Introduction to Lie Groups and Lie Algebras

Alexander Kirillov, Jr.

DEPARTMENT OF MATHEMATICS, SUNY AT STONY BROOK, STONY BROOK, NY 11794, USA

E-mail address: kirillov@math.sunysb.edu

URL: http://www.math.sunysb.edu/~kirillov

Contents

Contents

Chapter 1

Introduction

What are Lie groups and why do we want to study them? To illustrate this, let us start by considering this baby example.

Example 1.1. Suppose you have n numbers $a_1, \ldots, a_n$ arranged on a circle. You have a transformation A which replaces a_1 with $\frac{a_n + a_2}{2}$, a_2 with $\frac{a_1 + a_3}{2}$, and so on. If you do this sufficiently many times, will the numbers be roughly equal?

To answer this we need to look at the eigenvalues of A. However, explicitly computing the characteristic polynomial and finding the roots is a rather difficult problem. But we can note that the problem has rotational symmetry: if we denote by B the operator of rotating the circle by $2\pi/n$, i.e. sending $(a_1, \ldots, a_n)$ to $(a_2, a_3, \ldots, a_n, a_1)$, then $BAB^{-1} = A$. Since the operator B generates the group of cyclic permutation $\mathbb{Z}_n$, we will refer to this as the $\mathbb{Z}_n$ symmetry. So we might try to make use of this symmetry to find eigenvectors and eigenvalues of A.

Naive idea would be to just consider B-invariant eigenvectors. It is easy to see that there is only one such vector, and while it is indeed an eigenvector for A, this is not enough to diagonalize A. Instead, we can use the following well-known result from linear algebra: if linear operators A, B in a vector space V commute, and V_λ is an eigenspace for B, then $AV_\lambda \subset V_\lambda$. Thus, if the operator B is diagonalizable so that $V = \bigoplus V_\lambda$, then A preserves this decomposition and thus the problem reduces to diagonalizing A on each of V_λ separately, which is a much easier problem.

In this case, since $B^n = \text{id}$, its eigenvalues must be n-th roots of unity. Denoting by $\varepsilon = e^{2\pi i/n}$ the primitive n-th root of unity, it is easy to check that indeed, each n-th root of unity $\lambda = \varepsilon^k, k = 0 \ldots, n-1$, is an eigenvalue of B, with eigenvector $v_k = (1, \epsilon^k, \epsilon^{2k}, \ldots, \epsilon^{(n-1)k})$. Thus, in this case each eigenspace for B is one-dimensional, so each v_k must also be an eigenvector for A. Indeed, one immediately sees that $Av_k = \frac{\varepsilon^k + \varepsilon^{-k}}{2} v_k$.

This is the baby version of a real life problem. Consider $S^2 \subset \mathbb{R}^3$. Define the Laplace operator $\Delta_{sph} : C^\infty(S^2) \to C^\infty(S^2)$ by $\Delta_{sph} f = (\Delta \tilde{f})|_{S^2}$, where $\tilde{f}$ is the result of extending f to $\mathbb{R}^3 - \{0\}$ (constant along each ray), and Δ is the usual Laplace operator in $\mathbb{R}^3$. It is easy to see that Δ_{sph} is a second order differential operator on the sphere; one can write explicit formulas for it in the spherical coordinates, but they are not particularly nice.

For many applications, it is important to know the eigenvalues and eigenfunctions of Δ_{sph}. In particular, this problem arises in qunatum mechanics: the eigenvalues are related to the energy levels of a hydrogen atom in quantum mechanical description. Unfortunately, trying to find the eigenfunctions by brute force gives a second-order differential equation which is very difficult to solve.

But much as in the baby example, it is easy to notice that this problem also has some symmetry — namely, the group $SO(3, \mathbb{R})$ acting on the sphere by rotations. However, trying to repeat the approach used in the baby example (which had $\mathbb{Z}_n$-symmetry) immediately runs into the following two problems:

- $SO(3, \mathbb{R})$ is not a finitely generated group, so we can not use just one (or finitely many) operators B_i and consider their common eigenspaces.

- $SO(3, \mathbb{R})$ is not commutative, so different operators from $SO(3, \mathbb{R})$ can not be diagonalized simultaneously.

The goal of the theory of Lie groups is to give tools to deal with these (and similar) problems. In short, the answer to the first problem is that $SO(3, \mathbb{R})$ is in a certain sense finitely generated — namely, it is generated by three generators, "infinitesimal rotations" around x, y, z axes (see details in Example 3.10).

The answer to the second problem is that instead of decomposing the $C^\infty(S^2)$ into a direct sum of common eigenspaces for operators $B \in SO(3, \mathbb{R})$, we need to decompose it into "irreducible representations" of $SO(3, \mathbb{R})$. In order to do this, we need to develop the theory of representations of $SO(3, \mathbb{R})$. We will do this and complete the analysis of this example in Chapter 5.

Chapter 2

Lie Groups: Basic Definitions

2.1. Lie groups, subgroups, and cosets

Definition 2.1. A Lie group is a set G with two structures: G is a group and G is a (smooth, real) manifold. These structures agree in the following sense: multiplication and inversion are smooth maps.

A morphism of Lie groups is a smooth map which also preserves the group operation: $f(gh) = f(g)f(h), f(1) = 1$.

In a similar way, one defines complex Lie groups. However, unless specified otherwise, "Lie group" means a real Lie group.

Remark 2.2. The word "smooth" in the definition above can be understood in different ways: C^1, C^∞, analytic. It turns out that all of them are equivalent: every C^0 Lie group has a unique analytic structure. This is a highly non-trivial result (it was one of Hilbert's 20 problems), and we are not going to prove it (proof of a weaker result, that C^2 implies analyticity, is much easier and can be found in [**10**, Section 1.6]). In this book, "smooth" will be always understood as C^∞.

Example 2.3. The following are examples of Lie groups

(1) $\mathbb{R}^n$, with the group operation given by addition

(2) $\mathbb{R}^*, \times$
    $\mathbb{R}_+, \times$

(3) $S^1 = \{z \in \mathbb{C} : |z| = 1\}, \times$

(4) $\mathrm{GL}(n, \mathbb{R}) \subset \mathbb{R}^{n^2}$. Many of the groups we will consider will be subgroups of $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$.

(5) $\mathrm{SU}(2) = \{A \in \mathrm{GL}(2, \mathbb{C}) \mid A\bar{A}^t = 1, \det A = 1\}$. Indeed, one can easily see that

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

Writing $\alpha = x_1 + ix_2, \beta = x_3 + ix_4, x_i \in \mathbb{R}$, we see that $\mathrm{SU}(2)$ is diffeomorphic to $S^3 = \{x_1^2 + \cdots + x_4^2 = 1\} \subset \mathbb{R}^4$.

(6) In fact, all usual groups of linear algebra, such as $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{SL}(n, \mathbb{R})$, $O(n, \mathbb{R})$, $\mathrm{U}(n)$, $\mathrm{SO}(n, \mathbb{R})$, $\mathrm{SU}(n)$, $\mathrm{Sp}(2n, \mathbb{R})$ are Lie groups. This will be proved later (see Section 2.5).

Note that the definition of a Lie group does not require that G be connected. Thus, any finite group is a 0-dimensional Lie group. Since the theory of finite groups is complicated enough, it makes sense to separate the finite (or, more generally, discrete) part. It can be done as follows.

Theorem 2.4. *Let G be a Lie group. Denote by G^0 the connected component of unity. Then G^0 is a normal subgroup of G and is a Lie group itself. The quotient group G/G^0 is discrete.*

Proof. We need to show that G^0 is closed under the operations of multiplication and inversion. Since the image of a connected topological space under a continuous map is connected, the inversion map i must take G^0 to one component of G, that which contains $i(1) = 1$, namely G^0. In a similar way one shows that G^0 is closed under multiplication.

To check that this is a normal subgroup, we must show that if $g \in G$ and $h \in G^0$, then $ghg^{-1} \in G^0$. Conjugation by g is continuous and thus will take G^0 to some connected component of G; since it fixes 1, this component is G^0.

The fact that the quotient is discrete is obvious. $\square$

This theorem mostly reduces the study of arbitrary Lie groups to the study of finite groups and connected Lie groups. In fact, one can go further and reduce the study of connected Lie groups to connected simply-connected Lie groups.

Theorem 2.5. *If G is a connected Lie group then its universal cover $\tilde{G}$ has a canonical structure of a Lie group such that the covering map $p\colon \tilde{G} \to G$ is a morphism of Lie groups, and $\mathrm{Ker}\ p = \pi_1(G)$ as a group. Morever, in this case $\mathrm{Ker}\, p$ is a discrete central subgroup in $\tilde{G}$.*

Proof. The proof follows from the following general result of topology: if M, N are connected manifolds (or, more generally, nice enough topological spaces), then any continuous map $f\colon M \to N$ can be lifted to a map $\tilde{f}\colon \tilde{M} \to \tilde{N}$. Moreover, if we choose $m \in M, n \in N$ such that $f(m) = n$ and choose liftings $\tilde{m} \in \tilde{M}, \tilde{n} \in \tilde{N}$ such that $p(\tilde{m}) = m, p(\tilde{n}) = n$, then there is a unique lifting $\tilde{f}$ of f such that $\tilde{f}(\tilde{m}) = \tilde{n}$.

Now let us choose some element $\tilde{1} \in \tilde{G}$ such that $p(\tilde{1}) = 1 \in G$. Then, by the above theorem, there is a unique map $\tilde{i}\colon \tilde{G} \to \tilde{G}$ which lifts the inversion map $i\colon G \to G$ and satisfies $\tilde{i}(\tilde{1}) = \tilde{1}$. In a similar way one constructs the multiplication map $\tilde{G} \times \tilde{G} \to \tilde{G}$. Details are left to the reader.

Finally, the fact that $\mathrm{Ker}\, p$ is central follows from results of Exercise 2.2. $\square$

Definition 2.6. A Lie subgroup H of a Lie group G is a subgroup which is also a submanifold.

Remark 2.7. In this definition, the word "submanifold" should be understood as "imbedded submanifold". In particular, this means that H is locally closed but not necessarily closed; as we will show below, it will automatically be closed. It should also be noted that the terminology varies from book to book: some books use the word "Lie subgroup" for a more general notion which we will discuss later (see Definition 3.39).

Theorem 2.8.

(1) *Any Lie subgroup is closed in G.*

(2) *Any closed subgroup of a Lie group is a Lie subgroup.*

Proof. The proof of the first part is given in Exercise 2.1. The second part is much harder and will not be proved here. The proof uses the technique of Lie algebras and can be found, for example, in [**10**, Corollary 1.10.7]. We will give a proof of a weaker but sufficient for our purposes result later (see Section 3.6). $\square$

Corollary 2.9.

(1) *If G is a connected Lie group and U is a neighborhood of 1, then U generates G.*

(2) *Let $f\colon G_1 \to G_2$ be a morphism of Lie groups, with G_2 connected, and $f_*\colon T_1G_1 \to T_1G_2$ is surjective. Then f is surjective.*

Proof.   (1) Let H be the subgroup generated by U. Then H is open in G: for any element $h \in H$, the set $h \cdot U$ is a neighborhood of h in G. Since it is an open subset of a manifold, it is a submanifold, so H is a Lie subgroup. Therefore, by Theorem 2.8 it is closed, and is nonempty, so $H = G$.

(2) Given the assumption, the inverse function theorem says that f is surjective onto some neighborhood U of $1 \in G_2$. Since an image of a group morphism is a subgroup, and U generates G_2, f is surjective.

$\qquad\square$

As in the theory of discrete groups, given a subgroup $H \subset G$, we can define the notion of cosets and define the coset space G/H as the set of equivalence classes. The following theorem shows that the coset space is actually a manifold.

Theorem 2.10.

(1) *Let G be a Lie group of dimension n and $H \subset G$ a Lie subgroup of dimension k. Then the coset space G/H has a natural structure of a manifold of dimension $n - k$ such that the canonical map $p\colon G \to G/H$ is a fiber bundle, with fiber diffeomorphic to H. The tangent space at $\bar 1 = p(1)$ is given by $T_{\bar 1}(G/H) = T_1G/T_1H$.*

(2) *If H is a normal Lie subgroup then G/H has a canonical structure of a Lie group.*

Proof. Denote by $p\colon G \to G/H$ the canonical map. Let $g \in G$ and $\bar g = p(g) \in G/H$. Then the set $g\cdot H$ is a submanifold in G as it is an image of H under diffeomorphism $x \mapsto gx$. Choose a submanifold $M \subset G$ such that $g \in M$ and M is transversal to the manifold gH, i.e. $T_gG = (T_g(gH)) \oplus T_gM$ (this implies that $\dim M = \dim G - \dim H$). Let $U \subset M$ be a sufficiently small neighborhood of g in M. Then the set $UH = \{uh \mid u \in U, h \in H\}$ is open in G (which easily follows from inverse function theorem applied to the map $U \times H \to G$). Consider $\bar U = p(U)$; since $p^{-1}(\bar U) = UH$ is open, $\bar U$ is an open neighborhood of $\bar g$ in GH and the map $U \to \bar U$ is a homeomorphism. This gives a local chart for G/H and at the same time shows that $G \to G/H$ is a fiber bundle with fiber H. We leave it to the reader to show that transition functions between such charts are smooth and that the smooth structure does not depend on the choice of g, M.

This argument also shows that the kernel of the projection $p_*\colon T_gG \to T_{\bar g}(G/H)$ is equal to $T_g(gH)$. In particular, for $g = 1$ this gives an isomorphism $T_{\bar 1}(G/H) = T_1G/T_1H$. $\square$

Corollary 2.11.

(1) *If H is connected, then the set of connected components $\pi_0(G) = \pi_0(G/H)$. In particular, if $H, G/H$ are connected, then so is G.*

(2) *If G, H are connected, then there is an exact sequence of groups*

$$\pi_2(G/H) \to \pi_1(H) \to \pi_1(G) \to \pi_1(G/H) \to \{1\}$$

This corollary follows from more general long exact sequence of homotopy groups associated with any fiber bundle. We will later use it to compute fundamental groups of classical groups such as $\mathrm{GL}(n)$.

Finally, there is an analog of the standard homomorphism theorem for Lie groups.

Theorem 2.12. *Let $f\colon G_1 \to G_2$ be a morphism of Lie groups. Then $H = \mathrm{Ker}\, f$ is a normal Lie subgroup in G_1, and f gives rise to an injective morphism $G_1/H \to G_2$, which is an immersion of manifolds. If $\mathrm{Im}\, f$ is closed, then it is a Lie subgroup in G_2 and f gives an isomorphism of Lie groups $G_1/H \simeq \mathrm{Im}\, f$.*

The proof of this theorem will be given later (see Corollary 3.27). Note that it shows in particular that an image of f is a subgroup in G_2 which is an immersed submanifold; however, it may not be a Lie subgroup as the example below shows. Such more general kinds of subgroups are called immersed subgroups and will be discussed later (see Definition 3.39).

Example 2.13. Let $G_1 = \mathbb{R}$, $G = T^2 = \mathbb{R}^2/\mathbb{Z}^2$. Define the map $f\colon G_1 \to G_2$ by $f(t) = (t \mod \mathbb{Z}, \alpha t \mod \mathbb{Z})$, where α is some fixed irrational number. Then it is well-known that the image of this map is everywhere dense in T^2 (it is sometimes called *irrational winding* on the torus).

2.2. Action of Lie groups on manifolds and representations

The primary reason why Lie groups are so frequently used is that they usually appear as groups of symmetry of various geometric objects. In this section, we will show several examples.

Definition 2.14. An action of a Lie group G an a manifold M is an assignment to each $g \in G$ a diffeomorhism $\rho(g) \in \mathrm{Diff}\, M$ such that $\rho(1) = \mathrm{id}, \rho(gh) = \rho(g)\rho(h)$ and such that the map

$$G \times M \to M\colon (g, m) \mapsto \rho(g).m$$

is a smooth map.

Example 2.15.

(1) The group $\mathrm{GL}(n, \mathbb{R})$ (and thus, any its Lie subgroup) acts on $\mathbb{R}^n$.

(2) The group $\mathrm{O}(n, \mathbb{R})$ acts on the sphere $S^{n-1} \subset \mathbb{R}^n$. The group $\mathrm{U}(n)$ acts on the sphere $S^{2n-1} \subset \mathbb{C}^n$.

Closely related with the notion of a group acting on a manifold is the notion of a representation.

Definition 2.16. A representation of a Lie group G is a vector space V together with a group morphism $\rho\colon G \to \mathrm{End}(V)$. If V is finite-dimensional, we also require that the map $G \times V \to V\colon (g, v) \mapsto \rho(g).v$ be a smooth map, so that ρ is a morphism of Lie groups.

A morphism between two representations V, W is a linear map $f\colon V \to W$ which commutes with the action of G: $f\rho_V(g) = \rho_W(g)f$.

In other words, we assign to every $g \in G$ a linear map $\rho(g)\colon V \to V$ so that $\rho(g)\rho(h) = \rho(gh)$.

We will frequently use the shorter notation $g.m, g.v$ instead of $\rho(g).m$ in the cases when there is no ambiguity about the representation being used.

Remark 2.17. Note that we frequently consider representations on a *complex* vector space V, even for a real Lie group G.

