Lie groups, Lie algebras, and their representations

Gwyn Bellamy
gwyn.bellamy@glasgow.ac.uk
www.maths.gla.ac.uk/∼gbellamy

September 16, 2016

Abstract

These are the lecture notes for the 5M reading course "Lie groups, Lie algebras, and their representations" at the University of Glasgow, autumn 2015.

Contents

1 Introduction

Lie groups and Lie algebras, together called *Lie theory*, originated in the study of natural symmetries of solutions of differential equations. However, unlike say the finite collection of symmetries of the hexagon, these symmetries occurred in continuous families, just as the rotational symmetries of the plane form a continuous family isomorphic to the unit circle.

The theory as we know it today began with the ground breaking work of the Norwegian mathematician Sophus Lie, who introduced the notion of continuous transformation groups and showed the crucial role that Lie algebras play in their classification and representation theory. Lie's ideas played a central role in Felix Klein's grand "Erlangen program" to classify all possible geometries using group theory. Today Lie theory plays an important role in almost every branch of pure and applied mathematics, is used to describe much of modern physics, in particular classical and quantum mechanics, and is an active area of research.

You might be familiar with the idea that abstract group theory really began with Galois' work on algebraic solutions of polynomial equations; in particular, the generic quintic. But, in a sense, the idea of groups of transformations had been around for a long time already. As mentioned above, it was already present in the study of solutions of differential equations coming from physics (for instance in Newton's work). The key point is that these spaces of solutions are often stable under the action of a large group of symmetries, or transformations. This means that applying a given transformation to one solution of the differential equation gives us another solution. Hence one can quickly and easily generate new solutions from old one, "buy one, get one free".

As a toy example, consider a "black hole" centred at $(0,0)$ in $\mathbb{R}^2$. Then, any particle on $\mathbb{R}^2 \smallsetminus \{(0,0)\}$, whose position at time t is given by $(x(t), y(t))$, will be attracted to the origin, the strength of the attraction increasing as it nears $(0,0)$. This attraction can be encoded by the system of differential equations

$$\frac{dx}{dt} = \frac{-x}{(x^2+y^2)^{3/2}}, \quad \frac{dy}{dt} = \frac{-y}{(x^2+y^2)^{3/2}}. \tag{1}$$

In polar coordinates $(r(t), \theta(t))$, this becomes

$$\frac{dr}{dt} = \frac{-1}{r^2}. \tag{2}$$

The circle $S^1 = \{0 \leq \vartheta < 2\pi\}$ (this is a group, $\vartheta_1 \star \vartheta_2 = (\vartheta_1 + \vartheta_2) \bmod 2\pi$) acts on $\mathbb{R}^2$ by rotations:

$$\vartheta \cdot (x, y) = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\vartheta - y\sin\vartheta \\ x\sin\vartheta + y\cos\vartheta \end{pmatrix}.$$

Then it is clear from (2) that if $(x(t), y(t))$ is a solution to (1), then so too is

$$\vartheta \cdot (x(t), y(t)) = (x(t)\cos\vartheta - y(t)\sin\vartheta, x(t)\sin\vartheta + y(t)\cos\vartheta).$$

Thus, the group S^1 acts on the space of solutions to this system of differential equations.

The content of these lecture notes is based to a large extent on the material in the books [5] and [8]. Other sources that treat the material in these notes are [1], [2], [4], [9] and [7].

2 Manifolds - a refresher

In this course we will consider smooth real manifolds and complex analytic manifolds.

2.1 Topological terminology

Every manifold is in particular a topological space, therefore we will occasionally need some basic topological terms. We begin by recalling that a topological space is a pair $(X, \mathcal{T})$ where X is a set and $\mathcal{T}$ is a collection of subsets of X, the *open subsets* of X. The collection of open subsets $\mathcal{T}$ must include X and the emptyset $\emptyset$, be closed under arbitrary unions and finite intersections. Recall that a topological space M is *disconnected* if there exist two non-empty open and closed subsets $U, V \subset M$ such that $U \cap V = \emptyset$; otherwise M is *connected*. It is *path connected* if, for any two points $x, y \in X$, there is a path from x to y (every path connected space is connected, but the converse is not true). We say that X is *simply connected* if it is path connected and the fundamental group $\pi_1(X)$ of X is trivial i.e. every closed loop in X is homotopic to the trivial loop. The space X is *compact* if every open covering admits a finite subcover. The space X is said to be *Hausdorff* if for each pair of distinct points $x, y \in X$ there exists open sets $x \in U$, $y \in V$ with $U \cap V = \emptyset$.

For each $\ell \in \mathbb{R}_+$ and $x \in \mathbb{R}^n$, we denote by $B_\ell(x)$, resp. $B_\ell(x)^\circ$, the set of all points $y \in \mathbb{R}^n$ such that $||x - y|| \leq \ell$, resp. $||x - y|| < \ell$. We will always consider $\mathbb{R}^n$ as a topological space, equipped with the Euclidean topology i.e. a base for this topology is given by the open balls $B_\ell(x)^\circ$. Recall that the Heine-Borel Theorem states:

Theorem 2.1. *Let M be a subset of $\mathbb{R}^n$. Then, M is compact if and only if it is closed and contained inside a closed ball $B_\ell(0)$ for some $\ell \gg 0$.*

Remark 2.2. If M is some compact subspace of $\mathbb{R}^n$ and $\{X_i\}_{i \in \mathbb{N}}$ a Cauchy sequence in M, then the limit of this sequence exists in M. This fact is a useful way of showing that certain subspaces of $\mathbb{R}^n$ are not compact.

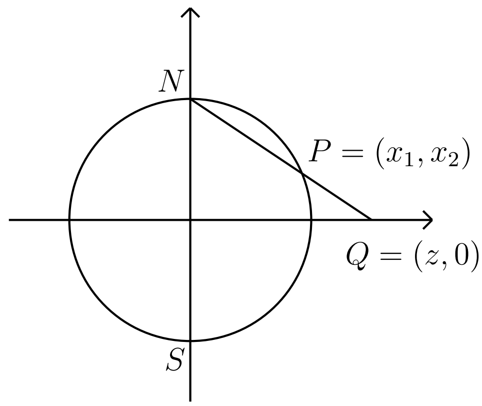
2.2 Manifolds

Recall that a real n-dimensional manifold is a Hausdorff topological space M with an atlas $\mathcal{A}(M) = \{\phi_i : U_i \to V_i \mid i \in I\}$, where $\{U_i \mid i \in I\}$ is an open cover of M, each chart $\phi_i : U_i \to V_i$ is a homeomorphism onto V_i, an open subset of $\mathbb{R}^n$, and the composite maps

$$\phi_i \circ \phi_j^{-1} : V_{i,j} \to V_{j,i}$$

are smooth morphisms, where $V_{i,j} = \phi_j(U_i \cap U_j) \subset \mathbb{R}^n$ and $V_{j,i} = \phi_i(U_i \cap U_j) \subset \mathbb{R}^n$.

Example 2.3. The circle $S^1 = \{(x_1, x_2) \mid x_1^2 + x_2^2 = 1\} \subset \mathbb{R}^2$ is a manifold. We can use stereographic projection to construct charts on S^1. Let $N = (0, 1)$ and $S = (0, -1)$ be the north and south poles respectively.



A line through N and P meets the x_1-axis in a unique point. This defines a map $S^1 \setminus \{N\} \to \mathbb{R}$. To get an explicit formula for this, the line through N and P is described by $\{t(x_1, x_2) + (1 - t)(0, 1) \mid t \in \mathbb{R}\}$. This implies that $z = \frac{x_1}{1-x_2}$. Therefore we define $U_1 = S^1 \setminus \{N\}$, $V_1 = \mathbb{R}$ and

$$\phi_1 : U_1 \to V_1, \quad \phi_1(x_1, x_2) = \frac{x_1}{1 - x_2}.$$

Similarly, if we perform stereographic projection from the south pole, then we get a chart

$$\phi_2 : U_2 \to V_2, \quad \phi_2(x_1, x_2) = \frac{x_1}{x_2 + 1},$$

where $U_2 = S^1 \setminus \{S\}$ and $V_2 = \mathbb{R}$.

For this to define a manifold, we need to check that the transition map $\phi_2 \circ \phi_1^{-1}$ is smooth. Since $U_1 \cap U_1 = S^1 \setminus \{N, S\}$, $\phi_1(U_1 \cap U_2) = \phi_2(U_1 \cap U_2) = \mathbb{R} \setminus \{0\}$. Hence $\phi_2 \circ \phi_1^{-1} : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$. A direct calculation shows that $\phi_2 \circ \phi_1^{-1}(z) = \frac{1}{z}$. This is clearly smooth.

A continuous function $f : M \to \mathbb{R}$ is said to be *smooth* if $f \circ \phi_i^{-1} : V_i \to \mathbb{R}$ is a smooth, i.e. infinitely differentiable, function for all charts ϕ_i. The space of all smooth functions on M is denoted $\mathcal{C}^\infty(M)$.

Example 2.4. The real projective space $\mathbb{RP}^n$ of all lines through 0 in $\mathbb{R}^{n+1}$ is a manifold. The set of points in $\mathbb{RP}^n$ are written $[x_0 : x_1 : \cdots : x_n]$, where $(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} \setminus \{0\}$ and $[x_0 : x_1 : \cdots : x_n]$ represents the line through 0 and $(x_0, x_1, \ldots, x_n)$. Thus, for each $\lambda \in \mathbb{R}^\times$,

$$[x_0 : x_1 : \cdots : x_n] = [\lambda x_0 : \lambda x_1 : \cdots : \lambda x_n].$$

The $n+1$ open sets $U_i = \{[x_0 : x_1 : \cdots : x_n] \in \mathbb{RP}^n \mid x_i \neq 0\}$ cover $\mathbb{RP}^n$. The maps $\phi_i : U_i \xrightarrow{\sim} \mathbb{R}^n$, $[x_0 : x_1 : \cdots : x_n] \mapsto (\frac{x_0}{x_i}, \dots, \widehat{x_i}, \dots, \frac{x_n}{x_i})$ define an atlas on $\mathbb{RP}^n$. Thus, it is a manifold.

Exercise 2.5. Check that the maps ϕ_i are well-defined i.e. only depends on the line through $(x_0, x_1, \dots, x_n)$. Prove that $\mathbb{RP}^n$ is a manifold by explicitly describing the maps $\phi_j \circ \phi_i^{-1} : \phi_i(U_i \cap U_j) \subset \mathbb{R}^n \to \phi_j(U_i \cap U_j) \subset \mathbb{R}^n$ and checking that they are indeed smooth map.

The following theorem is extremely useful for producing explicit examples of manifolds.

Theorem 2.6. *Let $m < n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$, $f = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$, a smooth map. Assume that u is in the image of f. Then $f^{-1}(u)$ is a smooth submanifold of $\mathbb{R}^n$, of dimension $n - m$ if and only if the differential*

$$d_v f = \left(\frac{\partial f_i}{\partial x_j} \right)_{j=1,\dots,n|_{x=v}}^{i=1,\dots,m} : \mathbb{R}^n \to \mathbb{R}^m$$

is a surjective linear map for all $v \in f^{-1}(u)$.

The proof of the above theorem is based on the inverse function theorem, which implies that for each v in the closed set $f^{-1}(u)$, there is some $\ell > 0$ such that $B_\ell(v) \cap f^{-1}(u) \simeq B_{\ell'}(v)$ for some $(n-m)$-dimensional ball $B_{\ell'}(v)$. This implies that $f^{-1}(u)$ is a submanifold of $\mathbb{R}^n$ of dimension $(n-m)$. A proof of this theorem can be found in [10, Corollary 1.29].

Exercise 2.7. By considering the derivatives of $f = x_1^2 + \cdots + x_n^2 - 1$, show that the $(n-1)$-sphere $S^{n-1} \subset \mathbb{R}^n$ is a manifold.

Exercise 2.8. Consider the smooth functions $f_1 = x_1^2 + x_2^3 - x_1 x_2 x_3 - 1$, $f_2 = x_1^2 x_2 x_3$ and $f_3 = (2x_2 - x_3^3)e^{x_1}$ from $\mathbb{R}^3$ to $\mathbb{R}$. For which of these functions is the differential $d_v f_i$ always surjective for all $v \in f_i^{-1}(0)$? For those that are not, the closed subset $f_i^{-1}(0)$ is *not* a submanifold of $\mathbb{R}^3$.

Remark 2.9. Since every point in a manifold admits an open neighborhood homeomorphic to an open subset of $\mathbb{R}^n$, for some n, they are "especially well-behaved" topological spaces. We mention, in particular, that a connected manifold is path connected and always admits a universal cover.

2.3 Tangent spaces

The tangent space at a point of manifold is extremely important, playing a key role when trying to compare manifolds, study functions on the manifold etc. Intuitively, I hope it is clear what the tangent space at a point should be - it's simply all vectors tangent to the manifold at that point. Unfortunately, to make this mathematically precise, some of the intuition is lost in the technicalities. Remarkably, there are several equivalent definitions of the tangent space $T_m M$ to M at the point $m \in M$. In this course we will see three of these definitions, and show that they are equivalent. The first definition, which we will take as "the definition" is in terms of point derivations.

Definition 2.10. Fix a point $m \in M$. A *point derivation* at m is a map $\nu : \mathcal{C}^\infty(M) \to \mathbb{R}$ such that the following two properties are satisfied

1. ν is linear i.e. $\nu(\alpha f + \beta g) = \alpha\nu(f) + \beta\nu(g)$ for $\alpha, \beta \in \mathbb{R}$ and $f, g \in \mathcal{C}^\infty(M)$.

2. It satisfies the derivation rule:

$$\nu(fg) = \nu(f)g(m) + f(m)\nu(g), \quad \forall\ f, g \in \mathcal{C}^\infty(M).$$

One can easily check that if ν, μ are point derivations at m and $\alpha \in \mathbb{R}$, then $\nu + \mu$ and $\alpha\nu$ are point derivations at m. Thus, the set of point derivations forms a vector space. The tangent space to M at m is defined to be the vector space $T_m M$ of all point derivations at m. To get some intuition for this notion, lets first consider the case where the manifold M is some just $\mathbb{R}^n$. For $a = (a_1, \dots, a_n) \in \mathbb{R}^n$, we define the point derivation v_a at u by

$$v_a(f) = a_1 \frac{\partial f}{\partial x_1}\Big|_{x=u} + \cdots + a_n \frac{\partial f}{\partial x_n}\Big|_{x=u}.$$

I claim that $a \mapsto v_a$ defines an isomorphism $\mathbb{R}^n \xrightarrow{\sim} T_u\mathbb{R}^n$. One can easily check that v_a is a point derivation. Also, $v_a(x_i) = a_i$, which implies that the map is injective. Therefore the only thing to check is that every point derivation at u can be written as v_a for some a. Let ν be an arbitrary point derivation and set $a_i = \nu(x_i) \in \mathbb{R}$. Then if $a = (a_1, \dots, a_n)$, we need to show that $\nu - v_a = 0$. It is certainly zero on each of the coordinate functions x_i. Locally, every smooth function has a Taylor expansion

$$f(x) = \sum_{k_1,\dots,k_n \geq 0} \frac{\partial^k f}{\partial x_1^{k_1} \partial x_2^{k_1} \cdots \partial x_n^{k_n}}\Big|_{x=u} (x_1 - u_1)^{k_1} \cdots (x_n - u_n)^{k_n}.$$

Using the derivation rule for point derivations it is easy to see that $\nu(f) - v_a(f) = 0$ too. Thus, we can think of $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$ as being a basis of $T_u\mathbb{R}^n$ for any $u \in \mathbb{R}^n$. In particular, it is clear that $\dim T_u\mathbb{R}^n = n$ for all $u \in \mathbb{R}^n$.

Remark 2.11. The tangent space $T_m M$ at m is a *locally property* of M i.e. it only sees what happens locally on M near m. This statement can be made precise as follows. Let $U \subset M$ be an open neighborhood of m. Then, the fact that any smooth function $f \in \mathcal{C}^\infty(M)$ restricts to a smooth function on U i.e. $f|_U \in \mathcal{C}^\infty(U)$ means that we can define a canonical map $T_m U \to T_m M$, $\nu \mapsto \widetilde{\nu}$ by

$$\widetilde{\nu}(f) := \nu(f|_U).$$

This map is an isomorphism. In order to prove this, one needs to use the existence of partitions of unity, which in this case imply that every function $g \in \mathcal{C}^\infty(U)$ can be extended to a smooth function on M i.e. for each $g \in \mathcal{C}^\infty(U)$ there exists $f \in \mathcal{C}^\infty(M)$ such that $f|_U = g$.

Since a manifold locally looks like $\mathbb{R}^n$, and the tangent space at m only sees what happens around m, it is not surprising that:

Proposition 2.12. *If M is an n-dimensional manifold, then* $\dim T_m M = n$ *for all* $m \in M$.

Proof. To prove the proposition, we will show that a chart $\varphi : U \to \mathbb{R}^n$ around m defines an isomorphism of vector spaces $\varphi_* : T_m M \xrightarrow{\sim} T_{\varphi(m)}\mathbb{R}^n$. The definition of φ_* is very simple. Given $\nu \in T_m M$ a point derivation and $f \in \mathcal{C}^\infty(\mathbb{R}^n)$ a function, we define

$$\varphi_*(\nu)(f) := \nu(f \circ \varphi).$$

We leave it to the reader to check that $\varphi_*(\nu) \in T_{\varphi(m)}\mathbb{R}^n$ is a point derivation. To show that it is an isomorphism, it suffices to note that we also have a map $\varphi^{-1} : \mathrm{Im}\ \varphi \to U$ and hence we can define a map $(\varphi^{-1})_* : T_{\varphi(m)}\mathbb{R}^n \to T_m$. Unpacking the definitions, one sees that $(\varphi^{-1})_* \circ \varphi_* = \mathrm{id}$ and $\varphi_* \circ (\varphi^{-1})_* = \mathrm{Id}$, as required. For instance,

$$\begin{aligned} (\varphi_* \circ (\varphi^{-1})_*)(\nu)(f) &= \varphi_*((\varphi^{-1})_*(\nu))(f) \\ &= (\varphi^{-1})_*(\nu)(f \circ \varphi) \\ &= \nu(f \circ \varphi \circ \varphi^{-1}) = \nu(f). \end{aligned}$$

□

The second definition of tangent space uses the notion of embedded curves. A *curve* on M is a smooth morphism $\gamma : (-\epsilon, \epsilon) \to M$, where $\epsilon \in \mathbb{R}_{>0} \cup \{\infty\}$. We say that γ is a curve *through* $m \in M$ if $\gamma(0) = m$. Given a curve γ through m, we can construct a point derivation $\overline{\gamma}$ at m by the simple rule

$$\overline{\gamma}(f) = \frac{d}{dt}(f \circ \gamma)\Big|_{t=0}.$$

The key point here is that $f \circ \gamma$ is just a function $(-\epsilon, \epsilon) \to \mathbb{R}$, which we can easily differentiate. Let's consider the case where $M = \mathbb{R}^n$. If $\rho : (-\epsilon, \epsilon) \to \mathbb{R}^n$ is a curve in $\mathbb{R}^n$ then we can differentiate

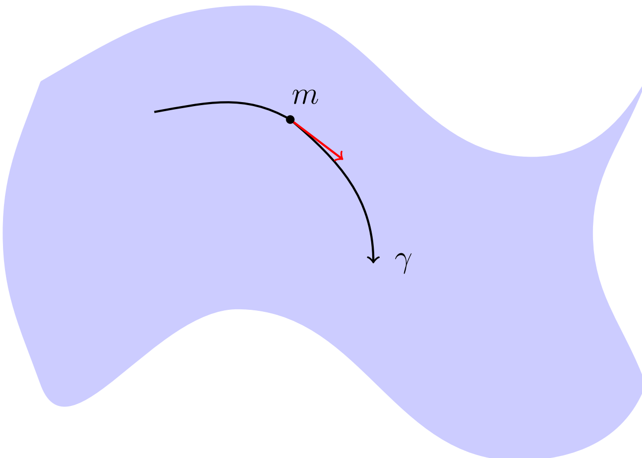


Figure 1: A curve through m and its point derivation.

it and get a vector $\rho'(0) \in T_{\rho(0)}\mathbb{R}^n = \mathbb{R}^n$. Concretely, if $\rho(t) = (\rho_1(t), \ldots, \rho_n(t))$, then $\rho'(0)$ is the point derivation

$$\rho'(0) = \sum_{i=1}^{n} \left(\frac{d\rho_i}{dt}|_{t=0} \right) \frac{\partial}{\partial x_i} \tag{3}$$

at $\rho(0)$. For instance, consider $\rho : \mathbb{R} \to \mathbb{R}^3$, $\rho(t) = (t^2, 3t, 2\sin t)$, then $\rho'(0) = 3\frac{\partial}{\partial x_2} + 2\frac{\partial}{\partial x_1}$.

Now the question becomes: when do two curves through m define the same point derivation? Well we see from the definition that if $\gamma_1 \sim \gamma_2$ if and only if

$$\frac{d}{dt}(f \circ \gamma_1)\Big|_{t=0} = \frac{d}{dt}(f \circ \gamma_2)\Big|_{t=0}, \quad \forall\ f \in \mathcal{C}^\infty(M),$$

then $\gamma_1 \sim \gamma_2 \ \Leftrightarrow\ \overline{\gamma}_1 = \overline{\gamma}_2 \in T_m M$. Denote by $[\gamma]$ the class of curves through m that are equivalent to γ. The claim is that, as a set at least, the tangent space $T_m M$ at m can be identified with the equivalence classes of curves through m, under the above equivalence relation. By construction, there is an injective map from the set of equivalence classes to $T_m M$. So we just need to show that it is surjective i.e.

Lemma 2.13. *For any $\nu \in T_m M$, there exists a curve γ through m such that $\overline{\gamma} = \nu$.*

Proof. Recall from the proof of Proposition 2.12 that, give a chart $\varphi : U \to \mathbb{R}^n$ around m, we constructed an isomorphism $\varphi_* : T_m M \xrightarrow{\sim} T_{\varphi(m)}\mathbb{R}^n$, where $(\varphi_*\nu)(f) = \nu(f \circ \varphi)$. Let's assume that we can find a curve $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}^n$ through $\varphi(m)$ such that $\overline{\gamma} = \varphi_*\nu$. Then let $\mu = \varphi^{-1} \circ \gamma$. We have

$$\begin{aligned}
\overline{\mu}(f) &= \frac{d}{dt}(f \circ \mu)\Big|_{t=0} \\
&= \frac{d}{dt}(f \circ \varphi^{-1}) \circ \gamma\Big|_{t=0} \\
&= \overline{\gamma}(f \circ \varphi^{-1}) \\
&= (\varphi_* \nu)(f \circ \varphi^{-1}) = \nu(f \circ \varphi^{-1} \circ \varphi) = \nu(f).
\end{aligned}$$

Thus, it suffices to assume that $M = \mathbb{R}^n$ and $m = (m_1, \dots, m_n) \in \mathbb{R}^n$. In this case we have seen that

$$\nu = a_i \frac{\partial}{\partial x_1} + \cdots + a_n \frac{\partial}{\partial x_n}$$

for some $a_i \in \mathbb{R}$. Let γ be the curve $\gamma(t) = (a_1 t + m_1, \dots, a_n t + m_n)$. Then $\gamma(0) = m$ and equation (3) shows that $\overline{\gamma} = \nu$. $\square$

Putting all the tangent spaces into a single family, we get

Definition 2.14. The *tangent bundle* of M is the set

$$TM = \{(m, v) \mid m \in M,\ v \in T_m M\}.$$

The tangent bundle TM is itself a manifold and comes equipped with smooth morphisms $i : M \hookrightarrow TM$, $i(m) = (m, 0)$ and $\pi : TM \to M$, $\pi(m, v) = m$.

Exercise 2.15. If $M = f^{-1}(0) \subset \mathbb{R}^n$, where $f : \mathbb{R}^n \to \mathbb{R}^m$, then the tangent space to M at m is the subspace $\mathrm{Ker}(d_m f : \mathbb{R}^n \to \mathbb{R}^m)$ of $T_m \mathbb{R}^n = \mathbb{R}^n$. Describe the tangent space to $S^{n-1} = f^{-1}(0)$ in $\mathbb{R}^n$ at $(1, 0, \dots, 0)$, where $f = x_1^2 + \cdots + x_n^2 - 1$.

The differential of a function $f \in \mathcal{C}^\infty(M)$, at a point m, is a linear map $d_m f : T_m M \to \mathbb{R}$. In terms of the first definition of a tangent space, $d_m f([\gamma]) = (f \circ \gamma)'(0)$. In terms of the second definition of a tangent space, if $X_i \in T_{\phi_i(m)} \mathbb{R}^n$, then $f \circ \phi_i^{-1} : V_i \to \mathbb{R}$. Differentiating this function gives $d_{\phi_i(m)}(f \circ \phi_i^{-1}) : \mathbb{R}^n \to \mathbb{R}$ and we define $(d_m f)([X_i]) = d_{\phi_i(m)}(f \circ \phi_i^{-1})(X_i)$. Of course, one must check that both these definitions are actually well-define i.e. independent of the choice of representative of the equivalence class.

Let $f : M \to N$ be a smooth map between manifolds M and N. Then, for each $m \in M$, f defines a linear map $d_m f : T_m M \to T_{f(m)} M$ between tangents spaces, given by $(d_m f)([\gamma]) = [f \circ \gamma]$. Since we get one such map for all points in m and they vary continuously over M, we actually get a smooth map

$$df : TM \to TN, \quad (df)(m, [\gamma]) = (f(m), d_m f([\gamma])).$$

The following fact, which is a consequence of the inverse function theorem, will be useful to us later. Let $f : M \to N$ be a smooth map between manifolds M and N such that the differential $d_m f$ is an isomorphism at every point $m \in M$. If N is simply connected and M connected then f is an isomorphism.

2.4 Vector fields

A vector field is a continuous family of vectors in the tangent bundle i.e. it is a rule that assigns to each $m \in M$ a vector field $X_m \in T_m M$ such that the family $\{X_m\}_{m \in M}$ varies smoothly on M. The notion of vector field will be crucial later in relating a Lie group to its Lie algebra.

Definition 2.16. A *vector field* on M is a smooth morphism $X : M \to TM$ such that $\pi \circ X = \mathrm{id}_M$. The space of all vector fields on M is denoted $\mathrm{Vect}(M)$.

The key point of defining a vector field is that one can differentiate functions along vector fields. Let X be a vector field on M and $f : M \to \mathbb{R}$ a smooth function. We define $X(f)(m) = (f \circ \gamma)'(0)$ for some (any) choice of curve through m such that $[\gamma] = X_m$.

Lemma 2.17. *The vector field X defines a map $\mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ satisfying the product rule*

$$X(fg) = X(f)g + fX(g), \quad \forall\ f, g \in \mathcal{C}^\infty(M). \tag{4}$$

Proof. Let f, g be smooth maps and $m \in M$. Then,

$$\begin{aligned} X(fg)(m) &= ((fg) \circ \gamma)'(0) = ((f \circ \gamma)(g \circ \gamma))'(0) \\ &= (f \circ \gamma)(0)(g \circ \gamma)'(0) + (f \circ \gamma)'(0)(g \circ \gamma)(0) \\ & f(m)X(g)(m) + X(f)(m)g(m). \end{aligned}$$

Hence $X(fg) = X(f)g + fX(g)$. □

A linear map $\mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ satisfying equation (4) is called a *derivation*. Thus, every vector field defines a derivation. The converse is also true - every derivation defines a unique vector field on M (we won't need this fact though). An equivalent definition of the action of a vector field is that $X(f)$ is the function on M whose value at m equals $(d_m f)(X_m)$. Once should think of vector fields, or derivations, as being continuous families of point derivations.

Exercise 2.18. Let M, N and P be manifolds and $f : M \to N$, $g : N \to P$ smooth maps. Show that the linear map $d_m(g \circ f) : T_m M \to T_{g(f(m))} P$ equals $(d_{f(m)} g) \circ (d_m f)$. Hint: Using the first definition of the tangent space, this is virtually a tautology.

Exercise 2.19. Since $S^2 \subset \mathbb{R}^3$, we can define the function $f : S^2 \to \mathbb{R}$ by saying that it is the restriction of $2x_1 - x_2^2 + x_1 x_3$. Recall the description of the tangent space $T_{(1,0,0)} S^2$ given in exercise 2.15. Describe $d_{(1,0,0)} f : T_{(1,0,0)} S^2 \to \mathbb{R}$.

Similarly, let $g : \mathbb{R}^3 \to \mathbb{R}^2$ be the function $f = (f_1(x_1, x_2, x_3), f_2(x_1, x_2, x_3))$, where $f_1(x_1, x_2, x_3) = x_2^3 x_1 - x_3 \sin x_1$ and $f_2(x_1, x_2, x_3) = e^{x_3} x_2 - \cos x_2$. What is is the linear map $d_{(0,\pi,1)} g : \mathbb{R}^3 \to \mathbb{R}^2$?

2.5 Integral curves

Let X be a vector field on a manifold M and fix $m \in M$. An *integral curve* (with respect to X) $\gamma : J \to M$ through m is a curve such that $\gamma(0) = m$ and $(d_x \gamma)(1) = X_{\gamma(x)}$ for all $x \in J$. Then γ is a solution to the equation

$$\frac{d}{dt}\gamma(t) \mid_{t=x} = X_{\gamma(x)}$$

for all $x \in J$. By choosing a chart containing m, the problem of finding an integral curve through m is easily seen to be equivalent to solving a system of linear, first order differential equations. Therefore the fundamental theorem of ordinary differential equations says that there exists some $\epsilon > 0$ and an integral curve $\gamma : (-\epsilon, \epsilon) \to M$. Moreover, γ is unique. One can try to make the open set $J \subset \mathbb{R}$ as large as possible. There is a unique largest open set on which γ exists; if J is this maximal set then $\gamma : J \to M$ is called the *maximal integral curve* for X through m.

