basis of eigenvectors and is therefore semisimple (diagonalizable). Similarly, $\operatorname{ad}(\rho(X_n))$ is nilpotent commuting with $\operatorname{ad}(\rho(X_s))$ and $\operatorname{ad}(\rho(X))$.

Accordingly, $\rho(X) = \rho(X_n) + \rho(X_s)$ is the Jordan decomposition of $\rho(X)$. But by the remarks above this is *the* Jordan decomposition of $\rho(X)$ relative to V. This means precisely that $\rho(X)_s = \rho(X_s)$ and $\rho(X_n) = \rho(X)_n$. ☐

Remark 5.12. There is another way to do this that uses the Killing form instead of complete reducibility, but it's a bit of a case of using a sledgehammer to crack a nut. An alternative approach to Theorem 5.10 not using Weyl's theorem is to prove that when $\mathfrak{g}$ is semisimple, every derivation D of $\mathfrak{g}$ is inner, that is, of the form $D = \operatorname{ad} X$ for some $X \in \mathfrak{g}$. Equivalently, $\operatorname{ad}\mathfrak{g} = \operatorname{Der}(\mathfrak{g})$.

Given that result, to prove Theorem 5.10 write $\operatorname{ad} X = x_s + x_n$ in $\mathfrak{gl}(\mathfrak{g})$ for some $X \in \mathfrak{g}$. As x_s and x_n are also derivations of $\mathfrak{g}$, then $x_s = \operatorname{ad} X_s$ and $x_n = \operatorname{ad} X_n$ for some $X_s, X_n \in \mathfrak{g}$. From the injectivity of ad, we get $X = X_s + X_n$ and $[X_s, X_n] = 0$. It's an easy exercise to see that X_s and X_n are semisimple and nilpotent, respectively. This gives us the Jordan decomposition of X.

Remark 5.13 (Important Examinable Material)**.** Last time we were talking about Jordan Decomposition, which is a recent Channel 4 documentary following the trials and tribulations of supermodel Jordan Price, wherein she is struck by a previously undetected case of leprosy. Most episodes focus on major reconstructive surgery, wherein her body parts are reattached. But unfortunately her doctors are so overworked that her knee is put back on backwards, so she has to walk around in a crablike fashion. This doesn't last too long, however, because soon her other knee becomes detached.

Previously, on *The Jordan Decomposition*, for a complex semisimple Lie algebra we have, for any $X \in \mathfrak{g}$, elements X_s and X_n in $\mathfrak{g}$ such that under any representation, $\rho(X)_s = X_s$ and $\rho(X)_n = X_n$. The power of this will become apparent in representation theory.

But to set that up, we need to generalize some of the facts about the representation theory of $\mathfrak{sl}_2$ to other Lie algebras.

Recall for $\mathfrak{g} = \mathfrak{sl}(2)$ we have a decomposition $\mathfrak{g} = \mathfrak{g}_0(H) \oplus \mathfrak{g}_2(H) \oplus \mathfrak{g}_{-2}(H)$, where $\mathfrak{g}_\lambda(H)$ denotes the generalized λ-eigenspace of $\operatorname{ad}(H)$. Here, $\mathfrak{g}_0(H) = \langle H \rangle$, $\mathfrak{g}_2(H) = \langle X \rangle$, and $\mathfrak{g}_{-2}(H) = \langle Y \rangle$.

Definition 5.14. A **Cartan subalgebra** of a semisimple complex Lie algebra $\mathfrak{g}$ is an abelian subalgebra consisting of ad-diagonalizable elements, which is maximal with respect to these properties.

Now, we could have just said diagonalizable elements, because we know there is an intrinsic notion of diagonalizability in Lie algebras, but for $\mathfrak{g}$ semisimple ad is a faithful representation anyway.

Definition 5.15. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$, and let $H \in \mathfrak{h}$. Then define the **centralizer** of H in $\mathfrak{g}$ as

$$\mathfrak{c}(H) = \mathfrak{c}_{\mathfrak{g}}(H) = \{X \in \mathfrak{g} \mid [X, H] = 0\}.$$

Lemma 5.16. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Suppose $H \in \mathfrak{h}$ such that the dimension of $\mathfrak{c}_{\mathfrak{g}}(H)$ is minimal over all elements $H \in \mathfrak{h}$. Then, $\mathfrak{c}_{\mathfrak{g}}(H) = \mathfrak{c}_{\mathfrak{g}}(\mathfrak{h}) = \bigcap_{X \in \mathfrak{h}} \mathfrak{c}_{\mathfrak{g}}(X)$.

Proof. Notice that for any $S \in \mathfrak{h}$, S is central in $\mathfrak{c}_{\mathfrak{g}}(H)$ if and only if $\mathfrak{c}_{\mathfrak{g}}(H) \subseteq \mathfrak{c}_{\mathfrak{g}}(S)$. We shall show that if S is not central, then a linear combination of S and H has a smaller centralizer in $\mathfrak{g}$, thus finding a contradiction.

First, we will construct a suitable basis for $\mathfrak{g}$. Start with a basis $\{c_1, \ldots, c_n\}$ of $\mathfrak{c}_{\mathfrak{g}}(H) \cap \mathfrak{c}_{\mathfrak{g}}(S)$. We know $\operatorname{ad} S$ acts diagonalizably on $\mathfrak{c}_{\mathfrak{g}}(H)$ because $S \in \mathfrak{h}$ is ad-diagonalizable. Therefore S commutes with every element of $\mathfrak{h}$, so we can extend this to a basis for $\mathfrak{c}_{\mathfrak{g}}(H)$ consisting of eigenvectors for $\operatorname{ad} S$, say by $\{x_1, \ldots, x_p\}$.

Similarly, we can extend $\{c_i\}$ to a basis of $\mathfrak{c}_{\mathfrak{g}}(S)$ of eigenvectors for $\operatorname{ad} H$ by adjoining $\{y_1, \ldots, y_q\}$. Then

$$\{c_1, \ldots, c_n, x_1, \ldots, x_p, y_1, \ldots, y_q\}$$

is a basis of $\mathfrak{c}_{\mathfrak{g}}(H) + \mathfrak{c}_{\mathfrak{g}}(S)$.

As $\operatorname{ad} S$ and $\operatorname{ad} H$ commute, we can complete to a basis of $\mathfrak{g}$, say by $\{w_1, \ldots, w_r\}$ of simultaneous eigenvectors for S and H.

Note that $[S, x_j] \neq 0$ because $x_j \in \mathfrak{c}_\mathfrak{g}(H) \backslash \mathfrak{c}_\mathfrak{g}(S)$, and also $[H, y_j] \neq 0$. Let $[H, w_i] = \theta_i w_i$ and $[s, w_i] = \sigma_i w_i$ with $\theta_i, \sigma_i \neq 0$. Thus if we choose $\lambda \neq 0$ such that $\lambda \neq {}^{-\sigma_\ell}/_{\theta_\ell}$ for any ℓ, w_j doesn't commute with $S + \lambda H$ for any j. Moreover, x_i and y_i don't commute with $S + \lambda H$ by construction, so the only things that commute with $S + \lambda H$ are linear combinations of the c_i – things that commute with both H and S. Therefore, $\mathfrak{c}_\mathfrak{g}(S + \lambda H) = \mathfrak{c}_\mathfrak{g}(S) \cap \mathfrak{c}_\mathfrak{g}(H)$.

Since S is not central in $\mathfrak{c}_\mathfrak{g}(H)$, $\mathfrak{c}_\mathfrak{g}(H) \nsubseteq \mathfrak{c}_\mathfrak{g}(S)$, so this is a subspace of smaller dimension. This is a contradiction, because $\dim \mathfrak{c}_\mathfrak{g}(H)$ was assumed to be the smallest possible. □

Lemma 5.17. Suppose H is any element of $\mathfrak{g}$. Then $[\mathfrak{g}_\lambda(H), \mathfrak{g}_\mu(H)] \subseteq \mathfrak{g}_{\lambda+\mu}(H)$. Additionally, if $\mathfrak{g}$ is a semisimple Lie algebra, then the restriction of the Killing form to $\mathfrak{g}_0(H)$ is nonzero, where H satisfies the hypotheses of the Lemma 5.16.

Proof. To show the first part, one proves by induction that

$$(\mathrm{ad}(H) - (\lambda+\mu)I)^k([X,Y]) = \sum_{j=0}^{k} \binom{k}{j} \left[(\mathrm{ad}(H) - \lambda I)^j X,\ (\mathrm{ad}(H) - \mu I)^{k-j} Y\right].$$

This just comes down to repeated application of the Jacobi identity. If $k = 1$, this is actually just the Jacobi identity.

Hence if $X \in \mathfrak{g}_\lambda(H)$ and $Y \in \mathfrak{g}_\mu(H)$, then we can take k sufficiently large (e.g. $k = 2\dim \mathfrak{g}$) such that either $(\mathrm{ad}(H) - \lambda I)^j X$ or $(\mathrm{ad}(H) - \mu I)^{k-j} Y$ vanishes, so $[X, Y]$ is in the generalized eigenspace of $\lambda + \mu$.

For the second statement, if $Y \in \mathfrak{g}_\lambda(H)$ with $\lambda \neq 0$, then $\mathrm{ad}\, Y$ maps each eigenspace to a different one. Furthermore, so does $\mathrm{ad}\, Y \circ \mathrm{ad}\, X$ for $X \in \mathfrak{g}_0(H)$. So this endomorphism $\mathrm{ad}\, Y \circ \mathrm{ad}\, X$ is traceless. Therefore, $B(X, Y) = 0$ for such X, Y. Therefore, $\mathfrak{g}_0(H)$ is perpendicular to all the other weight spaces for H.

But the Killing form is non-degenerate on $\mathfrak{g}$, so we should be able to find some Z such that $B(X, Z) \neq 0$. But this Z must be in $\mathfrak{g}_0(H)$, because all other weight spaces are perpendicular to $\mathfrak{g}_0(H)$. Hence, B is non-degenerate on $\mathfrak{g}_0(H)$. □

Theorem 5.18. Let $\mathfrak{h}$ be a Cartan subalgebra of a semisimple Lie algebra $\mathfrak{g}$. Then $\mathfrak{c}_\mathfrak{g}(\mathfrak{h}) = \mathfrak{h}$. The Cartan subalgebra is self-centralizing.

Proof. Choose $H \in \mathfrak{h}$ such that the dimension of $\mathfrak{c}_g(H)$ is minimal over all elements $H \in \mathfrak{h}$. Then $\mathfrak{c}_\mathfrak{g}(\mathfrak{h}) = \mathfrak{c}_\mathfrak{g}(H)$ by Lemma 5.16, so it suffices to show that $\mathfrak{c}_\mathfrak{g}(H) = \mathfrak{h}$.

Since $\mathfrak{h}$ is abelian, we have that $\mathfrak{h} \subseteq \mathfrak{c}_\mathfrak{g}(H)$.

Conversely, if $X \in \mathfrak{c}_\mathfrak{g}(H)$ has Jordan decomposition $X = X_s + X_n$, then X commutes with H implies that X_s commutes with H by Proposition 4.15.