Any action of the group G on a manifold M gives rise to several representations of G on various vector spaces associated with M:

(1) Representation of G on the (infinite-dimensional) space of functions $C^\infty(M)$ defined by

$$(2.1) \qquad\qquad\qquad (\rho(g)f)(m) = f(g^{-1}.m)$$

(note that we need g^{-1} rather than g to satisfy $\rho(g)\rho(h) = \rho(gh)$).

(2) Representation of G on the (infinite-dimensional) space of vector fields $\mathrm{Vect}(M)$ defined by

(2.2) $$(\rho(g).v)(m) = g_*(v(g^{-1}.m)).$$

In a similar way, we define the action of G on the spaces of differential forms and other types of tensor fields on M.

(3) Assume that $m \in M$ is a stationary point: $g.m = m$ for any $g \in G$. Then we have a canonical action of G on the tangent space $T_m M$ given by $\rho(g) = g_* \colon T_m M \to T_m M$, and similarly for the spaces $T_m^* M, \bigwedge^k T_m^* M$.

2.3. Orbits and homogeneous spaces

Let G act on a manifold M. Then for every point $m \in M$ we define its *orbit* by $\mathcal{O}_m = Gm = \{g.m \mid g \in G\}$.

Lemma 2.18. *Let M be a manifold with an action of G. Choose a point $m \in M$ and let $H = \mathrm{Stab}_G(m) = \{g \in G \mid g.m = m\}$. Then H is a Lie subgroup in G, and $g \mapsto g.m$ is an injective immersion $G/H \hookrightarrow M$ whose image coincides with the orbit $\mathcal{O}_m$.*

Proof. The fact that the orbit is in bijection with G/H is obvious. For the proof of the fact that H is a closed subgroup, we could just refer to Theorem 2.8. However, this would not help proving that $G/\mathrm{Stab}(m) \to M$ is an immersion. Both of these statements are easiest proved using the technique of Lie algebras; thus, we pospone the proof until later time (see Theorem 3.26). $\square$

Corollary 2.19. *The orbit $\mathcal{O}_m$ is an immersed submanifold in M, with tangent space $T_m \mathcal{O}_m = T_1 G / T_1 H$. If $\mathcal{O}_m$ is closed, then $g \mapsto g.m$ is a diffeomorphism $G/\mathrm{Stab}(m) \xrightarrow{\sim} \mathcal{O}_m$.*

An important special case is when the action of G is transitive, i.e. when there is only one orbit.

Definition 2.20. A G-homogeneous space is a manifold with a transitive action of G.

As an immediate corollary of Corollary 2.19, we see that each homogeneous space is diffeomorphic to a coset space G/H. Combining it with Theorem 2.10, we get the following result.

Corollary 2.21. *Let M be a G-homogeneous space and choose $m \in M$. Then the map $G \to M \colon g \mapsto gm$ is a fiber bundle over M with fiber $H = \mathrm{Stab}_G m$.*

Example 2.22.

(1) Consider the action of $\mathrm{SO}(n, \mathbb{R})$ on the sphere $S^{n-1} \subset \mathbb{R}^n$. Then it is a homogeneous space, so we have a fiber bundle

$$\mathrm{SO}(n-1, \mathbb{R}) \longrightarrow \mathrm{SO}(n, \mathbb{R})$$
$$\downarrow$$
$$S^{n-1}$$

(2) Consider the action of $\mathrm{SU}(n)$ on the sphere $S^{2n-1} \subset \mathbb{C}^n$. Then it is a homogeneous space, so we have a fiber bundle

$$\mathrm{SU}(n-1) \longrightarrow \mathrm{SU}(n)$$
$$\downarrow$$
$$S^{2n-1}$$

In fact, action of G can be used to define smooth structure on a set. Indeed, if M is a set (no smooth structure yet) with a transitive action of a Lie group G, then M is in bijection with G/H, $H = \mathrm{Stab}_G(m)$ and thus, by Theorem 2.10, M has a canonical structure of a manifold of dimension equal to $\dim G - \dim H$.

Example 2.23. Define a *flag* in $\mathbb{R}^n$ to be a sequence of subspaces

$$\{0\} \subset V_1 \subset V_2 \subset \cdots \subset V_n = \mathbb{R}^n, \qquad \dim V_i = i$$

Let $\mathcal{B}_n(\mathbb{R})$ be the set of all flags in $\mathbb{R}^n$. It turns out that $\mathcal{B}_n(\mathbb{R})$ has a canonical structure of a smooth manifold which is called the *flag manifold* (or sometimes *flag variety*). The easiest way to define it is to note that we have an obvious action of the group $\mathrm{GL}(n, \mathbb{R})$ on $\mathcal{B}_n(\mathbb{R})$. This action is transitive: by a change of basis, any flag can be identified with the standard flag

$$V^{st} = \big(\{0\} \subset \langle e_1 \rangle \subset \langle e_1, e_2 \rangle \subset \cdots \subset \langle e_1, \ldots, e_{n-1} \rangle \subset \mathbb{R}^n\big)$$

where $\langle e_1, \ldots, e_k \rangle$ stands for the subspace spanned by $e_1, \ldots, e_k$. Thus, $\mathcal{B}_n(\mathbb{R})$ can be identified with the coset $\mathrm{GL}(n, \mathbb{R})/B(n, \mathbb{R})$, where $B(n, \mathbb{R}) = \mathrm{Stab}\, V^{st}$ is the group of all invertible upper-triangular matrices. Therefore, $\mathcal{B}_n$ is a manifold of dimension equal to $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

Finally, we should say a few words about taking the quotient by the action of a group. In many cases when we have an action of a group G on a manifold M one would like to consider the quotient space, i.e. the set of all G-orbits. This set is commonly denoted by M/G. It has a canonical quotient topology. However, this space can be very singular, even if G is a Lie group; for example, it can be non-Hausdorff. For example, if $G = \mathrm{GL}(n, \mathbb{C})$ acting on the set of all $n \times n$ matrices by conjugation, then the set of orbits is described by Jordan canonical form. However, it is well-known that by a small perturbation, any matrix can be made diagonalizable. Thus, if X, Y are matrices with the same eigenvalues but different Jordan form, then any neighborhood of the orbit of X contains points from orbit of Y.

There are several ways of dealing with this problem. One of them is to impose additional requirements on the action, for example assuming that the action is proper. In this case it can be shown that M/G is indeed a Hausdorff topological space, and under some additional conditions, it is actually a manifold (see [**10**, Section 2]). Another approach, usually called Geometric Invariant Theory, is based on using the methods of algebraic geometry (see [**17**]). Both of these methods go beyond the scope of this book.

2.4. Left, right, and adjoint action

Important examples of group action are the following actions of G on itself:

> **Left action:** $L_g \colon G \to G$ is defined by $L_g(h) = gh$
>
> **Right action:** $R_g \colon G \to G$ is defined by $R_g(h) = hg^{-1}$
>
> **Adjoint action:** $\mathrm{Ad}_g \colon G \to G$ is defined by $\mathrm{Ad}_g(h) = ghg^{-1}$

One easily sees that left and right actions are transitive; in fact, each of them is simply transitive. It is also easy to see that the left and right actions commute and that $\mathrm{Ad}_g = L_g R_g$.

As mentioned above, each of these actions also defines the action of G on the spaces of functions, vector fields, forms, etc. on G. For simplicity, for a tangent vector $v \in T_m G$, we will frequently write just $gv \in T_{gm} G$ instead of technically more accurate but cumbersome notation $(Lg)_* v$. Similarly, we will write vg for $(R_{g^{-1}})_* v$. This is justified by Exercise 2.6, where it is shown that for matrix groups this notation agrees with usual multiplication of matrices.

Since the adjoint action preserves the identity element $1 \in G$, it also defines an action of G on the (finite-dimensional) space $T_1 G$. Slightly abusing the notation, we will denote this action also by

$$(2.3) \qquad \text{Ad}\, g \colon T_1 G \to T_1 G.$$

Definition 2.24. A vector field $v \in \text{Vect}(G)$ is *left-invariant* if $g.v = v$ for every $g \in G$, and right-invariant if $v.g = v$ for every $g \in G$. A vector field is called bi-invariant if it is both left- and right-invariant.

In a similar way one defines left- , right-, and bi-invariant differential forms and other tensors.

Theorem 2.25. *The map $v \mapsto v(1)$ (where 1 is the identity element of the group) defines an isomorphism of the vector space of left-invariant vector fields on G with the vector space $T_1 G$, and similarly for right-invariant vector spaces.*

Proof. It suffices to prove that every $x \in T_1 G$ can be uniquely extended to a left-invariant vector field on G. Let us define the extension by $v(g) = gx \in T_g G$. Then one easily sees that so defined vector field is left-invariant, and $v(1) = x$. This proves existence of extension; uniqueness is obvious. $\square$

Describing bi-invariant vector fields on G is more complicated: any $x \in T_1 G$ can be uniquely extended to a left-invariant vector field and to a right-invariant vector field, but these extensions may differ.

Theorem 2.26. *The map $v \mapsto v(1)$ defines an isomorphism of the vector space of bi-invariant vector fields on G with the vector space of invariants of adjoint action:*

$$(T_1 G)^{\text{Ad}\, G} = \{x \in T_1 G \mid \text{Ad}\, g(x) = x \text{ for all } g \in G\}$$

The proof of this result is left to the reader. Note also that a similar result holds for other types of tensor fields: covector fields, differential forms, etc.

2.5. Classical groups

In this section, we discuss the so-called classical groups, or various subgroups of the general linear group which are frequently used in linear algebra. Traditionally, the name "classical groups" is applied to the following groups:

- $\text{GL}(n, \mathbb{K})$ (here and below, $\mathbb{K}$ is either $\mathbb{R}$, which gives a real Lie group, or $\mathbb{C}$, which gives a complex Lie group)
- $\text{SL}(n, \mathbb{K})$
- $\text{O}(n, \mathbb{K})$
- $\text{SO}(n, \mathbb{K})$ and more general groups $\text{SO}(p, q; \mathbb{R})$.
- $\text{U}(n)$
- $\text{SU}(n)$
- $\text{Sp}(2n, \mathbb{K}) = \{A : \mathbb{K}^{2n} \to \mathbb{K}^{2n} \mid \omega(Ax, Ay) = \omega(x, y)\}$. Here $\omega(x, y)$ is the skew-symmetric bilinear form $\sum_{i=1}^{n} x_i y_{i+n} - y_i x_{i+n}$ (which, up to a change of basis, is the unique non-degenerate skew-symmetric bilinear form on $\mathbb{K}^{2n}$). Equivalently, one can write $\omega(x, y) = (Jx, y)$, where $(\, ,\,)$ is the standard symmetric bilinear form on $\mathbb{K}^n$ and

$$J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}.$$

Remark 2.27. There is some ambiguity with the notation for symplectic group: the group we denoted $\mathrm{Sp}(2n, \mathbb{K})$ in some books would be written as $\mathrm{Sp}(n, \mathbb{K})$. Also, it should be noted that there is a closely related compact group of quaternionic unitary transformations (see Exercise 2.7). This group, which is usually denoted simply $\mathrm{Sp}(n)$, is a "compact form" of the group $\mathrm{Sp}(2n, \mathbb{C})$ in the sense we will describe later (see Exercise 3.16). To avoid confusion, we have not included this group in the list of classical groups.

We have already shown that $\mathrm{GL}(n)$ and $\mathrm{SU}(2)$ are Lie groups. In this section, we will show that each of these groups is a Lie group and will find their dimensions.

Straightforward approach, based on implicit function theorem, is hopeless: for example, $\mathrm{SO}(n, \mathbb{K})$ is defined by n^2 equations in $\mathbb{K}^{n^2}$, and finding the rank of this system is not an easy task. We could just refer to the theorem about closed subgroups; this would prove that each of them is a Lie group, but would give us no other information — not even the dimension of G. Thus, we will need another approach.

Our approach is based on the use of exponential map. Recall that for matrices, the exponential map is defined by

$$(2.4) \qquad \exp(x) = \sum_0^\infty \frac{x^k}{k!}.$$

It is well-known that this power series converges and defines an analytic map $\mathfrak{gl}(n, \mathbb{K}) \to \mathfrak{gl}(n, \mathbb{K})$, where $\mathfrak{gl}(n)$ is the set of all $n \times n$ matrices. In a similar way, we define the logarithmic map by

$$(2.5) \qquad \log(1 + x) = \sum_1^\infty \frac{(-1)^{k+1} x^k}{k}.$$

So defined log is an analytic map defined in a neighborhood of $1 \in \mathfrak{gl}(n, \mathbb{K})$.

The following theorem summarizes properties of exponential and logarithmic maps. Most of the properties are the same as for numbers; however, there are also some differences due to the fact that multiplication of matrices is not commutative. All of the statements of this theorem apply equally well in real and complex cases.

Theorem 2.28.

(1) $\log(\exp(x)) = x$; $\exp(\log(X)) = X$ *whenever they are defined.*

(2) $\exp(x) = 1 + x + \ldots$ *This means* $\exp(0) = 1$ *and* $d \ \exp(0) = \mathrm{id}$.

(3) *If* $xy = yx$ *then* $\exp(x + y) = \exp(x) \exp(y)$. *If* $XY = YX$ *then* $\log(XY) = \log(X) + \log(Y)$ *in some neighborhood of the identity. In particular, for any* $x \in \mathfrak{gl}(n, \mathbb{K})$, $\exp(x) \exp(-x) = 1$, *so* $\exp x \in \mathrm{GL}(n, \mathbb{K})$.

(4) *For fixed* $x \in \mathfrak{gl}(n, \mathbb{K})$, *consider the map* $\mathbb{K} \to \mathrm{GL}(n, \mathbb{K}) \colon t \mapsto \exp(tx)$. *Then* $\exp((t+s)x) = \exp(tx) \exp(sx)$. *In other words, this map is a morphism of Lie groups.*

(5) *The exponential map agrees with change of basis and transposition:*
$\exp(AxA^{-1}) = A \exp(x) A^{-1}$, $\exp(x^t) = (\exp(x))^t$.

Full proof of this theorem will not be given here; instead, we just give a sketch. First two statements are just equalities of formal power series in one variable; thus, it suffices to check that they hold for $x \in \mathbb{R}$. Similarly, the third one is an identity of formal power series in two commuting variables, so it again follows from well-known equality for $x, y \in \mathbb{R}$. The fourth follows from the third, and the fifth follows from $(AxA^{-1})^n = Ax^n A^{-1}$ and $(A^t)^n = (A^n)^t$.

Note that group morphisms $\mathbb{R} \to G$ are frequently called *one-parameter subgroups* in G. This is not a quite accurate name, as the image may not be a Lie subgroup (see Theorem 2.12); however, the name is so widely used that it is too late to change it. Thus, we can reformulate part (4) of the theorem by saying that $\exp(tx)$ is a one-parameter subgroup in $\mathrm{GL}(n, \mathbb{K})$.

How does it help us to study various matrix groups? The key idea is that the logarithmic map identifies some neighborhood of the identity in $\mathrm{GL}(n, \mathbb{K})$ with some neighborhood of 0 in a vector space. It turns out that it also does the same for all of the classical groups.

Theorem 2.29. *For each classical group* $G \subset \mathrm{GL}(n, \mathbb{K})$, *there exists a vector space* $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{K})$ *such that for some some neighborhood* U *of* 1 *in* $\mathrm{GL}(n, \mathbb{K})$ *and some neighborhood* u *of* 0 *in* $\mathfrak{gl}(n, \mathbb{K})$ *the following maps are mutually inverse*

$$(U \cap G) \underset{exp}{\overset{\log}{\rightleftarrows}} (u \cap \mathfrak{g})$$

Before proving this theorem, note that it immediately implies the following important corollary.

Corollary 2.30. *Each classical group is a Lie group, with tangent space at identity* $T_1 G = \mathfrak{g}$ *and* $\dim G = \dim \mathfrak{g}$.

Let us prove this corollary first because it is very easy. Indeed, Theorem 2.29 shows that near 1, G is identified with an open set in a vector space. So it is immediate that near 1, G is smooth. If $g \in G$ then $g \cdot U$ is a neighborhood of g in $\mathrm{GL}(n, \mathbb{K})$, and $(g \cdot U) \cap G = g \cdot (U \cap G)$ is a neighborhood of g in G; thus, G is smooth near g.

For the second part, consider the differential of the exponential map $\exp_* \colon T_0 \mathfrak{g} \to T_1 G$. Since $\mathfrak{g}$ is a vector space, $T_0 \mathfrak{g} = \mathfrak{g}$, and since $\exp(x) = 1 + x + \ldots$, the derivative is the identity; thus, $T_0 \mathfrak{g} = \mathfrak{g} = T_1 G$.

Proof of Theorem 2.29. The proof is case by case; it can not be any other way, as "classical groups" are defined by a list rather than by some general definition.

$\mathrm{GL}(n, \mathbb{K})$: Immediate from Theorem 2.28; in this case, $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{K})$ is the space of all matrices.

$\mathrm{SL}(n, \mathbb{K})$: Suppose $X \in \mathrm{SL}(n, \mathbb{K})$ is close enough to identity. Then $X = \exp(x)$ for some $x \in \mathfrak{gl}(n, \mathbb{K})$. The condition that $X \in \mathrm{SL}(n, \mathbb{K})$ is equivalent to $\det X = 1$, or $\det \exp(x) = 1$. But it is well-known that $\det \exp(x) = \exp(\mathrm{tr}(x))$ (which is easy to see by finding a basis in which x is upper-triangular), so $\exp(x) \in \mathrm{SL}(n, \mathbb{K})$ if and only if $\mathrm{tr}(x) = 0$. Thus, in this case the statement also holds, with $\mathfrak{g} = \{x \in \mathfrak{gl}(n, \mathbb{K}) \mid \mathrm{tr}\, x = 0\}$.