Definition 2.20. A vector field X on M is said to be *complete* if, for all $m \in M$, the maximal integral curve through m with respect to X is defined on the whole of $\mathbb{R}$.

Exercise 2.21. Let X be the vector field $x_1 \frac{\partial}{\partial x_1} - (2x_1 + 1)\frac{\partial}{\partial x_2}$ on $\mathbb{R}^2$. Construct an integral curve γ for X through $(a, b) \in \mathbb{R}^2$. Is X complete?

2.6 Complex analytic manifolds

All the above definitions and results hold for complex analytic manifolds, where M is said to be a complex analytic manifold if the atlas consists of charts into open subsets of $\mathbb{C}^n$ such that the transition maps $\phi_i \circ \phi_j^{-1}$ are biholomorphic.

All maps between complex analytic manifolds are holomorphic e.g. holomorphic vector fields or $\mathcal{C}^{\text{hol}}(M)$, the spaces of holomorphic functions on M.

3 Lie groups and Lie algebras

In this section we introduce the stars of the show, Lie groups and Lie algebras.

3.1 Lie groups

Let (G, m, e) be a group, where $m : G \times G \to G$ is the multiplication map and $e \in G$ the identity element.

Definition 3.1. The group (G, m, e) is said to be a *Lie group* if G is a manifold such that both the multiplication map m, and inversion $g \mapsto g^{-1}$, are smooth maps $G \times G \to G$, and $G \to G$ respectively.

We drop the notation m and simply write gh for $m(g, h)$ if $g, h \in G$.

Example 3.2. Let $S^1 \subset \mathbb{C}$ be the unit circle. Multiplication on $\mathbb{C}$ restricts to $S^1 \times S^1 \to S^1$, making it a Lie group. It is the group of rotations of the real plane.

Example 3.3. The set of invertible $n \times n$ matrices $GL(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$ and hence a manifold. Matrix multiplication and taking inverse are smooth maps. Therefore $GL(n, \mathbb{R})$ is a Lie group.

Example 3.4. Every finite group can be considered as a zero-dimensional Lie group.

Example 3.5. Let $SO(3)$ denote the set of all 3 by 3 real matrices A such that $AA^T = 1$ and $\det(A) = 1$. It is clear that this set is a group under the usual multiplication of matrices.

3.2 Morphisms of Lie groups

A *morphism* $\phi : G \to H$ between Lie groups is a group homomorphism which is also a smooth map between manifolds.

Exercise 3.6. For $g \in G$, let $L_g : G \to G$ be the map $L_g(h) = gh$ of left multiplication.

1. Using the fact that $i_g : G \hookrightarrow G \times G$, $i_g(h) = (g, h)$, is a smooth map, show that $L_g : G \to G$ is a smooth map.

2. Using part (1), show that if $U \subset G$ is an open subset and $g \in G$, then gU is also open in G.

3. Similarly, show that if $C \subset G$ is a closed subset then gC is also closed in G.

Hints: In part 1. use the fact that the composite of two smooth maps is smooth. In parts 2. and 3. recall that a map $f : X \to Y$ between topological spaces is continuous if and only if $f^{-1}(U)$ is open, resp. $f^{-1}(C)$ is closed, in X for all $U \subset Y$ open, resp. $C \subset Y$ closed.

Assume that the Lie group G is connected. Then, as the following proposition shows, one can tell a great deal about the group by considering the various neighborhoods of the identity.

Proposition 3.7. *Let G be a connected Lie group and U an open neighborhood of the identity. Then the elements of U generate G.*

Proof. Recall that G connected implies that the only non-empty closed and open subset of G is G itself. Since the map $g \to g^{-1}$ is smooth, $U^{-1} = \{g^{-1} \mid g \in U\}$ is open in G. Thus, $U \cap U^{-1}$ is also open. It is non-empty because $e \in U \cap U^{-1}$. Replacing U by this intersection we may assume that $g^{-1} \in U$ if and only if $g \in U$. By exercise 3.6, if $g \in U$ then gU is open in U. Hence $U \cdot U = \bigcup_{g \in U} gU$ is open in G. This implies by induction that $H = \{g_1 \cdots g_k \mid g_i \in U\}$ is an open subset of G. But it's easy to check that H is also a subgroup of G. Therefore, to show that $H = G$, it suffices to show that H is closed in G.

Let $C = G \smallsetminus H$. Since H is open in G, C is closed. We assume that $C \neq \emptyset$. Notice that if $g \in C$, then $gH \subset C$ and hence $H \subset g^{-1}C$, a closed subset of G. Thus, H is contained in the intersection $C' = \bigcap_{g \in C} g^{-1}C$. The arbitrary intersection of closed sets is closed, thus C' is closed. Hence it suffices to show that $H = C'$. If $f \in C' \smallsetminus H$ then, in particular, $f \in (f^{-1})^{-1}(G \smallsetminus H)$ i.e. there exists $g \in G \smallsetminus H$ such that $f = (f^{-1})^{-1}g = fg$. But this implies that $g = e$ belongs to $G \smallsetminus H$; a contradiction. ☐

In particular, if $f : G \to H$ is a morphism of Lie groups with G connected then Proposition 3.7 implies that f is uniquely defined by what it does on neighborhoods of the identity. Taking smaller and smaller neighborhoods of e, one eventually "arrives" at the tangent space of G at e and the map $d_e f : T_e G \to T_e H$. Remarkably this linear map captures all the information about the original morphism f,

Theorem 3.8. *Let G and H be Lie groups, with G connected. Then a morphism $f : G \to H$ is uniquely defined by the linear map $d_e f : T_e G \to T_e H$.*

What this really means is that if $f, g : G \to H$ are morphisms of Lie groups then $f = g$ if and only if $d_e f = d_e g$. The proof of Theorem 3.8 is given in section 4; see Corollary 4.20.

Exercise 3.9. Let $G = \mathbb{R}^{\times} (=: \mathbb{R} \smallsetminus \{0\})$, where the multiplication comes from the usual multiplication on $\mathbb{R}$. Show that the map $\phi : G \to G$, $\phi(x) = x^n$ is a homomorphism of Lie groups. What is $T_e G$? Describe $d_e \phi : T_e G \to T_e G$.

Naturally, one can ask, as a converse to Theorem 3.8, which linear maps $T_eG \to T_eH$ extend to a homomorphism of groups $G \to H$? Surprisingly, there is a precise answer to this question. But before it can be given we will need to introduce the notion of Lie algebras and describe the Lie algebra that is associated to each Lie group.

3.3 The adjoint action

In this section we will define the Lie algebra of a Lie group. The idea is that geometric objects are inherently non-linear e.g. the manifold $M \subset \mathbb{R}^3$ defined by the non-linear equation $x^5+y^5-z^7=1$. The same applies to Lie groups. But humans don't seem to cope very well with non-linear objects. Therefore Lie algebras are introduced as linear approximations to Lie groups. The truly remarkable thing that makes Lie theory so successful is that this linear approximation captures a great deal (frankly unjustifiable) of information about the Lie group.

We begin with automorphisms. An *automorphism* ϕ of a Lie group is an invertible homomorphism $G \to G$.

Exercise 3.10. Show that the set $\mathrm{Aut}(G)$ of all automorphisms of G forms a group.

An easy way to cook up a lot of automorphisms of G is to make G act on itself by conjugation. Namely, for each $g \in G$, define $\mathrm{Ad}(g) \in \mathrm{Aut}(G)$ by $\mathrm{Ad}(g)(h) = ghg^{-1}$. This defines a map $G \to \mathrm{Aut}(G)$, $g \mapsto \mathrm{Ad}(g)$, called the *adjoint action*.

Exercise 3.11. Check that Ad is a group homomorphism.

If $\phi \in \mathrm{Aut}(G)$ belongs to the image of Ad, we say that ϕ is an *inner* automorphism of G. We may also consider Ad as a smooth map $G \times G \to G$, $\mathrm{Ad}(g,h) = ghg^{-1}$. The key reason for introducing the adjoint action is that it fixes the identity i.e. $\mathrm{Ad}(g)(e) = e$ for all $g \in G$.

If we return to the setting of Theorem 3.8, then the homomorphism $f : G \to H$ also sends $e \in G$ to $e \in H$. Moreover, the diagram

$$\begin{tikzcd} G \arrow[r, "f"] \arrow[d, "\mathrm{Ad}(g)"'] & H \arrow[d, "\mathrm{Ad}(f(g))"] \\ G \arrow[r, "f"'] & H \end{tikzcd} \tag{5}$$

is commutative; that is,

$$\mathrm{Ad}(f(g))(f(u)) = f(\mathrm{Ad}(g)(u)) \quad \text{for all } u \in G.$$

So we can being our linear approximation process by differentiating diagram (5) at the identity, to get a commutative diagram of linear maps

$$\begin{tikzcd} T_eG \arrow[r, "d_e f"] \arrow[d, "d_e \operatorname{Ad}(g)"'] & T_eH \arrow[d, "d_e \operatorname{Ad}(f(g))"] \\ T_eG \arrow[r, "d_e f"'] & T_eH \end{tikzcd} \tag{6}$$

However, if we want to check that $d_e f$ really is the differential of some homomorphism $f : G \to H$ then (6) doesn't really help because the right vertical arrow is $d_e \operatorname{Ad}(f(g))$ and we would need to know the value $f(g)$ for $g \neq e$ i.e. we still need to see what f is doing away from the identity element.

To overcome this problem, we will use the different interpretation of Ad as a map $G \times G \to G$, $(g, h) \mapsto ghg^{-1}$. Just as in diagram 5, we get a commutative diagram

$$\begin{tikzcd} G \times G \arrow[r, "f \times f"] \arrow[d, "\operatorname{Ad}"'] & H \times H \arrow[d, "\operatorname{Ad}"] \\ G \arrow[r, "f"'] & H. \end{tikzcd} \tag{7}$$

The temptation now is just to differentiate this diagram at (e, e). But this is not quite the right thing to do. Instead, we differentiate each entry of Ad, resp. of $f \times f$, separately to get a *bilinear map*. You may not have see the definition of bilinear before. As a remainder,

Definition 3.12. Let k be a field and U, V and W k-vector spaces. A map $b : U \times V \to W$ is said to be bilinear if both $b(u, -) : V \to W$ and $b(-, v) : U \to W$ are linear maps, $\forall\ u \in U, v \in V$, i.e.

$$b(\alpha u_1 + \beta u_2, \gamma v_1 + \delta v_2) = \alpha\gamma b(u_1, v_1) + \alpha\delta b(u_1, v_2) + \beta\gamma b(u_2, v_1) + \beta\delta b(u_2, v_2),$$

for all $u_1, u_2 \in U, v_1, v_2 \in V,\ \alpha, \beta, \gamma, \delta \in k$.

Remark 3.13. **The bidifferential**: if M, N and K are manifolds and $f : M \times N \to K$ is a smooth map, the the bidifferential $b_{(m,n)} f$ of f at $(m, n) \in M \times N$ is a bilinear map $T_m M \times T_n N \to T_{f(m,n)} K$. Informally, one first fixes $m \in M$ to get a smooth map $f_m : N \to K$. Differentiating at n, we get a linear map $d_n f_m : T_n N \to T_{f(m,n)} K$. Then we fix $w \in T_n N$ and define $f'_w : M \to TK$ by $f'(m) = d_n f_m(w)$. Differentiating again we get $d_m f'_w : T_m M \to T_{f(m,n)} K$ and hence $b_{(m,n)} f : T_m M \times T_n N \to T_{f(m,n)} K$ given by $(v, w) \mapsto (d_m f'_w)(v)$. If one differentiates along M first and then along N then we get the same bilinear map. This follows from the fact that $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$

for a smooth map $f : \mathbb{R}^n \to \mathbb{R}$.

Thus, we get a commutative diagram of bilinear maps

$$\begin{tikzcd}
T_eG \times T_eG \arrow[r, "b_{(e,e)}f\times f"] \arrow[d, "b_{(e,e)}\,\mathrm{Ad}"'] & T_eH \times T_eH \arrow[d, "b_{(e,e)}\,\mathrm{Ad}"] \\
T_eG \arrow[r, "d_ef"'] & T_eH.
\end{tikzcd} \tag{8}$$

Notice that only the biderivative of f at the identity appears in the above diagram. We have no need to know what f does away from the identity. To make the notation less cumbersome we fix $\mathfrak{g} = T_eG$ and $\mathfrak{h} = T_eH$. Then $[-,-]_G := b_{(e,e)}\,\mathrm{Ad} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ is a bilinear map. This is the *Lie bracket* on the vector space $\mathfrak{g}$. Thus, we have shown

Proposition 3.14. *Let $f : G \to H$ be a morphism of Lie groups. Then, the linear map $d_ef : \mathfrak{g} \to \mathfrak{h}$ preserves brackets i.e.*

$$(d_ef)([X,Y]_G) = [(d_ef)(X),(d_ef)(Y)]_H, \quad \forall\ X,Y \in \mathfrak{g}. \tag{9}$$

This leads us to the following key result, which is one of the main motivations in the definition of Lie algebras.

Theorem 3.15. *Let G and H be Lie groups, with G simply connected. Then a linear map $\mathfrak{g} \to \mathfrak{h}$ is the differential of a homomorphism $G \to H$ if and only if it preserves the bracket, as in (9).*

The proof of Theorem 3.15 will be given in section 4; see Theorem 4.18.

Example 3.16. As an example to keep some grasp on reality, we'll consider the case $G = GL(V)$, where V is some n-dimensional real vector space (so $V \simeq \mathbb{R}^n$). Then, for matrices $A, B \in GL(V)$, $\mathrm{Ad}(A)(B) = ABA^{-1}$ is really just naive matrix conjugation. For $G = GL(V)$, the tangent space $T_eGL(V)$ equal $\mathrm{End}(V)$, the space of all linear maps $V \to V$ (after fixing a basis of V, $\mathrm{End}(V)$ is just the space of all n by n matrices over $\mathbb{R}$ so that $\mathrm{End}(V) \simeq \mathbb{R}^{n^2}$). If $A \in GL(V)$ and $Y \in \mathrm{End}(V)$, then differentiating Ad with respect to B in the direction of Y gives

$$d_{(A,1)}\,\mathrm{Ad}(Y) = \lim_{\epsilon\to 0} \frac{A(1+\epsilon Y)A^{-1} - A(1)A^{-1}}{\epsilon} = AYA^{-1}.$$

Thus, $d_{(A,1)}\,\mathrm{Ad}$ is just the usual conjugation action of A on $\mathrm{End}(V)$. Next, for each $Y \in \mathrm{End}(V)$, we want to differential the map $A \mapsto d_{(A,1)}\,\mathrm{Ad}(Y) = AYA^{-1}$ at $1 \in GL(V)$. This will give a linear map $\mathrm{ad}(Y) : \mathrm{End}(V) \to \mathrm{End}(V)$. If $X, Y \in \mathrm{End}(V)$, then $(1 + \epsilon X)^{-1} = 1 - \epsilon X + \epsilon^2 X^2 - \cdots$. So,

$$\begin{aligned} \mathsf{ad}(Y)(X) &= \lim_{\epsilon \to 0} \frac{(1 + \epsilon X) Y (1 - \epsilon X + \epsilon^2 X^2 - \cdots) - Y}{\epsilon} \\ &= \lim_{\epsilon \to 0} \frac{\epsilon XY - \epsilon YX + \epsilon^2 \cdots}{\epsilon} = XY - YX. \end{aligned}$$

Thus, $b_{(1,1)} \mathrm{Ad}(X, Y) = [X, Y] = XY - YX$, the usual commutator of matrices.

As explained in remark 3.13, $[-,-]_G$ is defined by differentiating the map $\mathrm{Ad}'_Y : G \to \mathfrak{g}$ given by $\mathrm{Ad}'_Y(g) = (d_e \mathrm{Ad}(g))(Y)$. Swapping arguments, we may also consider the map $\mathsf{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ again given by $Y \mapsto (d_e \mathrm{Ad}(g))(Y)$.

Exercise 3.17. Show that $\mathsf{Ad}(g) \circ \mathsf{Ad}(h) = \mathsf{Ad}(gh)$ for all $g, h \in G$. Conclude that $\mathsf{Ad}(g)$ is an invertible linear map and hence defines a group homomorphism $\mathsf{Ad} : G \to GL(\mathfrak{g})$.

The map $\mathsf{Ad} : G \to GL(\mathfrak{g})$ is a morphism of Lie groups. Its differential at the identity is denoted $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$. Applying Proposition 3.14 to this situation gives the following lemma, which will be useful later.

Lemma 3.18. *The map* ad *preserves brackets i.e.* $\mathrm{ad}([X, Y]_G) = [\mathrm{ad}(X), \mathrm{ad}(Y)]_E$*, where* $[A, B]_E := AB - BA$ *is the bracket on* $\mathrm{End}(\mathfrak{g})$.

Remark 3.19. One can check that $\mathrm{ad}(Y)(X) = \mathrm{ad}(X)(Y) = [X, Y]_G$ for all $X, Y \in \mathfrak{g}$.

3.4 Lie algebras

We define here the second protagonist in the story - the Lie algebra. We've essential already seen above that for each Lie group G, the space $\mathfrak{g}$ is an example of a Lie algebra.

Definition 3.20. Let k be a field and $\mathfrak{g}$ a k-vector space. Then $\mathfrak{g}$ is said to be a *Lie algebra* if there exists a bilinear map $[-,-] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called the *Lie bracket*, such that

1. The Lie bracket is *anti-symmetric* meaning that

$$[X, Y] = -[Y, X], \quad \forall\ X, Y \in \mathfrak{g}.$$

2. The *Jacobi identity* is satisfied:

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0, \quad \forall\ X, Y, Z \in \mathfrak{g}.$$

Exercise 3.21. Assume that char $k \neq 2$. Show that the first axiom in the definition of a Lie algebra is equivalent to the condition $[X, X] = 0$ for all $X \in \mathfrak{g}$.

Example 3.22. Let M be a manifold. Then the space $\mathrm{Vect}(M)$ of vector fields on M is a Lie algebra via the rule

$$[X, Y] := X \circ Y - Y \circ X,$$

where $X \circ Y$ is the vector field that acts on $\mathcal{C}^\infty(M)$ by first applying Y and then applying X.

Example 3.23. Let V be a k-vector space and $\mathfrak{gl}(V) = \mathrm{End}(V)$ the space of all linear maps $V \to V$. Then, as we have already seen, $\mathfrak{gl}(V)$ is a Lie algebra, the *general linear Lie algebra*, with bracket $[F, G] = F \circ G - G \circ F$. If V is n-dimensional, then we may identify $\mathfrak{gl}(V)$ with $\mathfrak{gl}(n, k)$, the Lie algebra of $n \times n$ matrices, where the bracket of two matrices A and B is just the commutator $[A, B] = AB - BA$. The Lie algebra $\mathfrak{gl}(n, k)$ contains many interesting Lie subalgebras such as $\mathfrak{n}(n, k)$ the Lie algebra of all strictly upper triangular matrices or $\mathfrak{b}(n, k)$ the Lie algebra of upper triangular matrices.

Exercise 3.24. Prove that $\mathfrak{gl}(n, k)$ is a Lie algebra, and that $\mathfrak{b}(n, k)$ and $\mathfrak{n}(n, k)$ are Lie subalgebras of $\mathfrak{gl}(n, k)$.

3.5 The Lie algebra of a Lie group

Recall that we defined in section 3.3 a bracket $[-, -]$ on the tangent space $\mathfrak{g}$ at the identity of a Lie group G. As expected,

Proposition 3.25. *The pair* $(\mathfrak{g}, [-, -])$ *is a Lie algebra.*

Proof. The bracket is bilinear by construction. Therefore we need to check that it is anti-symmetric and satisfies the Jacobi identity. By exercise 3.21, to check that the first axiom holds it suffices to show that $[X, X] = 0$ for all $X \in \mathfrak{g}$.

Recall from the first definition of tangent spaces that an element $X \in \mathfrak{g}$ can be written $\gamma'(0)$ for some $\gamma : (-\epsilon, \epsilon) \to G$. Let $Y = \rho'(0)$ be another element in $\mathfrak{g}$. We can express the bracket of X and Y in terms of γ and ρ. First, for each $t \in (-\epsilon, \epsilon)$ and $g \in G$, $\mathrm{Ad}(\gamma(t))(g) = \gamma(t) g \gamma(t)^{-1}$. Then, taking $g = \rho(s)$ and differentiating ρ to get Y,

$$(d_e \,\mathrm{Ad}(\gamma(t)))(Y) = \left(\frac{d}{ds}\gamma(t)\rho(s)\gamma(t)^{-1}\right)|_{s=0}.$$

Thus,

$$\mathrm{ad}(X)(Y) = \left[\left(\frac{d}{ds}\gamma(t)\rho(s)\gamma(t)^{-1}\right)|_{s=0}\right]|_{t=0}.$$

In particular,

$$\begin{aligned}
[X,X] = \mathrm{ad}(X)(X) &= \left[\frac{d}{dt}\left(\frac{d}{ds}\gamma(t)\gamma(s)\gamma(t)^{-1}\right)|_{s=0}\right]|_{t=0} \\
&= \left[\frac{d}{dt}\left(\frac{d}{ds}\gamma(s)\right)|_{s=0}\right]|_{t=0} \\
&= \left[\frac{d}{dt}X\right]|_{t=0} = 0.
\end{aligned}$$

We have implicitly used the fact that $\gamma(t)\gamma(s) = \gamma(s)\gamma(t)$ for all s, t. This is proved in Lemma 4.6 below.

Using the anti-symmetric property of the bracket, the Jacobi identity is equivalent to the identity $[X,[Y,Z]] - [Y,[X,Z]] = [[X,Y],Z]$ for all $X, Y, Z \in \mathfrak{g}$. Recall that $\mathrm{ad}(X)(Y) = [X,Y]$. Therefore the above identity can be written $[\mathrm{ad}(X)\circ\mathrm{ad}(Y) - \mathrm{ad}(Y)\circ\mathrm{ad}(X)](Z) = \mathrm{ad}([X,Y])(Z)$, which would follow from the identity $\mathrm{ad}(X)\circ\mathrm{ad}(Y) - \mathrm{ad}(Y)\circ\mathrm{ad}(X) = \mathrm{ad}([X,Y])$ in $\mathrm{End}(\mathfrak{g})$. But this is exactly the statement of Lemma 3.18 that ad preserves brackets. $\square$

A Lie algebra $\mathfrak{h}$ is said to be *abelian* if $[X,Y] = 0$ for all $X, Y \in \mathfrak{h}$.

Exercise 3.26. Let H be an abelian Lie group. Show that its Lie algebra $\mathfrak{h}$ is abelian.

3.6 Ideals and quotients

Let $\mathfrak{g}$ be a Lie algebra. A subspace $\mathfrak{h}$ of $\mathfrak{g}$ is called a subalgebra if the bracket on $\mathfrak{g}$ restricts to a bilinear map $[-,-] : \mathfrak{h}\times\mathfrak{h} \to \mathfrak{h}$. This makes $\mathfrak{h}$ into a Lie algebra. A subalgebra $\mathfrak{l}$ is said to be an *ideal* if $[\mathfrak{l},\mathfrak{g}] \subset \mathfrak{l}$. If $\mathfrak{l}$ is an ideal of $\mathfrak{g}$ then the quotient vector space $\mathfrak{g}/\mathfrak{l}$ is itself a Lie algebra, with bracket

$$[X+\mathfrak{l}, Y+\mathfrak{l}] := [X,Y] + \mathfrak{l}.$$

Exercise 3.27. Let $\mathfrak{l} \subset \mathfrak{g}$ be an ideal. Check that the bracket on $\mathfrak{g}/\mathfrak{l}$ is well-defined and that $\mathfrak{g}/\mathfrak{l}$ is indeed a Lie algebra.

Exercise 3.28. Let $\mathfrak{l}$ be a Lie subalgebra of $\mathfrak{g}$ such that $[\mathfrak{g},\mathfrak{g}] \subset \mathfrak{l}$. Show that $\mathfrak{l}$ is an ideal and that the quotient $\mathfrak{g}/\mathfrak{l}$ is abelian.

Exercise 3.29. Show that $\mathfrak{n}(2,k)$ is a one-dimensional ideal in $\mathfrak{b}(2,k)$. More generally, show that $\mathfrak{n}(n,k)$ is an ideal in $\mathfrak{b}(n,k)$. What are the dimensions of $\mathfrak{b}(n,k)$ and $\mathfrak{n}(n,k)$? Is the quotient $\mathfrak{b}(n,k)/\mathfrak{n}(n,k)$ abelian?

3.7 Lie algebras of small dimension

Let $\mathfrak{g}$ be a one-dimensional Lie algebra. Then $\mathfrak{g} = k\{X\}$ for any $0 \neq X \in \mathfrak{g}$. What is the bracket on $\mathfrak{g}$? The fact that the bracket must be anti-symmetric implies that $[X, X] = 0$. Therefore the bracket is zero and $\mathfrak{g}$ is unique up to isomorphism.

When $\dim \mathfrak{g} = 2$, let X_1, X_2 be some basis of $\mathfrak{g}$. If the bracket on $\mathfrak{g}$ is not zero, then the only non-zero bracket can be $[X_1, X_2] = -[X_2, X_1]$ and hence $[\mathfrak{g}, \mathfrak{g}]$ is a one dimensional subspace of $\mathfrak{g}$. Let Y span this subspace. Let X be any element not in $[\mathfrak{g}, \mathfrak{g}]$ so that X, Y is also a basis of $\mathfrak{g}$. Then $[X, Y]$ must be a non-zero element in $[\mathfrak{g}, \mathfrak{g}]$, hence it is a multiple of Y. By rescaling X, we may assume that $[X, Y] = Y$. This uniquely defines the bracket on $\mathfrak{g}$ (and one can easily check that the bracket does indeed make $\mathfrak{g}$ into a Lie algebra). Thus, up to isomorphism there are only two Lie algebras of dimension two.

In dimension three there are many more examples, but it is possible to completely classify them. The most important three dimensional Lie algebra is $\mathfrak{sl}(2, \mathbb{C})$, the subalgebra of $\mathfrak{gl}(2, \mathbb{C})$ consisting of matrices of trace zero.

Exercise 3.30. Let

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

Show that $\{E, F, H\}$ is a basis of $\mathfrak{sl}(2, \mathbb{C})$. Calculate $[H, E], [H, F]$ and $[E, F]$.

4 The exponential map

In this section, we define the exponential map. This allows us to go back from the Lie algebra of a Lie group to the group itself.

4.1 Left-invariant vector fields

First we note that there is a natural action of G on $\mathcal{C}^\infty(G)$, the space of all smooth functions on G. Namely, given $f \in \mathcal{C}^\infty(G)$ and $g, h \in G$, define the new function $g \cdot f$ by

$$(g \cdot f)(h) = f(g^{-1}h). \tag{10}$$

Now let $X \in \mathrm{Vect}(G)$ be a vector field on G. Since X is uniquely defined by its action on $\mathcal{C}^\infty(G)$, we can define an action of G on $\mathrm{Vect}(G)$ by

$$(g \cdot X)(f) = g \cdot [X(g^{-1} \cdot f)], \quad \forall\, f \in \mathcal{C}^\infty(G). \tag{11}$$

It might seem a bit strange that I've just decided to put in the $(-)^{-1}$ into equation (10), but this is necessary to ensure that both $\mathcal{C}^\infty(G)$ and $\mathrm{Vect}(G)$ become *left G-modules*. The vector field X is said to be *left invariant* if $g \cdot X = X$ for all $g \in G$. We denote by $\mathrm{Vect}^L(G) \subset \mathrm{Vect}(G)$ the subspace of all left invariant vector fields.

Exercise 4.1. Show that $\mathrm{Vect}^L(G)$ is a Lie subalgebra of $\mathrm{Vect}(G)$.

Lemma 4.2. *Let G be a Lie group. The map $X \mapsto X_e$ defines an isomorphism of vector spaces $\mathrm{Vect}^L(G) \xrightarrow{\sim} \mathfrak{g}$.*

Proof. Let $X \in \mathrm{Vect}^L(G)$. We will show that X is unique defined by its value X_e at e. Let $g, h \in G$ and $f \in \mathcal{C}^\infty(G)$. First we consider the differential $d_h(g \cdot f) : T_hG \to \mathbb{R}$. Since $(g \cdot f)(h) = f(g^{-1}h)$, we have $g \cdot f = f \circ L_{g^{-1}}$, where $L_g : G \to G$ is defined by $L_g(h) = gh$. Thus, $d_h(g \cdot f) = d_{g^{-1}h}f \circ d_h L_{g^{-1}}$. Then,

$$\begin{aligned}
[(g \cdot X)(f)](h) &= [g \cdot (X(g^{-1} \cdot f))](h) = [X(g^{-1} \cdot f)](g^{-1}h) \\
&= [d_{g^{-1}h}(g^{-1} \cdot f)](X_{g^{-1}h}) \\
&= (d_h f) \circ (d_{g^{-1}h} L_g)(X_{g^{-1}h}).
\end{aligned}$$

Since X is assumed to be left invariant, this implies that $(d_h f)(X_h) = (d_h f) \circ (d_{g^{-1}h} L_g)(X_{g^{-1}h})$

for all $f \in \mathcal{C}^\infty(G)$. Hence

$$X_h = (d_{g^{-1}h} L_g)(X_{g^{-1}h}), \quad \forall\, g, h \in G. \tag{12}$$

In particular, taking $h = g$ shows that $X_g = (d_e L_g)(X_e)$ i.e. X is uniquely defined by X_e. Thus, the map $X \mapsto X_e$ is injective.

Conversely, for each $X \in \mathfrak{g}$, define the vector field $\widetilde{X}$ on G by $\widetilde{X}_g = (d_e L_g)(X)$. This means that $\widetilde{X}(f)(h) = (d_h f)(d_e L_h(X))$ for all $f \in \mathcal{C}^\infty(G)$. It is left as an exercise to check that $\widetilde{X}$ belongs to $\mathrm{Vect}^L(G)$. □

Exercise 4.3. Check that the vector field $\widetilde{X}$ defined in the proof of Lemma 4.2 is left invariant.