We know that X_s is semisimple, and commutes with H, so commutes with all elements of the Cartan subalgebra $\mathfrak{h}$ because $\mathfrak{c}_\mathfrak{g}(\mathfrak{h}) = \mathfrak{c}_\mathfrak{g}(H)$ by Lemma 5.16. But $\mathfrak{h}$ is the maximal abelian subalgebra consisting of semisimple elements. X_s is semisimple and commutes with everything in $\mathfrak{h}$, so must be in $\mathfrak{h}$.

Therefore $X_s \in \mathfrak{h}$. So we are done if $X_n = 0$.

For any $Y \in \mathfrak{c}_{\mathfrak{g}}(\mathfrak{h})$, we see by the above that Y_s is central in $\mathfrak{c}_{\mathfrak{g}}(\mathfrak{h})$, so $\operatorname{ad} Y_s$ acts by zero on $\mathfrak{c}_{\mathfrak{g}}(\mathfrak{h})$. Therefore, $\operatorname{ad} Y = \operatorname{ad} Y_n$ is nilpotent for arbitrary $Y \in \mathfrak{c}_{\mathfrak{g}}(\mathfrak{h})$, so every element of $\operatorname{ad} \mathfrak{c}_{\mathfrak{g}}(\mathfrak{h})$ is nilpotent. Then by the corollary to Engel's Theorem (Corollary 4.11), there is a basis of $\mathfrak{c}_{\mathfrak{g}}(\mathfrak{h})$ such that each $\operatorname{ad} Y$ is strictly upper triangular for $Y \in \mathfrak{c}_{\mathfrak{g}}(\mathfrak{h})$. Hence,

$$B(X, Y) = \operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y) = 0$$

for all $Y \in \mathfrak{c}_{\mathfrak{g}}(\mathfrak{h})$. But the Killing form is nondegenerate on restriction to $\mathfrak{c}_{\mathfrak{g}}(H) = \mathfrak{g}_0(H)$ by Lemma 5.17, so it must be that $\operatorname{ad} X = 0$. However, $\operatorname{ad} X = \operatorname{ad} X_n$ and ad is injective because $\mathfrak{g}$ is semisimple, so $X_n = 0$.

Therefore, for any $X \in c_{\mathfrak{g}}(H)$, $X = X_s$ and $X_s \in \mathfrak{h}$, so $\mathfrak{c}_{\mathfrak{g}}(H) \subseteq \mathfrak{h}$. $\square$

Previously on Cartan Subalgebras, we had maximal diagonalizable abelian subalgebras $\mathfrak{h}$ of $\mathfrak{g}$. We showed that $\mathfrak{c}_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$ and moreover $\mathfrak{h} = \mathfrak{c}_{\mathfrak{g}}(H)$ for some $H \in \mathfrak{h}$.

Remark 5.19. It's not clear that any two Cartan subalgebras have the same dimension. But in fact, it's true that they all have the same dimension, and moreover they are all centralizers of regular semisimple elements. Additionally, all Cartan subalgebras are conjugate under the adjoint action of G such that $\mathfrak{g}$ is the Lie algebra of G.

Definition 5.20. We say that an element of $\mathfrak{g}$ is **regular** if its centralizer dimension in $\mathfrak{g}$ is minimal.

Remark 5.21. The definition that we gave is not the original definition of Cartan subalgebra. Another useful one is that $\mathfrak{h}$ is a self-normalizing nilpotent subalgebra, that is, $\mathfrak{h}$ satisfies $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$. Then it is automatically maximal among nilpotent subalgebras. But then it is unclear when Cartan subalgebras exist, and remains unknown in many cases. Another definition is that $\mathfrak{h}$ is the centralizer of a maximal **torus**, where a torus is any abelian subalgebra consisting of semisimple elements.

Given a representation V for a semisimple Lie algebra $\mathfrak{g}$, we can decompose V into simultaneous eigenspaces for $\mathfrak{h}$ (since $\rho(\mathfrak{h})$ is still abelian and diagonalizable). Write $V = \bigoplus V_\alpha$ for these eigenspaces. For $v \in V_\alpha$ we have $Hv = \alpha(H)v$ for some function $\alpha \colon \mathfrak{h} \to \mathbb{C}$.

We can check that α is a linear function $\mathfrak{h} \to \mathbb{C}$, that is, $\alpha \in \mathfrak{h}^*$.

Definition 5.22. The vectors of eigenvalues α are called the **weights** of the representation V, and the V_α are the corresponding **weight spaces**.

Let's compare this to what we were doing with $\mathfrak{sl}(2)$. In $\mathfrak{sl}(2)$, we had weight spaces for just one element H, and the Cartan subalgebra $\mathfrak{h}$ was just spanned by H. So these α were really just the eigenvalues of H.

Example 5.23. Let's now consider $\mathfrak{g} = \mathfrak{sl}(3)$, the Lie algebra of traceless 3×3 matrices over $\mathbb{C}$. It's easy to check that for $\mathfrak{sl}(3)$, a Cartan subalgebra is

$$\mathfrak{h} = \left\{ \begin{bmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{bmatrix} \,\middle|\, a_1 + a_2 + a_3 = 0 \right\}.$$

Any other Cartan subalgebra is given by conjugating these matrices. Let's define some elements L_i of $\mathfrak{h}^*$ by

$$L_i\left(\begin{bmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{bmatrix}\right) = a_i.$$

For the standard representation $\mathfrak{sl}(3) \circlearrowright \mathbb{C}^3$, a basis of simultaneous eigenspaces for $\mathfrak{h}$ is just the standard basis $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. We have that

$$\begin{bmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{bmatrix} e_i = a_i e_i = L_i\left(\begin{bmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{bmatrix}\right) e_i,$$

so this representation decomposes as $V = V_{L_1} \oplus V_{L_2} \oplus V_{L_3}$, where $V_{L_i} = \langle e_i \rangle$.

Example 5.24. The previous example is a bit simple, so let's do something more interesting. Consider the adjoint representation, in which we have that $[H, E_{ij}] = (a_i - a_j)E_{ij}$ when

$$H = \begin{bmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{bmatrix},$$

and so the basis of simultaneous eigenspaces i s

$$\{E_{ij} \mid i \neq j\} \cup \left\{\begin{bmatrix} 1 & & \\ & -1 & \\ & & 0 \end{bmatrix}, \begin{bmatrix} 1 & & \\ & 0 & \\ & & -1 \end{bmatrix}\right\}$$

The representation decomposes as

$$V = \mathfrak{h} \oplus V_{L_1 - L_2} \oplus V_{L_1 - L_3} \oplus V_{L_2 - L_3} \oplus V_{L_2 - L_1} \oplus V_{L_3 - L_2} \oplus V_{L_3 - L_1},$$

where $V_{L_i - L_j} = \langle E_{ij} \rangle$.

Definition 5.25. Given a semisimple Lie algebra $\mathfrak{g}$ and a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, the **Cartan decomposition** of $\mathfrak{g}$ is given by

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \neq 0} \mathfrak{g}_\alpha,$$

where $\mathfrak{g}_\alpha$ is a weight space for the adjoint action of $\mathfrak{h}$ on $\mathfrak{g}$ with weight α. These nonzero weights are called **roots**.

Proposition 5.26. $\mathfrak{g}$ is a semisimple Lie Algebra, $\mathfrak{h}$ a Cartan subalgebra. Then

(1) $\mathfrak{g}_0 = \mathfrak{h}$;
(2) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$;
(3) the restriction of B to $\mathfrak{h}$ is non-degenerate;
(4) the roots $\alpha \in \mathfrak{h}^*$ span $\mathfrak{h}^*$;
(5) $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) \neq 0 \iff \alpha = -\beta$;
(6) if α is a root, then so is $-\alpha$;
(7) if $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$, then $B(H, [X, Y]) = \alpha(H)B(X, Y)$ for $H \in \mathfrak{h}$;
(8) $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \neq 0$.

Proof.

(1) Apply Theorem 5.18.

(2) This is a special case of Lemma 5.17, but it's important enough that we should do it again. This is what Fulton and Harris call the **fundamental calculation**. Let $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$.

$$\begin{aligned}[H,[X,Y]] &= [[H,X],Y] + [X,[H,Y]] \\ &= [\alpha(H)X, Y] + [X, \beta(H)Y] \\ &= (\alpha+\beta)(H)[X,Y].\end{aligned}$$

(3) Second part of Lemma 5.17 together with (1).

(4) If the roots don't span $\mathfrak{h}^*$, then in particular there is some functional δ_H that does not lie in the span of the roots. For this $H \in \mathfrak{h}$, $\alpha(H) = 0$ for all roots $\alpha \in \mathfrak{h}^*$. Since $\mathfrak{g}$ can be decomposed in terms of $\mathfrak{g}_\alpha$, we see that $[H,X] = 0$ for all $X \in \mathfrak{g}$, that is, $H \in Z(\mathfrak{g})$. But $\mathfrak{g}$ is semisimple, so $Z(\mathfrak{g}) = 0$ and $H = 0$ as required.

(5) We calculate

$$\begin{aligned}\alpha(H)B(X,Y) &= B([H,X],Y) \\ &= B(H,[X,Y]) \\ &= -B(H,[Y,X]) \\ &= -B([H,Y],X) = -\beta(H)B(X,Y)\end{aligned}$$

so $(\alpha(H)+\beta(H))B(X,Y) = 0$, so either $B(X,Y) = 0$ or $\alpha+\beta = 0$.

(6) If α is a root, but $-\alpha$ is not a root, then given any $X \in \mathfrak{g}_\alpha$, we have $B(X,Y) = 0$ for all $Y \in \mathfrak{g}$ by (5), but B is non-degenerate so it must be $X = 0$.

(7) $B(H,[X,Y]) = B([H,X],Y) = \alpha(H)B(X,Y)$.

(8) B is non-degenerate, so given $X \in \mathfrak{g}_\alpha$, there is some Y such that $B(X,Y) \neq 0$. Choose $H \in \mathfrak{h}$ such that $\alpha(H) \neq 0$, and then

$$B(H,[X,Y]) = \alpha(H)B(X,Y) \neq 0,$$

so $[X,Y] \neq 0$. □

Last time we introduced the Cartan decomposition of a semisimple Lie algebra $\mathfrak{g}$. This is all building up to finding a set of subalgebras of $\mathfrak{g}$, each isomorphic to $\mathfrak{sl}_2$.

Proposition 5.27.

(1) There is $T_\alpha \in \mathfrak{h}$, called **the coroot associated to** α, such that $B(T_\alpha, H) = \alpha(H)$, and $[X,Y] = B(X,Y)T_\alpha$ for $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$.

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

(3) $[[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}], \mathfrak{g}_\alpha] \neq 0$.

(4) If α is a root, $X_\alpha \in \mathfrak{g}_\alpha$, then we can find $Y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $s_\alpha = \langle X_\alpha, Y_\alpha, H_\alpha = [X_\alpha, Y_\alpha]\rangle \cong \mathfrak{sl}_2$.