$\mathrm{O}(n, \mathbb{K}), \mathrm{SO}(n, \mathbb{K})$: The group O_n is defined by $XX^t = I$. Then X, X^t commute. Writing $X = \exp(x), X^t = \exp(x^t)$ (since exponential map agrees with transposition), we see that x, x^t also commute, and thus $\exp(x) \in \mathrm{O}(n)$ implies $\exp(x) \exp(x^t) = \exp(x + x^t) = 1$, so $x + x^t = 0$; conversely, if $x + x^t = 0$, then x, x^t commute, so we can reverse the argument to get $\exp(x) \in \mathrm{O}(n, \mathbb{K})$. Thus, in this case the theorem also holds, with $\mathfrak{g} = \{x \mid x + x^t = 0\}$— the space of skew-symmetric matrices.

What about $\mathrm{SO}(n, \mathbb{K})$? In this case, we should add to the condition $XX^t = 1$ (which gives $x + x^t = 0$) also the condition $\det X = 1$, which gives $\mathrm{tr}(x) = 0$. However, this last condition is unnecessary, because $x + x^t = 0$ implies that all diagonal entries of x are zero. So both $\mathrm{O}(n, \mathbb{K})$ and $\mathrm{SO}(n, \mathbb{K})$ correspond to the same space of matrices $\mathfrak{g} = \{x \mid x + x^t = 0\}$. This might seem confusing until one realizes that $\mathrm{SO}(n, \mathbb{K})$ is exactly the connected component of identity in $\mathrm{O}(n, \mathbb{K})$; thus, neighborhood of 1 in $\mathrm{O}(n, \mathbb{K})$ coincides with the neighborhood of 1 in $\mathrm{SO}(n, \mathbb{K})$.

$\mathrm{U}(n), \mathrm{SU}(n)$**:** Similar argument shows that $\exp x \in \mathrm{U}(n) \iff x + x^* = 0$ (where $x^* = \bar{x}^t$) and $\exp x \in \mathrm{SU}(n) \iff x + x^* = 0, \operatorname{tr}(x) = 0$. Note that in this case, $x + x^*$ does not imply that x has zeroes on the diagonal: it only implies that the diagonal entries are purely imaginary. Thus, $\operatorname{tr} x = 0$ does not follow automatically from $x + x^* = 0$, so in this case the tangent spaces for $\mathrm{U}(n), \mathrm{SU}(n)$ are different.

$\mathrm{Sp}(2n, \mathbb{K})$**:** Similar argument shows that $\exp(x) \in \mathrm{Sp}(2n, \mathbb{K}) \iff x + Jx^t J^{-1} = 0$; thus, in this case the theorem also holds.

$\square$

The vector space $\mathfrak{g} = T_1 G$ is called the *Lie algebra* of the corresponding group G (this will be justified later, when we actually define an algebra operation on it). Traditionally, the Lie algebra is denoted by lowercase gothic letters: for example, the Lie algebra of group $\mathrm{SU}(n)$ is denoted by $\mathfrak{su}(n)$.

The following table summarizes results of Theorem 2.29. In addition, it also contains information about topological structure of classical Lie groups. Proofs of them can be found in exercsies.

G	$\mathrm{GL}(n, \mathbb{R})$	$\mathrm{SL}(n, \mathbb{R})$	$\mathrm{O}(n, \mathbb{R})$	$\mathrm{SO}(n, \mathbb{R})$	$\mathrm{U}(n)$	$\mathrm{SU}(n)$	$\mathrm{Sp}(2n, \mathbb{R})$
$\mathfrak{g}$	$\mathfrak{gl}(n, \mathbb{R})$	$\operatorname{tr} x = 0$	$x + x^t = 0$	$x + x^t = 0$	$x + x^* = 0$	$x + x^* = 0,\ \operatorname{tr} x = 0$	$x + Jx^t J^{-1} = 0$
$\dim G$	n^2	$n^2 - 1$	$\frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$	n^2	$n^2 - 1$	$n(2n+1)$
$\pi_0(G)$	$\mathbb{Z}_2$	$\{1\}$	$\mathbb{Z}_2$	$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$
$\pi_1(G)$	$\mathbb{Z}_2\ (n \geq 3)$	$\mathbb{Z}_2\ (n \geq 3)$	$\mathbb{Z}_2\ (n \geq 3)$	$\mathbb{Z}_2\ (n \geq 3)$	$\mathbb{Z}$	$\{1\}$	$\mathbb{Z}$

For complex classical groups, the Lie algebra and dimension are given by the same formula as for real groups. However, the topology of complex Lie groups is different and is given in the table below.

G	$\mathrm{GL}(n, \mathbb{C})$	$\mathrm{SL}(n, \mathbb{C})$	$\mathrm{O}(n, \mathbb{C})$	$\mathrm{SO}(n, \mathbb{C})$
$\pi_0(G)$	$\{1\}$	$\{1\}$	$\mathbb{Z}_2$	$\{1\}$
$\pi_1(G)$	$\mathbb{Z}$	$\{1\}$	$\mathbb{Z}_2$	$\mathbb{Z}_2$

Note that some of the classical groups are not simply-connected. As was shown in Theorem 2.5, in this case the universal cover has a canonical structure of a Lie group. Of special importance is the universal cover of $\mathrm{SO}(n, \mathbb{R})$ which is called the spin group and is denoted $\mathrm{Spin}(n)$; since $\pi_1(\mathrm{SO}(n, \mathbb{R})) = \mathbb{Z}_2$, this is a twofold cover.

Exercises

2.1. Let G be a Lie group and H — a Lie subgroup.
  (1) Let $\overline{H}$ be the closure of H in G. Show that $\overline{H}$ is a subgroup in G.
  (2) Show that each coset $Hx, x \in \overline{H}$, is open and dense in $\overline{H}$.
  (3) Show that $\overline{H} = H$, that is, every Lie subgroup is closed.

2.2. (1) Show that every discrete normal subgroup of a connected Lie group is central (hint: consider the map $G \to N \colon g \mapsto ghg^{-1}$ where h is a fixed element in N).
  (2) By applying part (a) to kernel of the map $\widetilde{G} \to G$, show that for any connected Lie group G, the fundamental group $\pi_1(G)$ is commutative.

2.3. Let $f \colon G_1 \to G_2$ be a morphism of connected Lie groups such that $f_* \colon T_1 G_1 \to T_1 G_2$ is an isomorphism (such a morphism is sometimes called *local isomorphism*). Show that f is a covering map, and $\operatorname{Ker} f$ is a discrete central subgroup.

2.4. Let $\mathcal{B}_n(\mathbb{C})$ be the set of all flags in $\mathbb{C}^n$ (see Example 2.23. Show that $\mathcal{B}_n(\mathbb{C}) = \mathrm{GL}(n,\mathbb{C})/B(n,\mathbb{C}) = \mathrm{U}(n)/T(n)$ where $B(n,\mathbb{C})$ is the group of invertible complex upper triangular matrices, and $T(n)$ is the group of diagonal unitary matrices (which is easily shown to be the n-dimensional torus $(\mathbb{R}/\mathbb{Z})^n$). Deduce from this that $\mathcal{B}_n(\mathbb{C})$ is a compact complex manifold and find its dimension over $\mathbb{C}$.

2.5. Let $G_{n,k}$ be the set of all dimension k subspaces in $\mathbb{R}^n$ (usually called the Grassmanian). Show that $G_{n,k}$ is a homogeneous space for the group $O(n)$ and thus can be identified with coset space $O(n)/H$ for appropriate H. Use it to prove that $G_{n,k}$ is a manifold and find its dimension.

2.6. Show that if $G = \mathrm{GL}(n,\mathbb{R}) \subset \mathrm{End}(\mathbb{R}^n)$ so that each tangent space is canonically identified with $\mathrm{End}(\mathbb{R}^n)$, then $(Lg)_* v = gv$ where the product in the right-hand side is the usual product of matrices, and similarly for the right action. Also, the adjoint action is given by $\mathrm{Ad}_g(v) = gvg^{-1}$.

2.7. Let $\mathrm{Sp}(n) = \mathrm{Sp}(2n,\mathbb{C}) \cap \mathrm{SU}(2n)$. Show that $\mathrm{Sp}(n)$ is a compact real Lie group, and compute its tangent space at 1. (This group is sometimes called the quaternionic unitary group. Its relation with the groups $\mathrm{Sp}(2n,\mathbb{R})$ and $\mathrm{Sp}(2n,\mathbb{C})$ will be clarified later — see Exercise 3.16.)

2.8. Using Example 2.22, show that for $n \geq 1$, we have $\pi_0(\mathrm{SU}(n+1)) = \pi_0(\mathrm{SU}(n))$, $\pi_0(\mathrm{U}(n+1)) = \pi_0(\mathrm{U}(n))$ and deduce from it that groups $\mathrm{U}(n)$, $\mathrm{SU}(n)$ are connected for all n. Similarly, show that for $n \geq 2$, we have $\pi_1(\mathrm{SU}(n+1)) = \pi_1(\mathrm{SU}(n))$, $\pi_1(\mathrm{U}(n+1)) = \pi_1(\mathrm{U}(n))$ and deduce from it that for $n \geq 2$, $\mathrm{SU}(n)$ is simply-connected and $\pi_1(\mathrm{U}(n)) = \mathbb{Z}$.

2.9. Using Gram-Schmidt orthogonalization process, show that $\mathrm{GL}(n,\mathbb{R})/O(n,\mathbb{R})$ is diffeomorphic to the space of upper-triangular matrices with positive entries on the diagonal. Deduce from this that $\mathrm{GL}(n,\mathbb{R})$ is homotopic (as a topological space) to $O(n,\mathbb{R})$.

2.10. Let L_n be the set of all Lagrangian subspaces in $\mathbb{R}^{2n}$ with the standard symplectic form ω defined in Section 2.5. (A subspace V is Lagrangian if $\dim V = n$ and $\omega(x,y) = 0$ for any $x, y \in V$.)

   Show that the group $\mathrm{Sp}(2n,\mathbb{R})$ acts transitively on L_n and use it to define on L_n a structure of a smooth manifold and find its dimension.

   The next series of problems is about the group $\mathrm{SU}(2)$ and its adjoint representation

2.11. Define a bilinear form on $\mathfrak{su}(2)$ by $(a,b) = \frac{1}{2}\mathrm{tr}(a\bar{b}^t)$. Show that this form is symmetric, positive definite, and invariant under the adjoint action of $\mathrm{SU}(2)$.

2.12. Define a basis in $\mathfrak{su}(2)$ by

$$i\sigma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad i\sigma_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

Show that the map

(2.6)
$$\varphi \colon \mathrm{SU}(2) \to \mathrm{GL}(3,\mathbb{R})$$
$$g \mapsto \text{matrix of } \mathrm{Ad}\, g \text{ in the basis } i\sigma_1, i\sigma_2, i\sigma_3$$

gives a morphism of Lie groups $\mathrm{SU}(2) \to \mathrm{SO}(3,\mathbb{R})$.

2.13. Let $\varphi \colon \mathrm{SU}(2) \to \mathrm{SO}(3,\mathbb{R})$ be the morphism defined in the previous problem. Compute explicitly the map of tangent spaces $\varphi_* \colon \mathfrak{su}(2) \to \mathfrak{so}(3,\mathbb{R})$ and show that φ_* is an isomorphism. Deduce from this that $\mathrm{Ker}\,\varphi$ is a discrete normal subgroup in $\mathrm{SU}(2)$, and that $\mathrm{Im}\,\varphi$ is an open subgroup in $\mathrm{SO}(3,\mathbb{R})$.

2.14. Prove that φ establishes an isomorphism $\mathrm{SU}(2)/\mathbb{Z}_2 \to \mathrm{SO}(3,\mathbb{R})$ and thus, since $\mathrm{SU}(2) \simeq S^3$, $\mathrm{SO}(3,\mathbb{R}) \simeq \mathbb{RP}^3$.

Chapter 3

Lie Groups and Lie algebras

3.1. Exponential map

We are now turning to the study of arbitrary Lie groups. Our first goal will be generalizing the exponential map $\exp\colon \mathfrak{g} \to G$, which proved so useful in the study of matrix groups (see Theorem 2.28), to general Lie algebras. We can not use power series to define it because we don't have multiplication in $\mathfrak{g}$. However, it turns out that there is still a way to define such a map so that most of the results about the exponential map for matrix groups can be generalized to arbitrary groups, and this gives us a key to studying the Lie groups. This definition is based on the notion of one-parameter subgroup (compare with Theorem 2.28).

Proposition 3.1. *Let G be a Lie group, $\mathfrak{g} = T_1 G$, and let $x \in \mathfrak{g}$. Then there exists a unique morphism of Lie groups $\gamma_x \colon \mathbb{R} \to G$ such that*

$$\dot{\gamma}_x(0) = x$$

where dot stands for derivative with respect to t. The map γ_x will be called the one-parameter subgroup corresponding to x.

Proof. We begin with uniqueness. The usual argument, used to compute derivative of e^x in calculus, shows that if $\gamma(t)$ is a one-parameter subgroup, then $\dot{\gamma}(t) = \gamma(t) \cdot \dot{\gamma}(0) = \dot{\gamma}(0) \cdot \gamma(t)$. This is immediate for matrix groups; for general groups, the same proof works if, as in Section 2.4, we interpret $\gamma(t) \cdot \dot{\gamma}(0)$ as $(L_{\gamma(t)})_* \dot{\gamma}(0)$ and similarly for the right action. This gives us a differential equation for γ: if v_x is a left-invariant vector field on G such that $v_x(1) = x$, then γ is an integral curve for v. This proves uniqueness of $\gamma_x(t)$.

For existence, let $\Phi^t\colon G \to G$ be the time t flow of the vector field v_x (a priori, it is only defined for small enough t). Since the vector field is left-invariant, the flow operator is also left-invariant: $\Phi^t(g_1 g_2) = g_1 \Phi^t(g_2)$. Now let $\gamma(t) = \Phi^t(1)$. Then $\gamma(t + s) = \Phi^{t+s}(1) = \Phi^s(\Phi^t(1)) = \Phi^s(\gamma(t) \cdot 1) = \gamma(t)\Phi^s(1) = \gamma(t)\gamma(s)$ as desired. This proves existence of γ for small enough t. The fact that it can be extended to any $t \in \mathbb{R}$ is obvious from $\gamma(t + s) = \gamma(t)\gamma(s)$. $\square$

The uniqueness immediately implies that $\gamma_x(\lambda t) = \gamma_{\lambda x}(t)$ for any $\lambda \in \mathbb{R}$. Indeed, $\gamma_x(\lambda t)$ is a one-parameter subgroup with $\frac{d\gamma_x(\lambda t)}{dt}|_{t=0} = \lambda x$. Thus, $\gamma_x(t)$ only depends on the product $tx \in \mathfrak{g}$. This justifies the following definition.

Definition 3.2. Let G be a Lie group, $\mathfrak{g} = T_1 G$. Then the exponential map $\exp \colon \mathfrak{g} \to G$ is defined by

$$\exp(x) = \gamma_x(1)$$

where $\gamma_x(t)$ is the one-parameter subgroup with tangent vector at 1 equal to x.

Example 3.3. For $G \subset \mathrm{GL}(n, \mathbb{R})$, it follows from Theorem 2.28 that this definition agrees with the exponential map defined by (2.4).

Example 3.4. Let $G = \mathbb{R}$, so that $\mathfrak{g} = \mathbb{R}$. Then for any $a \in \mathfrak{g}$, the corresponding one-parameter subgroup is $\gamma_a(t) = ta$, so the exponential map is given by $\exp(a) = a$.

Example 3.5. Let $G = S^1 = \mathbb{R}/\mathbb{Z} = \{z \in \mathbb{C} \mid |z| = 1\}$ (these two descriptions are related by $z = e^{2\pi i \theta}, \theta \in \mathbb{R}/\mathbb{Z}$). Then $\mathfrak{g} = \mathbb{R}$, and the exponential map is given by $\exp(a) = a \mod \mathbb{Z}$ (if we use $G = \mathbb{R}/\mathbb{Z}$ description) or $\exp(a) = e^{2\pi i a}$ (if we use $G = \{z \in \mathbb{C} \mid |z| = 1\}$).

Note that the construction of the one-parameter subgroup given in the proof of Proposition 3.1 immediately gives the following result, formal proof of which is left as an exercise to the reader.

Proposition 3.6.

    (1) *Let v be a left-invariant vector field on G. Then the time t flow of this vector field is given by $g \mapsto g \exp(tx)$, where $x = v(1)$.*

    (2) *Let v be a right-invariant vector field on G. Then the time t flow of this vector field is given by $g \mapsto \exp(tx)g$, where $x = v(1)$.*

The following theorem summarizes properties of the exponential map.

Theorem 3.7.

    (1) $\exp(x) = 1 + x + \ldots$ *(that is, $\exp(0) = 1$ and $\exp_*(0) \colon \mathfrak{g} \to T_1 G = \mathfrak{g}$ is the identity map).*

    (2) *The exponential map is a diffeomorphism of some neighborhood of 0 in $\mathfrak{g}$ with a neighborhood of 1 in G. The local inverse map will be denoted by $\log$.*

    (3) $\exp((t + s)x) = \exp(tx) \exp(sx)$.