Remark 4.4. One can actually show that the isomorphism $X \mapsto X_e$ of Lemma 4.2 is an isomorphism *of Lie algebras*.

Example 4.5. Let $G = (\mathbb{R}^\times, \cdot)$. Then $TG = \mathbb{R}^\times \times \mathbb{R}$ and $\mathrm{Vect}(G) = \{f\frac{\partial}{\partial x} \mid f \in \mathcal{C}^\infty(\mathbb{R}^\times)\}$. If $\alpha \in \mathbb{R}^\times$ and $f \in \mathcal{C}^\infty(G)$, then the group $\mathbb{R}^\times$ acts on $\mathcal{C}^\infty(\mathbb{R}^\times)$ by $(\alpha \cdot f)(p) = f(\alpha^{-1}p)$ e.g. $\alpha \cdot x^n = \alpha^n x^n$ since

$$(\alpha \cdot x^n)(\beta) = x^n(\alpha^{-1}\beta) = (\alpha^{-1}\beta)^n.$$

This implies that $\mathcal{C}^\infty(\mathbb{R}^\times)^G = \mathbb{R}$. Similarly,

$$\begin{aligned}\left(\alpha \cdot \left(f\frac{\partial}{\partial x}\right)\right)(x^n) &= \alpha \cdot \left(f\frac{\partial}{\partial x}\right)(\alpha^n x^n) \\ &= (\alpha \cdot f) n\alpha^n x^{n-1} \\ &= \left(\alpha(\alpha \cdot f)\frac{\partial}{\partial x}\right)(x^n).\end{aligned}$$

Thus, $\alpha \cdot \left(f\frac{\partial}{\partial x}\right) = \alpha(\alpha \cdot f)\frac{\partial}{\partial x}$; for instance $\alpha \cdot \left(x^n\frac{\partial}{\partial x}\right) = \alpha^{-n+1}x^n\frac{\partial}{\partial x}$. This implies that $\mathrm{Vect}^L(\mathbb{R}^\times) = \mathbb{R}\{x\frac{\partial}{\partial x}\}$.

4.2 One parameter subgroups

By Lemma 4.2 each element $X \in \mathfrak{g}$ defines a unique left invariant vector field $\nu(X)$ such that $\nu(X)_e = X$. Then, associated to $\nu(X)$ is an integral curve $\varphi_X : J \to G$ through e.

Lemma 4.6. *The integral curve φ_X is defined on the whole of $\mathbb{R}$. Moreover, it is a homomorphism of Lie groups $\mathbb{R} \to G$ i.e. $\varphi(s+t) = \varphi(s)\varphi(t)$ for all $s, t \in \mathbb{R}$.*

Proof. Choose $s, t \in J$ such that $s + t$ also belongs to J. Then we claim that $\varphi(s+t) = \varphi(s)\varphi(t)$. Fix s and let t vary in a some small open set $J_0 \subset J$ containing 0 such that $s+t$ still belongs to J.

Then $\alpha(t) = \varphi(s+t)$ and $\beta(t) = \varphi(s)\varphi(t)$ are curves in G such that $\alpha(0) = \beta(0)$. Differentiating, $\alpha(t)' = \varphi(s+t)' = \nu(X)_{\varphi(s+t)} = \nu(X)_{\alpha(t)}$. To calculate $\beta(t)'$, we first write $\beta(t) = (L_{\varphi(s)} \circ \varphi)(t)$. Then,

$$\begin{aligned}\beta(t)' &= (d_{\varphi(t)} L_{\varphi(s)})(\varphi'(t)) \\ &= (d_{\varphi(t)} L_{\varphi(s)}) \left(\nu(X)_{\varphi(t)}\right) \\ &= \nu(X)_{\varphi(s)\cdot\varphi(t)} = \nu(X)_{\beta(t)},\end{aligned}$$

where we have used (12) in the last equality but one. Thus, they are both integral curves for $\nu(X)$ through the point $\varphi(s)$. By uniqueness of integral curves, they are equal.

We can use the equality $\varphi(s+t) = \varphi(s)\varphi(t)$ to extend φ to the whole of $\mathbb{R}$: for each $s, t \in J$ such that $s+t$ is not in J, set $\varphi(s+t) = \varphi(s)\varphi(t)$. The uniqueness property of integral curves shows that this is well-defined. □

The homomorphism $\varphi_X : \mathbb{R} \to G$ is called the one-parameter subgroup associated to $X \in \mathfrak{g}$. The uniqueness of integral curves implies that it is the the unique homomorphism $\gamma : \mathbb{R} \to G$ such that $\gamma'(0) = X$.

Definition 4.7. The *exponential map* $\exp : \mathfrak{g} \to G$ is defined by $\exp(X) = \varphi_X(1)$.

The uniqueness of φ_X implies that $\varphi_{sX}(t) = \varphi_X(st)$ for all $s, t \in \mathbb{R}$ (since the derivatives with respect to t of $\varphi_{sX}(t)$ and $\varphi_X(st)$ at 0 are equal). Hence

$$\begin{aligned}(d_0 \exp)(X) &= \lim_{\epsilon \to 0} \frac{\exp(\epsilon X) - \exp(0)}{\epsilon} \\ &= \lim_{\epsilon \to 0} \frac{\varphi_{\epsilon X}(1) - \varphi_0(1)}{\epsilon} \\ &= \lim_{\epsilon \to 0} \frac{\varphi_X(\epsilon) - \varphi_X(0)}{\epsilon} \\ &= \varphi_X'(0) = X,\end{aligned}$$

where we have used the fact that $\varphi_0(t) = \varphi_X(0) = e$ for all $t \in \mathbb{R}$. Thus, the derivative $d_0 \exp$ of $\exp$ at $0 \in \mathfrak{g}$ is just the identity map.

In the case of $G = GL(n, \mathbb{R})$ and hence $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{R})$, the exponential map can be explicitly written down, it is just the usual exponential map

$$\exp(X) = \sum_{i=0}^{\infty} \frac{X^i}{i!}.$$

The same formula applies for any closed subgroup G of $GL(n,\mathbb{R})$. This function behaves much the same way as the exponential function $e^x : \mathbb{R} \to \mathbb{R}^\times$. However, it is *not* true that $\exp(X+Y)$ equals $\exp(X)\exp(Y)$ in general - see subsection 4.6 for more.

4.3 Cartan's Theorem

So far the only significant example of a Lie group we have is $GL(n,\mathbb{R})$. Cartan's Theorem allows us to cook up lots of new Lie groups. Our proof of Cartan's Theorem is based on the proof in [**?**]. The proof of the theorem involves some basic analysis, which we give as a separate lemma. Let S^{n-1} denote the unit sphere in $\mathbb{R}^n$. For each $x \in \mathbb{R}^n \smallsetminus \{0\}$, let $[x] = \frac{x}{\|x\|} \in S^{n-1}$ i.e. $[x]$ is the point on the unit sphere in $\mathbb{R}^n$ that also lies on the line through 0 and x.

Lemma 4.8. *Let $x_n, x \in \mathbb{R}^n \smallsetminus \{0\}$ with $\lim_{n\to\infty} x_n = 0$. Then, $\lim_{n\to\infty}[x_n] = [x]$ if and only if there exist positive integers a_n such that $\lim_{n\to\infty} a_n x_n = x$.*

Proof. If such integers exist then clearly $\lim_{n\to\infty}[x_n] = [x]$ since $[a_n x_n] = [x_n]$ for all n. Conversely, assume that $\lim_{n\to\infty}[x_n] = [x]$. To say that $\lim_{n\to\infty} a_n x_n = x$ means that the distance $\|a_n x_n - x\|$ between $a_n x_n$ and x becomes arbitrarily small as $n \to \infty$. For each n choose a positive integer a_n such that $|a_n - \frac{\|x\|}{\|x_n\|}| < 1$. If $v, w \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ then the triangle inequality implies that

$$\|\alpha v - w\| \leq \|\beta v - w\| + |\alpha - \beta| \cdot \|v\|.$$

In our situation, this implies that

$$\|a_n x_n - x\| \leq \left\| \frac{x_n \|x\|}{\|x_n\|} - x \right\| + \left| a_n - \frac{\|x\|}{\|x_n\|} \right| \cdot \|x_n\| \leq \left\| \frac{x_n \|x\|}{\|x_n\|} - x \right\| + \|x_n\|.$$

Now $\lim_{n\to\infty} x_n = 0$ implies that $\lim_{n\to\infty} \|x_n\| = 0$ and

$$\left\| \frac{x_n \|x\|}{\|x_n\|} - x \right\| = \frac{1}{\|x\|} \left\| \frac{x_n}{\|x_n\|} - \frac{x}{\|x\|} \right\| = \frac{1}{\|x\|} \|[x_n] - [x]\|$$

also tends to zero as n tends to infinity since $\lim_{n\to\infty}[x_n] = [x]$.

□

Theorem 4.9 (Cartan's Theorem)**.** *Let G be a Lie group and H a closed subgroup. Then H is a submanifold of G. Hence H is a Lie subgroup of G.*

Proof. We begin by noting that it suffices to construct some open neighborhood U of e in G such that $H \cap U$ is a closed submanifold of U. Assuming this, then for each $h \in H$, hU is an open neighborhood of h in G such that $H \cap hU = h(H \cap U)$ is a closed submanifold.

We fix a norm $\|\cdot\|$ on the vector space $\mathfrak{g}$. The key to the proof of Cartan's Theorem is to choose carefully a candidate subspace $W \subset \mathfrak{g}$ for the Lie algebra of H. The space W is defined to be all $w \in \mathfrak{g}$ such that either $w = 0$ or there exists a sequence $\{w_n\}$ in $\mathfrak{g}$ such that $w_n \neq 0$, $\exp(w_n) \in H$, $w_n \to 0$ and $[w_n] \to [w]$.

We will show:

1. $\exp(W) \subseteq H$.

2. W is a subspace of $\mathfrak{g}$.

3. There is an open neighborhood U of 0 in $\mathfrak{g}$ and a diffeomorphism $\phi : U \to \phi(U) \subset G$ with $\phi(0) = e$ such that $\phi(U \cap W) = \phi(U) \cap H$.

Assume (1)-(3) hold. Then, since W is a subspace of $\mathfrak{g}$, $U \cap W$ is clearly a submanifold of U. Therefore, part (3) implies that $\phi(U \cap W)$ is a submanifold of $\phi(U)$. Hence $\phi(U) \cap H$ is a submanifold of $\phi(U)$ as required.

Proof of part (1): Let $w \in W \smallsetminus \{0\}$, with $\{w_n\}$ as in the definition of W. By Lemma 4.8, there exist positive integers a_n such that $a_n w_n \to w$. Since $\exp(a_n w_n) = \exp(w_n)^{a_n} \in H$, and H is closed in G, the limit $\exp(w) = \lim_{n\to\infty} \exp(w_n)^{a_n}$ belongs to H.

Part (2): Since $[w] = [tw]$ for all $t \in \mathbb{R}^\times$, $tw \in W$ if $w \in W$. Thus, it suffices to show that if $v, w \in W$ then $v + w \in W$. We can assume without loss of generality that $v, w, v+w$ are non-zero. Recall that exp is a diffeomorphism in a neighborhood of 0. Therefore, for sufficiently small t, there exists a smooth curve $t \mapsto u(t)$ in $\mathfrak{g}$ such that

$$\exp(tv)\exp(tw) = \exp(u(t)), \tag{13}$$

and $u(0) = 0$. Equation (13) implies that $\exp(u(t)) \in H$ and

$$\lim_{n\to\infty} nu\left(\frac{1}{n}\right) = \lim_{\epsilon\to 0} \frac{u(\epsilon)}{\epsilon} = u'(0) = v + w.$$

Then, since $u\left(\frac{1}{n}\right) \to u(0) = 0$, $\exp\left(u\left(\frac{1}{n}\right)\right)$ belongs to H for all n, and $[u\left(\frac{1}{n}\right)] \to [v + w]$, we conclude that $v + w \in W$.

Part (3): Let V be a complement to W in $\mathfrak{g}$. We define

$$\phi : \mathfrak{g} = V \oplus W \to G, \quad (v, w) \mapsto \exp(v)\exp(w).$$

The differential $d_0\phi$ is just the identity map on $\mathfrak{g}$. Therefore there is some neighborhood U of 0 in $\mathfrak{g}$ such that ϕ is a diffeomorphism from U onto $\phi(U)$. Clearly, $\phi(U \cap W) \subset \phi(U) \cap \phi(W) \subset H \cap \phi(U)$. Therefore, we need to show that $H \cap \phi(U)$ is contained in $\phi(U \cap W)$.

Assume that this is not the case. Then, in every open neighbourhood U_n of 0 in $\mathfrak{g}$ there exist $(v_n, w_n) \in V \oplus W$ such that $\phi(v_n + w_n) \in H$ but $v_n \neq 0$. In particular, $\exp(v_n) \in H$. We take $U_1 \supset U_2 \supset \cdots$ such that $\bigcap_n U_n = \{0\}$, so that $\{v_n\}$ is a sequence converging to 0. Since the unit sphere S in V is compact, there exists some $v \in V \smallsetminus \{0\}$ and subsequence $\{v_n'\}$ of $\{v_n\}$ such that $[v_n'] \to [v]$ (this is the result from metric spaces saying that every sequence in a compact space has a convergent subsequence). But this implies that $v \in W$; a contradiction. $\square$

Remark 4.10. We've actually shown in the proof of Theorem 4.9 that there is some neighborhood U of e in G such that $U \cap H = \exp(\mathfrak{h}) \cap U$ for any closed subgroup H of G.

Example 4.11. As a consequence, the following are Lie groups

$$\begin{aligned}
B(n, \mathbb{R}) &= \{\text{All upper triangular, invertible matrices.}\}, \\
N(n, \mathbb{R}) &= \{A \in B(n, \mathbb{R}) \mid \text{Diagonal entries of } A \text{ all equal 1.}\}, \\
T(n, \mathbb{R}) &= \{\text{All diagonal matrices.}\}
\end{aligned}$$

since they are all subgroups of $GL(n, \mathbb{R})$ defined as the zeros of some polynomial equations.

Exercise 4.12. Using the criterion of example 2.6, show directly that $B(n, \mathbb{R})$, $N(n, \mathbb{R})$ and $T(n, \mathbb{R})$ are submanifolds of $GL(n, \mathbb{R})$.

The analogue of Cartan's Theorem holds for closed subgroups of $GL(n, \mathbb{C})$; they are complex Lie subgroups.

4.4 Simply connected Lie groups

We recall that a path connected topological space X is *simply connected* if the fundamental group $\pi_1(X)$ of X is trivial i.e. every closed loop in X is homotopic to the trivial loop. In general it is not true that a Lie group is uniquely defined by its Lie algebra i.e. it is possible to find non-isomorphic Lie groups whose Lie algebras are isomorphic. However, if we demand that our Lie group be simply connected, then:

Theorem 4.13 (S. Lie)**.** *Let $\mathfrak{g}$ be a finite dimensional, real Lie algebra. Then, there exists a unique simply connected Lie group G whose Lie algebra is $\mathfrak{g}$. Moreover, if G' is any other connected Lie group with Lie algebra $\mathfrak{g}$, then G' is a quotient of G.*

Recall that a covering map $f : M \to N$ is a map such that every $n \in N$ is contained in some open neighborhood U with $f^{-1}(U)$ a disjoint union of open sets, each mapping homeomorphically onto U. Before we prove the theorem, we require some preparatory results. A covering map always satisfies the *path lifting property*: if $\gamma : [0, 1] \to N$ is a path with $n = \gamma(0)$ and m is a lift of n, then there is a unique path $\tilde{\gamma} : [0, 1] \to M$ lifting γ such that $\tilde{\gamma}(0) = m$ (we say that $\tilde{\gamma}$ is a lift of γ if $f \circ \tilde{\gamma} = \gamma$). Using the path lifting property one can easily show that

Lemma 4.14. *Assume that M is simply connected and let $g : Z \to N$ be a smooth morphism from a simply connected manifold Z sending z to n. Then there exists a unique morphism $\tilde{g} : Z \to M$, sending z to m such that $f \circ \tilde{g} = g$.*

An easy, but important consequence of Lemma 4.14 is that every Lie group admits a simply connected cover.

Proposition 4.15. *Let G be a Lie group, H a connected manifold, and $\varphi : H \to G$ a covering map. Choose e', an element lying over the identity of G. Then there is a unique Lie group structure on H such that e' is the identity and φ is a map of Lie groups; and the kernel of φ is in the centre of H.*

Proof. We first assume that H is a simply connected manifold. Define the map $\alpha : H \times H \to G$ by $\alpha(h_1, h_2) = \varphi(h_1)\varphi(h_2)^{-1}$. By Lemma 4.14, there exists a unique map $\alpha' : H \times H \to H$ such that $\alpha'(e', e') = e'$ and $\varphi \circ \alpha' = \alpha$. Then we define $h^{-1} := \alpha'(e', h)$ and $h_1 \cdot h_2 = \alpha'(h_1, h_2^{-1})$ for $h, h_1, h_2 \in H$. This defines smooth morphisms $H \to H$ and $H \times H \to H$ resp. and one can use the uniqueness of the lift α to show that this makes H into a group.

□

The proof of the following proposition follows easily from the proof of Cartan's Theorem, Theorem 4.9.

Proposition 4.16. *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and $\mathfrak{h}$ a Lie subalgebra of $\mathfrak{g}$. Then the subgroup H generated by $\exp(\mathfrak{h})$ is a closed, Lie subgroup of G, whose Lie algebra is $\mathfrak{h}$.*

We also note:

Lemma 4.17. *Let G be a connected group and $\phi : G \to H$ a morphism of Lie groups. If $d_e\phi : \mathfrak{g} \to \mathfrak{h}$ is an isomorphism, then $d_g\phi : T_gG \to T_{\phi(g)}H$ is an isomorphism for all $g \in G$.*

Proof. This is simply the case of rewriting the map ϕ in a clever way. Recall that we have $L_g : G \to G$, $L_g(u) = gu$. Fix $g \in G$ and define $\varphi : G \to H$ by $\varphi(u) = \phi(g)\phi(u)$. We can rewrite this as $\varphi = L_{\phi(g)} \circ \phi$. This implies that $d_e\varphi = d_{\phi(g)}L_{\phi(g)} \circ d_e\phi$ is an isomorphism. On the other hand, φ also equals $u \mapsto \phi(gu)$, which we can write as $\varphi = \phi \circ L_g$. Therefore, $d_e\varphi = d_g\phi \circ d_eL_g$. Since both $d_e\varphi$ and d_eL_g are invertible linear maps, this implies that $d_g\phi$ is too. $\square$

Proof of Theorem 4.13. The key to Theorem 4.13 is Ado's Theorem, Theorem 7.27, which says that there exists some $n \gg 0$ such that $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R})$. Therefore Proposition 4.16 implies that there is some connected, closed Lie subgroup $G' \subset GL(n, \mathbb{R})$ with Lie $G' = \mathfrak{g}$. Let $\varphi : G \to G$ be the universal cover of G'. Proposition 4.15 says that we can endow G with the structure of a Lie group such that φ is a quotient of Lie groups. Moreover, $d_e\varphi : \mathfrak{g} \to \mathfrak{g}$ is the identity.

Thus, it suffices to show that if G' is another simply connected Lie group with Lie algebra $\mathfrak{g}$ then $G' \simeq G$. To show this, we consider the product $G \times G'$. It's Lie algebra is $\mathfrak{g} \oplus \mathfrak{g}$ and the diagonal copy $\mathfrak{g}_\Delta$ of $\mathfrak{g}$ in $\mathfrak{g} \oplus \mathfrak{g}$ is a Lie subalgebra. Therefore Proposition 4.16 implies that there is some connected, closed Lie subgroup $K \subset G \times G'$ such that Lie $K = \mathfrak{g}_\Delta$. The maps $\phi_1 : K \hookrightarrow G \times G' \twoheadrightarrow G$ and $\phi_2 : K \hookrightarrow G \times G' \twoheadrightarrow G'$ are homomorphisms of Lie groups whose differential at the identity is the identity map on $\mathfrak{g}$. Thus, we have maps ϕ_i between Lie groups whose differential is an isomorphism on Lie algebras. Hence, Lemma 4.17 implies that $d_k\phi_i$ is an isomorphism for all $k \in K$. Now, we have maps $\phi_i : K \to G, G'$ between connected manifolds, where G and G' are simply connected. As mentioned at the end of section 2.3, this implies that each ϕ_i is an isomorphism. Hence we may identify $K \simeq G \simeq G'$. $\square$

4.5 The proof of Theorems 3.15 and 3.8

Finally, we have the theorem that motivated the definition of a Lie algebra in the first place. It was stated as Theorem 3.15.

Theorem 4.18. *Let G and H be Lie groups with G simply connected, and let $\mathfrak{g}$ and $\mathfrak{h}$ be their Lie algebras. A linear map $\mu : \mathfrak{g} \to \mathfrak{h}$ is the differential of a morphism $\phi : G \to H$ if and only if μ is a map of Lie algebras.*

Proof. We showed in Proposition 3.14 that if μ is the differential of a homomorphism then it must be a map of Lie algebras. Therefore we need to show the existence of ϕ, knowing that μ is a map of Lie algebras.

We will deduce the theorem from Proposition 4.16, applied to $G \times H$. The Lie algebra of $G \times H$ is just $\mathfrak{g} \times \mathfrak{h}$. Let $\Gamma_\mu = \{(X, \mu(X)) \mid X \in \mathfrak{g}\} \subset \mathfrak{g} \times \mathfrak{h}$ be the graph of μ. Then the fact that μ is a map of Lie algebras is equivalent to saying that Γ_μ is a Lie subalgebra of $\mathfrak{g} \times \mathfrak{h}$. Let K be the subgroup of $G \times H$ generated by $\exp(\Gamma_\mu)$. By Proposition 4.16, K is a closed Lie subgroup of $G \times H$.

Projection from $G \times H$ onto G is a homomorphism of Lie groups. Therefore the composite $\eta : K \hookrightarrow G \times H \to G$ is also a homomorphism. The differential $d_e\eta$ is just the projection map from Γ_μ to $\mathfrak{g}$, which is an isomorphism. Thus, we have a map η between Lie groups whose differential is an isomorphism on Lie algebras. We note that since $\exp(\Gamma_\mu)$ is the image of a connected space under a continuous map, it is connected. Therefore, $\exp(\Gamma_\mu)$ is contained in K^0, the connected component of K containing e. Thus, $K = K^0$ is connected. Hence, Lemma 4.17 implies that $d_k\eta$ is an isomorphism for all k. Now, we have a map $\eta : K \to G$ between connected manifolds, where G is simply connected. As mentioned at the end of section 2.3, this implies that η is an isomorphism. Hence we may identify $G \simeq K$.

Since G is simply connected, Theorem 4.13 implies that this must be an isomorphism. Hence $G \simeq K \to H$ is a map of Lie groups whose differential is μ. $\square$

Remark 4.19. The map ϕ constructed in the proof of Theorem 4.18 is unique. To see this, assume we are given another map $\varphi : G \to H$ such that $d_e\varphi = \mu$. Let $\Gamma_\varphi \subset G \times H$ be the graph $\{(g, \varphi(g)) \in G \times H \mid g \in G\}$ of φ. Since φ is a homomorphism of Lie groups, Γ_φ is a Lie subgroup of $G \times H$. The Lie algebra of Γ_φ equals the Lie algebra Γ_μ of $K \subset G \times H$. This implies that $\exp(\Gamma_\mu)$ is contained in Γ_φ and hence $K \subset \Gamma_\varphi$. But both K and Γ_φ are connected Lie groups of the same dimension, hence they are equal. Since both φ and ϕ are defined to be projection from $K = \Gamma_\varphi$ onto H, $\varphi = \phi$.

Now the proof of Theorem 3.8 is an easy corollary.

Corollary 4.20. *Let G and H be Lie groups, with G connected. Then a morphism $f : G \to H$ is uniquely defined by the linear map $d_ef : T_eG \to T_eH$.*

Proof. By Theorem 4.13, there is a simply connected Lie group $\widetilde{G}$ and surjection $u : \widetilde{G} \to G$ such that $\mathrm{Lie}\,\widetilde{G} = \mathrm{Lie}\,G = \mathfrak{g}$ and $d_eu = \mathsf{id}_{\mathfrak{g}}$. Thus, $d_e(f \circ u) = d_ef$. By Theorem 4.18 and remark 4.19, there is a unique homomorphism $h : \widetilde{G} \to H$ such that $d_eh = d_e(f \circ u)$. This implies that $h = f \circ u$. If $g : G \to H$ was another map such that $d_eg = d_ef$, then $u \circ f = u \circ g$. But u is surjective, therefore this implies that $f = g$. $\square$

4.6 The Campbell-Hausdorff formula

Recall that $\exp(X)\exp(Y) \neq \exp(X + Y)$ in general. What we'd like is some "product" $X \star Y$ such that $\exp(X \star Y) = \exp(X)\exp(Y)$. If $g \in G$ is an element that is sufficiently close to the identity in G, then an inverse to exp is given by

$$\log(g) = \sum_{i=1}^{\infty} (-1)^{i+1} \frac{(g-e)^i}{i} \in \mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{R}).$$

We can try to use this to define $\star$ on $\mathfrak{gl}(n, \mathbb{R})$ by

$$X \star Y = \log(\exp(X)\exp(Y))$$

If we unpack this, being careful to remember that X and Y don't necessarily commute, then we get

$$\exp(X)\exp(Y) = e + (X+Y) + \left(\frac{X^2}{2} + XY + \frac{Y^2}{2}\right) + \cdots$$

and

$$\begin{aligned} X \star Y &= (X+Y) + \left(-\frac{(X+Y)^2}{2} + \left(\frac{X^2}{2} + XY + \frac{Y^2}{2}\right)\right) + \cdots \\ &= X + Y + \frac{1}{2}[X, Y] + \cdots \end{aligned}$$

We see that $X \star Y$ up to quadratic terms only depends on linear combinations of X, Y and brackets of X and Y. Remarkably, this is true for all higher terms too. The resulting formula is called the *Campbell-Hausdorff* formula.

Exercise 4.21. Calculate the degree three term of the Campbell-Hausdorff formula for $X \star Y$.

The key point of the Campbell-Hausdorff formula is that it shows that the product in $G \subset GL(n, \mathbb{R})$ can be described, at least in some neighborhood of the identity, completely in terms of the Lie bracket on $\mathfrak{g}$.

5 The classical Lie groups and their Lie algebras

In this section we describe the classical Lie groups. By Cartan's Theorem, every closed subgroup of $GL(n,\mathbb{R})$ or $GL(n,\mathbb{C})$ is a real (resp. complex) Lie group. Using this fact we can produce many interesting new examples.

5.1 The classical real Lie groups

Given a matrix A in $\mathrm{Mat}(n,\mathbb{R})$ or in $\mathrm{Mat}(n,\mathbb{C})$, we let A^T denote its transpose. Similarly, if $A \in \mathrm{Mat}(n,\mathbb{C})$ then $A^* = \overline{A^T} = \left(\overline{A}\right)^T$ is the *Hermitian conjugate* of A, where $\overline{A}$ is the matrix obtained by taking the complex conjugate of each entry e.g.

$$\begin{pmatrix} 2+i & -7i \\ 3 & -1+6i \end{pmatrix}^* = \begin{pmatrix} 2-i & 3 \\ 7i & -1-6i \end{pmatrix},$$

Finally we define the $2n \times 2n$ matrix

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

where $I_n \in GL(n,\mathbb{R})$ is the identity matrix.

Remark 5.1. If $v \in \mathbb{R}^{2n}$, thought of as a column vector, then $\omega(v,w) := v^T \cdot J_n \cdot w$ is a number. Thus, ω defines a bilinear form on $\mathbb{R}^{2n}$. It is skew-symmetric and non-degenerate. This form is the starting point of *symplectic geometry*, see [3] or [6]. This subject is the modern face of "Hamiltonian mechanics", first developed by the Irish mathematician William Hamilton in 1833 (the same mathematician who first conjured up the Quaternions).