Proof. (1) For existence, recall that $B|_\mathfrak{h}$ is nondegenerate, and hence induces an isomorphism $\mathfrak{h} \to \mathfrak{h}^*$ via $H \mapsto B(H, -)$. Define T_α to be the preimage of α under this map. Now compute

$$\begin{aligned} B(H, B(X,Y)T_\alpha) &= B(X,Y)B(H,T_\alpha) \\ &= \alpha(H)B(X,Y) \\ &= B(H,[X,Y]) \end{aligned}$$

the last line by Proposition 5.26(7). Now

$$B(H, B(X,Y)T_\alpha - [X,Y]) = 0,$$

and since H is arbitrary and B non-degenerate, then $B(X,Y)T_\alpha - [X,Y] = 0$.

(2) Suppose $\alpha(T_\alpha) = 0$. Take $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$. Then

$$[T_\alpha, X] = \alpha(T_\alpha)X = 0,$$

$$[T_\alpha, Y] = -\alpha(T_\alpha)Y = 0.$$

If $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$ with $B(X,Y) = 1$, then $[X,Y] = T_\alpha$ by part (1).

So we have a subalgebra, $\mathfrak{s} = \langle X, Y, T_\alpha\rangle$ with $\mathcal{D}(\mathfrak{s}) = \langle T_\alpha\rangle$. The adjoint representation $\operatorname{ad}\mathfrak{s}$ of this subalgebra is a solvable subalgebra of $\operatorname{ad}\mathfrak{g} \subseteq \mathfrak{gl}(\mathfrak{g})$. By Lie's Theorem, $\operatorname{ad}\mathfrak{s}$ consists of upper triangular matrices, so $\operatorname{ad}\mathcal{D}(\mathfrak{s})$ consists of strictly upper triangular matrices. Therefore, $\operatorname{ad} T_\alpha \in \operatorname{ad}\mathcal{D}(\mathfrak{s})$ is nilpotent. But $\operatorname{ad} T_\alpha$ is also semisimple, because $T_\alpha \in \mathfrak{h}$. Therefore, $\operatorname{ad} T_\alpha$ is both semisimple and nilpotent and must be zero. Hence, $T_\alpha = 0$.

(3) Take $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$ with $B(X,Y) \neq 0$. For $Z \in \mathfrak{g}_\alpha$, we have that

$$\begin{aligned} [[X,Y],Z] &= [B(X,Y)T_\alpha, Z] \\ &= B(X,Y)\alpha(T_\alpha)Z \end{aligned}$$

This is nonzero if Z is.

(4) Take $X_\alpha \in \mathfrak{g}_\alpha$. Find $Y_\alpha \in \mathfrak{g}_{-\alpha}$ such that

$$B(X_\alpha, Y_\alpha) = \frac{2}{\alpha(T_\alpha)}$$

Set

$$H_\alpha = \frac{2}{B(T_\alpha, T_\alpha)} T_\alpha.$$

Now check the $\mathfrak{sl}_2$ relations. We have that

$$[X_\alpha, Y_\alpha] = B(X_\alpha, Y_\alpha)T_\alpha = H_\alpha$$

$$[H_\alpha, X_\alpha] = \frac{2}{\alpha(T_\alpha)}[T_\alpha, X_\alpha] = 2X_\alpha$$

Similarly,

$$[H_\alpha, Y_\alpha] = -2Y_\alpha.$$

So this is isomorphic to $\mathfrak{sl}_2$. □

Proposition 5.28 ("Weights Add")**.** Let $\mathfrak{g}$ be semisimple with Cartan decomposition $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus \mathfrak{g}_\alpha$, and let V, W be $\mathfrak{g}$-modules with V_α, W_α the corresponding weight spaces. Then

(1) $\mathfrak{g}_\alpha V_\beta \subseteq V_{\alpha+\beta}$

(2) $V_\alpha \otimes W_\beta \subseteq (V \otimes W)_{\alpha+\beta}$

Lemma 5.29.

(1) If V is a finite-dimensional representation, then $V|_{\mathfrak{s}_\alpha}$ is a finite-dimensional representation of $\mathfrak{s}_\alpha$.

(2) If V is a representation for $\mathfrak{g}$, then

$$\sum_{n\in\mathbb{Z}} V_{\beta+n\alpha}$$

is an $\mathfrak{s}_\alpha$ submodule.

(3) $\beta(H_\alpha) \in \mathbb{Z}$ for all roots β and α, and $H_\alpha \in \mathfrak{h}$.

Proof.

(1) This follows from generic facts about restriction of representations.

(2) $\mathfrak{s}_\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. So this space is mapped to itself by $\mathfrak{s}_\alpha$.

(3) The eigenvalues of H_α on $V|_{\mathfrak{s}_\alpha}$ are integers, but each V_β is a set of eigenvectors on which H_α acts by the scalar $\beta(H_\alpha)$. Hence, $\beta(H_\alpha)$ is an integer. □

Proposition 5.30. The root spaces of $\mathfrak{g}_\alpha$ are 1-dimensional. The only roots proportional to α are $\pm\alpha$. In particular, twice a root is not a root.

Proof. For the first part, let α be a root. Let's assume that $\dim \mathfrak{g}_\alpha > 1$. Then let Y be a nonzero element of $\mathfrak{g}_{-\alpha}$. Then we can arrange that there is X_α such that $B(X_\alpha, Y) = 0$. We choose X_α by producing two independent elements of $\mathfrak{g}_\alpha$ and scaling appropriately and adding them together.

Now let Y_α be such that $\mathfrak{s}_\alpha = \langle X_\alpha, Y_\alpha, H_\alpha\rangle \cong \mathfrak{sl}_2$. We have

$$[X_\alpha, Y] = B(X_\alpha, Y)T_\alpha = 0.$$

So Y is killed by X_α, but $[H_\alpha, Y] = -2Y$, since $\mathfrak{g}_{-\alpha} \oplus \mathfrak{h} \oplus \mathfrak{g}_\alpha$ is a representation of $\mathfrak{s}_\alpha$, and $Y \in \mathfrak{g}_{-\alpha}$. So Y is in the -2 weight-space for H_α. But Y is killed by $\operatorname{ad} X_\alpha$. This is incompatible with the representation theory for $\mathfrak{sl}_2$, because $\operatorname{ad} X_\alpha$ should raise Y into the 0 weight-space; in particular, we should have that $[X_\alpha, Y] = H_\alpha$, yet $H_\alpha \neq 0$. This is a contradiction. Hence $\dim \mathfrak{g}_\alpha \leqslant 1$.

To see that the only roots proportional to α are $\pm\alpha$, assume that there is $\zeta \in \mathbb{C}$ with $\beta = \zeta\alpha$ and β is a root. Then $2\zeta = \zeta\alpha(H_\alpha) = \beta(H_\alpha) \in \mathbb{Z}$ by Lemma 5.29. Exchanging α and β, we see that $2\zeta^{-1} \in \mathbb{Z}$. The two equations $2\zeta, 2\zeta^{-1} \in \mathbb{Z}$ limits the possibilities to $\zeta \in \{\pm 1/2, \pm 1, \pm 2\}$.

We must exclude $\pm 1/2$ and ± 2 from these possibilities. Since the negative of a root is a root, we only need to check this for $\zeta = 1/2$ and $\zeta = 2$. Further, by exchanging α and β, we need only check the case that $\zeta = 2$.

So assume $\beta = 2\alpha$. Define

$$\mathfrak{a} = \mathfrak{h} \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \mathfrak{g}_{2\alpha} \oplus \mathfrak{g}_{-2\alpha}.$$

This is a representation for $\mathfrak{s}_\alpha \cong \mathfrak{sl}_2$. But if $X \in \mathfrak{g}_\alpha$ and $X_\alpha \in \mathfrak{s}_\alpha \cap \mathfrak{g}_\alpha$, then $[X_\alpha, X] = 0$ because $[X_\alpha, X] \in [\mathfrak{g}_\alpha, \mathfrak{g}_\alpha] = 0$ (we know that $[\mathfrak{g}_\alpha, \mathfrak{g}_\alpha] = 0$ since $\mathfrak{g}_\alpha$ is 1-dimensional). This again contradicts the representation theory of $\mathfrak{sl}_2$, because the highest weight space is $\mathfrak{g}_{2\alpha}$, yet not in the image of X_α. This is a contradiction. □

Proposition 5.31 (Facts about $\mathfrak{s}_\alpha$).

(1) $\mathfrak{s}_\alpha = \mathfrak{s}_{-\alpha}$.

(2) $H_\alpha = -H_{-\alpha}$.

Previously, we showed that the root spaces of a semisimple complex Lie algebra were 1-dimensional, and that $\mathfrak{g}$ is composed of copies of $\mathfrak{sl}_2$, given by

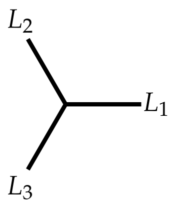
$$\mathfrak{s}_\alpha = \mathfrak{g}_{-\alpha} \oplus [\mathfrak{g}_{-\alpha}, \mathfrak{g}_\alpha] \oplus \mathfrak{g}_\alpha.$$

We would like to expand our theory of representations of $\mathfrak{sl}(2)$ to other Lie algebras, including weight diagrams. These will in general be difficult to draw, but at least for $\mathfrak{sl}(3)$ we can draw them in 2-dimensions.

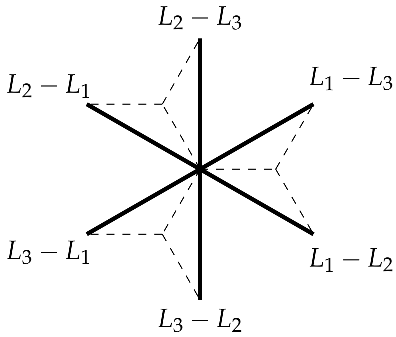
For $\mathfrak{sl}(3)$, recall that we had linear functionals L_1, L_2, L_3 spanning $\mathfrak{h}^*$, satisfying $L_1 + L_2 + L_3 = 0$. So we can represent the weights in the plane $\mathbb{C}[L_1, L_2, L_3]/\langle L_1 + L_2 + L_3 \rangle$.

Example 5.32.

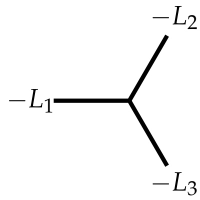
(1) Let $V = \mathbb{C}^3$ be the standard representation. Then the weights of V are L_1, L_2, and L_3.



(2) Let $V \cong \mathfrak{sl}(3)$ via the adjoint representation. Weights of V are $L_i - L_j$ for $i \neq j$.



(3) The dual representation of the standard representation has weights $-L_i$, and therefore the diagram



Definition 5.33. We define the **weight lattice**

$$\Lambda_W = \{\beta \in \mathfrak{h}^* \mid \beta(H_\alpha) \in \mathbb{Z} \text{ for all roots } \alpha\}.$$

We let $\mathrm{wt}(V)$ denote the set of weights in a representation V of $\mathfrak{g}$. And by Lemma 5.29(1), $\mathrm{wt}(V) \subseteq \Lambda_W$.