    (4) *For any morphism of Lie groups $\varphi \colon G_1 \to G_2$ and any $x \in \mathfrak{g}_1$, we have $\exp(\varphi_*(x)) = \varphi(\exp(x))$.*

    (5) *For any $X \in G, y \in \mathfrak{g}$, we have $X \exp(y) X^{-1} = \exp(\mathrm{Ad}\, X . y)$, where Ad is the adjoint action of G on $\mathfrak{g}$ defined by (2.3).*

Proof. The first statement is immediate from the definition; the second follows from the first one and inverse function theorem. The third statement is again immediate corollary of the definition ($\exp(tx)$ is a one-parameter subgroup in G).

Statement 4 follows from the uniqueness of one-parameter subgroup. Indeed, $\varphi(\exp(tx))$ is a one-parameter subgroup in G_2 with tangent vector at identity $\varphi_*(\exp_*(x)) = \varphi_*(x)$. Thus, $\varphi(\exp(tx)) = \exp(\varphi_*(x))$.

The last statement is a special case of the previous one: the map $Y \mapsto XYX^{-1}$ is a morphism of Lie groups $G \to G$. $\square$

Comparing this with Theorem 2.28, we see that we have many of the same results. Notable exception is that we have no analog of the statement that if $xy = yx$, then $\exp(x) \exp(y) = \exp(y) \exp(x)$. In fact the statement does not make sense for general groups, as product xy is not defined. A proper analog of this statement will be proved later (Theorem 3.33).

Remark 3.8. In general, the exponential map is not surjective — see Exercise 3.1. However, it can be shown that for compact Lie groups, the exponential map is surjective.

Proposition 3.9. *Let G_1, G_2 be connected Lie groups. Then any Lie group morphism $\varphi\colon G_1 \to G_2$ is uniquely determined by the linear map $\varphi_*\colon \mathfrak{g}_1 \to \mathfrak{g}_2$.*

Proof. By Theorem 3.7, φ_* determines φ in a neighborhood of identity in G_1. But by Corollary 2.9, any neighborhood of the identity generates G_1. $\square$

Example 3.10. Let $G = \mathrm{SO}(3, \mathbb{R})$. Then the corresponding Lie algebra is $\mathfrak{so}(3, \mathbb{R})$ consists of skew-symmetric 3×3 matrices. One possible choice of a basis in $\mathfrak{so}(3, \mathbb{R})$ is

$$(3.1) \qquad J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad J_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & & 0 \end{pmatrix}, \qquad J_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

We can explicitly describe the corresponding subgroups in G. Namely,

$$\exp(tJ_x) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}$$

is rotation around x-axis by angle t; similarly, J_y, J_z generate rotations around y, z axes. The easiest way to show this is to note that such rotations do form a one-parameter subgroup; thus, they must be of the form $\exp(tJ)$ for some $J \in \mathfrak{so}(3, \mathbb{R})$, and then compute the derivative to find J.

Since $\mathrm{SO}(3, \mathbb{R})$ is connected, by Theorem 3.7, elements of the form $\exp(tJ_x)$, $\exp(tJ_y)$, $\exp(tJ_z)$ generate a neighborhood of identity in $\mathrm{SO}(3, \mathbb{R})$ and thus, by Corollary 2.9, they generate the whole group $\mathrm{SO}(3, \mathbb{R})$. For this reason, it is not uncommon to refer to J_x, J_y, J_z as "infinitesimal generators" of $\mathrm{SO}(3, \mathbb{R})$. Thus, in a certain sense $\mathrm{SO}(3, \mathbb{R})$ is generated by three elements.

3.2. The commutator

So far, we have considered $\mathfrak{g} = T_1 G$ as a vector space with no additional structure. However, since the exponential map locally identifies G with $\mathfrak{g}$, the multiplication in G defines a certain operation in $\mathfrak{g}$. Namely, for sufficiently small $x, y \in \mathfrak{g}$, the product $\exp(x)\exp(y)$ will be close to $1 \in G$ and thus can be written in the form

$$\exp(x)\exp(y) = \exp(\mu(x, y))$$

for some smooth map $\mu\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ defined in a neighborhood of $(0, 0)$. The map μ is sometimes called the *group law in logarithmic coordinates*.

Lemma 3.11. *The Taylor series for μ is given by*

$$\mu(x, y) = x + y + \lambda(x, y) + \dots$$

where $\lambda\colon \mathfrak{g} \times \mathfrak{g}$ is a bilinear skew-symmetric map, and dots stand for the terms of order ≥ 3.

Proof. Any smooth map can be written in the form $\alpha_1(x) + \alpha_2(y) + Q_1(x) + Q_2(y) + \lambda(x, y) + \dots$, where α_1, α_2 are linear maps $\mathfrak{g} \to \mathfrak{g}$, Q_1, Q_2 are quadratic, and λ is bilinear. Letting $y = 0$, we see that $\mu(x, 0) = x$, which gives $\alpha_1(x) = x, Q_1(x) = 0$; similar argument shows that $\alpha_2(y) = y, Q_2(y) = 0$. Thus, $\mu(x, y) = x + y + \lambda(x, y) + \dots$.

To show that λ is skew-symmetric, it suffices to check that $\lambda(x, x) = 0$. But $\exp(x)\exp(x) = \exp(2x)$, so $\mu(x, x) = x + x$. $\square$

For reasons that will be clear in the future, it is traditional to denote $[x, y] = 2\lambda(x, y)$, so we have

(3.2) $$\exp(x) \exp(y) = \exp(x + y + \frac{1}{2}[x, y] + \dots)$$

for some bilinear skew-symmetric map $[\,,\,] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. This map is called the *commutator*.

Thus, we see that for any Lie group, its tangent space at identity $\mathfrak{g} = T_1 G$ has a canonical skew-symmetric bilinear operation, which appears as the lowest non-trivial term of the Taylor series for multiplication in G. This operation has the following properties.

Proposition 3.12.

(1) *For any morphism of Lie groups $\varphi \colon G_1 \to G_2$, the corresponding map of Lie algebras preserves the commutator:*

$$\varphi_*[x, y] = [\varphi_* x, \varphi_* y] \qquad \text{for any } x, y \in \mathfrak{g}_1$$

(2) *The adjoint action preserves the commutator:* $\mathrm{Ad}\, g([x, y]) = [\mathrm{Ad}\, g.x, \mathrm{Ad}\, g.y]$

(3)

(3.3) $$\exp(x) \exp(y) \exp(-x) \exp(-y) = \exp([x, y] + \dots)$$

where dots stand for terms of degree three and higher.

Proof. The first statement is immediate from definition of commutator (3.2) and the fact that every morphism of Lie groups commutes with the exponential map (Theorem 3.7). The second follows from the first and the fact that for any $g \in G$, the map $\mathrm{Ad}_g \colon G \to G$ is a morphism of Lie groups.

The last formula is proved by explicit computation using (3.2). $\qquad\square$

This theorem shows that the commutator in $\mathfrak{g}$ is closely related with the group commutator in G, which explains the name. It also shows that if G is commutative, then $[x, y] = 0$ for all $x, y \in \mathfrak{g}$. In this case we say that $\mathfrak{g}$ is a *commutative*, or *abelian*, Lie algebra.

Example 3.13. Let $G \subset \mathrm{GL}(n)$, so that $\mathfrak{g} \subset \mathfrak{gl}(n)$. Then the commutator is given by $[x, y] = xy - yx$. Indeed, using (3.3) and keeping only linear and bilinear terms, we can write $(1 + x + \dots)(1 + y + \dots)(1 - x + \dots)(1 - y + \dots) = 1 + [x, y] + \dots$ which gives $[x, y] = xy - yx$.

3.3. Adjoint action and Jacobi identity

So far, for a Lie group G, we have defined a bilinear operation on $\mathfrak{g} = T_1 G$, which is obtained from the multiplication on G. An obvious question is whether the associativity of multiplication gives rise to some identities for the commutator. In this section we will answer this question; as one might expect, the answer is "yes".

By results of Proposition 3.12, any map φ of Lie groups gives rise to a map φ_* of Lie algebras which preserves the commutator. Let us apply it to the adjoint action defined in Section 2.4, which can be considered as a morphism of Lie groups

(3.4) $$\mathrm{Ad} \colon G \to \mathrm{GL}(\mathfrak{g}).$$

Lemma 3.14. *Denote by* $\mathrm{ad} = \mathrm{Ad}_* \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ *the map of Lie algebras corresponding to the map* (3.4). *Then*

(1) $\mathrm{ad}\, x.y = [x, y]$

(2) $\mathrm{Ad}(\exp x) = \exp(\mathrm{ad}\, x)$ *as operators* $\mathfrak{g} \to \mathfrak{g}$.

Proof. By definition of Ad, we have $\operatorname{Ad} g.y = \frac{d}{dt} g \exp(ty) g^{-1}$. Thus, we see that ad is defined by

$$\operatorname{ad} x.y = \frac{d}{ds} \frac{d}{dt} \exp(sx) \exp(ty) \exp(-sx)|_{t=s=0}$$

On the other hand, by (3.3), $\exp(sx) \exp(ty) \exp(-sx) = \exp(ty + ts[x, y] + \dots)$. Combining these two results, we see that $\operatorname{ad} x.y = [x, y]$.

The second part is immediate from Theorem 3.7. $\qquad \square$

Theorem 3.15. *Let the commutator* $[\,,\,] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ *be defined by* (3.2). *Then it satisfies the following identity, called* Jacobi identity:

$$(3.5) \qquad\qquad [x, [y, z]] = [[x, y], z] + [y, [x, z]].$$

This identity can also be written in any of the following equivalent forms:

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$$
$$(3.6) \qquad\qquad \operatorname{ad} x.[y, z] = [\operatorname{ad} x.y, z] + [y, \operatorname{ad} x.z]$$
$$\operatorname{ad}[x, y] = \operatorname{ad} x \operatorname{ad} y - \operatorname{ad} y \operatorname{ad} x.$$

Proof. Since Ad is a morphism of Lie groups $G \to \operatorname{GL}(\mathfrak{g})$, by Proposition 3.12, $\operatorname{ad} \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ must preserve commutator. But the commutator in $\mathfrak{gl}(\mathfrak{g})$ is given by $[A, B] = AB - BA$ (see Example 3.13), so $\operatorname{ad}[x, y] = \operatorname{ad} x \operatorname{ad} y - \operatorname{ad} y \operatorname{ad} x$, which proves the last formula of (3.6).

Equivalence of all forms of Jacobi identity is left as an exercise to the reader (see Exercise 3.3). $\qquad \square$

Definition 3.16. A Lie algebra is a vector space with a bilinear map $[\,,\,] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which is skew-symmetric and satisfies Jacobi identity.

A morphism of Lie algebras is a linear map $f \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ which preserves the commutator.

Using this definition, we can summarize much of the results of the previous section in the following theorem.

Theorem 3.17. *Let G be a Lie group. Then $\mathfrak{g} = T_1 G$ has a canonical structure of a Lie algebra with commutator defined by* (3.2); *we will denote this Lie algebra by* $\operatorname{Lie}(G)$. *Every morphism of Lie groups $\varphi \colon G_1 \to G_2$ defines a morphism of Lie algebras $\varphi_* \colon \mathfrak{g}_1 \to \mathfrak{g}_2$, so we have a map* $\operatorname{Hom}(G_1, G_2) \to \operatorname{Hom}(\mathfrak{g}_1, \mathfrak{g}_2)$; *if G_1 is connected, then this map is injective:* $\operatorname{Hom}(G_1, G_2) \subset \operatorname{Hom}(\mathfrak{g}_1, \mathfrak{g}_2)$.

3.4. Subalgebras, ideals, and center

In the last section, we have shown that for every Lie group G the vector space $\mathfrak{g} = T_1 G$ has a canonical structure of a Lie algebra, and every morphism of Lie groups gives rise to a morphism of Lie algebras.

Continuing the study of this correspondence between groups and algebras, we define analogs of Lie subgroups and normal subgroups

Definition 3.18. Let $\mathfrak{g}$ be a Lie algebra. A subspace $\mathfrak{h} \subset \mathfrak{g}$ is called a *Lie subalgebra* if it is closed under commutator, i.e. for any $x, y \in \mathfrak{h}$, we have $[x, y] \in \mathfrak{h}$. A subspace $\mathfrak{h} \subset \mathfrak{g}$ is called an *ideal* if for any $x \in \mathfrak{g}, y \in \mathfrak{h}$, we have $[x, y] \in \mathfrak{h}$.

It is easy to see that if $\mathfrak{h}$ is an ideal, then $\mathfrak{g}/\mathfrak{h}$ has a canonical structure of a Lie algebra.

Theorem 3.19. *Let G be a Lie group with Lie algebra $\mathfrak{g}$.*

(1) *Let H be a Lie subgroup in G. Then $\mathfrak{h} = T_1 H$ is a Lie subalgebra in $\mathfrak{g}$.*

(2) *Let H be a normal Lie subgroup in G. Then $\mathfrak{h} = T_1 H$ is an ideal in $\mathfrak{g}$, and $\operatorname{Lie}(G/H) = \mathfrak{g}/\mathfrak{h}$. Conversely, if H is a connected Lie subroup in G and $\mathfrak{h} = T_1 H$ is an ideal in $\mathfrak{g}$, then H is normal.*

Proof. ??????????????????????????? □

3.5. Lie algebra of vector fields

In this section, we illustrate the theory developed above in the example of the group $\operatorname{Diff}(M)$ of diffeomorphisms of a manifold M. This is not, strictly speaking, a Lie group (it is infinite-dimensional), but in many ways it is similar to Lie groups. For example, it is easy to define what a smooth map from some group G to $\operatorname{Diff}(M)$ is: it is the same as an action of G on M by diffeomorphisms.

Ignoring the technical problem with infinite-dimensionality for now, let us try to see what is the natural analog of the Lie algebra for the group $\operatorname{Diff}(M)$. It should be the tangent space at the identity; thus, its elements are derivatives of one-parameter families of diffeomorphisms.

Let $\varphi^t \colon M \to M$ be a one-parameter family of diffeomorphisms. Then, for every point $m \in M$, $\varphi^t(m)$ is a curve in M and thus $\frac{d}{dt}\varphi^t(m) \in T_m M$ is a tangent vector to M at m. In other words, $\frac{d}{dt}\varphi^t$ is a vector field on M. Thus, it is natural to define the Lie algebra of $\operatorname{Diff}(M)$ to be the space $\operatorname{Vect}(M)$ of all smooth vector fields on M.

What is the exponential map? If $\xi \in \operatorname{Vect}(M)$ is a vector field, then $\exp(t\xi)$ should be a one-parameter family of diffeomorphisms whose derivative is vector field ξ. So this is the solution to a differential equation

$$\frac{d}{dt}\varphi^t(m)|_{t=0} = \xi(m).$$

In other words, φ^t is the flow of the vector field ξ for time t. We will denote it by

$$\exp(t\xi) = \Phi_\xi^t.$$

This may not be defined globally, but for the moment, let us ignore this problem.

What is the commutator $[\xi, \eta]$? By (3.3), we need to consider $\Phi_\xi^t \Phi_\eta^s \Phi_{-\xi}^t \Phi^s -\eta$. It is well-known that this might not be the identity (if a plane flies 500 miles north, then 500 miles west, then 500 miles south, then 500 miles east, then it does not necessarily lands at the same spot it started—because Earth is not flat). By analogy with (3.3), we expect that this expression can be written in the form $1 + ts[\xi, \eta] + \dots$ for some vector field $[\xi, \eta]$. This is indeed so, as the following proposition shows.

Proposition 3.20.

(1) *Let $\xi, \eta \in \operatorname{Vect}(M)$ be vector fields on M. Then there exists a unique vector field which we will denote by $[\xi, \eta]$ such that*

(3.7)
$$\Phi_\xi^t \Phi_\eta^s \Phi_{-\xi}^t \Phi_{-\eta}^s = \Phi_{[\xi,\eta]}^{ts} + \dots,$$

where dots stand for the terms of order 3 and higher in s, t.

(2) *The commutator (3.7) defines on the space of vector fields a structure of an (infinite-dimensional) Lie algebra.*

(3) *The commutator can also be defined by any of the following formulas:*

(3.8)
$$[\xi, \eta] = \frac{d}{dt}(\Phi^t(\xi))_* \eta$$

(3.9)
$$\partial_{[\xi,\eta]} f = \partial_\eta(\partial_\xi f) - \partial_\xi(\partial_\eta f), \qquad f \in C^\infty(M)$$

(3.10)
$$\left[\sum f_i \partial_i, \sum g_j \partial_j\right] = \sum_{i,j}(g_i \partial_i(f_j) - f_i \partial_i(g_j))\partial_j$$

where $\partial_\xi(f)$ *is the derivative of a function f in the direction of the vector field ξ, and* $\partial_i = \frac{\partial}{\partial x^i}$ *for some local coordinate system $\{x^i\}$.*

The first two parts are, of course, to be expected, by analogy with finite-dimensional situation. However, since $\mathrm{Diff}(M)$ is not a finite-dimensional Lie group, we can not just refer to Theorem 3.17 but need to give a separate proof. Such a proof (and also the proof of the last part) can be found in any good book on differential geometry, for example in [**23**].