Now we can define the real Lie groups

$$\begin{aligned}
SL(n,\mathbb{R}) &:= \{A \in GL(n,\mathbb{R}) \mid \det(A) = 1\}, \\
SO(n,\mathbb{R}) &:= \{A \in GL(n,\mathbb{R}) \mid \det(A) = 1,\ A^T \cdot A = 1\}, \\
O(n,\mathbb{R}) &:= \{A \in GL(n,\mathbb{R}) \mid A^T \cdot A = 1\}, \\
Sp(n,\mathbb{R}) &:= \{A \in GL(2n,\mathbb{R}) \mid A^T \cdot J_n \cdot A = J_n\}, \\
SU(n) &:= \{A \in GL(n,\mathbb{C}) \mid \det(A) = 1,\ A^* \cdot A = 1\}, \\
U(n) &:= \{A \in GL(n,\mathbb{C}) \mid A^* \cdot A = 1\};
\end{aligned}$$

these are the special linear group, special orthogonal group, orthogonal linear group, symplectic group, special unitary group and unitary group, respectively. Their Lie algebras are

$$\begin{aligned}
\operatorname{Lie} SL(n,\mathbb{R}) = \mathfrak{sl}(n,\mathbb{R}) &:= \{A \in \mathfrak{gl}(n,\mathbb{R}) \mid \operatorname{Tr}(A) = 0\},\\
\operatorname{Lie} SO(n,\mathbb{R}) = \operatorname{Lie} O(n,\mathbb{R}) = \mathfrak{o}(n,\mathbb{R}) &:= \{A \in \mathfrak{gl}(n,\mathbb{R}) \mid A^T + A = 0\},\\
\operatorname{Lie} Sp(n,\mathbb{R}) = \mathfrak{sp}(n,\mathbb{R}) &:= \{A \in \mathfrak{gl}(2n,\mathbb{R}) \mid A^T \cdot J_n + J_n \cdot A = 0\},\\
\operatorname{Lie} SU(n) = \mathfrak{su}(n) &:= \{A \in \mathfrak{gl}(n,\mathbb{C}) \mid \operatorname{Tr}(A) = 0,\ A^* + A = 0\},\\
\operatorname{Lie} U(n) = \mathfrak{u}(n) &:= \{A \in \mathfrak{gl}(n,\mathbb{C}) \mid A^* + A = 0\}.
\end{aligned}$$

Let us show that the Lie algebra of $SL(n,\mathbb{R})$ is $\mathfrak{sl}(n,\mathbb{R})$. If M is a submanifold of $\mathbb{R}^k$ defined, as in example 2.6, by the series of equations $f_1 = \cdots = f_r = 0$ and $m \in M$, then $T_m M$ is the subspace of $T_m\mathbb{R}^k = \mathbb{R}^k$ defined by

$$T_m M = \left\{ v \in \mathbb{R}^k \;\middle|\; \lim_{\epsilon\to 0} \frac{f_i(m+\epsilon v) - f_i(m)}{\epsilon} = 0,\ \forall\, i = 1,\dots,r \right\} = \operatorname{Ker}(d_m F : T_m\mathbb{R}^k \to T_0\mathbb{R}^r),$$

where $F = (f_1,\dots,f_r) : \mathbb{R}^k \to \mathbb{R}^r$. So, in our case, we have

$$\begin{aligned}
\mathfrak{sl}(n,\mathbb{R}) &= \left\{ A \in \mathfrak{gl}(n,\mathbb{R}) \;\middle|\; \lim_{\epsilon\to 0} \frac{(\det(I_n+\epsilon A) - 1) - (\det(I_n) - 1)}{\epsilon} = 0 \right\}\\
&= \left\{ A \in \mathfrak{gl}(n,\mathbb{R}) \;\middle|\; \lim_{\epsilon\to 0} \frac{\det(I_n+\epsilon A) - \det(I_n)}{\epsilon} = 0 \right\}.
\end{aligned}$$

Recall that an explicit formula for det is given by

$$\det(B) = \sum_{\sigma\in\mathfrak{S}_n} (-1)^\sigma b_{1,\sigma(1)} \cdots b_{n,\sigma(n)}, \tag{14}$$

where $\mathfrak{S}_n$ is the symmetric group on n letter. In the limit $\epsilon \to 0$ only the term ϵ^1 in the numerator will matter, all higher ϵ^i terms will go to zero. If we take an arbitrary term in the sum (14), corresponding to some $\sigma \neq 1$, then there must be i and j such that $\sigma(i) \neq i$ and $\sigma(j) \neq j$. For each such term in $\det(I_n + \epsilon A)$, $b_{i,\sigma(i)} b_{j,\sigma(j)}$ will contribute a ϵ^2. Thus,

$$\det(I_n + \epsilon A) = (1 + \epsilon a_{1,1}) \cdots (1 + \epsilon a_{n,n}) + \epsilon^2(\cdots),$$

and

$$\det(I_n + \epsilon A) - \det(I_n) = \epsilon \operatorname{Tr}(A) + \epsilon^2(\cdots).$$

Therefore, $A \in T_{I_n} SL(n,\mathbb{R})$ if and only if $\operatorname{Tr}(A) = 0$.

We'll also show Lie $Sp(n,\mathbb{R}) = \mathfrak{sp}(n,\mathbb{R})$, just to make sure we get the hang of things. This one is much easier:

$$\mathfrak{sp}(n,\mathbb{R}) = \left\{ A \in \mathfrak{gl}(2n,\mathbb{R}) \;\middle|\; \lim_{\epsilon\to 0} \frac{((I_{2n}+\epsilon A)^T J_n (I_{2n}+\epsilon A) - J_n) - ((I_{2n} J_n I_{2n} - J_n))}{\epsilon} = 0 \right\}.$$

But clearly,

$$((I_{2n}+\epsilon A)^T J_n (I_{2n}+\epsilon A) - J_n) - ((I_{2n} J_n I_{2n} - 1)) = \epsilon(A^T J_n + J_n A) + \epsilon^2(\cdots),$$

and hence $\mathfrak{sp}(n,\mathbb{R}) = \{A \in \mathfrak{gl}(2n,\mathbb{R}) \mid A^T J_n + J_n A = 0\}$.

Exercise 5.2. Show that Lie $SO(n,\mathbb{R}) =$ Lie $O(n,\mathbb{R}) = \mathfrak{o}(n,\mathbb{R})$ and Lie $SU(n) = \mathfrak{su}(n)$.

Exercise 5.3. What is the dimension of the Lie groups $SL(2,\mathbb{R})$, $Sp(2n,\mathbb{R})$, $O(n,\mathbb{R})$ and $SU(n)$. Hint: The dimension of a connected manifold is the same as the dimension of the tangent space $T_m M$ for any $m \in M$.

Exercise 5.4. If $e_{i,j}$ is the $n \times n$ matrix with a one in the (i,j)th position and zero elsewhere, give a formula in terms of Kronecker-delta symbols $\delta_{i,j}$ for the commutators $[e_{i,j}, e_{k,l}]$ in $\mathfrak{gl}(n,\mathbb{R})$.

5.2 The classical complex Lie groups

We list here the complex analogous of the above real Lie groups. There are no natural analogous to the unitary and special unitary groups.

$$\begin{aligned}
SL(n,\mathbb{C}) &:= \{A \in GL(n,\mathbb{C}) \mid \det(A) = 1\},\\
SO(n,\mathbb{C}) &:= \{A \in GL(n,\mathbb{C}) \mid \det(A) = 1,\ A^T \cdot A = 1\},\\
O(n,\mathbb{C}) &:= \{A \in GL(n,\mathbb{C}) \mid A^T \cdot A = 1\},\\
Sp(n,\mathbb{C}) &:= \{A \in GL(2n,\mathbb{C}) \mid A^T \cdot J_n \cdot A = J_n\}.
\end{aligned}$$

Their Lie algebras are

$$\begin{aligned}
\operatorname{Lie} SL(n,\mathbb{C}) = \mathfrak{sl}(n,\mathbb{C}) &:= \{A \in \mathfrak{gl}(n,\mathbb{C}) \mid \operatorname{Tr}(A) = 0\},\\
\operatorname{Lie} SO(n,\mathbb{C}) = \operatorname{Lie} O(n,\mathbb{C}) = \mathfrak{o}(n,\mathbb{C}) &:= \{A \in \mathfrak{gl}(n,\mathbb{C}) \mid A^T + A = 0\},\\
\operatorname{Lie} Sp(n,\mathbb{C}) = \mathfrak{sp}(n,\mathbb{C}) &:= \{A \in \mathfrak{gl}(2n,\mathbb{C}) \mid A^T \cdot J_n + J_n \cdot A = 0\},
\end{aligned}$$

5.3 The quaternions

The complex numbers $\mathbb{C}$ are a field, which can also be thought of as a two-dimensional vector space over $\mathbb{R}$. One can ask if there are other field that are finite dimensional vector spaces over $\mathbb{R}$. Strangely, the answer is no. However, if one considers *skew-fields* i.e. real vector spaces that are also rings (but not necessarily commutative) such that every non-zero element is invertible, then there does exist *one* other example.

The quaternions are a four-dimensional real vector space $\mathbb{H} = \mathbb{R} \oplus \mathbb{R}\mathbf{i} \oplus \mathbb{R}\mathbf{j} \oplus \mathbb{R}\mathbf{k}$ that are also a ring, where multiplication is defined by

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1.$$

Exercise 5.5. Show that every element in $\mathbb{H}^\times := \mathbb{H} \smallsetminus \{0\}$ is invertible. By giving explicit equations for multiplication, show that $\mathbb{H}^\times$ is a real Lie group.

The complex conjugation on $\mathbb{C}$ extends to a conjugation on $\mathbb{H}$ by $\overline{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$.

Exercise 5.6. Show that $u\overline{u} = a^2 + b^2 + c^2 + d^2$ if $u = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$.

As a complex vector space $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}\mathbf{j}$, so that $\overline{z + wb\mathbf{j}} = \overline{z} - w\mathbf{j}$ for $z, w \in \mathbb{C}$. The group $\mathbb{H}^\times$ acts on $\mathbb{H}$ on the right, $u \mapsto uv$ for $u \in \mathbb{H}$ and $v \in \mathbb{H}^\times$. The subgroup of $\mathbb{H}^\times$ consisting of all element v such that $v\overline{v} = 1$ is denoted $S^1(\mathbb{H})$.

Exercise 5.7.

1. Thinking of elements u in $\mathbb{H}$ as row vectors of length two, describe the action of $v \in \mathbb{H}^\times$ on u as a 2 by 2 complex matrix $A(v)$ so that $u \mapsto uA(v)$. Hint: First show that if $z \in \mathbb{C} \subset \mathbb{H}$, then $\mathbf{j}z = \overline{z}\mathbf{j}$.
2. Describe, as 2 by 2 matrices those elements belonging to $S^1(\mathbb{H})$.
3. Using part (2), construct an explicit isomorphism of Lie groups $S^1(\mathbb{H}) \xrightarrow{\sim} SU(2)$.
4. Show that $S^1(\mathbb{H}) \simeq SU(2) \simeq S^3$ (the 3-sphere in $\mathbb{R}^4$) as manifolds.

If you enjoy playing around with quaternions, you should take a look at John Baez' brilliant article on the octonions at

`www.ams.org/journals/bull/2002-39-02/S0273-0979-01-00934-X/home.html`.

5.4 Other exercises

Exercise 5.8. Construct an isomorphism between $GL(n, \mathbb{R})$ and a closed subgroup of $SL(n+1, \mathbb{R})$.

Exercise 5.9. Show that the map $\mathbb{C}^\times \times SL(n, \mathbb{C}) \to GL(n, \mathbb{C})$, $(\lambda, A) \mapsto \lambda A$, is surjective. Describe its kernel. Describe the corresponding homomorphism of Lie algebras.

6 Representation theory

In this section, we introduce representations of Lie algebras. For simplicity, we will assume that $\mathfrak{g}$ is a complex Lie algebra.

6.1 Representations of Lie algebras

Let $\mathfrak{g}$ be a Lie algebra and V a finite dimensional vector space. A representation of $\mathfrak{g}$ is a homomorphism of Lie algebras $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$.

Example 6.1. If $\mathfrak{g}$ is the Lie algebra of a Lie group, then Lemma 3.18 says that ad is a representation of $\mathfrak{g}$.

6.2 Modules

As in the literature, we will often use the equivalent language of modules. A $\mathfrak{g}$-module is a vector space V together with a *bilinear* action map $- \cdot - : \mathfrak{g} \times V \to V$ such that

$$[X, Y] \cdot v = X \cdot (Y \cdot v) - Y \cdot (X \cdot v) \quad \forall\ X, Y \in \mathfrak{g},\ v \in V.$$

Exercise 6.2. Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation. Show that V is a $\mathfrak{g}$-module with action map $X \cdot v = \rho(X)(v)$. Conversely, if V is a $\mathfrak{g}$-module, define $\rho : \mathfrak{g} \to \mathrm{End}(V)$ by $\rho(X)(v) = X \cdot v$. Show that ρ is actually a representation. Check that this defines a natural equivalence between $\mathfrak{g}$-representations and $\mathfrak{g}$-modules.

Remark 6.3. For those of you who are comfortable with the language of categories, both representations of a Lie algebra $\mathfrak{g}$ and the collection of all $\mathfrak{g}$-modules form categories; in fact they are *abelian* categories. Then exercise 6.2 is really saying that these two categories are equivalent.

6.3 Morphisms

A morphism of $\mathfrak{g}$-modules is a linear map $\phi : V_1 \to V_2$ such that $\phi(X \cdot v) = X \cdot \phi(v)$ for all $X \in \mathfrak{g}$ and $v \in V$ i.e. ϕ commutes with the action of $\mathfrak{g}$.

If $\phi : V_1 \to V_2$ is an invertible morphism of $\mathfrak{g}$-modules then ϕ is said to be an *isomorphism* of $\mathfrak{g}$-modules.

Exercise 6.4. Let $\phi : V_1 \to V_2$ be an isomorphism of $\mathfrak{g}$-modules. Show that $\phi^{-1} : V_2 \to V_1$ is also a morphism of $\mathfrak{g}$-modules.

6.4 Simple modules

Let V be a $\mathfrak{g}$-module. A subspace W of V is said to be a submodule of V if the action map $\cdot : \mathfrak{g} \times V \to V$ restricts to an action map $\cdot : \mathfrak{g} \times W \to W$. Equivalently, if $X \cdot w$ belongs to W for all $X \in \mathfrak{g}$ and $w \in W$. We say that W is a *proper* submodule of V if $0 \neq W \subsetneq V$.

Definition 6.5. A $\mathfrak{g}$-module is *simple* if it contains no proper submodules.

Given a particular Lie algebra $\mathfrak{g}$, one of the first things that one would want to work out as a representation theorist is a way to describe all the simple $\mathfrak{g}$-modules. This is often possible (but difficult).

6.5 New modules from old

We describe ways of producing new $\mathfrak{g}$-modules. Given two $\mathfrak{g}$-modules M and N, the space of $\mathfrak{g}$-module homomorphisms is denoted $\mathrm{Hom}_{\mathfrak{g}}(M, N)$. When $M = N$ we write $\mathrm{End}_{\mathfrak{g}}(M)$ for this space. Notice that $\mathrm{Hom}_{\mathfrak{g}}(M, N)$ is a subspace of $\mathrm{Hom}_{\mathbb{C}}(M, N)$ and $\mathrm{End}_{\mathfrak{g}}(M)$ is a subspace of $\mathrm{End}_{\mathbb{C}}(M)$.

Lemma 6.6 (Schur's Lemma)**.** *Let V be a simple, finite dimensional $\mathfrak{g}$-module. Then every $\mathfrak{g}$-module endomorphism of V is just a multiple of the identity i.e.* $\mathrm{End}_{\mathfrak{g}}(V) = \mathbb{C}$.

Exercise 6.7. 1. Prove Schur's lemma.

2. If V and W are non-isomorphic, simple $\mathfrak{g}$-modules, show that $\mathrm{Hom}_{\mathfrak{g}}(V, W) = 0$.

Let $\mathfrak{m}$ and $\mathfrak{l}$ be Lie algebras. Make $\mathfrak{m} \oplus \mathfrak{l}$ into a Lie algebra by

$$[(X_1, Y_1), (X_2, Y_2)] = ([X_1, X_2], [Y_1, Y_2]), \quad \forall\ X_i \in \mathfrak{m},\ Y_i \in \mathfrak{l}.$$

Exercise 6.8. If V is a $\mathfrak{m}$-module and W is a $\mathfrak{l}$-module, show that $V \otimes W$ is a $(\mathfrak{m} \oplus \mathfrak{l})$-module via

$$(X, Y) \cdot v \otimes w = (X \cdot v) \otimes (Y \cdot w), \quad \forall\ (X, Y) \in \mathfrak{m} \oplus \mathfrak{l},\ v \in V,\ w \in W.$$

Next we prove the following proposition, based on exercise 6.8.

Proposition 6.9. *Let V be a simple $\mathfrak{m}$-module and W a simple $\mathfrak{l}$-module. Then $V \otimes W$ is a simple $(\mathfrak{m} \oplus \mathfrak{l})$-module.*

Using only the Schur's lemma, the proof of proposition 6.9 is quite difficult. We break it into a series of lemmata.

Lemma 6.10. *Let V be a simple $\mathfrak{m}$-module and $v_1, v_2 \in V$ such that v_1 is not proportional to v_2. Then the smallest submodule of $V \oplus V$ containing (v_1, v_2) is $V \oplus V$.*

Proof. Let U be the smallest submodule of $V \oplus V$ containing (v_1, v_2). The inclusion map $i : U \hookrightarrow V \oplus V$ is a module homomorphism, as are the two projection maps $p_1, p_2 : V \oplus V \to V$. Therefore maps $p_1 \circ i, p_2 \circ i : U \to V$ are also homomorphisms. Since V is simple, U non-zero and $p_1 \circ i$ a non-zero map, it must be surjective. Now the kernel of $p_1 \circ i$ is contained in the kernel of p_1, which equals $\{0\} \oplus V$, a simple module. Therefore, either the kernel of $p_1 \circ i$ is $\{0\}$ and $U \simeq V$, or it is $\{0\} \oplus V$, in which case $U = V \oplus V$. Let's assume that the kernel of $p_1 \circ i$ is zero so that $p_1 \circ i : U \xrightarrow{\sim} V$. Its inverse will be written $\phi : V \to U$. The map $p_2 \circ i$ will also be an isomorphism $U \xrightarrow{\sim} V$. Hence $p_2 \circ i \circ \phi$ is an isomorphism $V \xrightarrow{\sim} V$. Since V is simple, Schur's lemma implies that $p_2 \circ i \circ \phi$ is some multiple of the identity map. But $p_2 \circ i \circ \phi$ applied to v_1 is $p_2 \circ i(v_1, v_2) = v_2$. By assumption, v_2 is not proportional to v_1, so $p_2 \circ i \circ \phi$ can't be a multiple of the identity. This is a contraction. Hence $U = V \oplus V$. □

Next, we prove a special case of the proposition.

Lemma 6.11. *Let V be a simple $\mathfrak{m}$-module and W a simple $\mathfrak{l}$-module. The smallest $(\mathfrak{m} \oplus \mathfrak{l})$-submodule of $V \otimes W$ containing a pure tensor $v \otimes w \neq 0$ is $V \otimes W$.*

Proof. The module $V \otimes W$ is spanned by all vectors $v' \otimes w'$ for $v' \in V$ and $w' \in W$. So it suffices to show that $v' \otimes w'$ is contained in the smallest submodule U of $V \otimes W$ containing $v \otimes w$. Since $(0, Y) \cdot v \otimes w = v \otimes Y \cdot w$ for all $Y \in \mathfrak{l} \subset \mathfrak{m} \oplus \mathfrak{l}$ and W is simple, $v \otimes W$ is the smallest $\mathfrak{l}$-submodule of $V \otimes W$ containing $v \otimes w$. Hence $v \otimes w' \in U$. Similarly, the smallest $\mathfrak{m}$-submodule of U containing $v \otimes w'$ is $V \otimes w'$. Hence $v' \otimes w' \in U$. □

Proof of Proposition 6.9. Let u be any non-zero element in $V \otimes W$ and U the smallest $(\mathfrak{m} \oplus \mathfrak{l})$-submodule of $V \otimes W$ containing u. Then we need to show that $U = V \otimes W$. If we knew that $0 \neq v \otimes w \in U$, then the result would follow from Lemma 6.11. Let $u = \sum_{i=1}^k v_i \otimes w_i$. After rewriting, we may assume that $v_1, \dots, v_k$ are linearly independent. Moreover, we may also assume that no pair w_{i_1}, w_{i_2} is proportional (if $w_{i_2} = \alpha w_{i_1}$, then $v_{i_1} \otimes w_1 + v_{i_2} \otimes w_{i_2} = (v_{i_1} + \alpha v_{i_2}) \otimes w_{i_1}$). Replacing u by another element of U if necessary, we may also assume that k is minimal satisfying the above properties. If $k = 1$ then we are done. So we assume that $k > 1$ and construct an element in U of "smaller length". Since $v_1, \dots, v_k$ are linearly independent, we can define an injective $\mathfrak{l}$-module homomorphism $\psi : W \oplus \cdots \oplus W \to V \otimes W$, $(w'_1, \dots, w'_k) \mapsto \sum_{i=1}^k v_i \otimes w'_i$. Then u is the image of $\mathbf{w} := (w_1, \dots, w_k)$ under this map. If $\mathbf{w}'$ is any element in the smallest $\mathfrak{l}$-submodule W' of $W \oplus \cdots \oplus W$ containing $\mathbf{w}$, then $\psi(\mathbf{w}')$ belongs to U. So it suffices to show that there is some non-zero element $\mathbf{w}'$ in W', with at least one coordinate 0. In this case $\psi(\mathbf{w}')$ will be a sum of less that k terms. We consider the projection of W' onto $W \oplus W \oplus \{0\} \oplus \cdots$. This is a $\mathfrak{l}$-submodule of $W \oplus W$ containing (w_1, w_2). But we assumed that w_1 and w_2 are not proportional. Hence Lemma 6.10 implies that this projection must be the whole of $W \oplus W$. In particular, $(w_1, 0)$ is in the image of the projection. Let $\mathbf{w}'$ be some element in W' projecting onto $(w_1, 0)$. This element has at least one coordinate zero, as required. □

Exercise 6.12. Let V and W be finite dimensional $\mathfrak{g}$-modules and $\mathrm{Hom}_{\mathbb{C}}(V, W)$ the space of linear maps from V to W. Show that the rule

$$(X \cdot f)(v) = X \cdot f(v) - f(X \cdot v), \quad \forall\ X \in \mathfrak{g},\ v \in V,\ f \in \mathrm{Hom}_{\mathbb{C}}(V, W)$$

makes $\mathrm{Hom}_{\mathbb{C}}(V, W)$ into a $\mathfrak{g}$-module.

6.6 Representations of $\mathfrak{sl}_2$

In this section we will completely describe the simple finite dimensional representations of $\mathfrak{sl}(2, \mathbb{C})$. Recall that $\mathfrak{g} := \mathfrak{sl}(2, \mathbb{C})$ has a basis $\{E, F, H\}$ such that $[H, E] = 2E$, $[H, F] = -2F$ and $[E, F] = H$. The relations imply that H is a semi-simple element. Let V be some finite dimensional $\mathfrak{g}$-module. We can decompose V into generalized eigenspaces with respect to H,

$$V = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha, \quad V_\alpha = \{v \in V \mid (H - \alpha)^N \cdot v = 0,\ \forall\ N \gg 0\}.$$

We say that V is a direct sum of *weight spaces* for H. If $v \in V$ belongs to a single V_α then we say that v is a *weight vector*.

Exercise 6.13. If $v \in V_\alpha$, show that $E \cdot v \in V_{\alpha+2}$ and $F \cdot v \in V_{\alpha-2}$.

Notice that if $V_\alpha \neq 0$ but $V_{\alpha+2} = 0$ then the exercise implies that $E \cdot v = 0$ for all $v \in V_\alpha$. A non-zero weight vector $v \in V$ is call *highest weight* if $E \cdot v = 0$ and $H \cdot v = \alpha v$.

Exercise 6.14. Let V be a finite dimensional $\mathfrak{g}$-module. Show that V contains a highest weight vector v_0. Set $v_{-1} = 0$ and $v_i = \frac{1}{i!} F^i \cdot v_0$ for $i \geq 0$. If v_0 has weight α, by induction on i show that

1. $H \cdot v_i = (\alpha - 2i) v_i$,
2. $E \cdot v_i = (\alpha - i + 1) v_{i-1}$,
3. $F \cdot v_i = (i + 1) v_{i+1}$.

Let V be as in exercise 6.14. Since $H \cdot v_i = (\alpha - 2i)v_i$ the vectors v_i are all linearly independent (they live in different weight spaces). But V is assumed to be finite dimensional, hence there exists some $N \gg 0$ such that $v_N \neq 0$ but $v_{N+1} = 0$. Consider equation (2) of exercise 6.14. We've said that $v_{N+1} = 0$ but $v_N \neq 0$. This implies that $\alpha - N = 0$ i.e. $\alpha = N$ is a positive integer. Since V is assumed to be simple, but contains v_0, the vectors $v_0, \ldots, v_N$ are a basis of V and each weight space V_{N-2i} for $i = 0, \ldots, N$ is one-dimensional with basis v_i. Moreover, $\dim V = N + 1$. Thus, we have shown:

Theorem 6.15. *Let V be a simple, finite dimensional $\mathfrak{sl}(2, \mathbb{C})$-module, with highest weight of weight α. Then,*

1. *α is a positive integer N.*
2. *$\dim V = N + 1$ and the non-zero weight spaces of V are $V_N, V_{N-2}, \ldots, V_{-N}$, each of which is one-dimensional.*
3. *Conversely, for any positive integer N there exists a unique simple $\mathfrak{sl}(2, \mathbb{C})$-module $V(N)$ with highest weight of weight N.*

Exercise 6.16. Recall from exercise 3.30 that the elements E, F and H can be written as particular 2×2 matrices. In representation theoretic terms, this means that $\mathfrak{sl}(2, \mathbb{C})$ has a natural two-dimensional representation, the "vectorial representation". Is this representation simple? If so, it is isomorphic to $V(N)$ for some N. What is N?

Exercise 6.17. We've also seen that $\mathfrak{sl}(2, \mathbb{C})$ acts on itself by the adjoint action $\mathrm{ad} : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{gl}(\mathfrak{sl}(2, \mathbb{C}))$. Is the adjoint representation simple? If so, what is N in this case? Explicitly write the highest weight vector in terms of E, F and H.

7 The structure of Lie algebras

In this section we introduce the notion of semi-simple, solvable and nilpotent Lie algebras. The main result of the section says that every Lie algebra is built up, in some precise way, from a semi-simple and a solvable Lie algebra (Levi's Theorem). Moreover, every semi-simple Lie algebra is a direct sum of simple Lie algebras. In order to simplify the proofs we will fix $k = \mathbb{C}$, the complex numbers (what is really needed for all of the results of this section to hold is that k be algebraically closed of characteristic zero).

7.1 A rough classification of Lie algebras

The *centre* $\mathfrak{z}(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$ is the subspace $\{X \in \mathfrak{g} \mid [X, Y] = 0, \forall\, Y \in \mathfrak{g}\}$. Notice that an abelian Lie algebra is precisely the same as a Lie algebra whose centre is the whole algebra. At the other extreme, a Lie algebra is called *simple* if it contains no proper ideals i.e. the only ideals in $\mathfrak{g}$ are $\mathfrak{g}$ and $\{0\}$.

Exercise 7.1. Show that $\mathfrak{z}(\mathfrak{g})$ is an ideal in $\mathfrak{g}$. Hence, if $\mathfrak{g}$ is simple, then $\mathfrak{z}(\mathfrak{g}) = 0$.

Example 7.2. The centre of $\mathfrak{gl}(n, \mathbb{C})$ is the one-dimensional ideal consisting of multiples of the identity matrix $\mathbb{C} \cdot I_n$. On the other hand, the Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ is a simple Lie algebra.

Let $\mathfrak{g}$ be a Lie algebra. The *lower central series* of $\mathfrak{g}$ is defined inductively by $\mathfrak{g}_1 = [\mathfrak{g}, \mathfrak{g}]$ and $\mathfrak{g}_k = [\mathfrak{g}, \mathfrak{g}_{k-1}]$. The *derived series* of $\mathfrak{g}$ is similarly inductively defined by $\mathfrak{g}^1 = [\mathfrak{g}, \mathfrak{g}]$ and $\mathfrak{g}^k = [\mathfrak{g}^{k-1}, \mathfrak{g}^{k-1}]$.

Definition 7.3. We say that $\mathfrak{g}$ is

- *nilpotent* if there exists some $n \gg 0$ such that $\mathfrak{g}_n = 0$,
- *solvable* if there exists some $n \gg 0$ such that $\mathfrak{g}^n = 0$,
- *semi-simple* if $\mathfrak{g}$ contains no proper, solvable ideals.

Exercise 7.4. Show that every nilpotent Lie algebra is solvable. Give an example of a solvable Lie algebra that is not nilpotent. Hint: Try $\dim \mathfrak{g} = 2$.

Exercise 7.5. Show that each piece $\mathfrak{g}_n$ of the lower central series is an ideal in $\mathfrak{g}$. Is the same true of the pieces $\mathfrak{g}^n$ of the derived series? Hint: what does the Jacobi identity tell you in this case?

Exercise 7.6. Let $\mathfrak{g}$ be a Lie algebra and I, J solvable ideals of $\mathfrak{g}$.

1. If $\mathfrak{g}/I$ is solvable, show that $\mathfrak{g}$ is solvable.

2. Using the fact that $(I+J)/I \simeq J/I \cap J$, show that $I+J$ is a solvable ideal of $\mathfrak{g}$.

Lemma 7.7. *Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Then $\mathfrak{g}$ contains a unique maximal solvable ideal. It is denoted* $\operatorname{rad}\mathfrak{g}$*, and called the solvable radical of* $\mathfrak{g}$*.*

Proof. Let $I = \sum_i I_i$ be the sum of all solvable ideals of $\mathfrak{g}$. Since $\mathfrak{g}$ is finite dimensional, I is certainly finite dimensional. Thus, there exists finitely many solvable ideals $I_1, \dots, I_k$ such that $I = I_1 + \cdots + I_k$. Inductively applying exercise 7.6, I is a solvable ideal. It is clearly the unique maximal one. $\square$

Notice that Lemma 7.7 implies that $\mathfrak{g}$ is semi-simple if and only if its solvable radical $\operatorname{rad}\mathfrak{g}$ is zero.

7.2 Engel's Theorem

Engel's Theorem is crucial in describing nilpotent Lie algebras. First, we begin with:

Exercise 7.8. Let V be a finite dimensional vector space and $X \in \mathfrak{gl}(V)$ a nilpotent endomorphism, $X^n = 0$ say. Show that $\operatorname{ad}(X)^{2n+1}(Y) = 0$ for all $Y \in \mathfrak{gl}(V)$.

Theorem 7.9. *Let $\mathfrak{n}$ be a subalgebra of $\mathfrak{gl}(V)$, for some finite dimensional vector space V. If $\mathfrak{n}$ consists of nilpotent endomorphisms and $V \neq 0$ then there exists some non-zero vector v such that* $\mathfrak{n}(v) = 0$.

Proof. The proof is by induction on $\dim \mathfrak{n}$. If $\dim \mathfrak{n} = 0$ or 1, the claim is clear.