There is additional symmetry arising from the subalgebras $\mathfrak{s}_\alpha \cong \mathfrak{sl}(2)$. For instance, the fact that the weight multiplicities are symmetric about the origin. So define hyperplanes

$$\Omega_\alpha = \{\beta \in \mathfrak{h}^* \mid \beta(H_\alpha) = 0\}.$$

Our symmetry amounts to saying that $\mathrm{wt}(V)$ is closed under reflections W_α across Ω_α.

More explicitly, to see that the weights are closed under these reflections, compute

$$W_\alpha(\beta) = \beta - \frac{2\beta(H_\alpha)}{\alpha(H_\alpha)}\alpha = \beta - \beta(H_\alpha)\alpha.$$

Take the submodule $Z = \sum_{n\in\mathbb{Z}} V_{\beta+n\alpha}$ for $\mathfrak{s}_\alpha$. Pick $v \in V_\beta$, say; then $H_\alpha v = \beta(H_\alpha)v$.

In Z, we must be able to find w such that

$$H_\alpha w = -\beta(H_\alpha)w.$$

Now

$$-\beta(H_\alpha) = \beta(H_\alpha) + m\alpha(H_\alpha) \implies -2\beta(H_\alpha) = m\alpha(H_\alpha)$$

$$\implies m = \frac{-2\beta(H_\alpha)}{\alpha(H_\alpha)} \implies m = -\beta(H_\alpha).$$

This implies that $\beta(H_\alpha) = -m$. Therefore, the element v of the β-weight-space $v \in V_\beta$ corresponds to $w \in V_{\beta+m\alpha} = V_{\beta-\beta(H_\alpha)\alpha}$ as required. In fact, we obtain an isomorphism $V_\beta \cong V_{\beta-\beta(H_\alpha)\alpha}$.

Remark 5.34 (Notation)**.** The integer

$$\beta(H_\alpha) = \frac{2\beta(H_\alpha)}{\alpha(H_\alpha)} = \frac{2\beta(T_\alpha)}{\alpha(T_\alpha)}$$

is often denoted by

$$\langle \beta, \alpha^\vee \rangle := \beta(H_\alpha),$$

and $\alpha^\vee$ is the coroot to α (in the case of Lie algebras, $\alpha^\vee = T_\alpha$ as we defined it).

The important thing to remember is that $\langle \beta, \alpha^\vee \rangle$ is the number of α's you need to take off β to reflect β in the hyperplane perpendicular to α.

Definition 5.35. Given a semisimple Lie algebra $\mathfrak{g}$, we define the **Weyl group** W as the group generated by the hyperplane reflections W_α,

$$W := \langle \{W_\alpha \mid \alpha \text{ is a root of } \mathfrak{g}\} \rangle$$

In fact, W is a finite group. Note that W preserves $\mathrm{wt}(V)$ for any representation V of $\mathfrak{g}$.

In order to generalize the idea of a highest weight vector as we had for $\mathfrak{sl}(2)$, it will be convenient to pick a complete ordering on Λ_W. In $\Lambda_W \otimes \mathbb{R}$, we choose a linear map $\ell\colon \Lambda_W \otimes \mathbb{R} \to \mathbb{R}$ satisfying $\alpha > \beta$ if and only if $\ell(\alpha) > \ell(\beta)$. To choose such than an ℓ, choose the gradient of ℓ irrational with respect to the weight lattice.

Example 5.36. In $\mathfrak{sl}(3)$,

$$\ell(\alpha) = \alpha \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} - 1 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

In this case, $\ell(L_1) = 1$, $\ell(L_2) = \frac{1}{\sqrt{2}} - 1$, $\ell(L_3) = -\frac{1}{\sqrt{2}}$. With this choice of ℓ,

$$L_1 > -L_3 > -L_2 > 0 > L_2 > L_3 > -L_1.$$

We can also check that

$$L_1 - L_3 > L_1 - L_2 > L_2 - L_3 > 0.$$

Definition 5.37. Given a semisimple Lie algebra $\mathfrak{g}$, denote by R the collection of roots, and define $R^+ = \{\alpha \in R \mid \alpha > 0\}$, and $R^- = \{\alpha \in R \mid \alpha < 0\}$.

Lemma 5.38. The subalgebras $\mathfrak{s}_\alpha$ span $\mathfrak{g}$ as a vector space.

Proof. We clearly get all root spaces $\mathfrak{g}_\alpha$ in this way, since $\mathfrak{g}_\alpha \subseteq \mathfrak{s}_\alpha$, so it's just a matter of checking that we get the whole of the Cartan. By Proposition 5.26, the dual $\mathfrak{h}^*$ is spanned by the roots. Now the Killing form gives an isomorphism between $\mathfrak{h} \to \mathfrak{h}^*$ under which $T_\alpha \mapsto \alpha$. But $T_\alpha \in \mathfrak{s}_\alpha$ for each α, as T_α is a multiple of H_α. □

Remark 5.39. The *Weight Lattice* is a game show derived from a Japanese concept wherein participants are suspended from a large metal lattice over the course of a week, while their families and friends must throw a sufficient quantity food to them so that they gain enough weight to touch the ground. The winners get a trip to the Bahamas, while the rest are humiliated for their fast metabolism.

Recall that the Weight Lattice is

$$\Lambda_W = \{\lambda \in \mathfrak{h}^* \mid \lambda(H_\alpha) \in \mathbb{Z} \text{ for all roots } \alpha\}.$$

Proposition 5.40. Let $\mathfrak{g}$ be semisimple, and let V be a finite-dimensional representation for $\mathfrak{g}$. Then

(1) V has a highest weight, λ say, such that $V_\lambda \neq 0$ and $V_\beta = 0$ for any $\beta > \lambda$ using the functional ℓ;

(2) If α is a highest weight, and $\beta \in R^+$ is a positive root, then $\mathfrak{g}_\beta V_\alpha = 0$.

(3) Given any nonzero $v \in V_\lambda$, where λ is a highest weight, then the subspace W generated by all vectors $Y_{\alpha_1} \cdots Y_{\alpha_k} v$ with $\alpha_i \in R^+$ and $Y_{\alpha_i} \in \mathfrak{g}_{-\alpha_i}$ for all $k \geqslant 0$ is an irreducible $\mathfrak{g}$-submodule.

(4) If V is irreducible, then $W = V$.

Proof.

(1) Just take λ maximal under the ordering subject to $V_\lambda \neq 0$. Such a weight space exists because we assumed that V is finite dimensional.

(2) Since $\mathfrak{g}_\beta V_\alpha \subseteq V_{\alpha+\beta}$ and $\ell(\alpha+\beta) = \ell(\alpha) + \ell(\beta) > \ell(\alpha)$ since $\beta \in R^+$, but α was a highest weight, so $V_{\alpha+\beta} = 0$.

(3) Let's first show that W is a submodule. By construction, W is stable under all $\mathfrak{g}_{-\alpha}$ for $\alpha \in R^+$. Also, v is a weight vector, hence stable under h, and since weights add, each $Y_{\alpha_1} \cdots Y_{\alpha_k} v$ is also a weight vector of weight $\lambda - \alpha_1 - \alpha_2 - \cdots - \alpha_k$. So W is stable under $\mathfrak{h}$. So it remains to show that W is stable under $X_\alpha \in \mathfrak{g}_\alpha$ for $\alpha \in R^+$.

Since $\mathfrak{g}_\alpha v = 0$ for all $\alpha \in R^+$, (as v is a highest weight vector), we proceed by induction on i, showing that $V_{(i)} = \langle Y_{\alpha_k} \cdots Y_{\alpha_1} v \mid 1 \leqslant k \leqslant i \rangle$ is stable under X. We have the result for $i = 0$ above.

Assume now that this holds for $i' < i$. Then calculate

$$X_\beta Y_{\alpha_i} \cdots Y_{\alpha_1} v = [X_\beta, Y_{\alpha_i}] Y_{\alpha_{i-1}} \cdots Y_{\alpha_1} v + Y_{\alpha_i} X_\beta Y_{\alpha_{i-1}} \cdots Y_{\alpha_1} v \tag{6}$$

Note that

$$Y_{\alpha_{i-1}} \cdots Y_{\alpha_1} V \in V_{(i-1)},$$

so

$$X_\beta Y_{\alpha_{i-1}} \cdots Y_{\alpha_1} v \in V_{(i-1)}$$

by induction, and therefore

$$Y_{\alpha_i} X_\beta Y_{\alpha_{i-1}} \cdots Y_{\alpha_1} v \in V_{(i)}.$$

This deals with the second term on the right hand side of (6). To deal with the first term on the right hand side of (6), notice that, as before,

$$Y_{\alpha_{i-1}} \cdots Y_{\alpha_1} v \in V_{(i-1)},$$

and $[X_\beta, Y_{\alpha_i}]$ is an element of a root space or in $\mathfrak{h}$. If $[X_\beta, Y_{\alpha_i}] \in \mathfrak{g}_\alpha$ for $\alpha \in R^+$, it follows that

$$[X_\beta, Y_{\alpha_i}] Y_{\alpha_{i-1}} \cdots Y_{\alpha_1} v \in V_{(i)}$$

by induction. Similarly if $[X_\beta, Y_{\alpha_i}] \in \mathfrak{g}_{-\alpha}$, or if $[X_\beta, Y_\alpha] \in \mathfrak{h}$.

This shows that W is in fact a submodule.

To see that W is irreducible, write $W = W_1 \oplus W_2$, and suppose the highest weight vector $v = v_1 + v_2$ with $v_1 \in W_1$, $v_2 \in W_2$. Then

$$Hv = \lambda(H) v = \lambda(H) v_1 + \lambda(H) v_2,$$

so under projection $\pi \colon W \to W_1$, we have that v_1 is also a highest weight vector for W_1, and similarly v_2 is a highest weight vector for W_2. So if $v_1, v_2 \neq 0$ then $\langle v_1, v_2 \rangle$ spans a subspace of W_λ with dimension larger than 1. This is a contradiction, since W_λ is 1-dimensional and generated by v.

(4) Since W is a non-zero submodule of an irreducible module, then $W = V$.

□

Proposition 5.41. $\mathfrak{g}$-modules are determined up to isomorphism by their highest weight. Let V and W be two irreducible representations with highest weight λ. Then $V \cong W$.

Proof. Let v, w be highest weight vectors for V and W, respectively. Let U be the submodule of $V \oplus W$ generated by $\mathfrak{g} \cdot (v, w)$. By Proposition 5.40(c), and the projections $U \to V$ and $U \to W$ are nonzero, they must be isomorphisms. □

What possibilities are there for highest weights?