Remark 3.21. In many books the definition of commutator of vector fields differs by sign from the one given here. Both versions define on the space of vector fields a structure of Lie algebra, so it is a matter of choice which of the definitions to use. However, in our opinion the definition here — which naturally arises from the multiplication in the group of diffeomorphisms — is more natural, so we use it. Thus, when using results from other books, be sure to double-check which definition of commutator they use for vector fields.

The reason for the appearance of the minus sign is that the action of a diffeomorphism $\Phi\colon M \to M$ on functions on M is given by $(\Phi f)(m) = f(\Phi^{-1}m)$ (note the inverse!); thus, the derivative $\partial_\xi f = -\frac{d}{dt}\Phi_\xi^t f$. For example, if $\xi = \partial_x$ is the constant vector field on $\mathbb{R}$, then the flow on points is given by $\Phi^t\colon x \mapsto x + t$, and on functions it is given by $(\Phi^t f)(x) = f(x - t)$, so $\partial_x f = -\frac{d}{dt}\Phi^t f$.

Theorem 3.22. *Let G be a finite dimensional Lie group acting on a manifold M, so we have a map $\rho\colon G \to \mathrm{Diff}(M)$. Then*

(1) *This action defines a linear map $\rho_* \colon \mathfrak{g} \to \mathrm{Vect}(M)$.*

(2) *The map ρ_* is a morphism of Lie algebras: $\rho_*[x,y] = [\rho_*(x), \rho_*(y)]$, where the commutator in the right-hand side is the commutator of vector fields.*

If $\mathrm{Diff}(M)$ were a Lie group, this result would be a special case of Proposition 3.12. Since $\mathrm{Diff}(M)$ is not a Lie group, we need to give a separate proof, suitably modifying the proof of Proposition 3.12. We leave this as an exercise to the reader.

We will refer to the map $\rho_* \colon \mathfrak{g} \to \mathrm{Vect}(M)$ as *action of $\mathfrak{g}$ by vector fields on M*.

Example 3.23. Consider the standard action of $\mathrm{GL}(n, \mathbb{R})$ on $\mathbb{R}^n$. Considering $\mathbb{R}^n$ as a manifold and forgetting the structure of a vector space, we see that each element $a \in \mathfrak{gl}(n, \mathbb{R})$ defines a vector field on $\mathbb{R}^n$. An easy calculation shows that this vector field is given by $v_a(x) = \sum a_{ij} x_j \partial_i$, where $x_1, \ldots x_n$ are the coordinates of a point x in the standard basis of $\mathbb{R}^n$, and $\partial_i = \frac{\partial}{\partial x_i}$.

Another important example is the action of G on itself by left multiplication.

Proposition 3.24. *Consider the action of a Lie group G on itself by left multiplication $L(g)\colon G \to G$ defined by $L(g).h = gh$. Then for every $x \in \mathfrak{g}$, the corresponding vector field $\xi = L_*(x) \in \mathrm{Vect}(G)$ is the* right-invariant *vector field such that $\xi(1) = x$.*

Proof. Consider the one-parameter subgroup $\exp(tx) \subset G$. By Proposition 3.6, for any $g \in G$, we have $L_*(x)(g) = \frac{d}{dt}|_{t=0}(\exp(tx)g) = xg$. $\qquad\square$

Corollary 3.25. *The isomorphism* $\mathfrak{g} \simeq \{$*right-invariant vector fields on* $G\}$ *defined in Theorem* 2.25 *is an isomorphism of Lie algebras.*

An analog of this statement for left-invariant fileds is given in Exercise 3.4

3.6. Stabilizers and the center

Having developed the basic theory of Lie algebras, we can now go back to proving various results about Lie groups which were announced in Chapter 2, such as proving that the stabilizer of a point is a Lie subgroup.

Theorem 3.26. *Let* G *be a Lie group acting on a manifold* M, *and let* $m \in M$.

(1) *The stabilizer* $H = \mathrm{Stab}(m) = \{g \in G \mid gm = m\}$ *is a Lie subgroup in* G, *with Lie algebra* $\mathfrak{h} = \{x \in \mathfrak{g} \mid \rho_*(x)(m) = 0\}$, *where* $\rho_*(x)$ *is the vector field on* M *corresponding to* x.

(2) *The map* $G/\mathrm{Stab}(m) \to M$ *given by* $g \mapsto g.m$ *is an immersion. Thus, the orbit* $\mathrm{O}_m = G \cdot m$ *is an immersed submanifold in* M, *with tangent space* $T_m M = \mathfrak{g}/\mathfrak{h}$.

Proof. As in the proof of Theorem 2.29, it suffices to show that in some neighborhood U of $1 \in G$ the intersection $U \cap H$ is a smooth manifold with tangent space $T_1 H = \mathfrak{h}$.

It easily follows from (3.9) that $\mathfrak{h}$ is closed under commutator, so it is a Lie subalgebra in $\mathfrak{g}$. Also, since for $h \in \mathfrak{h}$, the corresponding vector field $\xi = \rho_*(x)$ vanishes at m, we have $\rho(\exp(th))(m) = \Phi^t_\xi(m) = m$, so $\exp(th) \in \mathrm{Stab}(m)$.

Now let us choose some vector subspace (not a subalgebra!) $u \subset \mathfrak{g}$ which is complementary to $\mathfrak{h}$: $\mathfrak{g} = \mathfrak{h} \oplus u$. Since the kernel of the map $\rho_* \colon \mathfrak{g} \to T_m M$ is $\mathfrak{h}$, the restriction of this map to u is injective. By implicit function theorem, this implies that the map $u \to M \colon y \mapsto \rho(\exp(y))(m)$ is injective for sufficiently small $y \in u$, so $\exp(y) \in \mathrm{Stab}(m) \iff y = 0$.

Since in a sufficiently small neighborhood U of 1 in G, any element $g \in U$ can be uniquley written in the form $\exp(y)\exp(x)$, $y \in u, x \in \mathfrak{h}$ (which follows from inverse function theorem), and $\exp(y)\exp(x)m = \exp(y)m$ we see that $g \in \mathrm{Stab}(m) \iff g \in \exp(\mathfrak{h})$. Since $\exp\mathfrak{h}$ is a smooth submanifold in a neighborhood of $1 \in G$, we see that $\mathrm{Stab}(m)$ is smooth.

The same proof also shows that we have an isomorphism $T_1(G/\mathrm{Stab}(m)) = \mathfrak{g}/\mathfrak{h} \simeq u$, so injectivity of the map $\rho_* \colon u \to T_m M$ shows that the map $G/\mathrm{Stab}(m) \to M$ is an immersion. $\square$

This theorem immediately implies a number of corollaries. In particular, we get the following result which was announced in Theorem 2.12.

Corollary 3.27. *Let* $f \colon G_1 \to G_2$ *be a morphism of Lie groups, and* $f_* \colon \mathfrak{g}_1 \to \mathfrak{g}_2$—*the corresponding morphism of Lie algebras. Then* $\mathrm{Ker}\, f$ *is a Lie subgroup with Lie algebra* $\mathrm{Ker}\, f_*$, *and the map* $G_1/\mathrm{Ker}\, f \to G_2$ *is an immersion. If* $\Im f$ *is closed, then we have an isomorphism* $\Im f \simeq G_1/\mathrm{Ker}\, f$.

Proof. Consider the action of G_1 on G_2 given by $\rho(g).h = f(g)h$, $g \in G_1, h \in G_2$. Then the stabilizer of $1 \in G_2$ is exactly $\mathrm{Ker}\, f$, so by the previous theorem, it is a Lie group with Lie algebra $\mathrm{Ker}\, f_*$. $\square$

Corollary 3.28. *Let* V *be a representation of a group* G, *and* $v \in V$. *Then the stabilizer* $\mathrm{Stab}(v)$ *is a Lie subgroup in* G *with Lie algebra* $\{x \in \mathfrak{g} \mid x.v = 0\}$.

Example 3.29. Let V be a vector space with a bilinear form B, and let

$$\mathrm{O}(V, B) = \{g \in \mathrm{GL}(V) \mid B(g.v, g.w) = B(v, w) \text{ for all } v, w\}$$

be the group of symmetries of B. Then it is a Lie group with Lie algebra

$$\mathfrak{o}(V, B) = \{x \in \mathfrak{gl}(V) \mid B(x.v, w) + B(v, x.w) = 0 \text{ for all } v, w\}$$

Indeed, define the action of G on the space of bilinear forms by $(gF)(v, w) = F(g^{-1}.v, g^{-1}.w)$. Then $\mathrm{O}(V, B)$ is exactly the stabilizer of B, so by Corollary 3.28, it is a Lie group. Since the corresponding action of $\mathfrak{g}$ is given by $(xF)(v, w) = -F(x.v, w) - F(v, x.w)$ (which follows from Leibnitz rule), we get the formula for $\mathfrak{o}(V, B)$.

As special cases, we recover the usual groups $\mathrm{O}(n, \mathbb{R})$ and $\mathrm{Sp}(2n, \mathbb{R})$.

Example 3.30. Let A be a finite dimensional associative algebra. Then the group of all automorphisms of A

$$\mathrm{Aut}(A) = \{g \in \mathrm{GL}(A) \mid (ga) \cdot (gb) = g(a \cdot b) \text{ for all } a, b \in A\}$$

is a Lie group with Lie algebra

(3.11) $$\mathrm{Der}(A) = \{x \in \mathfrak{gl}(A) \mid (x.a)b + a(x.b) = x.(ab) \text{ for all } a, b \in A\}$$

(this Lie algebra is called the algebra of *derivations* of A).

Indeed, if we consider the space W of all linear maps $A \otimes A \to A$ and define the action of G by $(g.f)(a \otimes b) = gf(g^{-1}a \otimes g^{-1}b)$ then $\mathrm{Aut}\, A = \mathrm{Stab}(\mu)$, where $\mu \colon A \otimes A \to A$ is the multiplication. So by Corollary 3.28, $\mathrm{Aut}(A)$ is a Lie group with Lie algebra $\mathrm{Der}(A)$.

The same argument also shows that for a finite-dimensional Lie algebra $\mathfrak{g}$, the group

(3.12) $$\mathrm{Aut}(\mathfrak{g}) = \{g \in \mathrm{GL}(\mathfrak{g}) \mid [ga, gb] = g[a, b] \text{ for all } a, b \in \mathfrak{g}\}$$

is a Lie group with Lie algebra

(3.13) $$\mathrm{Der}(\mathfrak{g}) = \{x \in \mathfrak{gl}(\mathfrak{g}) \mid [x.a, b] + [a, x.b] = x.[a, b] \text{ for all } a, b \in \mathfrak{g}\}$$

called the Lie algebra of derivations of $\mathfrak{g}$. This algebra will play an important role in the future.

Finally, we can show that the center of G is a Lie subgroup.

Definition 3.31. Let $\mathfrak{g}$ be a Lie algebra. The center of $\mathfrak{g}$ is defined by

$$\mathfrak{z}(\mathfrak{g}) = \{x \in \mathfrak{g} \mid [x, y] = 0 \ \forall y \in \mathfrak{g}\}.$$

Obviously, $\mathfrak{z}(\mathfrak{g})$ is an ideal in $\mathfrak{g}$.

Theorem 3.32. *Let G be a connected Lie group. Then its center $Z(G)$ is a Lie subgroup with Lie algebra $\mathfrak{z}(\mathfrak{g})$.*

Proof. It follows from the identity $\exp(\mathrm{Ad}\, g.tx) = g \exp(tx) g^{-1}$ that g commutes with all elements of one-parameter subgroup $\exp(tx)$ iff $\mathrm{Ad}\, g.x = 0$. Since for a connected Lie group, elements of the form $\exp(tx)$ generate G, we see that $g \in Z(G) \iff \mathrm{Ad}\, g.x = 0$ for all $x \in \mathfrak{g}$. In other words, $Z(G) = \mathrm{Ker}\, \mathrm{Ad}$, where $\mathrm{Ad} \colon \to \mathrm{GL}(\mathfrak{g})$ is given by the adjoint action. Now the result follows from Corollary 3.27. $\qquad \square$

3.7. Campbell–Hausdorff formula

So far, we have shown that the multiplication in G defines the commutator in $\mathfrak{g} = T_1 G$. However, definition of commutator (3.2) only used the lowest non-trivial term of the group law in logarithmic coordinates. Thus, it might be expected that higher terms give more operations on $\mathfrak{g}$. However, it turns out that it is not so: the whole group law is completely determined by the lowest term, i.e. by the commutator. The following theorem gives the first indication of this.

Theorem 3.33. *Let* $x, y \in \mathfrak{g}$ *be such that* $[x, y] = 0$. *Then* $\exp(x) \exp(y) = \exp(x + y) = \exp(y) \exp(x)$.

Proof. The most instructive (but not the easiest; see Exercise 3.12) way of deducing this theorem is as follows. Let ξ, η be right-invariant vector fields corresponding to x, y respectively. By Corollary 3.25, $[\xi, \eta] = 0$. By (3.8), it implies that $\frac{d}{dt}(\Phi_\xi^t)_* \eta = 0$, which implies that $(\Phi_\xi^t)_* \eta = \eta$, i.e. the flow of ξ preserves field η. This, in turn, implies that Φ_ξ^t commutes with the flow of field η, so $\Phi_\xi^t \Phi_\eta^s \Phi_\xi^{-t} = \Phi_\eta^s$. Applying this to point $1 \in G$ and using Proposition 3.6, we get $\exp(tx) \exp(sy) \exp(-tx) = \exp(sy)$, so $\exp(tx), \exp(sy)$ commute for all values of s, t.

In particular, this implies that $\exp(tx) \exp(ty)$ is a one-parameter subgroup; computing the tangent vector at $t = 0$, we see that $\exp(tx) \exp(ty) = \exp(t(x + y))$. $\qquad\square$

In fact, similar ideas allow one to prove the following general statement, known as Campbell–Hausdorff formula.

Theorem 3.34. *For small enough* $x, y \in \mathfrak{g}$ *one has*

$$\exp(x) \exp(y) = \exp(\mu(x, y))$$

for some $\mathfrak{g}$-*valued function* $\mu(x, y)$ *which is given by the following series which converges in some neighborhood of* $(0, 0)$:

$$(3.14) \qquad\qquad \mu(x, y) = x + y + \sum_{n \geq 2} \mu_n(x, y)$$

where $\mu_n(x, y)$ *is a Lie polynomial in* x, y *of degree* n, *i.e. an expression consisting of commutators of* x, y, *their commutators, etc., of total degree* n *in* x, y. *This expression is universal: it does not depend on the Lie algebra* $\mathfrak{g}$ *or on the choice of* x, y.

It is possible to write the expression for μ explicitly (see, e.g., [**10**]). However, this is rarely useful, so we will only write the first several terms:

$$(3.15) \qquad\qquad \mu(x, y) = x + y + \frac{1}{2}[x, y] + \frac{1}{12}\big([x, [x, y]] + [y, [y, x]]\big) + \dots$$

The proof of this theorem is rather long. The key idea is writing the differential equation for the function $Z(t) = \mu(tx, y)$; the right-hand side of this equation will be a power series of the form $\sum a_n t^n (\operatorname{ad} x)^n y$. Solving this differential equation by power series gives the Campbell–Hausdorff formula. Details of the proof can be found, for example, in [**10**, Section 1.6].

Corollary 3.35. *The group operation in a connected Lie group* G *can be recovered from the commutator in* $\mathfrak{g} = T_1 G$.

Indeed, locally the group law is determined by Campbell–Hausdorff formula, and G is generated by a neighborhood of 1.

Note, however, that by itself this corollary does not allow us to recover the group G from its Lie algebra $\mathfrak{g}$: it only allows us to determine the group law provided that we already know the structure of G as a manifold.

3.8. Fundamental theorems of Lie theory

Let us summarize the results we have so far about the relation between Lie groups and Lie algebras:

(1) Every Lie group G defines a Lie algebra $\mathfrak{g} = T_1 G$, with commutator defined by (3.2); we will write $\mathfrak{g} = \mathrm{Lie}(G)$. Every morphism of Lie groups $\varphi \colon G_1 \to G_2$ defines a morphism of Lie algebras $\varphi_* \colon \mathfrak{g}_1 \to \mathfrak{g}_2$. For connected G_1, the map

$$\mathrm{Hom}(G_1, G_2) \to \mathrm{Hom}(\mathfrak{g}_1, \mathfrak{g}_2)$$

$$\varphi \mapsto \varphi_*$$

is injective. (Here $\mathrm{Hom}(\mathfrak{g}_1, \mathfrak{g}_2)$ is the set of Lie algebra morphisms.)

(2) As a special case of the previous, every Lie subgroup $H \subset G$ defines a Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$.

(3) The group law in connected Lie group G can be recovered from the commutator in $\mathfrak{g}$; however, we do not yet know whether we can also recover the geometry of G from $\mathfrak{g}$.

However, this still leaves a number of questions:

(1) Given a morphism of Lie algebras $\mathfrak{g}_1 \to \mathfrak{g}_2$, where $\mathfrak{g}_1 = \mathrm{Lie}(G_1), \mathfrak{g}_2 = \mathrm{Lie}(G_2)$, can this morphism be always lifted to a morphism of the Lie groups?

(2) Given a Lie subalgebra $\mathfrak{h} \subset \mathfrak{g} = \mathrm{Lie}(G)$, does there exist a corresponding Lie subgroup $H \subset G$?

(3) Can every Lie algebra be obtained as a Lie algebra of a Lie group?

It is clear that in order to have any chance of success, we must limit ourselves to the consideration of connected Lie groups. However, it turns out that even with this restriction the answer to the first two questions is negative, as is shown by the following counterexamples.