First we show that there is an ideal in $\mathfrak{n}$ of codimension one. Let $\mathfrak{l}$ be any maximal proper subalgebra of $\mathfrak{n}$. Since $[\mathfrak{l}, \mathfrak{l}] \subset \mathfrak{l}$, the algebra $\mathfrak{l}$ acts on $\mathfrak{n}/\mathfrak{l}$. Exercise 7.8 implies that $\operatorname{ad}(\mathfrak{l})$ consists of nilpotent endomorphisms in $\mathfrak{gl}(\mathfrak{n})$. Hence the image of $\mathfrak{l}$ in $\mathfrak{gl}(\mathfrak{n}/\mathfrak{l})$ also consists of nilpotent endomorphisms. By induction, this implies that there is some $0 \neq \bar{Y} \in \mathfrak{n}/\mathfrak{l}$ such that $\mathfrak{l} \cdot \bar{Y} = 0$ i.e. $[\mathfrak{l}, Y] \subset \mathfrak{l}$. Thus, $\mathfrak{l} \oplus \mathbb{C}\{Y\}$ is a subalgebra of $\mathfrak{n}$. This implies that $\mathfrak{l} \oplus \mathbb{C}\{Y\} = \mathfrak{n}$ and $\mathfrak{l}$ is an ideal in $\mathfrak{n}$.

Since $\dim \mathfrak{l} < \dim \mathfrak{n}$, induction implies that there exists some $0 \neq v \in V$ such that $\mathfrak{l} \cdot v = 0$. Let W be the subspace of all such vectors. Since $\mathfrak{n} = \mathfrak{l} \oplus \mathbb{C}\{Y\}$, it suffices to show that there is some $0 \neq w \in W$ such that $Y(w) = 0$. Let $w \in W$ and $X \in \mathfrak{l}$. Then,

$$XY(w) = YX(w) + [X, Y](w) = 0,$$

since $X, [X, Y] \in \mathfrak{l}$ implies that $X(w) = [X, Y](w) = 0$. Hence $Y(W) \subset W$. Since Y is a nilpotent endomorphism of V and preserves W, it is a nilpotent endomorphism of W. Thus, there exists some $0 \neq w \in W$ such that $Y(w) = 0$. $\square$

What is Engel's theorem really saying? Here are two important corollaries of his theorem.

Exercise 7.10. Let $\mathfrak{n}$ be a Lie algebra such that $\mathfrak{n}/\mathfrak{z}(\mathfrak{n})$ is nilpotent. Show that $\mathfrak{n}$ is nilpotent.

Corollary 7.11. *Let $\mathfrak{n}$ be a finite dimensional Lie algebra. If every element in $\mathfrak{n}$ is ad-nilpotent, then $\mathfrak{n}$ is nilpotent.*

Proof. We may consider $\mathfrak{l} := \mathrm{ad}(\mathfrak{n}) \subset \mathfrak{gl}(\mathfrak{n})$. Our hypothesis says that $\mathfrak{l}$ consists of nilpotent endomorphisms. Therefore, by Theorem 7.9, there exists $0 \neq m \in \mathfrak{n}$ such that $[\mathfrak{n}, m] = 0$ i.e. $\mathfrak{z}(\mathfrak{n}) \neq 0$. If $\mathfrak{n}$ consists of ad-nilpotent elements, then clearly $\mathfrak{n}/\mathfrak{z}(\mathfrak{n})$ does too. Hence, by induction, we may assume that $\mathfrak{n}/\mathfrak{z}(\mathfrak{n})$ is a nilpotent Lie algebra. Then the corollary follows from exercise 7.10. $\square$

Recall that $\mathfrak{n}(n,k)$ denotes the Lie subalgebra of $\mathfrak{gl}(n,k)$ consisting of all strictly upper-triangular matrices. In order to concretely understand the meaning of Engel's Theorem we introduce the notion of flags in V. A *flag* $V_\bullet$ in V is a collection

$$0 = V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_k \subset V$$

of nested subspaces of V. We say that the flag $V_\bullet$ is *complete* if $\dim V_i/V_{i-1} = 1$ i.e. it's not possible to insert any more subspaces into the flag. If one fixes a basis $\{e_1, \dots, e_n\}$ of V then the standard complete flag is $V_\bullet$, where $V_i = \mathrm{Span}\{e_1, \dots, e_i\}$. The following lemma is clear.

Lemma 7.12. *Let $V_\bullet$ be the standard complete flag in $\mathbb{C}^n$.*

1. *The endomorphism $X \in \mathfrak{gl}(n,\mathbb{C})$ belongs to $\mathfrak{n}(n,\mathbb{C})$ if and only if $X(V_i) \subset V_{i-1}$ for all i.*
2. *The endomorphism $X \in \mathfrak{gl}(n,\mathbb{C})$ belongs to $\mathfrak{b}(n,\mathbb{C})$ if and only if $X(V_i) \subset V_i$ for all i.*

Notice that every complete flag in V is the standard complete flag of V with respect to some basis of V.

Corollary 7.13. *Let $\mathfrak{l}$ be a subalgebra of $\mathfrak{gl}(V)$, for some finite dimensional vector space V. If $\mathfrak{l}$ consists of nilpotent endomorphisms then there exists a basis of V such that $\mathfrak{l} \subset \mathfrak{n}(n,k)$; where $n = \dim V$.*

Proof. By Lemma 7.12, it suffices to show that there exists a complete flag $V_\bullet$ of V such that $X(V_i) \subset V_{i-1}$ for all $X \in \mathfrak{l}$. Let v be as in Theorem 7.9 and set $V_1 = \mathbb{C}\{v\} \subset V$. Then $\mathfrak{l}$ acts on V/V_1, again by nilpotent endomorphisms, hence by induction, there is a complete flag $\overline{V}_\bullet = (0 \subset \overline{V}_1 \subset \cdots \subset \overline{V}_{n-1} = V/V_1)$ in V/V_1 such that $\mathfrak{l}(\overline{V}_i) \subset \overline{V}_{i-1}$ for all i. Let $V_i := \{v \in V \mid \overline{v} \in \overline{V}_{i-1}\}$, for $i = 2, \dots, n$. Then $\mathfrak{l}(V_i) \subset V_{i-1}$ as required. $\square$

You may think, based on Corollaries 7.11 and 7.13, that if $\mathfrak{l}$ is a nilpotent subalgebra of $\mathfrak{gl}(V)$, then one can always find a basis of V such that $\mathfrak{l} \subset \mathfrak{n}(n, \mathbb{C})$. But this is not the case. Consider for instance the subalgebra $\mathfrak{h}$ of $\mathfrak{gl}(n, \mathbb{C})$ consisting of all diagonal matrices. This is abelian and hence nilpotent. It is not possible to change the basis of $\mathbb{C}^n$ such that this algebra sits in $\mathfrak{n}(n, \mathbb{C})$. The point here is that even though $\mathfrak{h}$ is nilpotent, the element in $\mathfrak{h}$ are not nilpotent matrices.

7.3 Lie's Theorem

The analogy of Engel's Theorem for solvable Lie algebras is Lie's Theorem. Unfortunately, its proof is more involved than that of Engel's Theorem.

Theorem 7.14. *Let V be a finite dimensional k-vector space and $\mathfrak{s}$ a solvable subalgebra of $\mathfrak{gl}(V)$. If $V \neq 0$ then there exists a common eigenvector for all endomorphisms in $\mathfrak{s}$.*

We should first say what it actually means for $\mathfrak{s}$ to have a common eigenvector. This means that there exists $0 \neq v$ in V and scalars α_X for every $X \in \mathfrak{s}$ such that $X(v) = \alpha_X v$.

Proof. The structure of the proof of Lie's Theorem is identical to the proof of Engel's Theorem, but the justification of each step is slightly different.

The proof is again by induction on $\dim \mathfrak{s}$. If $\dim \mathfrak{s} = 1$ then the claim is trivial. Therefore we assume that $\dim \mathfrak{s} > 1$. Since $\mathfrak{s}$ is assumed to be solvable $[\mathfrak{s}, \mathfrak{s}]$ is a proper ideal of $\mathfrak{s}$; or is zero. In either case, the Lie algebra $\mathfrak{s}/[\mathfrak{s}, \mathfrak{s}]$ is abelian and hence every subspace of this Lie algebra is an ideal. Take a subspace of codimension one and let $\mathfrak{n}$ denote its pre-image in $\mathfrak{s}$. Then $\mathfrak{n}$ is an ideal in $\mathfrak{s}$ of codimension one. By induction, there exists a joint eigenvector $0 \neq v \in V$ for $\mathfrak{n}$ i.e. there is some linear function $\lambda : \mathfrak{n} \to \mathbb{C}$ such that $X(v) = \lambda(X)v$ for all $X \in \mathfrak{n}$.

Consider the subspace $W = \{w \in V \mid X(w) = \lambda(X)w \ \forall \ X \in \mathfrak{n}\}$ of V. We've shown that it's non-zero. Lemma 7.15 below says that W is a $\mathfrak{s}$-submodule of V. If we choose some $Y \in \mathfrak{s} \smallsetminus \mathfrak{n}$ (so that $\mathfrak{s} = \mathbb{C} \cdot Y \oplus \mathfrak{n}$ as a vector space), then $Y(W) \subset W$. Hence there exists some $0 \neq w \in W$ that is an eigenvector for Y; this w is an eigenvector for all elements in $\mathfrak{s}$. $\square$

Lemma 7.15. *Let $\mathfrak{n}$ be an ideal in a Lie algebra $\mathfrak{s}$. Let V be a representation of $\mathfrak{s}$ and $\lambda : \mathfrak{n} \to k$ a linear map. Set*

$$W = \{v \in V \mid X(v) = \lambda(v) \ \forall \ X \in \mathfrak{n}\}.$$

Then W is a $\mathfrak{s}$-subrepresentation of V i.e. $Y(W) \subset W$ for all $Y \in \mathfrak{s}$.

Proof. Let $w \in W$ be non-zero. To show that $Y(w)$ belongs to W, we need to show that $X(Y(w)) = \lambda(X)Y(w)$ for all $X \in \mathfrak{n}$. We have

$$X(Y(w)) = Y(X(w)) + [X,Y](w) \tag{15}$$
$$= \lambda(X)Y(w) + \lambda([X,Y])w \tag{16}$$

where we have used the fact that $[X,Y] \in \mathfrak{n}$ because $\mathfrak{n}$ is an ideal. Therefore we need to show that $\lambda([X,Y]) = 0$ i.e. $\lambda([\mathfrak{s},\mathfrak{n}]) = 0$.

The proof of this fact is a very clever trick using the trace of an endomorphism. Let U be the subspace of V spanned by all $w, Y(w), Y^2(w), \dots$. Clearly, $Y(U) \subset U$. We claim that U is also a $\mathfrak{n}$-submodule of V i.e. $X(U) \subset U$ for all $X \in \mathfrak{n}$. Certainly $X(w) = \lambda(X)w \in U$ and equation (16) implies that $X(Y(w)) \in U$. So we assume by induction that $X(Y^k(w)) \in U$ for all $k < n$. Then,

$$X(Y^n(w)) = Y(X(Y^{n-1}(w))) + [X,Y](Y^{n-1}(w)).$$

Since $X, [X,Y] \in \mathfrak{n}$, $X(Y^{n-1}(w))$ and $[X,Y](Y^{n-1}(w))$ belong to U. Thus, $X(Y^n(w))$ also belongs to U. In fact the above argument shows inductively that $X(Y^n(w)) = \lambda(X)Y^n(W) +$ terms involving only $Y^k(w)$ for $k < n$. Thus, there is a basis of U such that X is upper-triangular with $\lambda(X)$ on the diagonals. In particular, $\mathrm{Tr}(X|_U) = \lambda(X)\dim U$. This applies to $[X,Y]$ too, $\mathrm{Tr}([X,Y]|_U) = \lambda([X,Y])\dim U$. But the trace of the commutator of two endomorphisms is zero. Hence[1] $\lambda([X,Y]) = 0$. □

Notice that the proof of Lemma 7.15 shows that if V is a one-dimensional $\mathfrak{s}$-module, then $[\mathfrak{s},\mathfrak{s}] \cdot V = 0$.

Again, just as in corollary 7.13, Lie's Theorem, together with Lemma 7.12, immediately implies:

Corollary 7.16. *Let V be a finite dimensional k-vector space and $\mathfrak{s}$ a solvable subalgebra of $\mathfrak{gl}(V)$. Then, there exists a basis of V such that $\mathfrak{s}$ is a subalgebra of $\mathfrak{b}(n,k)$ i.e. we can simultaneously put every element of $\mathfrak{s}$ in upper triangular form.*

The proof of corollary 7.16 is essentially the same as the proof of corollary 7.13.

Corollary 7.17. *Let $\mathfrak{s} \subset \mathfrak{gl}(V)$ be a solvable Lie algebra. Then $[\mathfrak{s},\mathfrak{s}]$ consists of nilpotent endomorphisms in $\mathfrak{gl}(V)$. In particular, $[\mathfrak{s},\mathfrak{s}]$ is a nilpotent Lie algebra.*

[1] Notice that we are using here the fact that the characteristic of our field is zero. The lemma is false in positive characteristic (as is Lie's Theorem).

Proof. By Corollary 7.16, we may assume that $\mathfrak{s} \subset \mathfrak{b}(n,\mathbb{C})$. Then, $[\mathfrak{s},\mathfrak{s}] \subset [\mathfrak{b}(n,\mathbb{C}),\mathfrak{b}(n,\mathbb{C})] = \mathfrak{n}(n,\mathbb{C})$. Hence $[\mathfrak{s},\mathfrak{s}]$ consists of nilpotent endomorphisms of V. The second statement then follows from Corollary 7.11 of Engel's Theorem. $\square$

7.4 The Killing form

Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Recall that the adjoint representation defines a homomorphism $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. We can use the adjoint representation to define a particular bilinear form on $\mathfrak{g}$.

Definition 7.18. The *Killing form* on $\mathfrak{g}$ is the bilinear form $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ defined by $\kappa(X,Y) = \mathrm{Tr}(\mathrm{ad}(X) \circ \mathrm{ad}(Y))$.

Since $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ for any two square matrices, the Killing form is symmetric i.e. $\kappa(X,Y) = \kappa(Y,X)$.

Exercise 7.19. Let $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$ with the usual basis $\{E,F,H\}$.

1. Calculate explicitly the adjoint representation of $\mathfrak{g}$ in terms of the basis $\{E,F,H\}$.

2. Using part (1), calculate the Killing form on $\mathfrak{g}$.

Exercise 7.20. A bilinear form β on $\mathfrak{g}$ is said to be *associative* if $\beta([X,Y],Z) = \beta(X,[Y,Z])$ for all $X,Y,Z \in \mathfrak{g}$. Show that the Killing form is associative.

The following key (but difficult) result shows that the Killing form can be used to tell if a given Lie algebra is solvable or not.

Theorem 7.21 (Cartan's criterion)**.** *Let V be a finite dimensional vector space and $\mathfrak{g} \subset \mathfrak{gl}(V)$ a Lie algebra. If* $\mathrm{Tr}(X,Y) = 0$ *for all $X,Y \in \mathfrak{g}$ then $\mathfrak{g}$ is solvable.*

The proof of Cartan's criterion is tricky, so we'll skip it.

Exercise 7.22. Show that every non-zero solvable ideal of a finite dimensional Lie algebra contains a non-zero abelian ideal of $\mathfrak{g}$. Hint: if $\mathfrak{l}^n = 0$, consider $\mathfrak{l}^{n-1}$.

Exercise 7.23. Show that the Lie algebra $\mathfrak{s}$ is solvable if and only if its image $\mathrm{ad}(\mathfrak{s})$ in $\mathfrak{gl}(\mathfrak{s})$ is solvable.

Corollary 7.24. *Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Then $\mathfrak{g}$ is semi-simple if and only if its Killing form κ is non-degenerate.*

Proof. Let S be the radical of the Killing form κ.

First we assume that the solvable radical $\operatorname{rad}\mathfrak{g}$ of $\mathfrak{g}$ is zero. Since $\kappa(S,S) = 0$, Cartan's criterion together with exercise 8.12 implies that S is a solvable Lie algebra. But S is also an ideal in $\mathfrak{g}$. Hence it is contained in the solvable radical of $\mathfrak{g}$. Thus, it is zero.

Conversely, assume that $S = 0$. Since every solvable ideal of $\mathfrak{g}$ contains a non-zero abelian ideal, it suffices to show that every abelian ideal $\mathfrak{l}$ of $\mathfrak{g}$ is contained in S. Let $\mathfrak{l}$ be one such ideal. Let $x \in \mathfrak{l}$ and $y \in \mathfrak{g}$. Then $(\operatorname{ad} x) \circ (\operatorname{ad} y)$ is a map $\mathfrak{g} \to \mathfrak{g} \to \mathfrak{l}$, and hence $(\operatorname{ad} x \circ \operatorname{ad} y)^2$ maps $\mathfrak{g}$ to zero. Hence $\operatorname{ad} x \circ \operatorname{ad} y$ is a nilpotent endomorphism. This implies that

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

and hence $\mathfrak{l} \subset S$ as required. $\square$

Theorem 7.25. *Let $\mathfrak{g}$ be a finite dimensional, semi-simple Lie algebra. Then, there is a unique decomposition*

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$$

of $\mathfrak{g}$ into a direct sum of simple Lie algebras.

Proof. Let $\mathfrak{l}$ be any non-zero ideal in $\mathfrak{g}$ and let $\mathfrak{l}^\perp = \{X \in \mathfrak{g} \mid \kappa(X,Y) = 0 \,\forall\, Y \in \mathfrak{l}\}$. Then $\mathfrak{l}^\perp$ is also an ideal - check! The Killing form restricted to $\mathfrak{l} \cap \mathfrak{l}^\perp$ is clearly zero. Therefore Cartan's criterion says that it is a solvable Lie algebra. But since it is also the intersection of two ideals, it is an ideal in $\mathfrak{g}$. Thus, since $\mathfrak{g}$ is semi-simple, $\mathfrak{l} \cap \mathfrak{l}^\perp = 0$. Hence $\mathfrak{l} \oplus \mathfrak{l}^\perp \subset \mathfrak{g}$. But $\dim \mathfrak{l} + \dim \mathfrak{l}^\perp = \dim \mathfrak{g}$. Hence $\mathfrak{l} \oplus \mathfrak{l}^\perp = \mathfrak{g}$. Since any solvable ideal of $\mathfrak{l}$ (or $\mathfrak{l}^\perp$) would also be a solvable ideal in $\mathfrak{g}$, both $\mathfrak{l}$ and $\mathfrak{l}^\perp$ are semi-simple.

Applying the same argument to $\mathfrak{l}$ and $\mathfrak{l}^\perp$ gives a finer decomposition of $\mathfrak{g}$ into a direct sum of semi-simple algebras. Since $\mathfrak{g}$ is finite dimensional, this process can't continue indefinitely, so we get some decomposition of $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$ into a direct sum of simple Lie algebras.

To show uniqueness, let $\mathfrak{h}$ be some simple ideal in $\mathfrak{g}$. We must show that $\mathfrak{h} = \mathfrak{g}_i$ for some i. The space $[\mathfrak{g}, \mathfrak{h}]$ is an ideal in $\mathfrak{g}$; it is non-zero because the centre of $\mathfrak{g}$ is trivial. Therefore, since $\mathfrak{h}$ was assumed to be simple, $[\mathfrak{g}, \mathfrak{h}] = \mathfrak{h}$. But,

$$[\mathfrak{g}, \mathfrak{h}] = [\mathfrak{g}_1, \mathfrak{h}] \oplus \cdots [\mathfrak{g}_k, \mathfrak{h}] = \mathfrak{h}.$$

Thus, there must exists a unique i such that $[\mathfrak{g}_i, \mathfrak{h}] = \mathfrak{h}$. But since $\mathfrak{g}_i$ is also simple and $[\mathfrak{g}_i, \mathfrak{h}] \neq 0$, this implies that $[\mathfrak{g}_i, \mathfrak{h}] = \mathfrak{g}_i$. $\square$

In the proof of Theorem 7.25, we have shown that if $\mathfrak{g}$ is semi-simple then $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$.

7.5 Levi's Theorem and Ado's Theorem

In this final section we state, without proof, two further important structure theorems about Lie algebras.

Theorem 7.26 (Levi). *Let $\mathfrak{g}$ be a finite dimensional Lie algebra and $\mathfrak{r}$ its radical. Then there exists a semi-simple subalgebra $\mathfrak{l}$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{l}$ as an $\mathfrak{l}$-module.*

Most of the example of Lie algebras we will encounter in the course are subalgebras of $\mathfrak{gl}(V)$ for some finite dimensional vector space V. Ado's Theorem says that this is not a coincidence.

Theorem 7.27 (Ado). *Every finite dimensional Lie algebra admits a faithful, finite dimensional representation.*

That is, given any $\mathfrak{g}$, we can always find some finite dimensional vector space V such that $\mathfrak{g}$ is a subalgebra of $\mathfrak{gl}(V)$.

8 Complete reducibility

In this section, we state and prove Weyl's complete reducibility theorem for semi-simple Lie algebras.

Definition 8.1. A $\mathfrak{g}$-module V is said to be *completely reducible* if there is a decomposition $V = V_1 \oplus \cdots \oplus V_k$ of V into a direct sum of simple $\mathfrak{g}$-modules.

Weyl's complete reducibility theorem say:

Theorem 8.2. *Let $\mathfrak{g}$ be a simple Lie algebra. Then, every finite dimensional representation of $\mathfrak{g}$ is completely reducible.*

Remark 8.3. The decomposition of a completely reducible $\mathfrak{g}$-module need not be unique. For instance, consider the extreme example where $\mathfrak{g} = 0$ so that a $\mathfrak{g}$-module is just a vector space. Any such module is always completely reducible: this is just the statement that any finite dimensional vector space can be expressed as the direct sum of one-dimensional subspaces. But there are many ways to decompose a vector space into a direct sum of one-dimensional subspaces.

Weyl's Theorem is a truly remarkable, and to me surprising, result. It is also very useful. When $\mathfrak{g}$ is not semi-simple, the statement of the theorem is simply not true. Also, even when $\mathfrak{g}$ is semi-simple, the assumption that V is finite dimensional is crucial. It is false for infinite dimensional representations. It is very useful to rephrase the notion of complete reducibility in terms of complements.

Lemma 8.4. *A finite dimensional $\mathfrak{g}$-module V is completely reducible if and only if every submodule $W \subset V$ admits a complement i.e. there exists a submodule $W' \subset V$ such that $V = W \oplus W'$.*

Proof. Assume first that every submodule of V admits a complement. In order to use induction on $\dim V$, let us show that submodules inherit this property i.e. if $W \subset V$ is a submodule and $U \subset W$ another submodule then U admits a complement in W. We can certainly find a complement U'' to U in V. We claim that $U' := U'' \cap W$ is a complement to U in W. Being the intersection of two submodules, it is a submodule, and $U' \cap U = \{0\}$. So it suffices to show that $U' + U = W$. Let $w \in W$. Then $w = u + u''$, where $u \in U$ and $u'' \in U''$. But $u'' = w - u \in W$ since $u, w \in W$. Hence $u'' \in U'$ and $U + U' = W$. Therefore, it follows by induction on $\dim V$, that if every submodule of V admits a complement in V then V is completely reducible.

Now assume that $V = V_1 \oplus \cdots \oplus V_k$ for some simple submodules of V. Let W be an arbitrary submodule. Notice that if V' is a simple submodule of V then either $V' \oplus W$ is a submodule of V or $V' \subset W$ - if $V' \not\subset W$ then $V' \cap W$ is a proper submodule of V' hence $V' \cap W = 0$ and

$V' \oplus W \subset V$. Now, since $V = V_1 \oplus \cdots \oplus V_k$, there must exist i_1 such that $V_{i_1} \not\subset W$. Hence $V_{i_1} \oplus W \subset V$. If $V_i \oplus W \subsetneq V$ then again there is some i_2 such that $V_{i_2} \cap (V_{i_1} \oplus W) = 0$ i.e. $V_{i_2} \oplus V_{i_1} \oplus W \subset V$. Continuing in this way, we eventually get

$$V = V_{i_r} \oplus \cdots \oplus V_{i_1} \oplus W$$

for some $r \leq k$ i.e. $V_{i_r} \oplus \cdots \oplus V_{i_1}$ is a complement to W in V. □

We give a couple of counter-examples when $\mathfrak{g}$ is not semi-simple.

- If $\mathfrak{g} = \mathbb{C}\{X\}$ with the trivial bracket, the we can consider the two-dimensional $\mathfrak{g}$-module $V = \mathbb{C}\{e_1, e_2\}$, where $X \cdot e_2 = e_1$ and $X \cdot e_1 = 0$. Then $V_1 = \mathbb{C}\{e_1\}$ is a $\mathfrak{g}$-submodule of V. Any complimentary subspace to V in V_1 is of the form $V_2 = \mathbb{C}\{e_2 + \beta e_1\}$. But then $X \cdot (e_2 + \beta e_1) = e_1$ means that none of these subspaces is a $\mathfrak{g}$-submodule. So no decomposition $V = V_1 \oplus V_2$ exists.

- If we take $\mathfrak{g} = \mathfrak{n}(3, \mathbb{C})$ and $V = \mathbb{C}\{e_1, e_2, e_3\}$ the standard column vectors representations, then $X \cdot e_1 = 0$ for all $X \in \mathfrak{n}(3, \mathbb{C})$. So $V_1 = \mathbb{C}\{e_1\}$ is a $\mathfrak{g}$-submodule. But there can be no other $\mathfrak{g}$-submodule U such that $V = V_1 \oplus U$ as a $\mathfrak{g}$-module: assume that such a U exists and take a non-zero vector $u = \alpha e_1 + \beta e_2 + \gamma e_3$ in U. Then either $\beta \neq 0$ or $\gamma \neq 0$. We have

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot u = \beta e_1, \qquad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot u = \gamma e_1,$$

  which shows in either case that $U \cap V_1 \neq \{0\}$ - a contradiction.

In order to prove Weyl's Theorem, we will need the notion of the Casimir element in $\mathrm{End}(V)$. Let $\mathfrak{g}$ be a semi-simple Lie algebra. Let β be some symmetric, non-degenerate associative bilinear form on $\mathfrak{g}$. Then, if we fix a basis $X_1, \ldots, X_n$ of $\mathfrak{g}$, there exists a unique "dual basis" $Y_1, \ldots, Y_n$ of $\mathfrak{g}$ such that $\beta(X_i, Y_j) = \delta_{i,j}$. A $\mathfrak{g}$-module V is said to be *faithful* if the action morphism $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is injective.

Lemma 8.5. *Let $\mathfrak{g}$ be a simple Lie algebra and V a $\mathfrak{g}$-module. Then either V is faithful or $X \cdot v = 0$ for all $X \in \mathfrak{g}$ and $v \in V$.*

Proof. Let $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ be the action morphism. Then $\mathrm{Ker}\,\rho$ is an ideal in $\mathfrak{g}$. Since $\mathfrak{g}$ is simple it has no proper ideals. Therefore, either $\mathrm{Ker}\,\rho = \{0\}$ i.e. V is faithful, or $\mathrm{Ker}\,\rho = \mathfrak{g}$ i.e. $X \cdot v = 0$ for all $X \in \mathfrak{g}$ and $v \in V$. □

Let V be a faithful $\mathfrak{g}$-module. Since each $\rho(X)$, for $X \in \mathfrak{g}$, is an endomorphism of V, we can define $\beta_V(X,Y) = \mathrm{Tr}(\rho(X)\rho(Y))$. The fact that V is faithful implies, by Cartan's Criterion, that β_V is non-degenerate. The *Casimir* of V is defined to be

$$\Omega_V = \sum_{i=1}^{n} \rho(X_i)\rho(Y_i),$$

where the X_i and Y_i are dual basis with respect to the form β_V. The whole point of defining the Casimir is:

Lemma 8.6. *Let V be a faithful $\mathfrak{g}$-module. Then the Casimir Ω_V is an endomorphism of V commuting with the action of $\mathfrak{g}$ i.e. $[\Omega_V, \rho(X)] = 0$ in $\mathfrak{gl}(V)$ for all $X \in \mathfrak{g}$. Moreover, the trace $\mathrm{Tr}(\Omega_V)$ of Ω_V, as an endomorphism of V, equals $\dim \mathfrak{g}$.*

Proof. First let $X \in \mathfrak{g}$ and write $[X, X_i] = \sum_{j=1}^{n} a_{i,j} X_j$ for some $a_{i,j} \in \mathbb{C}$. Similarly, $[X, Y_i] = \sum_{j=1}^{n} b_{i,j} Y_j$. Then,

$$a_{i,j} = \sum_{k=1}^{n} a_{i,k}\beta_V(X_k, Y_j) = \beta_V([X, X_i], Y_j) = -\beta_V(X_i, [X, Y_j]) = \sum_{k=1}^{n} -b_{j,k}\beta_V(X_i, Y_k) = -b_{j,i}.$$

Now, using the fact that $[XY, Z] = [X, Z]Y + X[Y, Z]$ in $\mathrm{End}(V)$, we have

$$\begin{aligned}
[\Omega_V, \rho(X)] &= \left[\sum_{i=1}^{n} \rho(X_i)\rho(Y_i), X\right] \\
&= \sum_{i=1}^{n} [\rho(X_i), \rho(X)]\rho(Y_i) + \rho(X_i)[\rho(Y_i), \rho(X)] \\
&= \sum_{i=1}^{n} \rho([X_i, X])\rho(Y_i) + \rho(X_i)\rho([Y_i, X]) \\
&= -\sum_{i=1,j}^{n} a_{i,j}\rho(X_j)\rho(Y_i) + b_{i,j}\rho(X_i)\rho(Y_j) = 0
\end{aligned}$$

since $b_{j,i} = -a_{i,j}$.

We have

$$\mathrm{Tr}(\Omega_V) = \sum_{i=1}^{n} \mathrm{Tr}(\rho(X_i)\rho(Y_i)) = \sum_{i=1}^{n} \beta_V(X_i, Y_i) = \dim \mathfrak{g},$$

since $X_1, \ldots, X_n$ is a basis of $\mathfrak{g}$ and $\beta_V(X_i, Y_i) = 1$. $\square$

Exercise 8.7. Let V be a finite dimensional complex vector space and $X \in \mathrm{End}_{\mathbb{C}}(V)$. Prove that

$V = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha$, where $V_\alpha = \{v \in V \mid (X - \alpha)^N \cdot v = 0, \text{ for } N \gg 0.\}$.