Definition 5.42. Let $\mathbb{E} = \mathbb{R}\Lambda_W$. Then for α a root, define

$$\mathbb{E}_\alpha^+ = \{\beta \in \mathbb{E} \mid \beta(H_\alpha) > 0\},$$

$$\mathbb{E}_\alpha^- = \{\beta \in \mathbb{E} \mid \beta(H_\alpha) < 0\}.$$

$$\mathbb{E}_\alpha = \Omega_\alpha \sqcup \mathbb{E}_\alpha^+ \sqcup \mathbb{E}_\alpha^-$$

Recall that $\Omega_\alpha = \{\beta \in \mathfrak{h}^* \mid \beta(H_\alpha) = 0\}$.

Now number the positive roots by $\alpha_1, \ldots, \alpha_n$ and define

$$W_{(\pm,\pm,\ldots,\pm)} = \mathbb{E}_{\alpha_1}^\pm \cap \mathbb{E}_{\alpha_2}^\pm \cap \ldots \cap \mathbb{E}_{\alpha_n}^\pm.$$

In particular

$$\mathbb{R}^{>0}\Lambda_W^+ = W_{(+,+,\ldots,+)}$$

is called the **fundamental Weyl chamber** and

$$\Lambda_W^+ = \Lambda_W \cap \overline{W}_{(+,+,\ldots,+)}$$

is the set of **dominant weights**, where the bar denotes topological closure (these things lie in some $\mathbb{R}^n$).

Proposition 5.43. If λ is a highest weight for some finite-dimensional representation, then $\lambda \in \Lambda_W^+$.

Proof. Suppose $\beta \in \mathbb{E}_\alpha^-$ for some $\alpha \in R^+$, then $W_\alpha(\beta)$ is also a weight, and we have

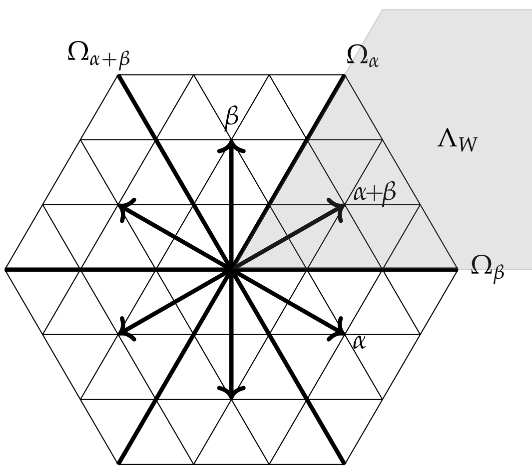
$$\ell(W_\alpha(\beta)) = \ell(\beta - \beta(H_\alpha)\alpha) = \ell(\beta) - \beta(H_\alpha)\ell(\alpha)$$

Note that $\beta(H_\alpha) < 0$ and $\ell(\alpha) > 0$, so we conclude that

$$\ell(W_\alpha(\beta)) = \ell(\beta) - \beta(H_\alpha)\ell(\alpha) > \ell(\beta).$$

So there is a higher weight in the representation. □

Example 5.44. Let's work this out in detail for $\mathfrak{sl}(3)$. The roots are $\alpha = L_1 - L_2$, $\beta = L_2 - L_3$, $\alpha + \beta = L_1 - L_3$, $-\alpha$, $-\beta$, $-\alpha - \beta$. We depicted these as



In this picture, $\mathbb{E}_\alpha^+$ is the half-plane bounded by Ω_α containing α, and $\mathbb{E}_\alpha^-$ is the half-plane bounded by Ω_α that contains $-\alpha$. Similarly for $\mathbb{E}_\beta^\pm$ and $\mathbb{E}_{\alpha+\beta}^\pm$. The fundamental Weyl Chamber is the region bounded by Ω_β and Ω_α containing $\alpha + \beta$.

The dominant weights are λ such that $\lambda(H_\mu) > 0$ for all $\mu \in R^+$. Claim that Λ_W^+ is generated by L_1 and $-L_3$. So clearly L_1 and $-L_3$ span $\mathfrak{h}^*$, and for any dominant weight λ, we require that

$$\lambda(H_\alpha) \in \mathbb{Z}_{\geqslant 0}$$

$$\lambda(H_\beta) \in \mathbb{Z}_{\geqslant 0}$$

$$\lambda(H_{\alpha+\beta}) \in \mathbb{Z}_{\geqslant 0}$$

The point is that once we know λ on H_α and H_β, then we know it on $H_{\alpha+\beta}$. If $\lambda(H_\alpha) = a$ and $\lambda(H_\beta) = b$, then $\lambda = aL_1 - bL_3$. To check this, $L_1(H_\alpha) = 1$ and $-L_3(H_\alpha) = 0$, $L_1(H_\beta) = 0$, $-L_3(H_\beta) = 1$.

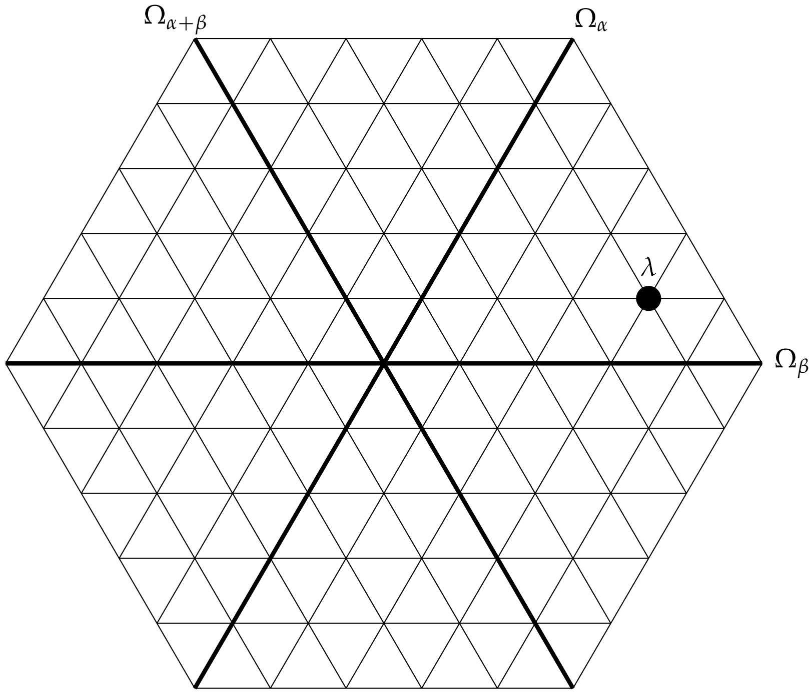
So any irreducible module is isomorphic to $\Gamma_{a,b}$ for some a, b, where $\Gamma_{a,b}$ has highest weight $aL_1 - bL_3$. Moreover, all such must exist.

And $\Gamma_{1,0} \cong V$ is the standard rep with highest weight L_1, and $\Gamma_{0,1}$ is it's dual V^*.

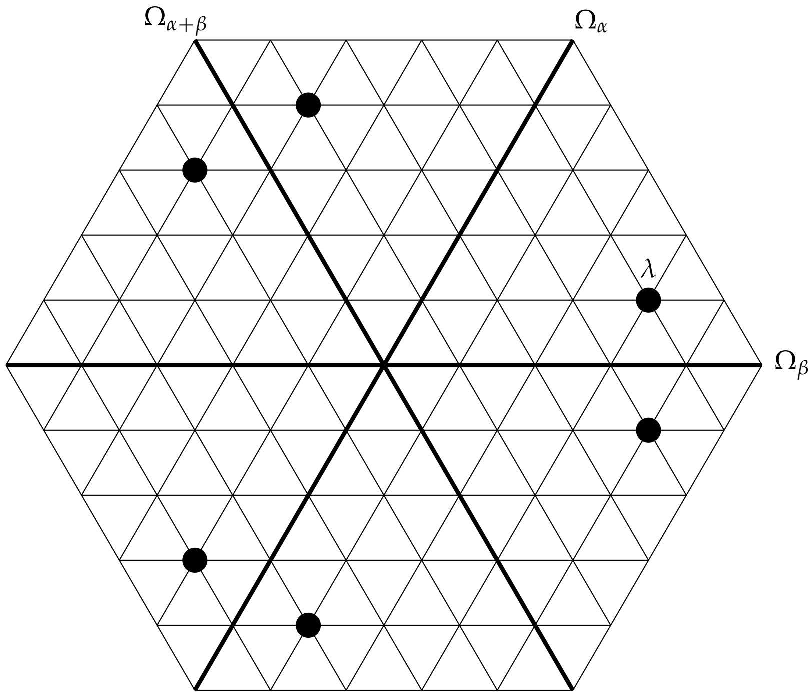
Moreover, $\Gamma_{a,b}$ must be contaınd in the tensor product $(\Gamma_{1,0})^{\otimes a} \otimes (\Gamma_{0,1})^{\otimes b}$.

Remark 5.45. Last time we were talking about *Weyl Chambers*, which is a 1990's adult entertainment film by BDSM specialists "Blood and Chains." For reasons of decency I can't go into the details.

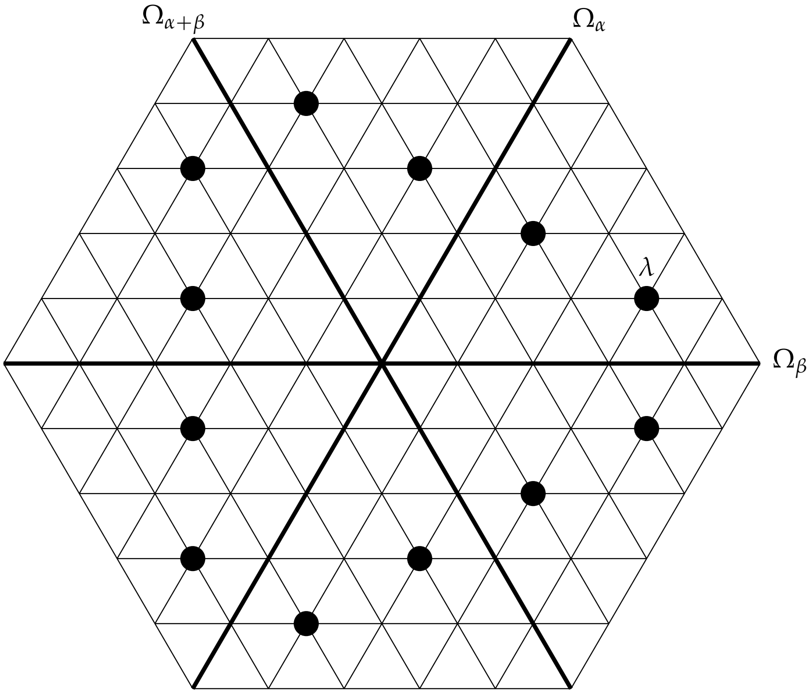
Example 5.46. Let's construct $\Gamma_{3,1}$. The representation is generated by the highest weight vector $\lambda = 3L_1 - L_3$.