Example 3.36. Let $G_1 = S^1 = \mathbb{R}/\mathbb{Z}, G_2 = \mathbb{R}$. Then the Lie algebras are $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathbb{R}$ with zero bracket. Consider the identity map $\mathfrak{g}_1 \to \mathfrak{g}_2 \colon a \mapsto a$. Then the corresponding morphism of Lie groups, if exists, should be given by $\theta \mapsto \theta$; on the other hand, it must also satisfy $f(\mathbb{Z}) = \{0\}$. Thus, this morphism of Lie algebras can not be lifted to a morphism of Lie groups.

Example 3.37. Let $G = T^2 = \mathbb{R}^2/\mathbb{Z}^2$; then $\mathfrak{g} = \mathbb{R}^2$ with zero bracket. Choose an irrational number α and consider the following subspace $\mathfrak{h} = \{(t, \alpha t) \mid t \in \mathbb{R}\} \subset \mathfrak{g}$. Since $\mathfrak{g}$ is commutative, $\mathfrak{h}$ is automatically a one-dimensional Lie subalgebra. However, there is no Lie subgroup corresponding to $\mathfrak{h}$: the image of $\mathfrak{h}$ under the exponential map $\mathbb{R}^2 \to \mathbb{R}^2/\mathbb{Z}^2$ is everywhere dense on the torus (see Example 2.13).

It turns out, however, that it is possible to fix these problems and reformulate questions 1 and 2 above so that the answer is positive.

Theorem 3.38. *If G_1 is a connected, simply connected Lie group then* $\mathrm{Hom}(G_1, G_2) = \mathrm{Hom}(\mathfrak{g}_1, \mathfrak{g}_2)$.

In order to be able to positively answer question 2 (about correspondence between subgroups and subalgebras), we need to extend our definition of Lie subgroup to allow for things like irrational winding of the torus.

Definition 3.39. An immersed subgroup in a Lie group G is a subset $H \subset G$ such that

(1) H is a subgroup of G

(2) H is an immersed submanifold: $H = \mathrm{Im}(i)$ for some injective immersion $i \colon \tilde{H} \hookrightarrow G$.

Remark 3.40. The terminology is not quite standard: the same object we called an immersed subgroup in other books can be called "analytic subgroup" or sometimes even a "Lie subgroup".

Informally, this means that we can change topology of H so that H with this new topology is a manifold (we denote it by $\tilde{H}$) and the inclusion $\tilde{H} \hookrightarrow G$ is an immersion. Recall that an immersion

is a smooth map of manifolds $f\colon X \to Y$ such that each point $x \in X$ has a neighborhood $U \ni x$ such that $f\colon U \to f(U)$ is a diffeomorphism.

Example 3.41. Let $\varphi\colon G_1 \to G_2$ be a morphism of Lie groups such that $\varphi_*\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is injective. Then φ is an immersion, and $\varphi(G_1)$ is an immersed subgroup in G_2. In particular, every one-parameter subgroup in G is an immersed subgroup.

One easily sees that immersed subgroups have the following properties.

Proposition 3.42. *Let $i\colon \tilde{H} \hookrightarrow G$ be an immersed subgroup. Then*

(1) *$\tilde{H}$ is a Lie group, and i is a morphism of Lie groups*

(2) *$\mathfrak{h} = i_*(T_1\tilde{H})$ is a Lie subalgebra in $\mathfrak{g}$ (it will be denoted by $\mathrm{Lie}(H)$).*

(3) *If $H = i(\tilde{H})$ is closed in G, then H is a Lie subgroup.*

It turns out that this generalization solves all our problems.

Theorem 3.43. *There is a bijection between connected immersed subgroups $H \subset G$ and subalgebras $\mathfrak{h} \subset \mathfrak{g}$. The correspondence is given by $H \to \mathfrak{h} = \mathrm{Lie}(H) = T_1H$.*

Theorems 3.38 and 3.43 are the two fundamental theorems of Lie theory. There is also a third fundamental theorem which will be discussed later (Theorem 3.48).

Our next goal is proving these two theorems. We start by noticing that it is easy to derive the first theorem from the second one.

Proposition 3.44. *Theorem* 3.43 *implies Theorem* 3.38.

Proof. We already know that any morphism of Lie groups defines a morphism of Lie algebras and that for connected G_1, the map $\mathrm{Hom}(G_1, G_2) \to \mathrm{Hom}(\mathfrak{g}_1, \mathfrak{g}_2)$ is injective (see Theorem 3.17). Thus, it remains to show that it is surjective, i.e. that every morphism of Lie algebras $f\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ can be lifted to a morphism of Lie groups $\varphi\colon G_1 \to G_2$ with $\varphi_* = f$.

Define $G = G_1 \times G_2$. Then the Lie algebra of G is $\mathfrak{g}_1 \times \mathfrak{g}_2$. Let $\mathfrak{h} = \{(x, f(x)) \mid x \in \mathfrak{g}_1\} \subset \mathfrak{g}$. This is a subalgebra: it is obviously a subspace, and $[(x, f(x)), (y, f(y))] = ([x, y], [f(x), f(y)]) = ([x, y], f([x, y]))$ (the last identity uses that f is a morphism of Lie algebras). By Theorem 3.43, there is a corresponding immersed subgroup $H = i(\tilde{H}) \subset G_1 \times G_2$. Composing this embedding with the projection $p\colon G_1 \times G_2 \to G_1$, we get a morphism of Lie groups $\pi\colon \tilde{H} \to G_1$, and $\pi_*\colon \mathfrak{h} = \mathrm{Lie}(\tilde{H}) \to \mathfrak{g}_1$ is an isomorphism. By results of Exercise 2.3, π is a covering map. On the other hand, G_1 is simply-connected, and $\tilde{H}$ is connected, so π must be an isomorphism. This implies that $\tilde{H} \to G_1 \times G_2$ is actually a closed submanifold, so $H = \tilde{H}$ is a Lie subgroup, and projection $H \to G_1$ is an isomorphism; thus, we have an inverse map $G_1 \to H$.

Now construct the map $\varphi\colon G_1 \to G_2$ as a composition $G_1 \to H \subset G_1 \times G_2 \to G_2$. By definition, it is a morphism of Lie groups, and $\varphi_*\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is the composition $x \mapsto (x, f(x)) \mapsto f(x)$. Thus, we have lifted f to a morphism of Lie groups. $\qquad\square$

Remark 3.45. In fact, the arguments above can be reversed to deduce Theorem 3.43 from Theorem 3.38. For example, this is the way these theorems are proved in [**18**].

We can now prove Theorem 3.43. The proof we give here is based on the notion of integrable distribution. For reader's convenience, we give the basic definitions here; details can be found in [**23**] or [**25**].

A k-dimensional *distribution* on a manifold M is a k-dimensional subbundle $\mathcal{D} \subset TM$. In other words, at every $p \in M$ we have a k-dimensional subspace $\mathcal{D}_p \subset T_pM$, which smoothly depends on

p. This is a generalization of well-known notion of direction field, commonly used in the theory of differential equations.

An integral manifold for a distribution $\mathcal{D}$ is a k-dimensional submanifold $X \subset M$ such that at every point $p \in X$, we have $T_p X = \mathcal{D}_p$. Again, this is a straightforward generalization of the notion of an integral curve. However, for $k > 1$, existence of integral manifolds (even locally) is not automatic. We say that a distribution $\mathcal{D}$ is *completely integrable* if for every $p \in M$, locally there exists an integral manifold containing p (it is easy to show that such an integral manifold is unique). In this case, $\mathcal{D}$ defines a foliation of M.

The following theorem gives a necesary and sufficient criterion of integrability of a distribution.

Theorem 3.46 (Frobenius integrability criterion). *Let $\mathcal{D}$ be a distribution on M. We say that a vector field ξ is tangent to $\mathcal{D}$ if, for every point $p \in M$, $\xi(p) \in \mathcal{D}_p$. Then $\mathcal{D}$ is completely integrable if and only if the space of vector fields tangent to $\mathcal{D}$ is closed under commutator: for any fields ξ, η tangent to $\mathcal{D}$, the commutator $[\xi, \eta]$ is also tangent to $\mathcal{D}$.*

Proof of this theorem can be found in many books on differential geometry, such as [**23**], and will not be repeated here.

Using this theorem, we can now prove Theorem 3.43.

Proof of Theorem 3.43. Let G be a Lie group, and $\mathfrak{h} \subset \mathfrak{g}$ be a Lie subalgebra. We want to show that there exists an immersed subgroup $H = i(\tilde{H}) \subset G$ with $\mathrm{Lie}(H) = \mathfrak{h}$.

Notice that if such an H exists, then at every point $p \in H$, $T_p H = (T_1 H)p = \mathfrak{h} \cdot p$. Thus, H will be an integral manifold of the distribution $\mathcal{D}^{\mathfrak{h}}$ defined by $\mathcal{D}_p^{\mathfrak{h}} = \mathfrak{h} \cdot p$. Let us use this to construct H.

Lemma 3.47. *For every point $g \in G$, there is locally an integral manifold of the distrubution $\mathcal{D}^{\mathfrak{h}}$ containing g, namely $H^0 \cdot g$, where $H^0 = \exp u$ for some neighborhood u of 0 in $\mathfrak{h}$.*

This lemma can be easily proved using Frobenius theorem. Indeed, the distribution $\mathcal{D}^{\mathfrak{h}}$ is generated by right-invariant vector fields corresponding to elements of $\mathfrak{h}$. Since $\mathfrak{h}$ is closed under $[\,,\,]$, and commutator of right invariant vector fields coincides with the commutator in $\mathfrak{g}$ (Corollary 3.25), this shows that the space of fields tangent to $\mathcal{D}^{\mathfrak{h}}$ is closed under the commutator, and thus $\mathcal{D}^{\mathfrak{h}}$ is completely integrable.

To get an explicit description of the integral manifold, note that by Proposition 3.6, the curve $e^{tx}g$ for $x \in \mathfrak{h}$ is the integral curve for a right invariant vector field corresponding to x and thus this curve must be in the integral manifold. Thus, for small enough $x \in \mathfrak{h}$, $\exp(x)g$ is in the integral manifold passing through g. Comparing dimensions we get the statement of the lemma.

Alternatively, this lemma can also be proved without use of Frobenius theorem but using Campbell–Hausdorff formula instead.

Now that we have proved the lemma, we can construct the immersed subgroup H as the maximal connected integral manifold containing 1. This doesn't have to be a submanifold in G; however, it is not too difficult to show that H is an immersed submanifold (see [**25**]).

The only thing which remains to be shown is why H is a subgroup. To prove it, notice that by the definition of H and the lemma, a point $g \in H$ if and only if we can find a sequence of points $g_0 = 1, g_1, \ldots, g_n = g$ such that $g_{i+1} \in H^0 g_i$. Thus, $g \in H$ iff $g = \exp(x_1) \ldots \exp(x_n)$ for some $x_i \in u \subset \mathfrak{h}$. Thus, H is a subgroup.

$\square$

Thus, we have proved the first two fundamental theorems about relation between Lie groups and Lie algebras. To complete the picture, here is the final theorem.

Theorem 3.48 (Lie's third theorem)**.** *Any finite-dimensional Lie algebra is isomorphic to a Lie algebra of a Lie group.*

Proof. Proof of this theorem is rather complicated. The basic idea is to show that any Lie algebra is isomorphic to a subalgebra in $\mathfrak{gl}(n)$ (this statement is known as Ado theorem), after which we can use Theorem 3.43. However, the proof of Ado theorem is long and requires a lot of structure theory of Lie algebras, some of which will be given in the subsequent chapters. The simplest case is when Lie algebra has no center (that is, $\operatorname{ad} x \neq 0$ for all x), then $x \mapsto \operatorname{ad} x$ gives an embedding $\mathfrak{g} \subset \mathfrak{gl}(\mathfrak{g})$. Proof of general case can be found, e.g., in [**11**]. □

Combining this theorem with the previous ones, we immediately get some useful corollaries.

Corollary 3.49. *For any Lie algebra $\mathfrak{g}$, there is a unique (up to isomorphism) connected simply-connected Lie group G with $\operatorname{Lie}(G) = \mathfrak{g}$. Any other connected Lie group G' with Lie algebra $\mathfrak{g}$ must be of the form G/Z for some discrete central subgroup $Z \subset G$.*

Proof. By Theorem 3.48, there is a Lie group with Lie algebra $\mathfrak{g}$. Taking the universal cover of the connected component of identity in this group (see Theorem 2.5), we see that there exists a connected, simply-connected G with $\operatorname{Lie}(G) = \mathfrak{g}$. By Theorem 3.43, if G' is another connected Lie group with Lie algebra $\mathfrak{g}$, then there is a group homomorphism $G \to G'$ which is locally an isomorphism; thus, by results of Exercise 2.3, $G' = G/Z$ for some discrete central subgroup Z.

Uniqueness of simply-connected group G now follows from $\pi_1(G/Z) = Z$. □

This corollary can be reformulated as follows.

Corollary 3.50. *The categories of finite-dimensional Lie algebras and connected, simply-connected Lie groups are equivalent.*

3.9. Complex and real forms

An interesting application of the correspondence between Lie groups and Lie algebras is the interplay between real and complex Lie algebras and groups.

Definition 3.51. Let $\mathfrak{g}$ be a real Lie algebra. Its *complexification* is the complex Lie algebra $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} = \mathfrak{g} \oplus i\mathfrak{g}$ with the obvious commutator. In this situation, we will also say that $\mathfrak{g}$ is a real form of $\mathfrak{g}_{\mathbb{C}}$.

In some cases, complexification is obvious: for example, if $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{R})$, then $\mathfrak{g}_{\mathbb{C}} = \mathfrak{sl}(n, \mathbb{C})$. The following important example, however, is less obvious.

Example 3.52. Let $\mathfrak{g} = \mathfrak{u}(n)$. Then $\mathfrak{g}_{\mathbb{C}} = \mathfrak{gl}(n, \mathbb{C})$.

Indeed, this immediately follows from the fact that any complex matrix can be uniquely written as a sum of skew-hermitian (i.e., from $\mathfrak{u}(n)$) and a hermitian ($i\mathfrak{u}(n)$).

These notions can be extended to Lie groups. For simplicity, we only consider the case of connected groups.

Definition 3.53. Let G be a connected complex Lie group, $\mathfrak{g} = \operatorname{Lie}(G)$ and let $K \subset G$ be a real Lie subgroup in G such that $\mathfrak{k} = \operatorname{Lie}(K)$ is a real form of $\mathfrak{g}$. Then K is called a *real form* of G

It can be shown (see Exercise 3.15) that if $\mathfrak{g} = \operatorname{Lie}(G)$ is the Lie algebra of a connected complex Lie group G, then every real form $\mathfrak{k} \subset \mathfrak{g}$ can be obtained from a real form $K \subset G$ of the Lie group.

Going in the opposite direction, from a real Lie group to a complex one, is more subtle: there are real Lie groups that can not be obtained as real forms of a complex Lie group. It is still possible, however, to define a complexification $G_{\mathbb{C}}$ for any real Lie group G; however, in general G is not a subgroup of $G_{\mathbb{C}}$. Detailed discussion of this can be found in [**8**, Section I.7].

Example 3.54. The group $G = \mathrm{SU}(n)$ is a compact real form of the complex group $\mathrm{SL}(n, \mathbb{C})$.

This operation, which is trivial at the level of Lie algebras, is highly non-trivial at the level of Lie groups. Lie groups G and $G_{\mathbb{C}}$ may be quite different: for example, $\mathrm{SU}(n)$ is compact while $\mathrm{SL}(n, \mathbb{C})$ is not. On the other hand, it is natural to expect — and is indeed so, as we will show later — that $\mathfrak{g}$ and $\mathfrak{g}_{\mathbb{C}}$ share many algebraic properties, such as semisimplicity. Thus, we may use, for example, compact group $\mathrm{SU}(n)$ to prove some results for non-compact group $\mathrm{SL}(n, \mathbb{C})$. Moreover, since $\mathfrak{sl}(n, \mathbb{R})_{\mathbb{C}} = \mathfrak{sl}(n, \mathbb{C})$, this will also give us results about non-compact real group $\mathrm{SL}(n, \mathbb{R})$. We will give an application of this to the study of representations of Lie groups in the next chapter.

3.10. Example: $\mathfrak{so}(3, \mathbb{R})$, $\mathfrak{su}(2)$, and $\mathfrak{sl}(2, \mathbb{C})$.

In this section, we bring together various explicit formulas related to Lie algebras $\mathfrak{so}(3, \mathbb{R})$, $\mathfrak{su}(2)$, $\mathfrak{sl}(2, \mathbb{C})$. Most of these results have appeared before in various examples and exercises; this section brings them together for reader's convenience. This section contains no proofs: they are left to the reader as exercises.

Basis and commutation relations. A basis in $\mathfrak{so}(3, \mathbb{R})$ is given by matrices

$$(3.16) \qquad J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad J_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad J_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The corresponding one-parameter subgroups in $\mathrm{SO}(3, \mathbb{R})$ are rotations: $\exp(tJ_x)$ is rotation by angle t around x-axis, and similarly for y, z.