Lemma 8.4 implies that, in order to prove Weyl's Theorem, it suffices to show that if V is a $\mathfrak{g}$-module and W a submodule, then there exists a complementary $\mathfrak{g}$-submodule W' to W in V. That is, $V = W \oplus W'$ as $\mathfrak{g}$-modules.

Exercise 8.8. Let $\mathfrak{g}$ be a semi-simple Lie algebra and $V = \mathbb{C} \cdot e$ a one-dimensional $\mathfrak{g}$-module. Show that $X \cdot e = 0$ for all $X \in \mathfrak{g}$. Hint: use the fact that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$.

We begin by proving Weyl's Theorem in a special case. The general case reduces easily to this special case.

Lemma 8.9. *Let V be a $\mathfrak{g}$-module and W a submodule of codimension one. Then there exists a one-dimensional complementary $\mathfrak{g}$-submodule W' to W in V.*

Proof. The proof is by induction on $\dim V$. The case $\dim V = 1$ is vacuous.

By Lemma 8.5, either V is faithful or $X \cdot v = 0$ for any $X \in \mathfrak{g}$ and $v \in V$. In the latter case, one can take W' to be any complementary subspace to W in V (in this case every subspace of V is a submodule). Therefore we assume that V is a faithful $\mathfrak{g}$-module. By Lemma 8.6, the Casimir Ω_V is a $\mathfrak{g}$-module endomorphism of V. Therefore V will decompose into a direct sum of $\mathfrak{g}$-submodules

$$V = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha,$$

where $V_\alpha = \{v \in V \mid (\Omega_V - \alpha)^N(v) = 0, \ N \gg 0\}$. Since W is a $\mathfrak{g}$-submodule and Ω_V is expressed in terms of the $\rho(X)$, it maps W into itself. Therefore $W = \bigoplus_{\alpha \in \mathbb{C}} W_\alpha$, where

$$W_\alpha = \{w \in W \mid (\Omega_V - \alpha)^N(w) = 0, \ N \gg 0\} = V_\alpha \cap W.$$

Claim 8.10. We have $V_\alpha = W_\alpha$ for all $\alpha \neq 0$ and $\dim V_0/W_0 = 1$.

Proof. Since each V_α is a $\mathfrak{g}$-module and W_α a submodule, the quotient is a $\mathfrak{g}$-module. We have $V/W = \bigoplus_{\alpha \in \mathbb{C}} V_\alpha/W_\alpha$. Since $\dim V/W = 1$, there is exactly one α for which $V_\alpha \neq W_\alpha$. For this α, $\dim V_\alpha/W_\alpha = 1$. Now exercise 8.8 says that $X \cdot [v] = 0$ for all $X \in \mathfrak{g}$ and $[v] \in V_\alpha/W_\alpha$. Since Ω_V is expressed in terms of the $\rho(X)$, Ω_V acts as zero on the quotient. On the other hand, Ω acts with generalized eigenvalue α on V_α/W_α. Hence $\alpha = 0$. This completes the proof of the claim. $\square$

To complete the proof of the lemma, it suffices to show by induction that $\dim V_0 < \dim V$ since $W_0 \subset V_0$ is a submodule of codimension one. But Lemma 8.6 says that $\mathrm{Tr}(\Omega_V) = \dim \mathfrak{g} \neq 0$. Hence there exists at least one $\alpha \neq 0$ such that $V_\alpha \neq 0$. $\square$

Finally, we are in a position to prove Weyl's Theorem in complete generality. Thus, let V be a $\mathfrak{g}$-module and W a proper submodule. Recall from exercise 6.12 that $\mathrm{Hom}_{\mathbb{C}}(V, W)$ is a $\mathfrak{g}$-module, where

$$(X \cdot f)(v) = X \cdot f(v) - f(X \cdot v), \quad \forall\, X \in \mathfrak{g},\ v \in V,\ f \in \mathrm{Hom}_{\mathbb{C}}(V, W). \tag{17}$$

Define

$$U = \{f \in \mathrm{Hom}_{\mathbb{C}}(V, W) \mid f|_W = \lambda \mathrm{Id}_W \text{ for some } \lambda \in \mathbb{C}.\}$$

and $U' = \{f \in \mathrm{Hom}_{\mathbb{C}}(V, W) \mid f|_W = \mathrm{Id}_W\}$. Then it is an easy exercise, left to the reader, that U' and U are $\mathfrak{g}$-submodules of $\mathrm{Hom}_{\mathbb{C}}(V, W)$. There is a natural map $U \to \mathrm{Hom}_{\mathbb{C}}(W, W)$ given by $f \mapsto f|_W - \mathrm{Id}_W$. Clearly the image is a one-dimensional subspace of $\mathrm{Hom}_{\mathbb{C}}(W, W)$; this is actually a homomorphism of $\mathfrak{g}$-modules. Moreover, the kernel is precisely U'. Thus, U' has codimension one in U. Lemma 8.9 says that there is a complementary one-dimensional submodule U'' to U' in U. Choose $0 \neq \phi \in U''$. Then $\phi|_W = \mathrm{Id}_W$.

Claim 8.11. The map ϕ is a homomorphism of $\mathfrak{g}$-modules.

Proof. Since U'' is one-dimensional, $X \cdot \phi = 0$ for all $X \in \mathfrak{g}$. Equation (17) implies that this means that ϕ is a homomorphism of $\mathfrak{g}$-modules. □

The fact that $\phi|_W = \mathrm{Id}_W$ means that ϕ is a surjective map. The kernel $\mathrm{Ker}\,\phi$ is the complementary $\mathfrak{g}$-submodule to W in V.

Exercise 8.12. Let $\mathfrak{g} = \mathbb{C}\{x, y\}$ with $[x, y] = y$ be the unique non-abelian solvable 2-dimensional Lie algebra. Show that the adjoint representation of $\mathfrak{g}$ is not completely reducible.

8.1 Generalizing to semi-simple Lie algebras

It is possible to generalize Weyl's Theorem to semi-simple Lie algebras.

Exercise 8.13. Let V be a $\mathfrak{m} \oplus \mathfrak{l}$-module and W a $\mathfrak{l}$-module. Show that

$$(X \cdot f)(w) = (X, 0) \cdot f(w), \quad \forall\, X \in \mathfrak{m},\ f \in \mathrm{Hom}_{\mathfrak{l}}(W, V),\ w \in W,$$

makes $\mathrm{Hom}_{\mathfrak{l}}(W, V)$ into a $\mathfrak{m}$-module.

Theorem 8.14. *Let $\mathfrak{g}$ be a semi-simple Lie algebra. Then, every finite dimensional representation of $\mathfrak{g}$ is completely reducible.*

Proof. Recall from Theorem 7.25 that $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$, where each $\mathfrak{g}_i$ is simple. We prove the claim by induction on k. The case $k = 1$ is Theorem 8.2.

Write $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{m}$, where $\mathfrak{m} = \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_k$. We may assume that every finite dimensional $\mathfrak{m}$-module is completely reducible. Since $\mathfrak{g}_1$ is simple, $V = V_1^{n_1} \oplus \cdots \oplus V_r^{n_r}$, where the V_i are simple, *pairwise non-isomorphic* $\mathfrak{g}_1$-modules. We define a map $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_i \to V$ by $(\phi, v) \mapsto \phi(v)$. By exercise 8.13, $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V)$ is a $\mathfrak{m}$-module. Hence $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_i$ is a $\mathfrak{g}$-module. Then the map $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_i \to V$ is a homomorphism of $\mathfrak{g}$-modules. This extends to an isomorphism of $\mathfrak{g}$-modules

$$(\mathrm{Hom}_{\mathfrak{g}_1}(V_1, V) \otimes V_1) \oplus \cdots \oplus (\mathrm{Hom}_{\mathfrak{g}_1}(V_r, V) \otimes V_r) \xrightarrow{\sim} V.$$

Since each $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V)$ is a completely reducible $\mathfrak{m}$-module, we can decompose each $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_i$ into a direct sum of simple $\mathfrak{g}$-modules. $\square$

Exercise 8.15. In the proof of Theorem 8.14, show that the map $\mathrm{Hom}_{\mathfrak{g}_1}(V_i, V) \otimes V_i \to V$ is a homomorphism of $\mathfrak{g}$-modules. Show that $(\mathrm{Hom}_{\mathfrak{g}_1}(V_1, V) \otimes V_1) \oplus \cdots \oplus (\mathrm{Hom}_{\mathfrak{g}_1}(V_r, V) \otimes V_r) \to V$ is an isomorphism.

9 Cartan subalgebras and Dynkin diagrams

In this section we begin on the classification of simple, complex Lie algebras. Throughout, $\mathfrak{g}$ will be a simple Lie algebra over $\mathbb{C}$.

9.1 Cartan subalgebras

Let V be a finite dimensional vector space. Recall that an element $A \in \mathrm{End}_{\mathbb{C}}(V)$ is called semi-simple, resp. nilpotent, if A can be diagonalized, resp. $A^n = 0$ for some $n \gg 0$. In this context, Jordan's decomposition theorem can be stated as

Proposition 9.1. *Let $A \in \mathrm{End}_{\mathbb{C}}(V)$. Then there is a unique decomposition $A = A_s + A_n$, where A_s is semi-simple and A_n is nilpotent such that $[A_s, A_n] = 0$.*

Now let $\mathfrak{g}$ be a finite dimensional Lie algebra. We say that $X \in \mathfrak{g}$ is *semi-simple*, resp. *nilpotent*, if $\mathrm{ad}(X) \in \mathrm{End}(\mathfrak{g})$ is semi-simple, resp. is nilpotent.

Definition 9.2. A Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is called a *Cartan subalgebra* if every element of $\mathfrak{h}$ is semi-simple and it is a maximal subalgebra with these properties i.e. if $\mathfrak{h}'$ is another subalgebra of $\mathfrak{g}$ consisting of semi-simple elements and $\mathfrak{h} \subset \mathfrak{h}'$ then $\mathfrak{h} = \mathfrak{h}'$.

Since $\{0\}$ is a subalgebra of $\mathfrak{g}$ consisting of semi-simple elements and $\mathfrak{g}$ is finite dimensional, it is contained in at least one Cartan subalgebra i.e. Cartan subalgebras exist.

Lemma 9.3. *Let $\mathfrak{h}$ be a subalgebra of $\mathfrak{g}$ consisting of semi-simple elements. Then $\mathfrak{h}$ is abelian.*

Proof. Let $X \in \mathfrak{h}$. We need to show that $\mathrm{ad}_{\mathfrak{h}}(X) = 0$. Since $\mathrm{ad}_{\mathfrak{g}}(X)$ is semi-simple and $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{h}) \subset \mathfrak{h}$, $\mathrm{ad}_{\mathfrak{h}}(X)$ is also semi-simple. So it suffices to show that $\mathrm{ad}_{\mathfrak{h}}(X)$ has no non-zero eigenvalues i.e. if $Y \in \mathfrak{h} \smallsetminus \{0\}$ such that $\mathrm{ad}(X)(Y) = aY$, then $a = 0$. Assume that $a \neq 0$.

Since $Y \in \mathfrak{h}$ too, we may diagonalize $\mathrm{ad}(Y)$. That is, there exists a basis $X_1, \dots, X_n$ of $\mathfrak{h}$ and $\alpha_i \in \mathbb{C}$ such that $\mathrm{ad}(Y)(X_i) = \alpha_i X_i$. We may assume that $X_1 = Y$ and hence $\alpha_1 = 0$. Therefore, there exist unique $u_i \in \mathbb{C}$ such that $X = u_1 Y + u_2 X_2 + \cdots u_n X_n$. But then,

$$\mathrm{ad}(Y)(X) = -aY = \alpha_2 u_2 X_2 + \cdots \alpha_n u_n X_n.$$

Since $Y, X_2, \dots, X_n$ is a basis of $\mathfrak{h}$, this implies that $a = 0$; a contradiction. $\square$

In particular, Lemma 9.3 implies that every Cartan subalgebra of $\mathfrak{g}$ is abelian. For any subalgebra $\mathfrak{m}$ of $\mathfrak{g}$, we denote by $N_{\mathfrak{g}}(\mathfrak{m})$ the normalizer $\{X \in \mathfrak{g} \mid [X, \mathfrak{m}] \subset \mathfrak{m}\}$ of $\mathfrak{m}$ in $\mathfrak{g}$.

Exercise 9.4. Show that the normalizer $N_{\mathfrak{g}}(\mathfrak{m})$ of a subalgebra $\mathfrak{m}$ is itself a subalgebra of $\mathfrak{g}$. Moreover, show that $N_{\mathfrak{g}}(\mathfrak{m}) = \mathfrak{g}$ if and only if $\mathfrak{m}$ is an ideal in $\mathfrak{g}$.

A key property of Cartan subalgebras is that they equal their normalizers i.e. if $X \in \mathfrak{g}$ such that $[X, \mathfrak{h}] \subset \mathfrak{h}$, then $X \in \mathfrak{h}$.

Proposition 9.5. *The normalizer of a Cartan subalgebra* $\mathfrak{h}$ *is* $\mathfrak{h}$ *itself.*

Since each element h in $\mathfrak{h}$ is semi-simple, $\mathfrak{g}$ will decompose into a direct sum of eigenspaces $\mathfrak{g} = \bigoplus_\alpha \mathfrak{g}_\alpha$, where $[h, X] = \alpha X$ for all $X \in \mathfrak{g}_\alpha$. In fact, since $\mathfrak{h}$ is abelian, we can *simultaneously* decompose $\mathfrak{g}$ into eigenspaces for all $h \in \mathfrak{h}$. This means that

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \mathfrak{h}^* \smallsetminus \{0\}} \mathfrak{g}_\alpha$$

where $\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \mid [h, X] = \alpha(h)X, \ \forall h \in \mathfrak{h}\}$. Since $\mathfrak{g}$ is finite dimensional, there are only finitely many $\alpha \in \mathfrak{h}^* \smallsetminus \{0\}$ such that $\mathfrak{g}_\alpha \neq 0$. This set $R \subset \mathfrak{h}^*$ is called the set of *roots* of $\mathfrak{g}$. Since $\mathfrak{h}$ is abelian, $\mathfrak{h} \subset \mathfrak{g}_0$; this is actually an equality $\mathfrak{g}_0 = \mathfrak{h}$ by Proposition 9.5.

Corollary 9.6. *Let* $\mathfrak{h}$ *be a Cartan subalgebra of* $\mathfrak{g}$ *and* $\alpha, \beta \in R \cup \{0\}$*. The Killing form defines a non-degenerate pairing between* $\mathfrak{g}_\alpha$ *and* $\mathfrak{g}_{-\alpha}$*, and* $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ *if* $\alpha + \beta \neq 0$*. In particular,* $\kappa|_{\mathfrak{h}}$ *is non-degenerate.*

Proof. Let $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$. Then,

$$\alpha(h)\kappa(X, Y) = \kappa([h, X], Y) = -\kappa([X, h], Y) = -\kappa(X, [h, Y]) = -\beta(h)\kappa(X, Y).$$

Thus, $(\alpha(h) + \beta(h))\kappa(X, Y) = 0$ for all $h \in \mathfrak{h}$. If $\alpha + \beta \neq 0$, then $\mathrm{Ker}(\alpha + \beta)$ is a proper subset of $\mathfrak{h}$. Hence there exists some $h \in \mathfrak{h}$ such that $\alpha(h) + \beta(h) \neq 0$. Thus, if $\beta \neq -\alpha$, we have $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$. However, we know that the Killing form is non-degenerate on $\mathfrak{g}$. Therefore, it must define a non-degenerate pairing between $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$. ☐

Since κ is non-degenerate on $\mathfrak{h}$, there is a *canonical* identification $\eta : \mathfrak{h} \xrightarrow{\sim} \mathfrak{h}^*$ given by $\eta(t) = \kappa(t, -)$. Under this identification, we denote by t_α the element in $\mathfrak{h}$ corresponding to $\alpha \in \mathfrak{h}^*$ i.e. $t_\alpha := \eta^{-1}(\alpha)$.

9.2 α-strings through β

Next we'll try to understand a bit more this decomposition of $\mathfrak{g}$ with respect to a fixed Cartan subalgebra $\mathfrak{h}$. The key point is that the Killing form κ is non-degenerate on both $\mathfrak{g}$ and $\mathfrak{h}$.

Lemma 9.7. *Let R be the roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$.*

1. *If $\alpha \in R$ then $-\alpha \in R$.*

2. *The set of roots in R span $\mathfrak{h}^*$.*

Proof. Part (1) is a direct consequence of Corollary 9.6. If R does not span $\mathfrak{h}^*$ then there exists some $h \in \mathfrak{h}$ such that $\alpha(h) = 0$ for all $\alpha \in R$ (this follows from the fact that if $U \subseteq \mathfrak{h}^*$ is a subspace then $\dim U + \dim U^{\perp} = \dim \mathfrak{h}$, where $U^{\perp} = \{h \in \mathfrak{h} \mid u(h) = 0, \ \forall u \in U\}$). But this means that $[h, \mathfrak{g}_\alpha] = 0$ for all α. Since $[h, \mathfrak{h}] = 0$ too, this implies that $[h, \mathfrak{g}] = 0$ i.e. h belongs to the centre $\mathfrak{z}(\mathfrak{g})$ of $\mathfrak{g}$. But $\mathfrak{g}$ is semi-simple so $\zeta(\mathfrak{g}) = 0$. □

The following proposition tells us that every semi-simple $\mathfrak{g}$ is "built up" from a collection of copies of $\mathfrak{sl}(2, \mathbb{C})$ that interact in some complex way.

Proposition 9.8. *Let R be the roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. For each $\alpha \in R$:*

1. *There exist elements E_α in $\mathfrak{g}_\alpha$ and F_α in $\mathfrak{g}_{-\alpha}$ such that E_α, F_α and $H_\alpha := [E_\alpha, F_\alpha]$ span a copy $\mathfrak{s}_\alpha$ of $\mathfrak{sl}(2, \mathbb{C})$ in $\mathfrak{g}$.*

2. *$H_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$, where t_α was defined above.*

Proof. We begin by showing that

$$[X, Y] = \kappa(X, Y)t_\alpha, \quad \forall\ X \in \mathfrak{g}_\alpha,\ Y \in \mathfrak{g}_{-\alpha}. \tag{18}$$

Since $[X, Y] - \kappa(X, Y)t_\alpha$ belongs to $\mathfrak{h}$ and $\kappa|_{\mathfrak{h}}$ is non-degenerate, it suffices to show that $\kappa([X, Y] - \kappa(X, Y)t_\alpha, h) = 0$ for all $h \in \mathfrak{h}$. But

$$\kappa([X, Y], h) = \kappa(X, [h, Y]) = \alpha(h)\kappa(X, Y) = \kappa(t_\alpha, h)\kappa(X, Y).$$

Hence

$$\kappa([X, Y] - \kappa(X, Y)t_\alpha, h) = \kappa(t_\alpha, h)\kappa(X, Y) - \kappa(t_\alpha, h)\kappa(X, Y) = 0,$$

as required.

Claim 9.9. $\kappa(t_\alpha, t_\alpha) = \alpha(t_\alpha)$ is non-zero.

Proof. Assume otherwise, then $[t_\alpha, X] = [t_\alpha, Y] = 0$ for all $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$. By Corollary 9.6 we can choose X, Y such that $\kappa(X, Y) = 1$. Then $\mathfrak{s} = \mathbb{C}\{X, Y, t_\alpha\} \simeq \mathrm{ad}_{\mathfrak{g}}(\mathfrak{s})$ is a *solvable* subalgebra of $\mathfrak{g}$. By Corollary 7.17, this implies that $\mathbb{C}\,\mathrm{ad}_{\mathfrak{g}}(t_\alpha) = [\mathrm{ad}_{\mathfrak{g}}(\mathfrak{s}), \mathrm{ad}_{\mathfrak{g}}(\mathfrak{s})]$ consists of nilpotent endomorphisms. That is, $\mathrm{ad}_{\mathfrak{g}}(t_\alpha)$ is nilpotent. Since $t_\alpha \in \mathfrak{h}$, $\mathrm{ad}_{\mathfrak{g}}(t_\alpha)$ is also semi-simple i.e. $t_\alpha = 0$. This is a contradiction. Hence $\alpha(t_\alpha) \neq 0$. This completes the proof of the claim. $\square$

Now choose any $0 \neq E_\alpha \in \mathfrak{g}_\alpha$ and find $F_\alpha \in \mathfrak{g}_{-\alpha}$ such that $\kappa(E_\alpha, F_\alpha) = \frac{2}{\kappa(t_\alpha, t_\alpha)}$. Set $H_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$. Then the fact that $[X, Y] = \kappa(X, Y)t_\alpha$ implies that $\{E_\alpha, F_\alpha, H_\alpha\}$ is an $\mathfrak{sl}_2$-triple. $\square$

Corollary 9.10. *For each $\alpha \in R$, we have $\dim \mathfrak{g}_\alpha = 1$ and $\pm\alpha$ are the only multiplies of α in R.*

Proof. We will consider the decomposition of $\mathfrak{g}$ as a $\mathfrak{s}_\alpha$-module. We already know what the simple $\mathfrak{s}_\alpha$-modules look like; they are the $V(n)$ described in section 6.6.

Let M be the space spanned by all $\mathfrak{g}_{c\alpha}$, where $c \in \mathbb{C}$, and $\mathfrak{h}$. Then M is a $\mathfrak{s}_\alpha$-submodule of $\mathfrak{g}$. The classification of simple $\mathfrak{sl}(2, \mathbb{C})$-modules, together with Weyl's complete reducibility theorem implies that the weights of H_α on M are all integers; if $c\alpha \in R$ then $c\alpha(H_\alpha) = 2c \in \mathbb{Z}$. Now $\mathfrak{h} = \mathrm{Ker}\,\alpha \oplus \mathbb{C}H_\alpha$. This implies that $\mathrm{Ker}\,\alpha$ is a $\mathfrak{s}_\alpha$-submodule on which it acts trivially. Thus, we can decompose $M = \mathfrak{s}_\alpha \oplus \mathrm{Ker}\,\alpha \oplus M'$ for some $\mathfrak{s}_\alpha$-submodule M'. Notice that $M_0 = \mathfrak{h}$, so $M'_0 = 0$ i.e. all the H_α weights in M' are odd. This implies already that $\dim \mathfrak{g}_\alpha = 1$. Also, we see that 2α cannot be a root in R i.e. for any root in R, twice that root is never a root. Now if $\frac{1}{2}\alpha$ were a root then $2\frac{1}{2}\alpha = \alpha$ cannot possibly be a root. So actually $\frac{1}{2}\alpha$ isn't a root either. One can deduce from this that $\pm\alpha$ are the only multiplies of α in R. $\square$

Finally, we turn to asking how $\mathfrak{s}_\alpha$ acts on those weight spaces $\mathfrak{g}_\beta$ for $\beta \neq \pm\alpha$. By studying this action, we will deduce:

Lemma 9.11. *Let $\alpha, \beta \in R$, $\beta \neq \pm\alpha$. Then, $\beta(H_\alpha) \in \mathbb{Z}$ and $\beta - \beta(H_\alpha)\alpha \in R$.*

Proof. The space $\mathfrak{g}_\beta$ will be part of some simple $\mathfrak{s}_\alpha$-module $V(n)$ say. Since $E_\alpha \cdot \mathfrak{g}_\beta \subset \mathfrak{g}_{\beta+\alpha}$ and $F_\alpha \cdot \mathfrak{g}_\beta \subset \mathfrak{g}_{\beta-\alpha}$, there will be some $r, s \geq 0$ such that $V(n) = \mathfrak{g}_{\beta+rs} \oplus \cdots \oplus \mathfrak{g}_{\beta-s\alpha}$. Notice that all the weights $\beta + r\alpha, \beta + (r-1)\alpha, \ldots, \beta - s\alpha$ are roots in R. Moreover, the fact that $\dim \mathfrak{g}_\beta = 1$ implies that if $\beta + t\alpha \in R$ then $r \geq t \geq -s$ (otherwise consider the $\mathfrak{s}_\alpha$-module generated by $\mathfrak{g}_{\beta+t\alpha}$). We call $\beta + r\alpha, \beta + (r-1)\alpha, \ldots, \beta - s\alpha$ the *α-string through* β. In this case $n = (\beta + r\alpha)(H_\alpha) = \beta(H_\alpha) + 2r$. Now, as a subspace of $V(n)$, $\mathfrak{g}_\beta = V(n)_k$, where $k = \beta(H_\alpha)$. Thus, $\beta(H_\alpha) \in \mathbb{Z}$. Moreover, $V(n)_{-k}$ will also be a non-zero weight space $\mathfrak{g}_\gamma$ for some $\gamma \in R$. What's this γ? Well, it must be in the α-string through β so $\gamma = \beta + q\alpha$ for some q and $\gamma(H_\alpha) = -\beta(H_\alpha)$. Since $\alpha(H_\alpha) = 2$, we see that $\gamma = \beta - \beta(H_\alpha)\alpha$. $\square$

Recall that we have used the Killing form to identify $\mathfrak{h}^*$ with $\mathfrak{h}$, $\alpha \leftrightarrow t_\alpha$. Therefore we can define a non-degenerate symmetric bilinear form $(-,-)$ on $\mathfrak{h}^*$ by $(\alpha, \beta) = \kappa(t_\alpha, t_\beta)$. Notice that

$(\alpha, \alpha) = \kappa(t_\alpha, t_\alpha)$. Then, Lemma 9.11 says

$$\frac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z}, \quad \beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha \in R, \quad \forall\ \alpha, \beta \in R.$$

9.3 Root systems

We can axiomatize some of the properties of the set R. This leads to the notion of a root system. Let E be some real finite dimensional vector space (in our examples E with be a real subspace of our Cartan $\mathfrak{h}$ such that $R \subset E$, $\dim_{\mathbb{R}} E = \dim_{\mathbb{C}} \mathfrak{h}$ and $\mathfrak{h} = \mathbb{C} \otimes_{\mathbb{R}} E$), equipped with a positive definite symmetric bilinear form $(-, -)$. Such a vector space is called an Euclidean space. A linear transformation $M : E \to E$ is said to *orthogonal* if $(M(v), M(w)) = (v, w)$ for all $v, w \in E$.

Exercise 9.12. By fixing an orthonormal basis $x_1, \ldots, x_n$ of E, show that one can identify the group of orthogonal transformations of E with

$$\{M \in \mathfrak{gl}(n, \mathbb{R}) \mid M^T M = \mathrm{Id}\}.$$

Deduce that $\det(M) = \pm 1$.

A *reflection* of E is an orthogonal transformation s such that the subspace $\mathrm{Fix}(s) = \{x \in E \mid s(x) = x\}$ has codimension one in E. Similarly, a *rotation* of E is an orthogonal transformation r such that the subspace $\mathrm{Fix}(r) = \{x \in E \mid r(x) = x\}$ has codimension two in E. For each non-zero $\alpha \in E$, we define $H_\alpha = \{x \in E \mid (x, \alpha) = 0\}$ to be the orthogonal hyperplane to α.

Exercise 9.13. Let s be a reflection and $\alpha \in E$ such that $\mathrm{Fix}(s) = H_\alpha$. Show that

1. $s(\alpha) = -\alpha$,
2. For all $x \in E$,
$$s(x) = x - \frac{2(x, \alpha)}{(\alpha, \alpha)}\alpha.$$
3. Deduce that $s^2 = \mathrm{Id}$.

Hint: First show $E = \mathbb{R}\{\alpha\} \oplus H_\alpha$ and $s_\alpha(\alpha) \in \mathbb{R}\{\alpha\}$.

The above exercise shows that s is uniquely defined by α. Therefore, we can associate to each $\alpha \in E \smallsetminus \{0\}$ a reflection s_α.

Exercise 9.14. Let M be an orthogonal transformation of E, with $\dim E = 2$. Fixing an orthonormal basis of E, show that there is a unique $\theta \in [0, 2\pi)$ such that

$$M = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \text{or} \quad M = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

Deduce that every orthogonal transformation of E is either a rotation or reflection. How can one easily distinguish the two cases?

Definition 9.15. A subset R of E is called a *root system* if

1. R is finite, spans E and doesn't contain zero.

2. If $\alpha \in R$ then the only multiplies of α in R are $\pm\alpha$.

3. If $\alpha \in R$ then the reflection s_α maps R to itself.

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

Of course, this definition is chosen precisely so that Lemma 9.7, Corollary 9.10 and Lemma 9.11 imply that:

Proposition 9.16. *Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\mathfrak{h}$ a Cartan subalgebra. The decomposition of $\mathfrak{g}$ into $\mathfrak{h}$ weight spaces defines a root system R, in the abstract sense of Definition 9.15.*

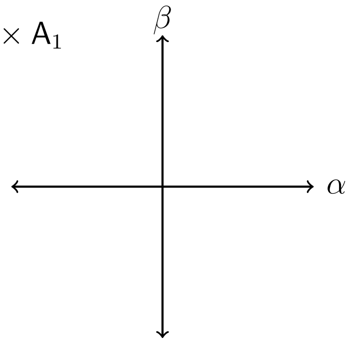
Exercise 9.17. Let $\alpha \neq \pm\beta \in R$, for some root system R. Show that $s_\alpha s_\beta$ is a rotation of E. Hint: decompose $E = \mathbb{R}\{\alpha, \beta\} \oplus H_\alpha \cap H_\beta$ and consider $s_\alpha s_\beta$ acting on $\mathbb{R}\{\alpha, \beta\}$. Use exercise 9.14.