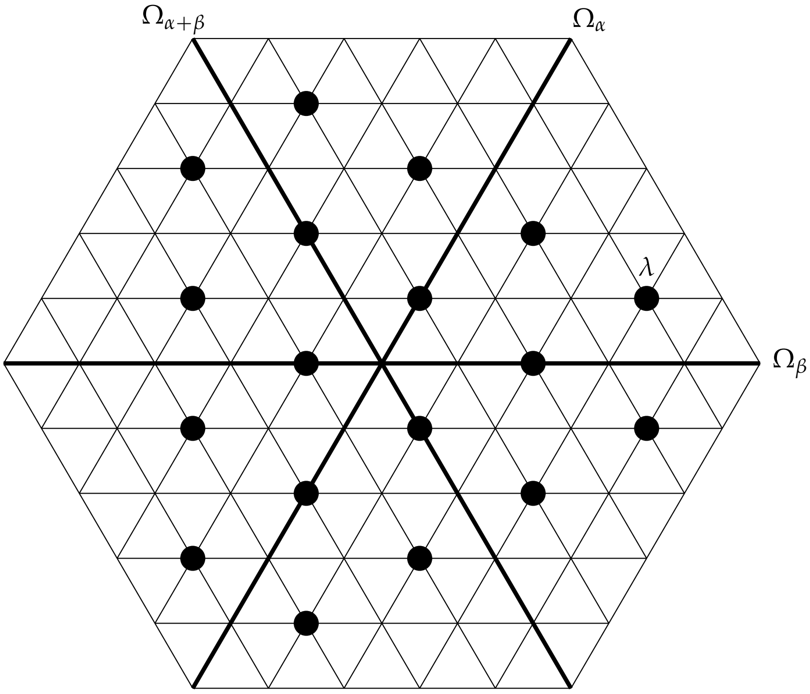
The weights are stable under the reflection in the hyperplanes Ω_{α_i}, so we reflect in these hyperplanes to find other roots.



Once we've done so, we know that a weight μ and it's reflection over any hyperplane Ω_{α_i} forms a representation of a copy of $\mathfrak{sl}(2)$, so we should fill in all the steps in-between these weights as the weight spaces of that $\mathfrak{sl}(2)$ representation.



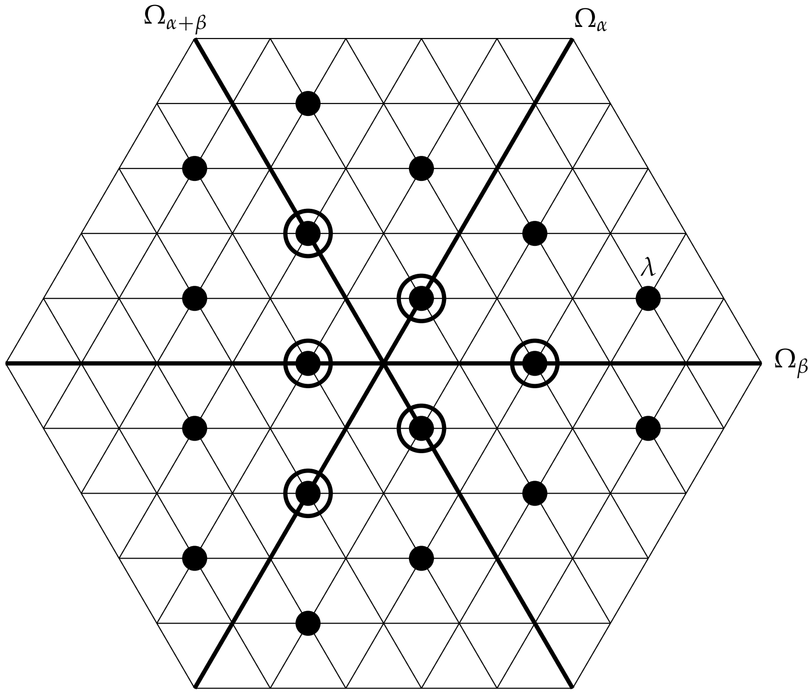
This forms a border of the weight-space of the representation, and we can use the same rule again to fill in the dots inside the borders; along any line parallel to the roots, we have another $\mathfrak{sl}(2)$ representation with highest and lowest weights on the border of the weight space, so we fill in all the even steps in-between.



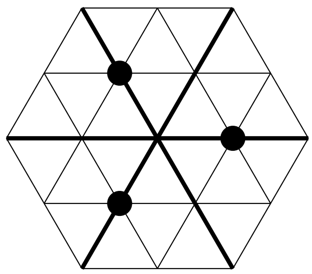
We don't yet know the multiplicities of the weights (= dimensions of the weight spaces), but we can use the following rule.

Fact 5.47 (Rule of Thumb)**.** Multiplicites of weights of irreducible $\mathfrak{sl}(3)$ representations increase by one when moving in towards the origin in the weights space from a hexagon-border, and remain stable when moving in towards the origin in a triangle-border.

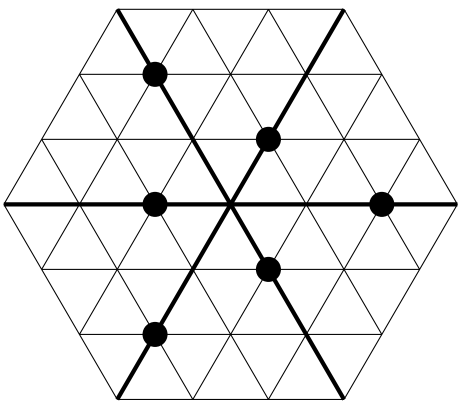
We draw concentric circles for each multiplicity past the first. So the representation $\Gamma_{3,1}$ has the weight diagram as below.



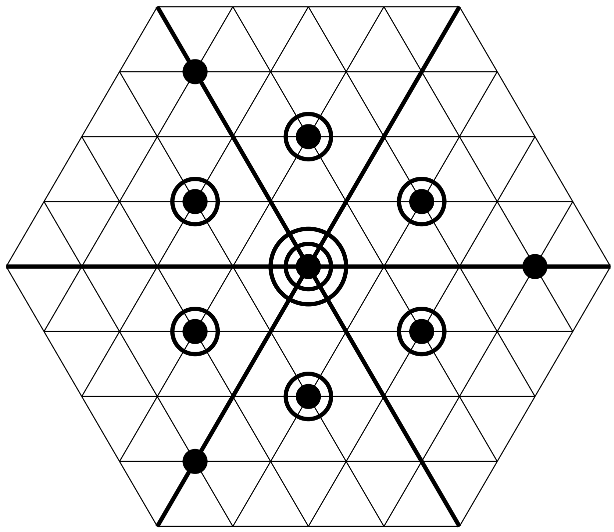
Example 5.48. Suppose V is the standard representation $V \cong \Gamma_{1,0}$.



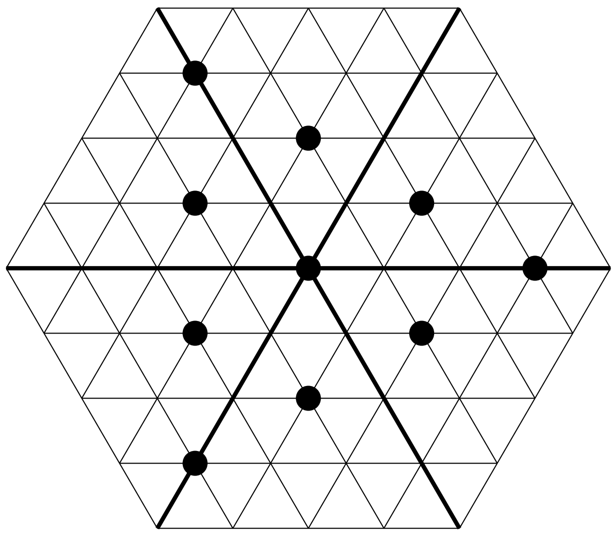
Then $\mathrm{Sym}^2(V)$ has weights the sums of distinct pairs in $\Gamma_{1,0}$, and we see that this is $\Gamma_{2,0}$ when we compare the weight diagrams for $\mathrm{Sym}^2(V)$ and $\Gamma_{2,0}$. Hence, $\mathrm{Sym}^2(V) \cong \Gamma_{2,0}$ is irreducible – as in the weight diagram below.



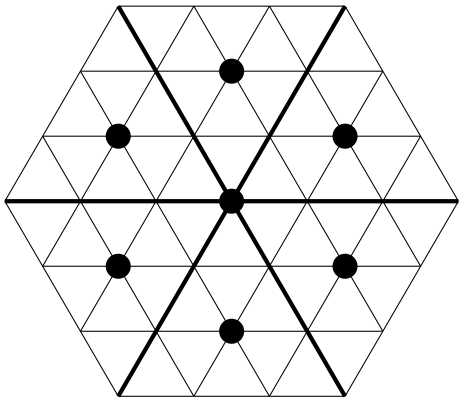
By adding all the weights in V^* with weights of $\Gamma_{2,0}$, we get weights of $\mathrm{Sym}^2(V) \otimes V^*$, we get the following diagram



Note that $\Gamma_{3,0}$ has the weight diagram



And $\Gamma_{1,1}$ has the weight diagram



Taking $\Gamma_{3,0}$ from $\mathrm{Sym}^2(V) \otimes V^*$ we are left with a weight diagram for $\Gamma_{1,1} \oplus \Gamma_{0,0}$. Therefore, $\mathrm{Sym}^2(V) \otimes V^* \cong \Gamma_{3,0} \oplus \Gamma_{1,1} \oplus \Gamma_{0,0}$.

Remark 5.49. Clearly you can see this is a wonderful source of exam questions (*hint hint*). In fact, this has many applications in physics, where decomposing the tensor product of two representations into irreducible direct summands corresponds to what comes out of the collision of two particles, so it's not surprising that there are many algorithms and formulas for this kind of thing.

Let's summarize what we now know from our investigation of representations of $\mathfrak{sl}(3)$. Let $\mathfrak{g}$ be a semisimple complex Lie algebra, $\mathfrak{h}$ its Cartan subalgebra, R its set of roots, Λ_W the weight lattice, W the Weyl group with accompanying reflecting hyperplanes Ω_α. Pick a linear functional ℓ with irrational slope with respect to the weight lattice, and Λ_W^+ the dominant weights. We also get $R = R^+ \sqcup R^-$.

Definition 5.50. Let α be a positive root which is not expressible as the sum of two positive roots. Then we say α is a **simple root**.

Definition 5.51. The **rank of** $\mathfrak{g}$ is the dimension of the Cartan subalgebra $\mathfrak{h}$.

Fact 5.52.

(1) Under $\mathbb{Z}$, the simple roots generate all roots, i.e. if S is the set of simple roots, $\mathbb{Z}S \cap R = R$.

(2) The number of simple roots is equal to the rank of $\mathfrak{g}$.

(3) Any root is expressible as $\omega \cdot \alpha$ for $\omega \in W$, α a simple root.

(4) The Weyl group is generated by reflections W_α for all simple roots α.

(5) The Weyl group acts simply transitively on the set of decompositions of R into positive and negative parts. (The action has only one orbit, and if the action of any element σ has a fixed point, then σ is the identity of W).