The commutation relations are given by

$$(3.17) \qquad [J_x, J_y] = J_z, \qquad [J_y, J_z] = J_x, \qquad [J_z, J_x] = J_y.$$

A basis in $\mathfrak{su}(2)$ is given by so-called Pauli matrices multiplied by i:

$$(3.18) \qquad i\sigma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad i\sigma_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

The commutation relations are given by

$$(3.19) \qquad [i\sigma_1, i\sigma_2] = 2i\sigma_3, \qquad [i\sigma_2, i\sigma_3] = 2i\sigma_1, \qquad [i\sigma_3, i\sigma_1] = 2i\sigma_2.$$

Since $\mathfrak{sl}(2, \mathbb{C}) = \mathfrak{su}(2) \otimes \mathbb{C}$, the same matrices can also be taken as a basis of $\mathfrak{sl}(2, \mathbb{C})$. However, it is customary to use the following basis in $\mathfrak{sl}(2, \mathbb{C})$:

$$(3.20) \qquad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In this basis, the commutation relations are given by

$$(3.21) \qquad [e, f] = h, \qquad [h, e] = 2e, \qquad [h, f] = -2f.$$

Invariant bilinear form. Each of these Lie algebras has an $\operatorname{Ad} G$-invariant symmetric bilinear form. In each case, it can be defined by $(x, y) = - \operatorname{tr}(xy)$ (of course, it could also be defined without the minus sign). For $\mathfrak{so}(3, \mathbb{R})$, this form can also be rewritten as $(x, y) = \operatorname{tr}(xy^t)$; for $\mathfrak{su}(n)$, as $(x, y) = \operatorname{tr}(x\bar{y}^t)$, which shows that in these two cases this form is positive definite. In terms of bases defined above, it can be written as follows:

- $\mathfrak{so}(3, \mathbb{R})$: elements J_x, J_y, J_z are orthogonal to each other, and $(J_x, J_x) = (J_y, J_y) = (J_z, J_z) = 2$
- $\mathfrak{su}(2)$: elements $i\sigma_k$ are orthogonal, and $(i\sigma_k, i\sigma_k) = 2$.
- $\mathfrak{sl}(2, \mathbb{C})$: $(h, h) = -2$, $(e, f) = (f, e) = -1$, all other inner products are zero.

Isomorphisms. We have an isomorphism of Lie algebras $\mathfrak{su}(2) \xrightarrow{\sim} \mathfrak{so}(3, \mathbb{R})$ given by

$$i\sigma_1 \mapsto 2J_x$$
$$(3.22) \qquad i\sigma_2 \mapsto 2J_y$$
$$i\sigma_3 \mapsto 2J_z.$$

It can be lifted to a morphism of Lie groups $\mathrm{SU}(2) \to \mathrm{SO}(3, \mathbb{R})$, which is a twofold cover (see Exercise 2.12).

The inclusion $\mathfrak{su}(2) \subset \mathfrak{sl}(2, \mathbb{C})$ gives an isomorphism $\mathfrak{su}(2)_{\mathbb{C}} \simeq \mathfrak{sl}(2, \mathbb{C})$. In terms of basis, it is given by

$$i\sigma_1 \mapsto e - f$$
$$(3.23) \qquad i\sigma_2 \mapsto i(e + f)$$
$$i\sigma_3 \mapsto ih.$$

Combining these two isomorphisms, we get an isomorphism $\mathfrak{so}(3, \mathbb{R})_{\mathbb{C}} = \mathfrak{so}(3, \mathbb{C}) \xrightarrow{\sim} \mathfrak{sl}(2, \mathbb{C})$

$$J_x \mapsto \frac{1}{2}(e - f)$$
$$(3.24) \qquad J_y \mapsto \frac{i}{2}(e + f)$$
$$J_z \mapsto \frac{ih}{2}.$$

Exercises

3.1. Consider the group $\mathrm{SL}(n, \mathbb{R})$. Show that the element $X = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$ is not in the image of the exponential map. (Hint: if $X = \exp(x)$, what are the eigenvalues of x?).

3.2. Let $f \colon \mathfrak{g} \to G$ be any smooth map such that $f(0) = 1, f_*(0) = \mathrm{id}$; we can view such a map as a local coordinate system near $1 \in G$. Show that the the group law writen in this coordinate system has the form $f(x)f(y) = f(x + y + B(x, y) + \dots)$ for some bilinear map $B \colon \mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$ and that $B(x, y) - B(y, x) = [x, y]$.

3.3. Show that all forms of Jacobi identity given in (3.5), (3.6) are equivalent.

3.4. Show that if we denote, for $x \in \mathfrak{g}$, by ξ_x the *left-invariant* vector field on G such that $\xi_x(1) = x$ (cf. Theorem 2.25), then $[\xi_x, \xi_y] = -\xi_{[x, y]}$.

3.5. (1) Prove that $\mathbb{R}^3$, considered as Lie algebra with the commutator given by the cross-product, is isomorphic (as a Lie algebra) to $\mathfrak{so}(3, \mathbb{R})$.

(2) Let $\varphi\colon \mathfrak{so}(3,\mathbb{R}) \to \mathbb{R}^3$ be the isomorphism of part (1). Prove that under this isomorphism, the standard action of $\mathfrak{so}(3)$ on $\mathbb{R}^3$ is identified with the action of $\mathbb{R}^3$ on itself given by the cross-product:

$$a \cdot \vec{v} = \varphi(a) \times \vec{v}, \qquad a \in \mathfrak{so}(3), \vec{v} \in \mathbb{R}^3$$

where $a \cdot \vec{v}$ is the usual multiplication of a matrix by a vector.

This problem explains common use of cross-products in mechanics (see, e.g. [**1**]): angular velocities and angular momenta are actually elements of Lie algebra $\mathfrak{so}(3,\mathbb{R})$ (to be precise, angular momenta are elements of the dual vector space, $(\mathfrak{so}(3,\mathbb{R}))^*$, but we can ignore this difference). To avoid explaining this, most textbooks write angular velocities as vectors in $\mathbb{R}^3$ and use cross-product instead of commutator. Of course, this would completely fail in dimensions other than 3, where $\mathfrak{so}(n,\mathbb{R})$ is not isomorphic to $\mathbb{R}^n$ even as a vector space.

3.6. Let P_n be the space of polynomials with real coefficients of degree $\leq n$ in variable x. The Lie group $G = \mathbb{R}$ acts on P_n by translations of the argument: $\rho(t)(x) = x + t, t \in G$. Show that the corresponding action of the Lie algebra $\mathfrak{g} = \mathbb{R}$ is given by $\rho(a) = a\partial_x, a \in \mathfrak{g}$ and deduce from this the Taylor formula for polynomials:

$$f(x+t) = \sum_{n \geq 0} \frac{(t\partial_x)^n}{n!} f.$$

3.7. Let G be the Lie group of all maps $A\colon \mathbb{R} \to \mathbb{R}$ having the form $A(x) = ax + b, a \neq 0$. Describe explicitly the corresponding Lie algebra. [There are two ways to do this problem. The easy way is to embed $G \subset \mathrm{GL}(2,\mathbb{R})$, which makes the problem trivial. More straightforward way is to explicitly construct some basis in the tangent space, construct the corresponding one-parameter subgroups, and compute the commutator using (3.3). The second way is recommended to those who want to understand how the correspondence between Lie groups and Lie algebras works.]

3.8. Let $\mathrm{SL}(2,\mathbb{C})$ act on $\mathbb{CP}^1$ in the usual way:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} (x : y) = (ax + by : cx + dy).$$

This defines an action of $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$ by vector fields on $\mathbb{CP}^1$. Write explicitly vector fields corresponding to h, e, f in terms of coordinate $t = x/y$ on the open cell $\mathbb{C} \subset \mathbb{CP}^1$.

3.9. Let G be a Lie group with Lie algebra $\mathfrak{g}$, and $\mathrm{Aut}(\mathfrak{g})$, $\mathrm{Der}(\mathfrak{g})$ be as defined in Example 3.30.
   (1) Show that $g \mapsto \mathrm{Ad}\, g$ gives a morphism of Lie groups $G \to \mathrm{Aut}(G)$; similarly, $x \mapsto \mathrm{ad}\, x$ is a morphism of Lie algebras $\mathfrak{g} \to \mathrm{Der}\,\mathfrak{g}$. (The automorphisms of the form $\mathrm{Ad}\, g$ are called *inner* automorphisms; the derivations of the form $\mathrm{ad}\, x, x \in \mathfrak{g}$ are called inner derivations.)
   (2) Show that for $f \in \mathrm{Der}\,\mathfrak{g}$, $x \in \mathfrak{g}$, one has $[f, \mathrm{ad}\, x] = \mathrm{ad}\, f(x)$ as operators in $\mathfrak{g}$, and deduce from this that $\mathrm{ad}(\mathfrak{g})$ is an ideal in $\mathrm{Der}\,\mathfrak{g}$.

3.10. Let $\{H_\alpha\}_{\alpha \in A}$ be some family of Lie subgroups in G, with the Lie algebras $\mathfrak{h}_\alpha = \mathrm{Lie}(H_\alpha)$. Let $H = \bigcap_\alpha H_\alpha$. Without using the theorem about closed subgroup, show that H is a Lie group with Lie algebra $\mathfrak{h} = \bigcap_\alpha \mathfrak{h}_\alpha$.

3.11. Let J_x, J_y, J_z be the basis in $\mathfrak{so}(3,\mathbb{R})$ described in Section 3.10. The standard action of $\mathrm{SO}(3,\mathbb{R})$ on $\mathbb{R}^3$ defines an action of $\mathfrak{so}(3,\mathbb{R})$ by vector fields on $\mathbb{R}^3$. Abusing the language, we will use the same notation J_x, J_y, J_z for the corresponding vector fields on $\mathbb{R}^3$. Let $\Delta_{sph} = J_x^2 + J_y^2 + J_z^2$; this is a second order differential operator on $\mathbb{R}^3$, which is usually called the *spherical Laplace operator*, or the *Laplace operator on the sphere*.
   (1) Write Δ_{sph} in terms of $x, y, z, \partial_x, \partial_y, \partial_z$.
   (2) Show that Δ_{sph} is well defined as a differential operator on a sphere $S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$, i.e., if f is a function on $\mathbb{R}^3$ then $(\Delta_{sph}f)|_{S^2}$ only depends on $f|_{S^2}$.

(3) Show that the usual Laplace operator $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$ can be written in the form $\Delta = \frac{1}{r^2}\Delta_{sph} + \Delta_{radial}$, where Δ_{radial} is a differential operator written in terms of $r = \sqrt{x^2 + y^2 + z^2}$ and $r\partial_r = x\partial_x + y\partial_y + z\partial_z$.

(4) Show that Δ_{sph} is rotation invariant: for any function f and $g \in \mathrm{SO}(3, \mathbb{R})$, $\Delta_{sph}(gf) = g(\Delta_{sph}f)$. (Later we will describe a better way of doing this.)

3.12. Give an alternative proof of Theorem 3.33, using Lemma 3.14.

3.13. (1) Let $\mathfrak{g}$ be a three-dimensional real Lie algebra with basis x, y, z and commutation relations $[x, y] = z$, $[z, x] = [z, y] = 0$ (this algebra is called *Heisenberg algebra*). Without using Campbell-Hausdorff formula, show that in the corresponding Lie group, one has $\exp(tx)\exp(sy) = \exp(tsz)\exp(sy)\exp(tx)$ and construct explicitly the connected, simply connected Lie group corresponding to $\mathfrak{g}$.

(2) Generalize the previous part to the Lie algebra $\mathfrak{g} = V \oplus \mathbb{R}z$, where V is a real vector space with non-degenerate skew-symmetric form ω and the commutation relations are given by $[v_1, v_2] = \omega(v_1, v_2)z$, $[z, v] = 0$.

3.14. This problem is for readers familiar with mathematical formalism of classical mechanics.

Let G be a real Lie group and A— a positive definite symmetric bilinear form on $\mathfrak{g}$; such a form can also be considered as a linear map $\mathfrak{g} \to \mathfrak{g}^*$.

(1) Let us extend A to a left invariant metric on G. Consider mechanical system describing free motion of a particle on G, with kinetic energy given by $A(\dot{g}, \dot{g})$ and zero potential energy. Show that equations of motion for this system are given by Euler's equations:

$$\dot{\Omega} = \mathrm{ad}^*\, v.\Omega$$

where $v = g^{-1}\dot{g} \in \mathfrak{g}$, $\Omega = Av \in \mathfrak{g}^*$, and ad^* is the coadjoint action:

$$\langle \mathrm{ad}^*\, x.f, y \rangle = -\langle f, \mathrm{ad}\, x.y \rangle \qquad x, y \in \mathfrak{g},\ f \in \mathfrak{g}^*.$$

(For $G = \mathrm{SO}(3, \mathbb{R})$, this system describes motion of a solid body rotating around its center of gravity — so called Euler's case of rotation of a solid body. In this case, A describes the body's moment of inertia, v is angular velocity, and Ω is angular momentum, both measured in the moving frame. Details can be found in [**1**]).

(2) Using the results of the previous part, show that if A is a bi-invariant metric on G, then one-parameter subgroups $\exp(tx), x \in \mathfrak{g}$ are geodesics for this metric.

3.15. Let G be a complex connected Lie group, with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$, and let $\mathfrak{k} \subset \mathfrak{g}$ be a real form of $\mathfrak{g}$.

(1) Define $\mathbb{R}$-linear map $\theta\colon \mathfrak{g} \to \mathfrak{g}$ by $\theta(x+iy) = x-iy$, $x, y \in \mathfrak{k}$. Show that θ is an automorphism of $\mathfrak{g}$ (considered as a real Lie algebra), and that it can be uniquely lifted to an automorphism $\theta\colon G \to G$ of the group G (considered as a real Lie group). ????

(2) Let $K = G^\theta$. Show that K is a real Lie group with Lie algebra $\mathfrak{k}$.

3.16. Let $\mathrm{Sp}(n)$ be the unitary quaternionic group defined in Exercise 2.7. Show that $\mathfrak{sp}(n)_{\mathbb{C}} = \mathfrak{sp}(2n, \mathbb{C})$. Thus $\mathrm{Sp}(n)$ is a compact real form of $\mathrm{Sp}(2n, \mathbb{C})$.

3.17. Let $\mathfrak{so}(p, q) = \mathrm{Lie}(\mathrm{SO}(p, q))$. Show that its complexification is $\mathfrak{so}(p, q)_{\mathbb{C}} = \mathfrak{so}(p + q, \mathbb{C})$.

Chapter 4

Representations of Lie Groups and Lie Algebras

In this section, we will discuss representation theory of Lie groups and Lie algebras — as far as it can be discussed without using the structure theory of semisimple Lie algebras. Unless specified otherwise, all Lie groups, algebras, and representations are finite-dimensional, and all representations are complex. Lie groups and Lie algebras can be either real or complex; unless specified otherwise, all results are valid both for real and complex case.

4.1. Basic definitions

Let us start by recalling basic definitions.

Definition 4.1. A representation of a Lie group G is a vector space V together with a morphism $\rho\colon G \to \mathrm{GL}(V)$.

A representation of a Lie algebra $\mathfrak{g}$ is a vector space V together with a morphism $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$.

A morphism between two representations V, W of the same group G is a linear map $f\colon V \to W$ which commutes with the action of G: $f\rho(g) = \rho(g)f$. In a similar way one defines a morphism of representations of a Lie algebra. The space of all G-morphisms (respectively, $\mathfrak{g}$-morphisms) between V and W will be denoted by $\mathrm{Hom}_G(V, W)$ (respectively, $\mathrm{Hom}_{\mathfrak{g}}(V, W)$).

Remark 4.2. Morphisms between representations are also frequently called *intertwining operators* because they "intertwine" action of G in V and W.

Notion of representation is completely parallel to the notion of module over an associative ring or algebra; the difference of terminology is due to historical reasons. In fact, it is also usual to use the word "module" rather than "representation" for Lie algebras: a module over Lie algebra $\mathfrak{g}$ is the same as a representation of $\mathfrak{g}$.

Note that in this definition we didn't specify whether V and G, $\mathfrak{g}$ are real or complex. Usually if G (respectively, $\mathfrak{g}$) is complex, then V should also be taken a complex vector space. However, it also makes sense to take complex V even if G is real: in this case we require that the morphism $G \to \mathrm{GL}(V)$ be smooth, considering $\mathrm{GL}(V)$ as $2n^2$-dimensional real manifold. Similarly, for real Lie algebras we can consider complex representations requiring that $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be $\mathbb{R}$-linear.

Of course, we could also restrict ourselves to consideration of real representations of real groups. However, it turns out that introduction of complex representations significantly simplifies the theory even for real groups and algebras. Thus, from now on, all representations will be complex unless specified otherwise.

The first important result about representations of Lie groups and Lie algebras is the following theorem.

Theorem 4.3. *Let G be a Lie group with Lie algebra $\mathfrak{g}$.*

(1) *Every representation $\rho\colon G \to \mathrm{GL}(V)$ defines a representation $\rho_*\colon \mathfrak{g} \to \mathfrak{gl}(V)$, and every morphism of representations of G is automatically a morphism of representations of $\mathfrak{g}$.*

(2) *If G is connected, simply-connected, then $\rho \mapsto \rho_*$ gives an equivalence of categories of representations of G and representations of $\mathfrak{g}$. In particular, every representation of $\mathfrak{g}$ can be uniquely lifted to a representation of G, and $\mathrm{Hom}_G(V, W) = \mathrm{Hom}_{\mathfrak{g}}(V, W)$.*

Indeed, part (1) is a special case of Theorem 3.17, and part (2) follows from Theorem 3.38.