9.4 The angle between roots

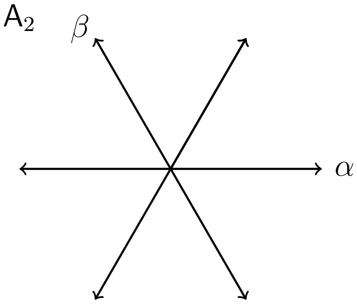
Recall that if $\alpha, \beta \in E$ are non-zero vectors, then the angle θ between α and β can be calculated using the cosine formula $||\alpha|| \cdot ||\beta|| \cos\theta = (\alpha, \beta)$. Thus,

$$\langle \beta, \alpha \rangle = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} = 2\frac{||\beta||}{||\alpha||} \cos\theta,$$

A$_1 \times$ A$_1$



A$_2$



and hence $\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = 4\cos^2\theta$. If $\alpha, \beta \in R$, then the fourth axiom implies that $4\cos^2\theta$ is a positive integer. Since $0 \le \cos^2\theta \le 1$, we have $0 \le \langle \beta, \alpha \rangle \langle \alpha, \beta \rangle \le 4$. Hence the only possible values of $\langle \alpha, \beta \rangle$ are:

$\langle \beta, \alpha \rangle$	$\langle \alpha, \beta \rangle$	θ
0	0	$\frac{\pi}{2}$
1	1	$\frac{\pi}{3}$
-1	-1	$\frac{2\pi}{3}$
1	2	$\star$
-1	-2	$\frac{3\pi}{4}$
1	3	$\frac{\pi}{6}$
-1	-3	$\star\star$
1	4	0
-1	-4	π

(19)

What about $\langle \beta, \alpha \rangle = \langle \alpha, \beta \rangle = \pm 2$?

Exercise 9.18. What are the angles θ in $(\star)$ and $(\star\star)$?

From this, it is possible to describe the rank two root systems.

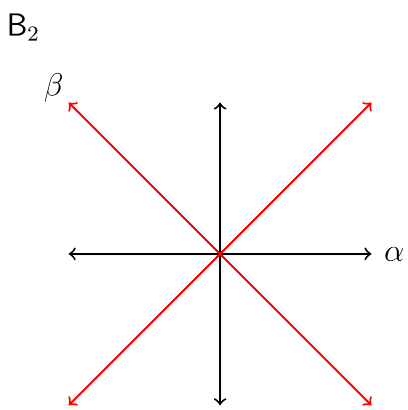
Exercise 9.19. Write out all the roots in $R(B_2)$ as linear combinations of the roots α and β.

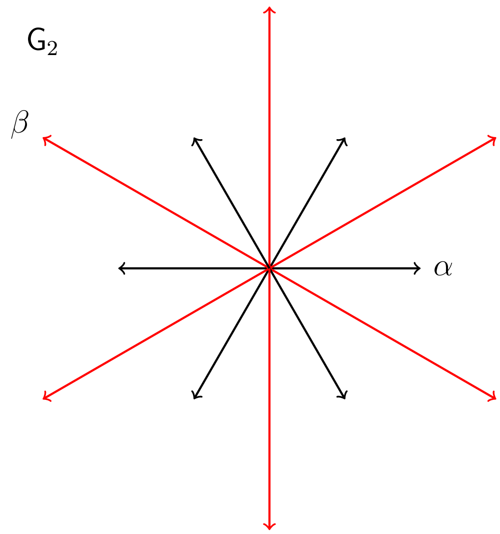
9.5 The Weyl group

The subgroup of $GL(E)$ generated by all reflections $\{s_\alpha \mid \alpha \in R\}$ is called the *Weyl group* of R.

Lemma 9.20. *The Weyl group W of R is finite.*

Proof. By axiom (3) of a root system every s_α maps the finite set R into itself. Therefore, there is a group homomorphism $W \to \mathfrak{S}_N$, where $N = |R|$. This map is injective: if $w \in W$ such that





its action on R is trivial then, in particular, w fixes a basis of E. Hence w acts trivially on E i.e. $w = 1$. $\square$

Exercise 9.21. Let $E = \{x = \sum_{i=1}^{n+1} x_i \epsilon_i \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i = 0\}$, where $\{\epsilon_1, \ldots, \epsilon_{n+1}\}$ is the standard basis of $\mathbb{R}^{n+1}$ with $(\epsilon_i, \epsilon_j) = \delta_{i,j}$. Let $R = \{\epsilon_i - \epsilon_j \mid 1 \leq i \neq j \leq n+1\}$.

1. Show that R is a root system.
2. Construct a set of simple roots for R.
3. By considering the action of the reflections $s_{\epsilon_i - \epsilon_j}$ on the basis $\{\epsilon_1, \ldots, \epsilon_{n+1}\}$ of $\mathbb{R}^{n+1}$, show that the Weyl group of R is isomorphic to $\mathfrak{S}_{n+1}$.

9.6 Simple roots

If R is a root system in E then by the first axiom of root systems, there is some subset of R that forms a basis of E.

Definition 9.22. Let R be a root system. A *set of simple roots* for R is a subset $\Delta \subset R$ such that

1. Δ is a basis of E.
2. Each $\beta \in R$ can be written as $\beta = \sum_{\alpha \in \Delta} m_\alpha \alpha$, where all m_α are positive integers, or all are negative integers.

The problem with the above definition is that it's not at all clear that a given root system contains a set of simple roots. However, one can show:

Theorem 9.23. *Let R be a root system.*

1. *There exists a set Δ of simple roots in R.*

2. *The group W is generated by $\{s_\alpha \mid \alpha \in \Delta\}$.*

3. *For any two sets of simple roots Δ, Δ', there exists a $w \in W$ such that $w(\Delta) = \Delta'$.*

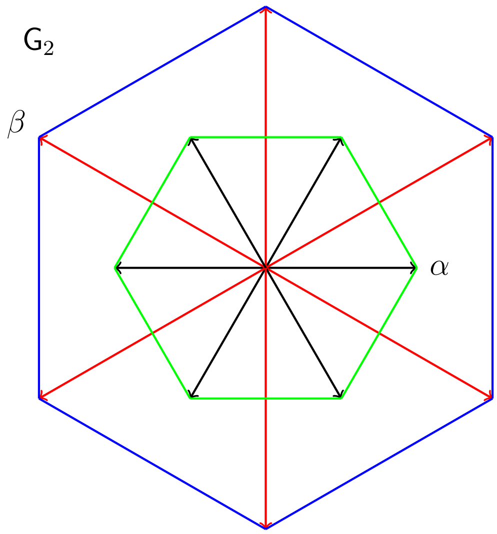
In fact the Weyl group W acts simply transitively on the collections of all sets of simple roots. The construction of all Δ's is quite easy, the difficulty is in showing that the sets constructed are indeed sets of simple roots and that the properties of Theorem 9.23 are satisfied. Recall that H_α denotes the hyperplane in E perpendicular to α. The open subset $E \smallsetminus \bigcup_{\alpha \in R} H_\alpha$ is a union of connected components. The components are called the Weyl chambers of R. If P is one of these chambers then choose some $\gamma \in P$. Notice that $(\gamma, \alpha) \neq 0$ for all $\alpha \in R$. We define $R(\gamma)^+$ to be all α such that $(\gamma, \alpha) > 0$ and $R(\gamma)^-$ similarly, so that $R = R(\gamma)^+ \sqcup R(\gamma)^-$. The sets $R(\gamma)^\pm$ are independent of the choice of γ. Then we say that $\alpha \in R(\gamma)^+$ is *decomposable* if there exist $\beta_1, \beta_2 \in R(\gamma)^+$ such that $\alpha = \beta_1 + \beta_2$. Otherwise α is said to be *indecomposable*. The set of all indecomposable roots in $R(\gamma)^+$ is denoted $\Delta(\gamma)$. Then, $\Delta(\gamma)$ is a set of simple roots in R and the construction we have described is in fact a bijection between the Weyl chambers of R and the collection of sets of simple roots. In particular, the Weyl group acts simply transitively on the set of Weyl chambers.

Example 9.24. The root system of type G_2 has 12 Weyl chambers. This implies that the Weyl group has order 12. On the other hand, the picture in figure 9.24 shows that the group generated by the reflections s_α and s_β is the set of symmetries either hexagon. This group (which must be contained in the Weyl group) has order 12 too, thus $W(G_2)$ equals the group of symmetries of the hexagon. It is called the dihedral group of order 12.

Exercise 9.25. How many Weyl chambers are there for the root system of type B_2? For each chamber describe the corresponding set of simple roots.

9.7 Cartan matrices and Dynkin diagrams

We have already extracted from our simple Lie algebra $\mathfrak{g}$ a root system, which is some abstract set of vectors in a real vector space. From this root system we will extract something even simpler - a finite graph. Remarkably, this graph, called the *Dynkin diagram* of $\mathfrak{g}$ contains all the information



needed to recover the Lie algebra $\mathfrak{g}$. It is this fact that allows us to completely classify the simple Lie algebras.

We proceed as follows. Give a roots system R, fix a set of simple roots Δ and order them: $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$. Then define the *Cartan matrix* of Δ to be the $\ell \times \ell$ matrix whose (i,j)th entry is $\langle \alpha_i, \alpha_j \rangle$.

Example 9.26. The rank two root systems have Cartan matrix:

$$\mathsf{A}_1 \times \mathsf{A}_1 \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}; \quad \mathsf{A}_2 \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}; \quad \mathsf{B}_2 \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}; \quad \mathsf{G}_2 \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}.$$

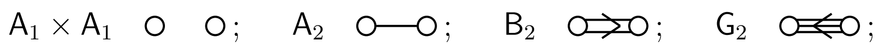
The following proposition implies that a root system R is uniquely defined up to isomorphism by its Cartan matrix.

Proposition 9.27. *Let $R \subset E$ and $R' \subset E'$ be two root systems, and $\Delta \subset R$, $\Delta' \subset R'$ sets of simple roots. Assume that there exists a bijection $\Phi : \Delta \to \Delta'$ such that $\langle \alpha_i, \alpha_j \rangle = \langle \Phi(\alpha_i), \Phi(\alpha_j) \rangle$ for all i, j. Then there is an isomorphism $\phi : E \xrightarrow{\sim} E'$ such that ϕ defines a bijection $R \to R'$, whose restriction to Δ equals Φ. This induces an isomorphism of Weyl groups $W \simeq W'$.*

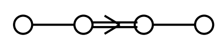
Proof. Since Δ and Δ' define a basis of E and E' respectively, Φ extends uniquely to an isomorphism $\phi : E \to E'$. For each $\alpha \in \Delta$, $\phi \circ s_\alpha \circ \phi^{-1} = s_{\phi(\alpha)}$ since s_α is uniquely defined by what it does on Δ. By Theorem 9.23, W is generated by the reflections in Δ. Therefore $\phi \circ W \phi^{-1} = W'$. Theorem 9.23 also says that for each $\alpha \in R$, there is some $w \in W$ such that $w(\alpha) \in \Delta$. Therefore $\phi(\alpha) = \phi(w)^{-1}(\phi(w(\alpha))) \in R'$ i.e. $\phi(R) \subset R$. Similarly, $\phi^{-1}(R') \subset R$. Thus, $\phi(R) = R'$. $\square$

Now given a Cartan matrix, define the Dynkin diagram of R to be the graph with ℓ vertices labeled by the ℓ simples roots $\alpha_1, \ldots, \alpha_\ell$ and with $\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle$ edges between vertex α_i and α_j. Finally to encode whether a simple root is long or short, we decorate the double and triple edges with an arrow pointing *towards* the shorter roots.

Example 9.28. The rank two root systems have Dynkin diagrams:



Exercise 9.29. Calculate the Cartan matrix associated to the Dynkin diagram

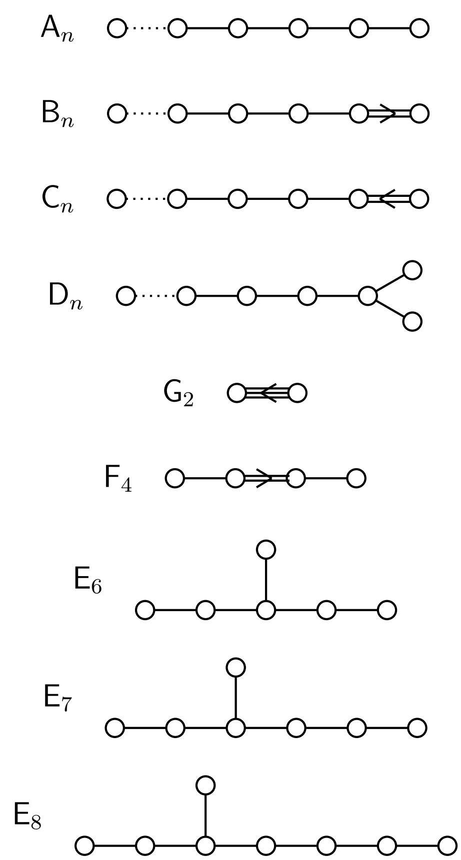
F_4 

Exercise 9.30. What is the Dynkin diagram associated to the root system of exercise 9.21?

10 The classification of simple, complex Lie algebras

Based solely on the definition of root system and Cartan matrix, it is possible to completely classify the possible Dynkin diagrams that can arise. The proof of the classification theorem involves only elementary combinatorial arguments, but it is rather long and tedious.

Theorem 10.1. *Let D be a connected Dynkin diagram. Then D belongs to the list:*



The Dynkin diagrams of type A_n, B_n, C_n and D_n correspond to the classical complex Lie algebras $\mathfrak{sl}(n+1,\mathbb{C})$, $\mathfrak{so}(2n+1,\mathbb{C})$, $\mathfrak{sp}(2n,\mathbb{C})$ and $\mathfrak{so}(2n,\mathbb{C})$ respectively.

Of course one has to then show that the Dynkin diagrams of Theorem 10.1 do come from some root system. This is done by explicitly constructing the root system in each case (for the classical types one can do this just by choosing a Cartan subalgebra and explicitly decomposing the Lie algebra with respect to the Cartan subalgebra).

10.1 Constructing Lie algebras from roots systems

Let R be a root system. Serre described how to construct a semi-simple Lie algebra from R such that R is the root system of $\mathfrak{g}$. This is done by giving $\mathfrak{g}$ in terms of generators and relations. Let $\Delta \subset R$ be a choice of simple roots, $\Delta = \{\alpha_1, \dots, \alpha_\ell\}$. Then $\mathfrak{g}$ will be generated by elements $H_1, \dots, H_\ell, E_1, \dots, E_\ell$ and $F_1, \dots, F_\ell$. Now we need to give all nessecary relations amongst the generators $H_1, \dots, F_\ell$, in addition to the anti-symmetry and Jacobi relation that all Lie algebras satisfy.

Theorem 10.2 (Serre)**.** *Let $\mathfrak{g}$ be the complex Lie algebra generated by $H_1, \dots, H_\ell, E_1, \dots, E_\ell$ and $F_1, \dots, F_\ell$ and satisfying the relations*

1. $[H_i, H_j] = 0$, *for all* $1 \le i, j \le \ell$;
2. $[H_i, E_j] = \langle \alpha_i, \alpha_j \rangle E_j$, *and* $[H_i, F_j] = -\langle \alpha_i, \alpha_j \rangle F_j$, *for all* $1 \le i, j \le \ell$;
3. $[E_i, F_i] = H_i$ *and* $[E_i, F_j] = 0$ *for all* $i \neq j$;
4. $\mathrm{ad}(E_i)^{-\langle \alpha_i, \alpha_j \rangle + 1}(E_j) = 0$ *for all* $i \neq j$;
5. $\mathrm{ad}(F_i)^{-\langle \alpha_i, \alpha_j \rangle + 1}(F_j) = 0$ *for all* $i \neq j$.

Then $\mathfrak{g}$ is a semi-simple, finite dimensional Lie algebra with root system R.

If $\mathfrak{g}$ is as in Theorem 10.2 then are a basis for $H_1, \dots, H_\ell$ a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. Since $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha$ and we have shown that $\dim \mathfrak{g}_\alpha = 0$, it follows that $\dim \mathfrak{g} = \ell + |R|$. The number ℓ is called the *rank* of $\mathfrak{g}$.

Exercise 10.3. Let $H_i = e_{i,i} - e_{i+1,i+1}$, $E_i = e_{i,i+1}$ and $F_i = e_{i+1,i}$ in $\mathfrak{gl}(n, \mathbb{C})$. If R is the root system of exercise 9.21, show that the satisfy the relations of Theorem 10.2. Show that they also generate the Lie algebra $\mathfrak{sl}(n+1, \mathbb{C})$. Thus, R is the root system of $\mathfrak{sl}(n+1, \mathbb{C})$.

Exercise 10.4. Write out the explicit relations for $\mathfrak{sl}(n+1, \mathbb{C})$ by calculating the integers $\langle \alpha_i, \alpha_j \rangle$.

Exercise 10.5. We define the $2n \times 2n$ matrix

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

where $I_n \in GL(n, \mathbb{C})$ is the identity matrix. Concretely, the *symplectic Lie algebra* $\mathfrak{sp}(2n)$ is defined to be the set of all $A \in \mathfrak{gl}(2n, \mathbb{C})$ such that $A^T \cdot J_n + J_n \cdot A = 0$. Let $\mathfrak{h} \subset \mathfrak{sp}(2n)$ denote the set of diagonal matrices.

(a) Check that $\dim \mathfrak{h} = n$.

(b) Writing $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where $a, b, c, d \in \mathfrak{gl}(n, \mathbb{C})$, show that $A \in \mathfrak{sp}(2n)$ if and only if $b^T = b$, $c^T = c$ and $a^T = -d$. What is the dimension of $\mathfrak{sp}(2n)$?

(c) Using part (b), decompose $\mathfrak{sp}(2n)$ into a direct sum of weight spaces for $\mathfrak{h}$. Deduce that $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{sp}(2n)$.

(d) (Harder) Using (c), write down the root system of $\mathfrak{sp}(2n)$.

(e) Using (d), compute the Dynkin diagram for $\mathfrak{sp}(2n)$.

10.2 The classification

Just as in Lie's classification theorem, Theorem 4.13, it is known that $\mathfrak{g}$ is always the Lie algebra of some complex Lie group G (though in general there are several Lie groups whose Lie algebra is $\mathfrak{g}$). We've seen that every simple $\mathfrak{g}$ gives rise to a Dynkin diagram. Conversely, given a Dynkin diagram, Theorem 10.2 says that we can construct a simple $\mathfrak{g}$ with that Dynkin diagram. We would like to show that this gives a bijection between the isomorphism classes of simple Lie algebras and Dynkin diagrams. In extracting the Dynkin diagram from $\mathfrak{g}$ we made two choices. Firstly, we choose a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. We need to show that the corresponding root system is essential independent of this choice of Cartan.

Theorem 10.6. *Any two Cartan subalgebra of $\mathfrak{g}$ are conjugate by an element of G.*

Theorem 10.6 says that if I'm given two Cartan subalgebras $\mathfrak{h}$ and $\mathfrak{h}'$ of $\mathfrak{g}$ then there exists some $g \in G$ such that $\mathfrak{h}' = \mathrm{Ad}(g)(\mathfrak{h})$. Thus, $\mathrm{Ad}(g)$ is an automorphism of $\mathfrak{g}$ sending the $\mathfrak{h}$-weight decomposition of $\mathfrak{g}$ into the $\mathfrak{h}'$-weight decomposition. This means that the isomorphism $\mathrm{Ad}(g) : \mathfrak{h} \to \mathfrak{h}'$ (which is an isometry since the Killing form satisfies $\kappa(\mathrm{Ad}(g)(X), \mathrm{Ad}(g)(Y)) = \kappa(X, Y)$) maps the root system $R(\mathfrak{h})$ bijectively onto $R(\mathfrak{h}')$. The second choice we made was a choice of simple roots Δ in R. We've seen in Theorem 9.23 that for any two $\Delta, \Delta' \subset R$ there is an element w in the Weyl group W such that $w(\Delta) = \Delta'$. Thus, the worst that can happen is that the Dynkin diagram of Δ differs by an automorphism from the Dynkin diagram of Δ'. Hence, up to Dynkin automorphisms, the Dynkin diagram is uniquely defined by $\mathfrak{g}$.

Summarizing, the classification result says:

Theorem 10.7. *A simple Lie algebra $\mathfrak{g}$ is uniquely defined, up to isomorphism, by its Dynkin diagram. Moreover, for each connected Dynkin diagram, there exists a simple Lie algebra with that Dynkin diagram.*

Hence, up to isomorphism, there are four infinite series (type A_n, B_n, C_n and D_n) of simple Lie algebras - these are the Lie algebras of classical type, and five exceptional Lie algebras (type G_2, F_4, E_6, E_7 and E_8).

11 Weyl's character formula

We've seen in the previous two sections that the key to classifying semi-simple complex Lie algebras is to chose a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ and decompose $\mathfrak{g}$, via the adjoint representation, as a $\mathfrak{h}$-module. We can apply the same idea to $\mathfrak{g}$-modules i.e. given a $\mathfrak{g}$-module V, we can consider it as a $\mathfrak{h}$-module and ask for its decomposition. First, we describe the classification of simple, finite dimensional $\mathfrak{g}$-modules.

11.1 Highest weight modules

Fix a set of simple roots Δ so that $R = R^+ \sqcup R^-$. Let $\mathfrak{n}_+ = \bigoplus_{\alpha \in R^+} \mathfrak{g}_\alpha$. Since $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$, the subspace $\mathfrak{n}_+$ is a nilpotent Lie subalgebra. The subalgebra $\mathfrak{n}_-$ is defined similarly, so that $\mathfrak{g} = \mathfrak{n}_+ \oplus \mathfrak{h} \oplus \mathfrak{n}_-$. Since we have fixed a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, every finite dimensional $\mathfrak{g}$-module admits a weight space decomposition.

$$V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda$$

where $V_\lambda = \{v \in V \mid (H - \lambda(H))^N \cdot v = 0, \ \forall \ N \gg 0, H \in \mathfrak{h}\}$. In fact, since $\mathfrak{g}$ is semi-simple, $V_\lambda = \{v \in V \mid H \cdot v = \lambda(H)v, \ \forall \ H \in \mathfrak{h}\}$. To prove this, we first note that Weyl's complete reducibility Theorem implies that it is enough to check this for V simple. In that case, consider the subspace $\bigoplus_{\lambda \in \mathfrak{h}^*} V'_\lambda$, where $V'_\lambda = \{v \in V \mid H \cdot v = \lambda(H)v, \ \forall \ H \in \mathfrak{h}\} \subset V_\lambda$. Since $[H, X] = \alpha(H)X$ for all $X \in \mathfrak{g}_\alpha$ and $H \in \mathfrak{h}$, the subspace $\bigoplus_{\lambda \in \mathfrak{h}^*} V'_\lambda$ is a $\mathfrak{g}$-submodule. Therefore, the fact that V is simple implies that this subspace is the whole of V.

Definition 11.1. Let V be a finite-dimensional $\mathfrak{g}$-module. A weight vector $v \in V_\lambda$ is said to be *highest weight* if $\mathfrak{n}_+ \cdot v = 0$.

Lemma 11.2. *Let V be a finite dimensional $\mathfrak{g}$-module. Then V contains a highest weight vector.*

Proof. Let $\mathfrak{b}_+ = \mathfrak{h} \bigoplus_{\alpha \in R^+} \mathfrak{g}_\alpha = \mathfrak{h} \oplus \mathfrak{n}_+$. Then, since $[\mathfrak{h}, \mathfrak{n}_+] = \mathfrak{n}_+$ and $\mathfrak{n}_+$ is nilpotent, $\mathfrak{b}_+$ is a solvable subalgebra of $\mathfrak{g}$. Lie's Theorem implies that V contains a common eigenvector v for $\mathfrak{b}_+$ i.e. there exists $\lambda : \mathfrak{b}_+ \to \mathbb{C}$, a linear functional, such that $X \cdot v = \lambda(X)v$ for all $X \in \mathfrak{b}_+$. It suffices to show that $\lambda(X) = 0$ for all $X \in \mathfrak{n}_+$. Let $X \in \mathfrak{g}_\alpha \subset \mathfrak{n}_+$. Then $\alpha \neq 0$ implies that there exists some $H \in \mathfrak{h}$ such that $\alpha(H) \neq 0$. This implies that $\frac{1}{\alpha(H)}[H, X] = X$ and hence $[\mathfrak{h}, \mathfrak{n}_+] = \mathfrak{n}_+$. But then

$$\lambda(X)v = X \cdot v = \frac{1}{\alpha(H)}\lambda(H)\lambda(X)v - \frac{1}{\alpha(H)}\lambda(X)\lambda(H)v = 0.$$

Hence $\lambda(X) = 0$, as required i.e. v is a highest weight vector. $\square$

If $\Delta = \{\beta_1, \ldots, \beta_\ell\}$, then by Proposition 9.8, there exist $\mathfrak{sl}_2$-triples $\{E_i, F_i, H_i\}$ in $\mathfrak{g}$ such that E_i is a basis of $\mathfrak{g}_{\beta_i}$. Since the elements in Δ are a basis of $\mathfrak{h}^*$, the set $\{H_1, \ldots, H_\ell\}$ is a basis of $\mathfrak{h}$. The set of all $\lambda \in \mathfrak{h}^*$ such that $\lambda(H_i) \in \mathbb{Z}$, for all i, is denoted P. It is a $\mathbb{Z}$-lattice in $\mathfrak{h}^*$, and is called the *weight lattice* of $\mathfrak{g}$. Let $P^+ = \{\lambda \in P \mid \lambda(H_i) \in \mathbb{Z}_+ \ \forall\ i\}$; elements of P^+ are called *dominant integral weights*. Notice that

$$P = \{\lambda \in \mathfrak{h}^* \mid \langle \lambda, \beta_i \rangle \in \mathbb{Z},\ \forall\ i\}, \quad P^+ = \{\lambda \in \mathfrak{h}^* \mid \langle \lambda, \beta_i \rangle \in \mathbb{Z}_{\geq 0},\ \forall\ i\}.$$

As an example, consider the simple Lie algebra $\mathfrak{g}(\mathsf{B}_3)$ of type B_3. If E is the three dimensional real vector space with basis $\epsilon_1, \epsilon_2, \epsilon_3$ then $\{\beta_1 = \epsilon_1 - \epsilon_2, \beta_2 = \epsilon_2 - \epsilon_3, \beta_3 = \epsilon_3\}$ is a set of simple roots and $R = \{\pm\epsilon_i, \pm(\epsilon_i \pm \epsilon_j) \mid 1 \leq i \neq j \leq 3\}$ is the set of all roots.

Exercise 11.3. Describe the set of positive roots with respect to $\{\beta_1, \beta_2, \beta_3\}$. What is the dimension of $\mathfrak{g}(\mathsf{B}_3)$?

A weight $\lambda = \lambda_1\epsilon_1 + \lambda_2\epsilon_2 + \lambda_3\epsilon_3$ will be dominant integral if and only if $\langle \lambda, \beta_1 \rangle = \lambda_1 - \lambda_2$, $\langle \lambda, \beta_2 \rangle = \lambda_2 - \lambda_3$, and $\langle \lambda, \beta_3 \rangle = 2\lambda_3$ are all positive integers i.e. $\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, 2\lambda_3 \in \mathbb{Z}_{\geq 0}$. The Weyl group W of type B_3 is generated by the three reflections

$$s_{\beta_1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad s_{\beta_2} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad s_{\beta_3} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It has $3!2^3 = 48$ elements. The $3!$ comes from the fact that s_{β_1} and s_{β_2} generate a subgroup isomorphic to $\mathfrak{S}_3$.

In the example of $R = \{\epsilon_i - \epsilon_j \mid 1 \leq i \neq j \leq n+1\}$, the root system of type A_n, an element $\lambda_1(\epsilon_1 - \epsilon_2) + \cdots + \lambda_n(\epsilon_n - \epsilon_{n+1})$ belongs to P^+ if and only if $2\lambda_i - \lambda_{i-1} - \lambda_{i+1} \in \mathbb{Z}_{\geq 0}$ for all i.

Lemma 11.4. *Let V be a finite dimensional $\mathfrak{g}$-module.*

1. *If $\mu \in \mathfrak{h}^*$ such that $V_\mu \neq 0$, then $\mu \in P$.*

2. *If $v \in V_\lambda$ is a highest weight vector, then $\lambda \in P^+$.*

Exercise 11.5. Prove Lemma 11.4. Hint: first check that the lemma holds for $\mathfrak{sl}_2$-modules. Next, note that it suffices to check for each $i = 1, \ldots, \ell$ that $V_\mu \neq 0$ implies $\mu(H_i) \in \mathbb{Z}$ and for (2) that $\lambda(H_i) \geq 0$. Deduce this from the $\mathfrak{sl}_2$ case.

A $\mathfrak{g}$-module M is called *highest weight* if it is generated by some highest weight vector $m \in M_\lambda$.

11.2 Verma modules

Next we will define certain "universal" highest weight modules. These are called *Verma modules.* Let $Y_1, \ldots, Y_N$ be a weight basis of $\mathfrak{n}_-$. Since $\mathfrak{n}_-$ is nilpotent, we may order the Y_i so that $[Y_i, Y_j] \in \mathbb{C}\{Y_k \mid k > i\}$ for all $i \leq j$. Choose some $\lambda \in \mathfrak{h}^*$ and let $\mathbb{C}\{v_\lambda\}$ be the one-dimensional $\mathfrak{b}_+$-module such that $\mathfrak{n}_+ \cdot v_\lambda = 0$ and $H \cdot v_\lambda = \lambda(H) v_\lambda$ for all $H \in \mathfrak{h}$. Let $\Delta(\lambda)$ be the (infinite dimensional!) vector space with basis given by $Y_{i_1} Y_{i_2} \cdots Y_{i_k} v_\lambda$, where $(i_1 \leq i_2 \leq \cdots \leq i_k)$ is a k-tuple of element from $\{1, \ldots, N\}$. We say that the length of the basis element $Y_{i_1} Y_{i_2} \cdots Y_{i_k} v_\lambda$ is k.