(6) The elements H_α such that α is a simple root generate the lattice

$$\mathbb{Z}\{H_\alpha \mid \alpha \in R\} \subseteq \mathfrak{h}.$$

(7) Define the **fundamental dominant weights** ω_α for each simple root α by the property that $\omega_\alpha(H_\beta) = \delta_{\alpha\beta}$ for α, β simple roots. They generate the weight lattice Λ_W.

(8) The set $\mathbb{Z}_{\geqslant 0}\{\omega_\alpha\}$ is precisely the set of dominant weights.

(9) Every representation has a dominant highest weight, and there exists one and only one representation with this highest weight up to isomorphism.

(10) The set of weights of a representation is stable under the Weyl group, and moreover we can use $\mathfrak{sl}(2)$-theory to establish the set of weights (but maybe not the multiplicities) in a given representation.

(11) The multiplicities are not obvious.

The fact that the multiplicities are not obvious is the motivation for the next section.

5.1 Multiplicity Formulae

Let's define an inner product on $\mathfrak{h}^*$ via $(\alpha, \beta) = B(T_\alpha, T_\beta)$. Recall that $T_\alpha \in \mathfrak{h}$ is dual to $\alpha \in \mathfrak{h}^*$, with $B(T_\alpha, H) = \alpha(H)$ for any $H \in \mathfrak{h}$.

Proposition 5.53 (Freudenthal's Multiplicity Formula)**.** Given a semisimple Lie algebra $\mathfrak{g}$ and irreducible Γ_λ with highest weight λ, then

$$c(\mu)n_\mu(\Gamma_\lambda) = 2\sum_{\alpha\in R^+}\sum_{k\geqslant 1}(\mu + k\alpha)\,\alpha n_{\mu+k\alpha}(\Gamma_\lambda)$$

where $n_\mu(\Gamma_\lambda) = \dim(\Gamma_\lambda)_\mu$ and $c(\mu)$ is defined by

$$c(\mu) = \|\lambda + \rho\|^2 - \|\mu + \rho\|^2$$

where

$$\rho = \frac{1}{2}\sum_{\alpha\in R^+}\alpha.$$

Remark 5.54. *Freudenthal's Multiplicity Formula* is a sequel to *Complete Reducibility*, wherein Rick Moranis plays the mad scientist Freudenthal, who invents a chemical formula that duplicates DNA. Unfortunately, the bad guy (played by Bill Murray) gets a hold of this formula and takes a shower in it, making thousands of Bill Murrays. He then manages to infiltrate the Pentagon and get the nuclear codes, and the planet is destroyed within hours.

Remark 5.55. Today we'll be talking about *Root Systems*, which is an upcoming indie film about the fallout from the Fukashima Nuclear Reactor. Some ginger from near the plant mutates and starts to grow out of control. And since it's a major component of Japanese cuisine, it wants to take revenge on the people who've been eating it for so long. At first it just pops out of the ground and squirts hot ginger at people's faces, but it has more diabolical intentions. Eventually, it finds an underground internet cable and starts sending messages to the world's leaders. To show that it means business, it deletes all cat videos from the internet. To try and stop the mutant ginger, some samurai warriors, the X-men, Batman and Captain America are sent to destroy it. But they're ultimately unsuccessful, and the ginger takes over the world.

6 Classification of Complex Semisimple Lie Algebras

In this section, we can basically forget everything we've talked about so far and distill the information about Lie algebras into a few basic facts that will be all that we need to classify the Lie algebras.

Definition 6.1. Let $\mathbb{E} = \mathbb{R}^n$ for some $n \in \mathbb{N}$, equipped with an inner product $(\ ,\ )$. Then a **root system** on $\mathbb{E}$ is a finite set R such that

(R1) R spans $\mathbb{E}$ as a vector space.

(R2) $\alpha \in R \iff -\alpha$ a root, but $k\alpha$ is not a root for all $k \neq \pm 1$.

(R3) for $\alpha \in R$, the reflection W_α in the hyperplane $\alpha^\perp$ perpindicular to α is a map from $R \to R$.

(R4) For roots $\alpha, \beta \in R$, the real number $n_{\beta\alpha} = 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}$ is an integer.

Exercise 6.2. Show that the root system of a Lie algebra forms an abstract root system.

Remark 6.3. Note that for $\mathfrak{g}$ semisimple, $n_{\beta\alpha}$ would be $\beta(H_\alpha)$, and

$$W_\alpha(\beta) = \beta - \beta(H_\alpha)\alpha = \beta - n_{\beta\alpha}\alpha.$$

What are the possibilities for $n_{\beta\alpha}$? Turns out there are very few possibilities. We have that

$$n_{\beta\alpha} = 2\cos\theta\frac{\|\beta\|}{\|\alpha\|},$$

where θ is the angle between α and β. Hence, $n_{\alpha\beta}n_{\beta\alpha} = 4\cos^2\theta \in \mathbb{Z}$. Since $|\cos^2\theta| \leqslant 1$, we see that

$$n_{\beta\alpha}n_{\alpha\beta} \in \{0, 1, 2, 3, 4\}.$$

So $n_{\alpha\beta}$ is an integer between -4 and 4, since $n_{\beta\alpha}$ is also an integer. If $\beta \neq \pm\alpha$ then $n_{\beta\alpha}$ lies between -3 and 3.

Furthermore, $n_{\beta\alpha}$ has the same sign as $n_{\alpha\beta}$, and if $|n_{\alpha\beta}|, |n_{\beta\alpha}| > 1$, then $|n_{\alpha\beta}| = |n_{\beta\alpha}| = 2$, and so $\cos^2\theta = 1 \implies \alpha = \pm\beta$.

We may assume that $n_{\alpha\beta} = \pm 1$. So what are the options for $n_{\beta\alpha}$?

$$\begin{array}{c|ccccccc} n_{\beta\alpha} & 3 & 2 & 1 & 0 & -1 & -2 & -3 \\ n_{\alpha\beta} & 1 & 1 & 1 & 0 & -1 & -1 & -1 \\ \cos(\theta) & \sqrt{3}/2 & \sqrt{2}/2 & 1/2 & 0 & -1/2 & -\sqrt{2}/2 & -\sqrt{3}/2 \\ \theta & \pi/6 & \pi/4 & \pi/3 & \pi/2 & 2\pi/3 & 3\pi/4 & 5\pi/6 \\ \frac{\|\beta\|}{\|\alpha\|} & \sqrt{3} & \sqrt{2} & 1 & * & 1 & \sqrt{2} & \sqrt{3} \end{array} \tag{7}$$

Some consequences of this table are the following facts. This and the next proposition should settle any outstanding proofs owed for Fact 5.52.

Fact 6.4.

(1) In an α-string through β, there are at most 4 elements.

(2) If α and β are roots, $\alpha \neq \pm\beta$, then $(\beta, \alpha) > 0$ implies that $\alpha - \beta$ is a root. If $(\beta, \alpha) < 0$, then $\alpha + \beta$ is a root. If $(\beta, \alpha) = 0$ then $\beta + \alpha$ and $\alpha - \beta$ are either both roots or both non-roots.

(3) As with $\mathbb{R}\Lambda_W$, we can find $\ell \colon \mathbb{E} \to \mathbb{R}$ irrational with respect to $\mathbb{Z}R$ such that it separates R into $R = R^+ \sqcup R^-$. With respect to this ordering, we say that a positive root α is **simple** if $\alpha \neq \beta + \gamma$ for any $\beta, \gamma \in R^+$.

(4) If α, β are distinct simple roots, then $\alpha - \beta$ and $\beta - \alpha$ are not roots.

(5) If α, β are distinct simple roots, then then angle between them is obtuse, $(\alpha, \beta) < 0$.

(6) The set of simple roots S is linearly independent.

(7) Every positive root is a nonnegative integral combination of the simple roots.

Proof. (1) Let α, β be roots, with $\alpha \neq \pm\beta$. Then consider an α-string through β given by $\{\beta - p\alpha, \beta - (p+1)\alpha, \ldots, \beta + q\alpha\}$. We have

$$W_\alpha(\beta + q\alpha) = W_\alpha(\beta) + qW_\alpha(\alpha).$$

The left hand side is $\beta - p\alpha$, and the right hand side is $\beta - n_{\beta\alpha}\alpha - q\alpha$, so

$$\beta - p\alpha = \beta - n_{\beta\alpha}\alpha - q\alpha.$$

So $p - q = n_{\beta\alpha}$, so $|p - q| \leqslant 3$. So there are at most 4 elements in this string.

Relabelling, we may assume $p = 0$, so q is an integer no more than 3.

(2) To see this, we inspect Table 7. Either $n_{\alpha\beta}$ or $n_{\beta\alpha}$ is ± 1, without loss, say $n_{\beta\alpha} = \pm 1$. Then $W_\beta(\alpha) = \beta - n_{\beta\alpha}\alpha$. Then by the previous fact, Fact 6.4(1), we also get that all weights in the interior of an α-string through β are roots.

(3) Same thing we did before.

(4) If either $\alpha - \beta$ or $\beta - \alpha$ is a root, then both are. So one of them is a positive root. If say $\alpha - \beta$ was a positive root, then $\alpha = \beta + (\alpha - \beta)$ is not a simple root.

(5) If not, then either $\alpha - \beta$ or $\beta - \alpha$ is a root by inspection of Table 7. This is in contradiction to Fact 6.4(4).

(6) Assume that $\sum_i n_i \alpha_i = 0$, and renumber so that the first k-many n_i are positive. Then let

$$v = \sum_{i=1}^{k} n_i \alpha_i = -\sum_{j=k+1}^{m} n_j \alpha_j.$$

Now consider the inner product of v with itself.

$$0 \leqslant (v, v) = -\sum_{i=1}^{k} \sum_{j=k+1}^{n} n_i n_j (\alpha_i, \alpha_j)$$

Note that $n_i \geqslant 0$, $n_j \leqslant 0$, and $(\alpha_i, \alpha_j) \leqslant 0$ by Lemma 6.5(5), so the right hand side is $\leqslant 0$;

$$0 \leqslant (v, v) = -\sum n_i n_j (\alpha_i, \alpha_j) \leqslant 0.$$

So it must be $v = 0$. But

$$0 = \ell(0) = \ell(v) = \ell\left(\sum_{i=1}^{k} n_i \alpha_i\right) \geqslant 0$$

So the n_i are all zero for $1 \leqslant i \leqslant k$, and similarly for $k + 1 \leqslant j \leqslant m$.

(7) Assume not. Then there is $\beta \in R^+$ with $\ell(\beta)$ minimal such that $\beta \notin \mathbb{Z}S$. But since β is not simple, $\beta = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in R^+$, and $\ell(\beta_1), \ell(\beta_2) < \ell(\beta)$. But by minimality of β, β_1 and β_2 are expressible as sums of simple roots so also is β. □

Now recall that the Weyl group $W = \langle W_\alpha \mid \alpha \in R \rangle$ injects into $S_{|R|}$ so in particular, W is finite.

Lemma 6.5. Let $W_0 = \langle W_\alpha \mid \alpha \in S \rangle$, where S is the set of simple roots of a root system R. Then every positive root is sent by elements of W_0 to a simple root, and furthermore $W = W_0$.