Lemma 11.6. *The space $\Delta(\lambda)$ has a natural $\mathfrak{g}$-module structure.*

Proof. The proof is by induction on the length of the basis element. The only element of length zero is v_λ. We define $\mathfrak{n}_+ \cdot v_\lambda = 0$, $H \cdot v_\lambda = \lambda(H) v_\lambda$ and $Y_i \cdot v_\lambda = Y_i v_\lambda$. Now we assume that we've defined the action of $\mathfrak{g}$ on all basis elements of length less than k. First we consider $\Delta(\lambda)$ as a module over $\mathfrak{n}_-$ i.e. over the Y_i. We'd like to define this in the most stupid way: $Y_i \cdot (Y_{i_1} Y_{i_2} \cdots Y_{i_k} v_\lambda) = Y_i Y_{i_1} Y_{i_2} \cdots Y_{i_k} v_\lambda$. But if $i > i_1$, this won't be a basis element because the subscripts $i, i_1, \ldots$ have to be ordered. So we set

$$Y_i \cdot (Y_{i_1} Y_{i_2} \cdots Y_{i_k} v_\lambda) = \begin{cases} Y_i Y_{i_1} Y_{i_2} \cdots Y_{i_k} v_\lambda & \text{if } i \leq i_1 \\ Y_{i_1} \cdot (Y_i \cdot (Y_{i_2} \cdots Y_{i_k} v_\lambda)) + [Y_i, Y_{i_1}] \cdot (Y_{i_2} \cdots Y_{i_k} v_\lambda) & \text{if } i > i_1. \end{cases}$$

Using the fact that we have ordered the Y_i so that $[Y_i, Y_j] \in \mathbb{C}\{Y_k \mid k > i\}$ for all $i \leq j$, show that the above rule makes $\Delta(\lambda)$ into a $\mathfrak{n}_-$-module. The action of $\mathfrak{b}_+$ is easier to define:

$$X \cdot (Y_{i_1} Y_{i_2} \cdots Y_{i_k} v_\lambda) = Y_{i_1} \cdot (X \cdot (Y_{i_2} \cdots Y_{i_k} v_\lambda)) + [X, Y_1] \cdot (Y_{i_2} \cdots Y_{i_k} v_\lambda).$$

It is a direct check to see that the above rules do indeed make $\Delta(\lambda)$ into a $\mathfrak{g}$-module. This check is not so obvious since one really needs to check that all the relations of Theorem 10.2 hold. $\square$

Each of the basis elements $Y_{i_1} Y_{i_2} \cdots Y_{i_k} v_\lambda$ is a weight vector with weight $-\alpha_{i_1} - \cdots - \alpha_{i_k}$, where $Y_i \in \mathfrak{g}_{-\alpha_i}$ and $\alpha_i \in R^+$. Here is the key fact about Verma modules that we need.

Lemma 11.7. *Let $\lambda \in \mathfrak{h}^*$. The Verma module $\Delta(\lambda)$ has a unique simple quotient $V(\lambda)$ i.e. if $\Delta(\lambda) \twoheadrightarrow V$ and $\Delta(\lambda) \twoheadrightarrow V'$ for some simple $\mathfrak{g}$-modules V and V' then $V \simeq V'$.*

Proof. Let's call the maps $\Delta(\lambda) \twoheadrightarrow V$ and $\Delta(\lambda) \twoheadrightarrow V'$, ϕ and ϕ' respectively. First, we claim that $\phi(v_\lambda)$ is a highest weight vector in V, that generates V as a $\mathfrak{g}$-module, and similarly for $\phi'(v_\lambda) \in V'$.

Since ϕ is surjective, V is spanned by $\phi(Y_{i_1}Y_{i_2}\cdots Y_{i_k}v_\lambda)$. But notice that our rule for the $\mathfrak{g}$-action on $\Delta(\lambda)$ means that $Y_{i_1}Y_{i_2}\cdots Y_{i_k}v_\lambda = Y_{i_1}\cdot(Y_{i_2}\cdot(\cdots(Y_{i_k}\cdot v_\lambda)\cdots)$ and hence

$$\phi(Y_{i_1}Y_{i_2}\cdots Y_{i_k}v_\lambda) = Y_{i_1}\cdot(Y_{i_2}\cdot(\cdots(Y_{i_k}\cdot\phi(v_\lambda)\cdots).$$

Thus, $\phi(v_\lambda)$ generates V. If $X \in \mathfrak{b}_+$ then $X\cdot\phi(v_\lambda) = \phi(X\cdot v_\lambda)$, so it is clear that $\phi(v_\lambda)$ is also a highest weight vector.

Let $M = \sum_{M'\subset\Delta(\lambda)} M'$ be the sum of all $\mathfrak{g}$-submodules of $\Delta(\lambda)$ such that $M'\cap\Delta(\lambda)_\lambda = 0$ (equivalently, the λ weight space M'_λ of M' is zero). Then M is a $\mathfrak{g}$-submodule of $\Delta(\lambda)$ such that $M_\lambda = 0$ i.e. it is a *proper* submodule. To show that V and V' are isomorphic, it suffices to show that $\operatorname{Ker}\phi = \operatorname{Ker}\phi' = M$. Since $\phi(v_\lambda)\neq 0$ and the λ-weight space of $\Delta(\lambda)$ is spanned by v_λ, ϕ restricts to an isomorphism $\Delta(\lambda)_\lambda \xrightarrow{\sim} V_\lambda$. Thus, $\operatorname{Ker}\phi\cap\Delta(\lambda)_\lambda = 0$ and hence $\operatorname{Ker}\phi\subset M$. This means that the *non-zero* $\mathfrak{g}$-module $\Delta(\lambda)/M$ is a quotient of $V = \Delta(\lambda)/\operatorname{Ker}\phi$. But V is simple. Hence $\Delta(\lambda)/M = V$ and $M = \operatorname{Ker}\phi$. The same argument applies to ϕ'. □

Exercise 11.8. Let V be a finite dimensional $\mathfrak{g}$-module with highest weight vector v of weight $\lambda\in P^+$. Show that $v_\lambda\mapsto v$ extends uniquely to a $\mathfrak{g}$-module homomorphism $\Delta(\lambda)\to V$. This explains why $\Delta(\lambda)$ is the *universal* highest weight module with weight λ. Hint: read carefully through the proof of Lemma 11.7.

11.3 The classification

Notice that we have already shown that if $\lambda\notin P^+$ then there can be no finite dimensional $\mathfrak{g}$-module with highest weight of weight λ. Thus, the simple quotient $V(\lambda)$ of $\Delta(\lambda)$ is infinite dimensional in this case.

Theorem 11.9. *Let V be a simple, finite dimensional $\mathfrak{g}$-module.*

1. *There is some $\lambda\in P^+$ such that $V\simeq V(\lambda)$.*

2. *If $\lambda,\mu\in\mathfrak{h}^*$, $\lambda\neq\mu$ then $V(\lambda)\not\simeq V(\mu)$.*

Proof. By Lemma 11.2, there exists at least one highest weight vector, v say, in V. If λ is the weight of v then we have shown that $\lambda\in P^+$. By exercise 11.8, there exists a non-zero homomorphism $\Delta(\lambda)\to V$, $v_\lambda\mapsto v$. But then Lemma 11.7 says that $V\simeq V(\lambda)$ as required.

Now take $\lambda,\mu\in\mathfrak{h}^*$, $\lambda\neq\mu$. Since isomorphisms will map highest weight vectors to highest weight vectors, it suffices to show that $V(\lambda)$ does not contain a highest weight vector of weight μ. We know that $V(\lambda)$ is a quotient of $\Delta(\lambda)$ and all non-zero weight spaces in $\Delta(\lambda)$ have weight

$\lambda - \alpha_{i_1} - \cdots - \alpha_{i_k}$ for some $\alpha_i \in R^+$. Since $V(\lambda)$ is a quotient of $\Delta(\lambda)$, the same applies to it. There are two cases to consider: the first is where there exist $i_1, \dots, i_k$ such that $\lambda - \mu = \alpha_{i_1} + \cdots + \alpha_{i_k}$ and the second, when we can't find any such α_i. In the second case, there can't be any highest weight in $V(\lambda)$ of weight μ so there's nothing to show.

So we assume $\lambda - \mu = \alpha_{i_1} + \cdots + \alpha_{i_k}$ and there is some $v \in V(\lambda)$ of weight μ. Then, by exercise 11.8, there is a non-zero map $\Delta(\mu) \to V(\lambda)$, $v_\mu \mapsto v$. It is surjective. Hence the weights of $V(\lambda)$ must also be of the form $\mu - \alpha_{j_1} - \cdots - \alpha_{j_l}$ for some j's. In particular $\lambda = \mu - \alpha_{j_1} - \cdots - \alpha_{j_l}$ and hence

$$\alpha_{i_1} + \cdots + \alpha_{i_k} = -\alpha_{j_1} - \cdots - \alpha_{j_l}. \tag{20}$$

But recall that we have fixed simple roots $\beta_1, \dots, \beta_\ell$. This means that every $\alpha \in R^+$ can be written as $\alpha = \sum_{i=1}^{\ell} n_i \beta_i$ for some positive integers n_i. Moreover, $\{\beta_1, \dots, \beta_\ell\}$ are a basis of $\mathfrak{h}^*$ so such an expression is unique. This means that $\alpha_{i_1} + \cdots + \alpha_{i_k}$ can also be uniquely written as $\sum_{i=1}^{\ell} n_i \beta_i$ for some positive integers n_i. But it also means that $-\alpha_{j_1} - \cdots - \alpha_{j_l}$ can be uniquely written as $\sum_{i=1}^{\ell} -m_i \beta_i$ for some positive integers m_i. Thus, equation 20 implies that $\lambda = \mu$, contradicting our initial assumptions.

□

In fact, it is not so difficult to show (see e.g. [8, Theorem 22.2]) that $V(\lambda)$ is finite dimensional for all $\lambda \in P^+$. Thus, $\lambda \mapsto V(\lambda)$ defines a bijection

$$P^+ \xrightarrow{\sim} \{ \text{ Isomorphism classes of simple, f.d. } \mathfrak{g}\text{-modules.} \}.$$

11.4 Weyl's formula

We end with Weyl's beautiful character formula for the character of a simple, finite dimensional $\mathfrak{g}$-module. First we should explain what we mean by the character $\mathrm{Ch}(V)$ of a finite dimensional $\mathfrak{g}$-module. We know that V decomposes as the direct sum of its weight spaces. Just as the dimension of a vector space is a good invariant of that space (in fact it uniquely defines it up to isomorphism), the character of a $\mathfrak{g}$-module, which records the dimension of all the weight spaces of the module is a good invariant of the module (again, for $\mathfrak{g}$ semi-simple and V finite dimensional, the character $\mathrm{Ch}(V)$ uniquely defines V up to isomorphism).

We consider formal linear combinations $\sum_{\lambda \in P} m_\lambda e^\lambda$, where $m_\lambda \in \mathbb{Z}$. One can add these expressions in the obvious way. But it is also possible to multiply then: first we define $e^\lambda \cdot e^\mu = e^{\lambda + \mu}$, then we extend this rule by linearity to all expressions. Now given V, define

$$\mathrm{Ch}(V) = \sum_{\lambda \in P} (\dim V_\lambda) e^\lambda.$$

Define $\rho = \frac{1}{2}\sum_{\alpha\in R^+} \alpha$; ρ is called the half sum of positive roots. In the example of B_3, considered in the previous section, $\rho = \frac{5}{2}\epsilon_1 + \frac{3}{2}\epsilon_2 + \frac{1}{2}\epsilon_3$. Finally, we need one extra ingredient in order to give Weyl's formula. Recall that we can define the sign function on permutations in the symmetric group; a permutation is even if and only if it can be written as a product of an even number of transpositions. This generalizes to an arbitrary Weyl group. Fix a set of simple roots Δ and recall from Theorem 9.23 that W is generated by the set $\{s_\alpha \mid \alpha \in \Delta\}$. We define the length $\ell(\sigma)$ of $\sigma \in W$ to be the length of σ, written as a minimal product of elements from $\{s_\alpha \mid \alpha \in \Delta\}$. Then $\mathrm{sgn}(\sigma) := (-1)^{\ell(\sigma)}$, generalizing the sign function on symmetric groups.

Theorem 11.10 (Weyl)**.** *Let $V(\lambda)$ be the simple $\mathfrak{g}$-module with highest weight $\lambda \in P^+$. Then*

$$\mathrm{Ch}\ V(\lambda) = \frac{\sum_{w\in W} \mathrm{sgn}(w) e^{w(\lambda+\rho)}}{\sum_{w\in W} \mathrm{sgn}(w) e^{w(\rho)}}.$$

To illustrate how to use Weyl's character formula, we'll consider the simple Lie algebra $\mathfrak{sl}(3,\mathbb{C})$ of type A_2. As per usual, it is easier when dealing with root systems of type A to consider the simple $\mathfrak{sl}(n,\mathbb{C})$-module V as a simple $\mathfrak{gl}(n,\mathbb{C})$-module. This can be done by making the central element $\mathrm{Id} \in \mathfrak{gl}(n,\mathbb{C})$ act by a fixed scalar (which we are free to choose).

We have $\rho = \epsilon_1 - \epsilon_3$ and

$$\sum_{w\in\mathfrak{S}_3} \mathrm{sgn}(w) e^{w(\rho)} = x_1x_3^{-1} - x_2x_3^{-1} - x_1^{-1}x_3 - x_1x_2^{-1} + x_1^{-1}x_2 + x_2^{-1}x_3.$$

Let's consider the simple $\mathfrak{gl}(3,\mathbb{C})$-module with highest weight $(2,1,0)$. I claim that the character of $V((2,1,0))$ equals $x_1^2x_2 + x_1^2x_3 + x_2^2x_1 + x_2^2x_3 + x_3^2x_1 + x_3^2x_2 + 2x_1x_2x_3$. To show this, we note that

$$\sum_{w\in\mathfrak{S}_3} \mathrm{sgn}(w) e^{w(3,1,-1)} = x_1^3x_2x_3^{-1} - x_2^3x_1x_3^{-1} - x_3^3x_2x_1^{-1} - x_1^3x_3x_2^{-1} + x_2^3x_3x_1^{-1} + x_3^3x_1x_2^{-1}.$$

Then, one checks, by explicitly multiplying, that

$$\left(\sum_{w\in\mathfrak{S}_3} \mathrm{sgn}(w) e^{w(\rho)}\right) \mathrm{Ch}\ V((2,1,0)) = \sum_{w\in\mathfrak{S}_3} \mathrm{sgn}(w) e^{w(3,1,-1)}.$$

Exercise 11.11. Using Weyl's formula, show that Ch $V((1,1,1)) = x_1x_2x_3$, where $V((1,1,1))$ is the irreducible $\mathfrak{gl}(3,\mathbb{C})$-module with highest weight $(1,1,1)$.

Exercise 11.12. Let E be the two dimensional real vector space with basis ϵ_1, ϵ_2 then $\{\beta_1 = \epsilon_1 - \epsilon_2, \beta_2 = \epsilon_2\}$ is a set of simple roots and $R = \{\pm\epsilon_i, \pm(\epsilon_i \pm \epsilon_j) \mid 1 \le i \ne j \le 2\}$ is the set of all roots for B_2.

(a) Show that a weight $\lambda = \lambda_1\epsilon_1 + \lambda_2\epsilon_2$ will be dominant integral if and only if $\lambda_2 \in \frac{1}{2}\mathbb{Z}_{\ge 0}$ and $\lambda_1 - \lambda_2 \in \mathbb{Z}_{\ge 0}$. Also show that $\rho = \frac{3}{2}\epsilon_1 + \frac{1}{2}\epsilon_2$.

(b) The Weyl group W of type B_2 is generated by the two reflections

$$s_{\beta_1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad s_{\beta_2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Show that W has eight elements - list them.

(c) Finally, calculate $\mathrm{Ch}(V(1,0))$, $\mathrm{Ch}(V(1,1))$ and $\mathrm{Ch}(V(3,1))$. What is the dimension of these modules?

Exercise 11.13. The root system for $\mathfrak{sl}_2$ (i.e. type A_1) is $\{\pm 2\epsilon_1\}$ and if $\Delta = \{2\epsilon_1\}$, then $P^+ = n\epsilon_1$, where $n \in \mathbb{Z}_{\ge 0}$. What is the Weyl group? For all $n \ge 0$, calculate Ch $V(n)$. How does this compare to Theorem 6.15?

12 Appendix: Quotient vector spaces

12.1 The definition

Let $\mathbb{F}$ be a field and V a vector space. Let W be a subspace of V. Then, we form the *quotient space*

$$V/W = \{[v] \mid v \in V, \text{ and } [v] = [v'] \text{ iff } v - v' \in W. \}.$$

Equivalently, since V is an abelian group and W a subgroup V/W is the set of W-cosets in V.

Lemma 12.1. *The quotient V/W is a vector space.*

Proof. It is already clear that it is an abelian group. So we just need to define $\alpha[v]$ for $\alpha \in \mathbb{F}$ and $[v] \in V/W$, and check that this makes V/W into a vector space.

We define $\alpha[v] = [\alpha v]$. To see that this is well-define, let $[v] = [v']$. Then $v - v' \in W$ and hence $\alpha(v - v') = \alpha v - \alpha v' \in W$ and thus $[\alpha v] = [\alpha v']$. Hence it is well-defined.

To make sure that this makes V/W into a vector space, we need to check that $\alpha([v] + [v']) = \alpha[v] + \alpha[v']$. But

$$\alpha([v] + [v']) = \alpha([v + v']) = [\alpha(v + v')] = [\alpha v + \alpha v'] = \alpha[v] + \alpha[v'].$$

$\square$

One way to think about the quotient space V/W (though I would argue, the wrong way!) is in terms of complimentary subspaces. Recall that a compliment to W is a subspace W' of V such that

- $W' + W = V$
- $W' \cap W = \{0\}$.

We normally write this as $V = W \oplus W'$.

There is a canonical map $V \to V/W$, $v \mapsto [v]$. It is a linear map.

Lemma 12.2. *Let W' be a compliment to W in V. Then the quotient map defines an isomorphism $\phi : W' \to V/W$.*

Proof. To be precise, ϕ is the composite $W' \hookrightarrow V \to V/W$. Being the composite of two linear maps, it is linear. To show that it is surjective, take $[v] \in V/W$. Since $W' + W = V$, we can find

$w' \in W'$ and $w \in W$ such that $v = w + w'$. This implies that $v - w' = w \in W$ and hence $[v] = [w']$ in V/W i.e. $\phi(w') = [v]$. So ϕ is surjective.

Now assume that $w' \in \operatorname{Ker}\phi$. This means that $[w'] = [0]$ i.e. $w' \in W$. But then $w' \in W' \cap W = \{0\}$. Hence $w' = 0$, which means that $\operatorname{Ker}\phi = \{0\}$. So ϕ is injective. $\square$

Notice that, if V is finite dimensional, then

$$\dim V = \dim(W + W') = \dim W + \dim W' - \dim W \cap W' = \dim W + \dim W'.$$

But $\phi : W' \xrightarrow{\sim} V/W$. Hence $\dim W' = \dim V/W$. Thus, we see that

$$\dim V/W = \dim V - \dim W.$$

12.2 Basis of V/W

Assume now that V is finite dimensional.

Lemma 12.3. *Let $v_1, \dots, v_k \in V$ such that $[v_1], \dots, [v_k]$ is a basis of V/W and let $w_1, \dots, w_\ell$ be a basis of W. Then $v_1, \dots, v_k, w_1, \dots, w_\ell$ is a basis of V.*

Proof. We need to show that $v_1, \dots, v_k, w_1, \dots, w_\ell$ span V and are linearly independent.

First, we show spanning. If $v \in V$, then the fact that $[v_1], \dots, [v_k]$ is a basis of V/W implies that there exists $\alpha_1, \dots, \alpha_k \in \mathbb{F}$ such that $[v] = \alpha_1[v_1] + \cdots + \alpha_k[v_k]$. This means that $[v - \alpha_1 v_1 - \cdots - \alpha_k v_k] = [0]$ i.e. $v - \alpha_1 v_1 - \cdots - \alpha_k v_k \in W$. Since $w_1, \dots, w_\ell$ are a basis of W, we can find $\beta_1, \dots, \beta_\ell \in \mathbb{F}$ such that

$$v - \alpha_1 v_1 - \cdots - \alpha_k v_k = \beta_1 w_1 + \cdots \beta_\ell w_\ell$$

i.e. $v = \alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 w_1 + \cdots \beta_\ell w_\ell$. So $v_1, \dots, v_k, w_1, \dots, w_\ell$ span V.

Next, assume that $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_\ell \in F$ such that $\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 w_1 + \cdots \beta_\ell w_\ell = 0$. This means that

$$[\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 w_1 + \cdots \beta_\ell w_\ell] = \alpha_1[v_1] + \cdots + \alpha_k[v_k] = [0]$$

in V/W. But $[v_1], \dots, [v_k]$ are a basis of V/W. Hence $\alpha_1, \dots, \alpha_k = 0$. Thus, $\beta_1 w_1 + \cdots \beta_\ell w_\ell = 0$. But $w_1, \dots, w_\ell$ is a basis of W. Hence $\beta_1, \dots, \beta_\ell = 0$ too. $\square$

Notice that we didn't really have to prove linear independence in the above proof. If we know that $v_1, \dots, v_k, w_1, \dots, w_\ell$ span, then the fact that $\dim V = \dim V/W + \dim W$ implies that $\dim V = k + l$ and hence any spanning set of V with $k + l$ elements must be a basis.

Lemma 12.4. *If $v_1, \dots, v_k \in V$ such that $[v_1], \dots, [v_k]$ is a basis of V/W, then $W' = \mathbb{C}\{v_1, \dots, v_k\}$ is a compliment to W in V.*

Proof. Since $[v_1], \dots, [v_k]$ is a basis of V/W, the map $\phi : W' \hookrightarrow V \to V/W$, $v_i \mapsto [v_i]$ is surjective. But $\dim W' = \dim V/W = k$. Hence ϕ is an isomorphism. The kernel of ϕ equals $W \cap W'$ hence, since ϕ is an isomorphism, we have $W \cap W' = \{0\}$.

If $v \in V$, then we can find $w' \in W'$ such that $\phi(w') = [v]$ i.e. $v - w' \in W$. So there must be $w \in W$ such that $v - w' = w$. Hence $v = w' + w$. So we have shown that $V = W + W'$. □

12.3 Endomorphisms of V/W

Some times linear maps $V \to V$ also define for us linear maps $V/W \to V/W$. Let $X : V \to V$ be a linear map.

Lemma 12.5. *If $X(W) \subseteq W$, then X defines a linear map $\overline{X} : V/W \to V/W$, by $\overline{X}([v]) = [X(v)]$.*

Proof. First we show that $\overline{X}$ is well-defined and then check that it is linear. So we need to check that if $[v] = [v']$ then $\overline{X}([v]) = \overline{X}([v'])$. But $[v] = [v']$ iff $v - v' \in W$. Since $X(W) \subseteq W$, we have $X(v - v') = X(v) - X(v') \in W$ and hence

$$\overline{X}([v]) = [X(v)] = [X(v')] = \overline{X}([v']),$$

as required.

So now we just check that it is linear. This is easy:

$$\begin{aligned}
\overline{X}(\alpha[v] + \beta[v']) &= \overline{X}([\alpha v + \beta v']) \\
&= [X(\alpha v + \beta v')] \\
&= [\alpha X(v) + \beta X(v')] \\
&= \alpha[X(v)] + \beta[X(v')] \\
&= \alpha\overline{X}([v]) + \beta\overline{X}([v']).
\end{aligned}$$

□

We'll do an example just to get a feel for things. Let's take $V = \mathbb{C}^4$ and

$$W = \left\{ \begin{pmatrix} \star \\ \star \\ 0 \\ 0 \end{pmatrix} \right\}, \quad W' = \left\{ \begin{pmatrix} 0 \\ 0 \\ \star \\ \star \end{pmatrix} \right\},$$

where W' is a compliment to W in V. Since $v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ are a basis of W', $[v_1]$ and $[v_2]$ are a basis of V/W. Let

$$X = \begin{pmatrix} 4 & 2 & -1 & 6 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Then, one can check that $X(W) \subset W$. So now

$$X \begin{pmatrix} 0 \\ 0 \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} -\alpha + 6\beta \\ 6\beta \\ 2\alpha - 3\beta \\ \beta \end{pmatrix} = \begin{pmatrix} -\alpha + 6\beta \\ 6\beta \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2\alpha - 3\beta \\ \beta \end{pmatrix}.$$

Thus, $\overline{X}([v_1]) = 2[v_1]$ and $\overline{X}([v_2]) = -3[v_1] + [v_2]$. This implies that

$$\overline{X} = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}.$$

12.4 Bilinear pairings

Let V be a $\mathbb{F}$-vector space. A bilinear pairing is a map

$$(-,-) : V \times V \to \mathbb{F}$$

such that

1. $(u+v,w) = (u,w) + (v,w)$ and $(u, v+w) = (u,v) + (u,w)$ for all $u,v,w \in V$.
2. $(\alpha u, \beta v) = \alpha\beta(u,v)$ for all $u, v \in V$ and $\alpha, \beta \in \mathbb{F}$.

The bilinear form is said to be *symmetric* if $(u,v) = (v,u)$ for all $u, v \in V$. We do not assume $(-,-)$ is symmetric in general. The *radical* of $(-,-)$ is defined to be

$$\text{rad } (-,-) = \{u \in V \mid (u,v) = 0 \ \forall \ v \in V\}.$$

We say that the bilinear form $(-,-)$ is *non-degenerate* if $\text{rad } (-,-) = \{0\}$.

The *dual* of a vector space V is defined to be the space V^* of all linear maps $\lambda : V \to \mathbb{F}$.

Lemma 12.6. *If V is finite dimensional, then V^* is finite dimensional and* $\dim V = \dim V^*$.

Proof. Let $n = \dim V$. Fixing a basis of V, every element of V^* can be uniquely written as a $1 \times n$-matrix i.e. a row vector. It is clear that the space of row vectors of length n is n-dimensional. □

The following lemma shows the relationship between dual vector spaces and bilinear forms.

Lemma 12.7. *If V is equipped with a non-degenerate bilinear form $(-,-)$, then the form defines a canonical isomorphism $\phi : V \to V^*$ given by*

$$\phi(u)(v) := (u,v).$$

Proof. For fixed u, the fact that $(-,-)$ is bilinear implies that $\phi(u)$ is a linear map i.e. $\phi(u) \in V^*$. Moreover, bilinearity also implies that $\phi : V \to V^*$ is a linear map. The kernel of ϕ is clearly $\text{rad } (-,-)$. Therefore, since we have assumed that $(-,-)$ is non-generate, ϕ is injective. On the other hand, Lemma 12.6 says that $\dim V = \dim V^*$. Therefore ϕ must be an isomorphism. □

If W is a subspace of V then the *perpendicular* of W with respect to V is defined to be $W^\perp := \{u \in V \mid (u,w) = 0 \ \forall \ w \in W\}$.

Lemma 12.8. *Let V be a finite dimensional vector space, $(-,-)$ a non-degenerate bilinear form on V and W a subspace of V. Then we have a canonical isomorphism*

$$W^\perp \xrightarrow{\sim} (V/W)^*$$

and $\dim W^\perp + \dim W = \dim V$.

Proof. We define $\psi : W^\perp \to (V/W)^*$ by $\psi(u)([v]) = (u, v)$. Let us check that this is well-defined i.e. $\psi(u)([v]) = \psi(u)([v'])$ if $[v] = [v']$. We have $[v] = [v']$ if and only if $v - v' \in W$. But then $(u, v - v') = 0$ for all $u \in W^\perp$. Hence $(u, v) = (u, v')$ and $\psi(u)([v]) = \psi(u)([v'])$.

Next we show that $(V/W)^*$ can be identified with the subspace U of V^* consisting of all λ such that $W \subset \operatorname{Ker} \lambda$ i.e. $\lambda(W) = 0$. If $\lambda \in U$ then we can define $\lambda' \in (V/W)^*$ by $\lambda'([v]) = \lambda(v)$. Just as in the previous paragraph, it is easy to check that this is well-defined. Now let $v_1, \dots, v_\ell, w_1, \dots, w_k$ be the basis of V defined in Lemma 12.3. Given $\nu \in (V/W)^*$, we define $\lambda(v_i) = \nu([v_i])$ and $\lambda(w_j) = 0$ for all i, j. Then λ extends uniquely by linearity to an element of V^*. Since $\lambda(w_j) = 0$ for all j, λ belongs to U. By construction, $\lambda' = \nu$. Hence $U \xrightarrow{\sim} (V/W)^*$.

Finally, we note that, if ϕ is the isomorphism of Lemma 12.7, then

$$\phi^{-1}(U) = \{u \in V \mid \phi(u)(W) = 0\} = \{u \in V \mid (u, w) = 0 \ \forall\ w \in W\} = W^\perp.$$

Since $\phi(u)' = \psi(u)$ for all $u \in \phi^{-1}(U)$, ψ is an isomorphism.

The dimension formula follows from the fact that $\dim V = \dim V/W + \dim W$ and $\dim V = \dim V^*$ by Lemma 12.6. $\square$

References

[1] A. Baker. *Matrix groups*. Springer Undergraduate Mathematics Series. Springer-Verlag London Ltd., London, 2002. An introduction to Lie group theory.

[2] R. W. Carter. *Lie algebras of finite and affine type*, volume 96 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2005.

[3] A. C. de Silva. Lecture on symplectic geometry. `http://www.math.ist.utl.pt/~acannas/Books/lsg.pdf`, 2006.

[4] K. Erdmann and M. J. Wildon. *Introduction to Lie algebras*. Springer Undergraduate Mathematics Series. Springer-Verlag London Ltd., London, 2006.

[5] W. Fulton and J. Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[6] V. Guillemin and S. Sternberg. *Symplectic techniques in physics*. Cambridge University Press, Cambridge, second edition, 1990.

[7] B. C. Hall. *Lie groups, Lie algebras, and representations*, volume 222 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2003. An elementary introduction.

[8] J. E. Humphreys. *Introduction to Lie Algebras and Representation Theory*. Springer-Verlag, New York, 1972. Graduate Texts in Mathematics, Vol. 9.

[9] A. Kirillov, Jr. *An introduction to Lie groups and Lie algebras*, volume 113 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008.

[10] D. Mond. Lecture notes for MA455 manifolds. `http://homepages.warwick.ac.uk/~masbm/manifolds.html`, 2008.