Proof. Let $\alpha \in R^+$. To prove that α is sent by elements of W_0 to a member of S, define the **height** of α by $\mathrm{ht}(\alpha) = \sum_i n_i$ such that $\alpha = \sum_i n_i \alpha_i$ for $\alpha_i \in S$.

First claim that there is $\gamma \in S$ such that $(\gamma, \alpha) > 0$. If not,

$$(\alpha, \alpha) = \sum_i n_i (\alpha, \alpha_i) < 0$$

because $n_i > 0$, $(\alpha, \alpha_i) < 0$ for all i. This is a contradiction, because $(\alpha, \alpha) \geqslant 0$.

So $\mathrm{ht}(W_\gamma(\alpha)) = \mathrm{ht}(\alpha - n_{\alpha\gamma}\gamma) < \mathrm{ht}(\alpha)$, and so we're done by induction.

Finally, to show that $W_0 = W$, it's an exercise to check that for $g \in W$, we have $gW_\alpha g^{-1} = W_{g\alpha}$. It suffices to show that W_α for $\alpha \in R^+$ is in W_0, since W is generated by such W_α. Let $\sigma \in W_0$ be the element sending the simple root α_i to α, say $\sigma\alpha_i = \alpha$. Then

$$W_\alpha = W_{\sigma\alpha_i} = \sigma W_{\alpha_i} \sigma^{-1} \in W_0. \qquad \square$$

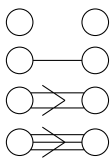
Remark 6.6. Recall:

- If α, β are simple roots then $(\alpha, \beta) \leqslant 0$. In fact, the angle between them is $\pi/2, 2\pi/3, 3\pi/4$ or $5\pi/6$.
- The set of simple roots is linearly independent and every positive root is a nonnegative integral combination of simple roots.
- Every root is conjugate to a simple root under W.
- $W = W_0 = \langle W_\alpha \mid \alpha \in S \rangle$.

To classify root systems (and thereby semisimple Lie algebras), we will classify the **Dynkin diagrams.**

Definition 6.7. A **Dynkin diagram** consists of a collection of nodes, one for each simple root, and some lines between them indicating the angle between them. Furthermore, we put an arrow to indicate which of the two roots is longer.

If there are just two nodes,



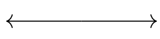
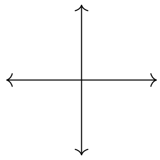
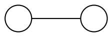
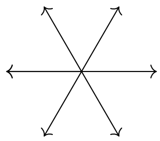
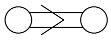
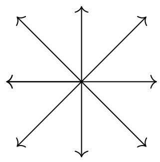
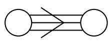
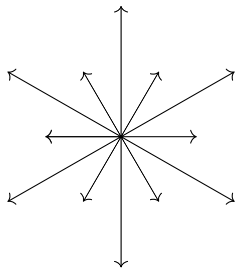
$\theta = \pi/2$

$\theta = 2\pi/3$

$\theta = 3\pi/4$

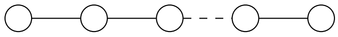
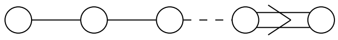
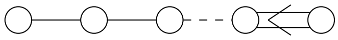
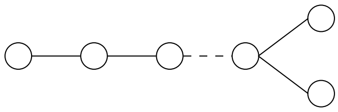
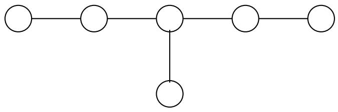
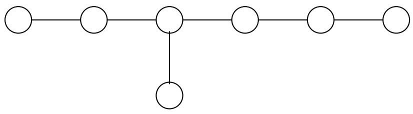
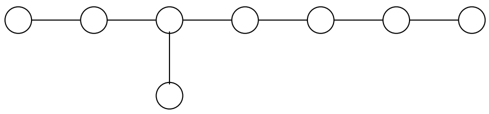
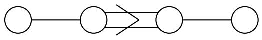
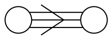
$\theta = 5\pi/6$

In the simplest cases,

Diagram	Lie Algebra	Root System
	$\mathfrak{sl}(2)$	
	$\mathfrak{sl}(2) \times \mathfrak{sl}(2)$	
	$\mathfrak{sl}(3)$	
	$\mathfrak{sp}(4)$	
	G_2	

Definition 6.8. A root system is **irreducible** if it's Dynkin diagram is connected.

Theorem 6.9 (Classification Theorem)**.** The Dynkin diagrams of irreducible root systems are as follows:

Type		Diagram
A_n	$(n \geqslant 1)$	
B_n	$(n \geqslant 2)$	
C_n	$(n \geqslant 3)$	
D_n	$(n \geqslant 4)$	
E_6		
E_7		
E_8		
F_4		
G_2		

The families $A_\ell, B_\ell, C_\ell, D_\ell$ are the Lie algebras of **classical type**. The others, E_6, E_7, E_8, F_4 and G_2 are **exceptional type**.

Type	Lie Algebra	Simple Roots
A_n	$\mathfrak{sl}(n+1)$	$L_1 - L_2, \dots, L_n - L_{n+1}$
B_n	$\mathfrak{so}(2n+1)$	$L_1 - L_2, L_2 - L_3, \dots, L_{n-1} - L_n, L_n$
C_n	$\mathfrak{sp}(2n)$	$L_1 - L_2, L_2 - L_3, \dots, L_{n-1} - L_n, 2L_n$
D_n	$\mathfrak{so}(2n)$	$L_1 - L_2, L_2 - L_3, \dots, L_{n-1} - L_n, L_{n-1} + L_n$

To prove the Classification Theorem, we will consider Coxeter diagrams.

Definition 6.10. Define a **Coxeter diagram** to be a Dynkin diagram without the arrows (so in effect, we assume all root lengths are 1).

Proof of Theorem 6.9. Let e_i, $i = 1, \dots, n$ denote the simple root vectors.Coming from the Coxeter diagram, we know that $(e_i, e_i) = 1$ and if $i \neq j$, $(e_i, e_j) = 0, -\frac{1}{2}, -\frac{\sqrt{2}}{2}$, or $-\sqrt{3}/2$, if the number of edges between them are $0, 1, 2, 3$, respectively. Hence, $4(e_i, e_j)^2$ is the number of edges between e_i and e_j.

Now we classify the admissible diagrams, that is, the possible Coxeter diagrams coming form valid Dynkin diagrams. This is done in the following steps:

(1) Clearly, any (connected) subdiagram of an admissible diagram is admissible. So we consider only connected diagrams. If the diagrams are not connected, then the connected components are themselves simple Lie algebras.

(2) There are at most $(n-1)$ pairs of connected vertices. In particular, there are no loops.

Proof. If e_i and e_j are connected, then $2(e_i, e_j) \leqslant -1$. Hence

$$0 < \left(\sum e_i, \sum e_i\right) = \sum_i (e_i, e_i) + 2\sum_{i<j}(e_i, e_j) = n - \big(\text{\# of edges}\big) \qquad \square$$

(3) No node has more than three edges coming into it.

Proof. Label the central node e_1, and suppose $e_2, \ldots, e_n$ are connected to it. By (2), there are no loops, so none of $e_2, \ldots, e_n$ are connected to any other. Hence $\{e_i \mid i = 1, \ldots, n\}$ is orthonormal. By Gram-Schmidt, extend to an orthonormal basis by adding some e_{n+1} with

$$\operatorname{span}\{e_2, \ldots, e_{n+1}\} = \operatorname{span}\{e_1, \ldots, e_n\}.$$

We must have that $(e_{n+1}, e_1) \neq 0$. So let

$$e_1 = \sum_{i=2}^{n+1} (e_1, e_i) e_i.$$

Then

$$1 = (e_1, e_1) = \sum_{i=2}^{n+1} (e_1, e_i)^2.$$

So if there is an edge e_1 to e_i, we have $4(e_1, e_i)^2 \geqslant 1$.

$$\big(\text{\# of edges out of } e_1\big) \leqslant 4\sum_{i=2}^{n} (e_1, e_i)^2 < 4$$

and the result follows from the admissible values for (e_i, e_j). $\square$

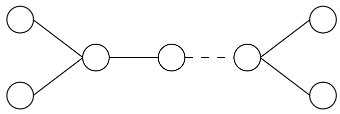
(4) (Shrinking Lemma) In any admissible diagram, we can shrink any string of the form $\circ - \circ - \cdots - \circ$ down to one node to get another admissible diagram.

Proof. Let $e_1, \ldots, e_r$ be vectors along the string and replace with $e = e_1 + \ldots + e_r$. Then

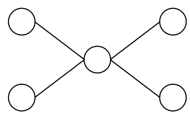
$$(e,e) = \left(\sum_i (e_i, e_i)\right) + 2\left((e_1,e_2) + (e_2,e_3) + \ldots + (e_{r-1}, e_r)\right) = r - (r-1) = 1.$$

And for each other e_k in the diagram but not in the string that we are shrinking, (e, e_k) satisfies the desired conditions, since (e, e_k) is either (e_1, e_k) in the case that e_k is a neighbor of e_1 or (e_r, e_k) in the case that e_k is a neighbor of e_r. □

(5) Immediately, from (3) and (4), we now see that G_2 is the only connected Dynkin diagram with a triple bond. Moreover, we cannot have

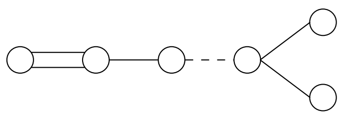


since this would imply



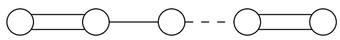
is a valid diagram, which is disallowed by (3).

And we also can't have



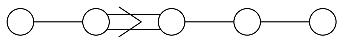
either.

We can also exclude



by (3).

(6) So there are a few other things that we have to rule out to complete the classification, namely



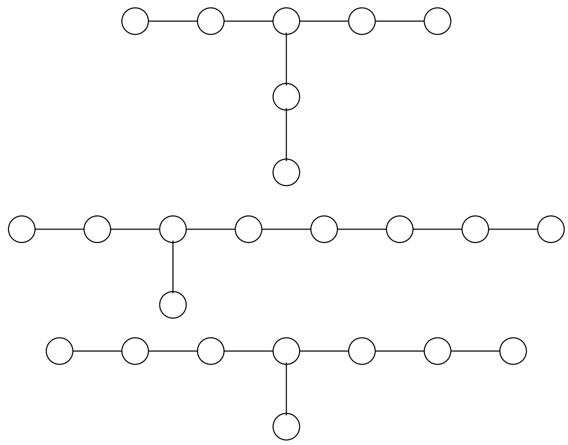
(8)

Proof. To rule out (8), let $v = e_1 + 2e_2$ and $w = 3e_3 + 2e_4 + e_5$. Then we calculate that

$$(v,w)^2 = \|v\|^2 \|w\|^2.$$

So if θ is the angle between v and w, then $(\cos\theta)^2 = 0$, so v and w are linearly dependent. This is a contradiction because the e_i are supposed to be linearly independent. □

(7) Similar considerations rule out the following:



□