This is an important result, as Lie algebras are, after all, finite dimensional vector spaces, so they are easier to deal with. For example, this theorem shows that a representation of SU(2) is the same as a representation of $\mathfrak{su}(2)$, i.e. a vector space with three endomorphisms X, Y, Z, satisfying commutation relations $XY - YX = Z$, $YZ - ZY = X$, $ZX - XZ = Y$.

This theorem can also be used to describe representations of a group which is connected but not simply-connected: indeed, by Corollary 3.49 any such group can be written as $G = \tilde{G}/Z$ for some simply-connected group $\tilde{G}$ and a discrete central subgroup $Z \subset G$. Thus, representations of G are the same as representations of $\tilde{G}$ satisfying $\rho(Z) = \mathrm{id}$. An important example of this is when $G = \mathrm{SO}(3, \mathbb{R})$, $\tilde{G} = \mathrm{SU}(2)$ (see Exercise 4.1).

Lemma 4.4. *Let $\mathfrak{g}$ be a real Lie algebra, and $\mathfrak{g}_{\mathbb{C}}$ its complexification as defined in Definition 3.51. Then any complex representation of $\mathfrak{g}$ has a unique structure of representation of $\mathfrak{g}_{\mathbb{C}}$, and $\mathrm{Hom}_{\mathfrak{g}}(V, W) = \mathrm{Hom}_{\mathfrak{g}_{\mathbb{C}}}(V, W)$. In other words, categories of complex representations of $\mathfrak{g}, \mathfrak{g}_{\mathbb{C}}$ are equivalent.*

Proof. Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be the representation of $\mathfrak{g}$. Extend it to $\mathfrak{g}_{\mathbb{C}}$ by $\rho(x+iy) = \rho(x) + i\rho(y)$. We leave it to the reader to check that so defined ρ is complex-linear and agrees with the commutator. $\quad\square$

Example 4.5. The categories of representations of $\mathrm{SL}(2, \mathbb{C})$, $\mathrm{SU}(2)$, $\mathfrak{sl}(2, \mathbb{C})$ and $\mathfrak{su}(2)$ are all equivalent.

This, in particular, allows us to reduce the problem of study of representations of a non-compact Lie group $\mathrm{SL}(2, \mathbb{C})$ to the study of representations of a compact Lie group $\mathrm{SU}(2)$. This is useful because, as we will show below, representation theory of compact Lie groups is especially nice.

Remark 4.6. This only works for finite-dimensional representations; theory of infinite-dimensional representations of $\mathrm{SL}(2, \mathbb{C})$ is very different from that of $\mathrm{SU}(2)$.

The following are some examples of representations which can be defined for any Lie group G (and thus, for any Lie algebra $\mathfrak{g}$).

Example 4.7. Trivial representation: $V = \mathbb{C}$, $\rho(g) = \mathrm{id}$ for any $g \in G$ (respectively, $\rho(x) = 0$ for any $x \in \mathfrak{g}$).

Example 4.8. Adjoint representation: $V = \mathfrak{g}$, $\rho(g) = \mathrm{Ad}\, g$ (respectively, $\rho(x) = \mathrm{ad}\, x$). See (2.3), Lemma 3.14 for definition of Ad, ad.

4.2. Operations on representations

In this section, we discuss basic notions of representation theory of Lie groups and Lie algebras, giving examples of representations, operations on representations such as direct sum and tensor product, and more.

Subrepresentations and quotients.

Definition 4.9. Let V be a representation of G (respectively $\mathfrak{g}$). A subrepresentation is a vector subspace $W \subset V$ stable under the action: $\rho(g)W \subset W$ for all $g \in G$ (respectively, $\rho(x)W \subset W$ for all $x \in \mathfrak{g}$).

It is easy to check that if G is a connected Lie group with Lie algebra $\mathfrak{g}$, then $W \subset V$ is a subrepresentaion for G if and only if it is a subrepresentation for $\mathfrak{g}$.

It is trivial to check that if $W \subset V$ is a subrepresentation, then the quotient space V/W has a canonical sructure of a representation. It will be called *factor representation*, or the *quotient representation.*

Direct sum and tensor product.

Lemma 4.10. *Let V, W be representations of G (respectively, $\mathfrak{g}$). Then there is a canonical structure of a representation on V^*, $V \oplus W$, $V \otimes W$.*

Proof. Action of G on $V \oplus W$ is given by $\rho(g)(v + w) = \rho(g)v + \rho(g)w$, and similarly for $\mathfrak{g}$.

For tensor product, we define $\rho(g)(v \otimes w) = \rho(g)v \otimes \rho(g)w$. However, action of $\mathfrak{g}$ is trickier: indeed, naive definition $\rho(x)(v \otimes w) = \rho(x)v \otimes \rho(x)w$ does not define a representation (it is not even linear in x). Instead, if we write $x = \dot{\gamma}(0)$ for some one-parameter family $\gamma(t)$ in a Lie group G with $\gamma(0) = 1$, then

$$\rho(x)(v \otimes w) = \frac{d}{dt}|_{t=0}(\gamma(t)v \otimes \gamma(t)w) = (\dot{\gamma}(0)v \otimes \gamma(0)w) + (\gamma(0)v \otimes \dot{\gamma}(t)w) = \rho(x)v \otimes w + v \otimes \rho(x)w$$

by using Leibnitz rule. Thus, we define $\rho(x)(v \otimes w) = \rho(x)v \otimes w + v \otimes \rho(x)w$. It is easy to show, even without using the Lie group G, that so defined action is indeed a representation of $\mathfrak{g}$ on $V \otimes W$.

To define action of $G, \mathfrak{g}$ on V^*, we require that the natural pairing $V \otimes V^* \to \mathbb{C}$ be a morphism of representations, considering $\mathbb{C}$ as the trivial representation. This gives, for $v \in V, v^* \in V^*$, $\langle \rho(g)v, \rho(g)v^* \rangle = \langle v, v^* \rangle$, so action of G in V^* is given by $\rho_{V^*}(g) = \rho(g^{-1})^t$, where for $A \colon V \to V$, we denote by A^t the adjoint operator $V^* \to V^*$.

Similarly, for the action of $\mathfrak{g}$ we get $\langle \rho(x)v, v^* \rangle + \langle v, \rho(x)v^* \rangle = 0$, so $\rho_{V^*}(x) = -(\rho_V(x))^t$. $\qquad \square$

As an immediate corollary, we see that for a representation V, any tensor space $V^{\otimes k} \otimes (V^*)^{\otimes l}$ has a canonical structure of a representation.

Example 4.11. Let V be a representation of G (respectively, $\mathfrak{g}$). Then the space $\mathrm{End}(V) \simeq V \otimes V^*$ of linear operators on V is also a representation, with the action given by $g \colon A \mapsto \rho_V(g)A\rho_V(g^{-1})$ for $g \in G$ (respectively, $x \colon A \mapsto \rho_V(x)A - A\rho_V(x)$ for $x \in \mathfrak{g}$). More generally, the space of linear maps $\mathrm{Hom}(V, W)$ between two representations is also a representation with the action defined by $g \colon A \mapsto \rho_W(g)A\rho_V(g^{-1})$ for $g \in G$ (respectively, $x \colon A \mapsto \rho_W(x)A - A\rho_V(x)$ for $x \in \mathfrak{g}$).

Similarly, the space of bilinear forms on V is also a representation, with action given by

$$gB(v, w) = B(g^{-1}v, g^{-1}w), \quad g \in G$$
$$xB(v, w) = -(B(x.v, w) + B(v, x.w)), \quad x \in \mathfrak{g}.$$

Proof of these formulas is left to the reader as an exercise.

Invariants.

Definition 4.12. Let V be a representation of a Lie group G. A vector $v \in V$ is called *invariant* if $\rho(g)v = v$ for all $g \in G$. The subspace of invariant vectors in V is denoted by V^G.

Similarly, let V be a representation of a Lie algebra $\mathfrak{g}$. A vector $v \in V$ is called *invariant* if $\rho(x)v = 0$ for all $x \in \mathfrak{g}$. The subspace of invariant vectors in V is denoted by $V^{\mathfrak{g}}$.

We leave it to the reader to prove that if G is a connected Lie group with the Lie algebra $\mathfrak{g}$, then for any representation V of G, we have $V^G = V^{\mathfrak{g}}$.

Example 4.13. Let V, W be representations and $\mathrm{Hom}(V, W)$ be the space of linear maps $V \to W$, with the action of G defined as in Example 4.11. Then $(\mathrm{Hom}(V, W))^G = \mathrm{Hom}_G(V, W)$ is the space of intertwining operators. In particular, this shows that $V^G = (\mathrm{Hom}(\mathbb{C}, V))^G = \mathrm{Hom}_G(\mathbb{C}, V)$, with $\mathbb{C}$ considered as a trivial representation.

Example 4.14. Let B be a bilinear form on a representation V. Then B is invariant under the action of G defined in Example 4.11 iff

$$B(gv, gw) = B(v, w)$$

for any $g \in G$, $v, w \in V$. Similarly, B is invariant under the action of $\mathfrak{g}$ iff

$$B(x.v, w) + B(v, x.w) = 0$$

for any $x \in \mathfrak{g}$, $v, w \in V$.

4.3. Irreducible representations

One of the main problems of the representation theory is the problem of classification of all representations of a Lie group or a Lie algebra. In this generality, it is an extremely difficult problem and for a general Lie group, no satisfactory answer is known. We will later show that for some special classes of Lie groups (namely compact Lie groups and semisimple Lie groups, to be defined later) this problem does have a good answer.

The most natural approach to this problem is to start by studying simplest possible representations, which would serve as building blocks for more complicated representations.

Definition 4.15. A non-zero representation V of G or $\mathfrak{g}$ is called *irreducible* or *simple* if it has no subrepresentations other than 0, V. Otherwise V is called *reducible*.

Example 4.16. Space $\mathbb{C}^n$, considered as a representation of $\mathrm{SL}(n, \mathbb{C})$, is irreducible.

If a representation V is not irreducible (such representations are called *reducible*), then it has a non-trivial subrepresentation W and thus, V can be included in a short exact sequence $0 \to W \to V \to V/W \to 0$; thus, in a certain sense it is built out of simpler pieces. The natural question is whether this exact sequence splits, i.e. whether we can write $V = W \oplus (V/W)$ as a representation. If so then repeating this process, we can write V as a direct sum of irreducible representations.

Definition 4.17. A representation is called completely reducible (or *semisimple*) if it is isomorphic to a direct sum of irreducible representations: $V \simeq \bigoplus V_i$, V_i irreducible.

In such a case one usually groups together isomorphic summands writing $V \simeq \bigoplus n_i V_i$, $n_i \in \mathbb{Z}_+$, where V_i are pairwise non-isomorphic irreducible representations. The numbers n_i are called *multiplicities*.

However, as the following example shows, not every representation is completely reducible.

Example 4.18. Let $G = \mathbb{R}$, so $\mathfrak{g} = \mathbb{R}$. Then a representation of $\mathfrak{g}$ is the same as a vector space V with a linear map $\mathbb{R} \to \text{End}(V)$; obviously, every such map is of the form $t \mapsto tA$ for some $A \in \text{End}(V)$ which can be arbitrary. The corresponding representation of the group $\mathbb{R}$ is given by $t \mapsto \exp(tA)$. Thus, classifying representations of $\mathbb{R}$ is equivalent to classifying linear maps $V \to V$ up to a change of basis. Such a classification is known (Jordan normal form) but non-trivial.

If v is an eigenvector of A then the one-dimensional space $\mathbb{C}v \subset V$ is invariant under A and thus a subrepresentation in V. Since every operator in a complex vector space has an eigenvector, this shows that every representation of $\mathbb{R}$ is reducible, unless it is one-dimensional. Thus, the only irreducible representations of $\mathbb{R}$ are one-dimensional.

Now one easily sees that writing a representation given by $t \mapsto \exp(tA)$ as a direct sum of irreducible ones is equivalent to diagonalizing A. So a representation is completely reducible iff A is diagonalizable. Since not every linear operator is diagonalizable, not every representation is completely reducible.

Thus, more modest goals of the representation theory would be answering the following questions:

(1) For a given Lie group G, classify all irreducible representations of G.

(2) For a given representation V of a Lie group G, given that it is completely reducible, find explicitly the decomposition of V into direct sum of irreducibles.

(3) For which Lie groups G all representations are completely reducible?

One tool which can be used in decomposing representations into direct sum is the use of central elements.

Lemma 4.19. *Let V be a representation of G (respectively, $\mathfrak{g}$) and $A \colon V \to V$ a diagonalizable intertwining operator. Let $V_\lambda \subset V$ be the eigenspace for A with eigenvalue λ. Then each V_λ is a subrepresentation, so $V = \bigoplus V_\lambda$ as a representation of G (respectively $\mathfrak{g}$).*

The proof of this lemma is trivial and is left to the reader. As an immediate corollary, we get the following result.

Lemma 4.20. *Let V be a representation of G and let $Z \in Z(G)$ be a central element of G such that $\rho(Z)$ is diagonalizable. Then as a representation of G, $V = \bigoplus V_\lambda$ where V_λ is the eigenspace for $\rho(Z)$ with eigenvalue λ. Similar result also holds for central elements in $\mathfrak{g}$.*

Of course, there is no guarantee that V_λ will be an irreducible representation; moreover, in many cases the Lie groups we consider have no central elements at all.

Example 4.21. Consider action of $\text{GL}(n, \mathbb{C})$ on $\mathbb{C}^n \otimes \mathbb{C}^n$. Then the permutation operator $P \colon v \otimes w \mapsto w \otimes v$ commutes with the action of $\text{GL}(n, \mathbb{C})$, so the subspaces $S^2\mathbb{C}^n, \Lambda^2\mathbb{C}^n$ of symmetric and skew-symmetric tensors (which are exactly the eigenspaces of P) are $\text{GL}(n, \mathbb{C})$-invariant $\mathbb{C}^n \otimes \mathbb{C}^n = S^2\mathbb{C}^n \oplus \Lambda^2\mathbb{C}^n$ as a representation. In fact, both $S^2\mathbb{C}^n, \Lambda^2\mathbb{C}^n$ are irreducible (this is not obvious but can be proved by a lengthy explicit calculation; later we will discuss better ways of doing this). Thus, $\mathbb{C}^n \otimes \mathbb{C}^n$ is completely reducible.

4.4. Intertwining operators and Schur lemma

Recall that an intertwining operator is a linear map $V \to W$ which commutes with the action of G. Such operators frequently appear in various applications. A typical example is a quantum mechanics, where we have a vector space V (describing the space of states of some mechanical system) and the

Hamiltonian operator $H\colon V \to V$. Then saying that this whole system has symmetry described by a group G is the same as saying that we have an action of G on V which leaves H invariant, i.e. $gHg^{-1} = H$ for any $g \in G$. This exactly means that H is an intertwining operator. A baby example of such a system was described in the introduction; more realistic example (hydrogen atom in quantum mechanics) will be described in Section 5.2.

These examples motivate the study of intertwining operators. For example, does G-invariance of an operator helps computing eigenvalues and eigenspaces? It worked in the baby example of the introduction; what about the general case?

The first result in this direction is the following famous lemma.

Lemma 4.22 (Schur Lemma)**.**

(1) *Let V be an irreducible complex representation of G. Then the space of intertwining operators $\mathrm{Hom}_G(V, V) = \mathbb{C}\,\mathrm{id}$: any endomorphism of an irreducible representation of G is constant.*

(2) *If V and W are irreducible complex representations which are not isomorphic then $\mathrm{Hom}_G(V, W) = 0$.*

Similar result holds for representations of a Lie algebra $\mathfrak{g}$.

Proof. We note that if $\Phi\colon V \to W$ is an intertwining operator, then $\mathrm{Ker}\,\Phi, \mathrm{Im}\,\Phi$ are subrepresentations in V, W respectively. If V is irreducible, either $\ker \Phi = V$ (in which case $\Phi = 0$) or $\ker \Phi = 0$, so Φ is injective. Similarly, if W is irreducible, either $\mathrm{Im}\,\Phi = 0$ (so $\Phi = 0$) or $\mathrm{Im}\,\Phi = W$, Φ is surjective. Thus, either $\Phi = 0$ or Φ is an isomorphism.

Now part (2) follows immediately: since V, W are not isomorphic, Φ must be zero. To prove part (1), notice that the above argument shows that any non-zero intertwiner $V \to V$ is invertible. Now let λ be an eigenvalue of Φ. Then $\Phi - \lambda\,\mathrm{id}$ is not invertible. On the other hand, it is also an intertwiner, so it must be zero. Thus, $\Phi = \lambda\,\mathrm{id}$. $\square$

Example 4.23. Consider the group $\mathrm{GL}(n, \mathbb{C})$. Since $\mathbb{C}^n$ is irreducible as a representation of $\mathrm{GL}(n, \mathbb{C})$, every operator commuting with $\mathrm{GL}(n, \mathbb{C})$ must be scalar. Thus, the center $Z(\mathrm{GL}(n, \mathbb{C})) = \{\lambda\,\mathrm{id}, \lambda \in \mathbb{C}^\times\}$; similarly, the center of the Lie algebra is $\mathfrak{z}(\mathfrak{gl}(n, \mathbb{C})) = \{\lambda\,\mathrm{id}, \lambda \in \mathbb{C}\}$.

Since $\mathbb{C}^n$ is also irreducible as a representation of $\mathrm{SL}(n, \mathbb{C})$, $\mathrm{U}(n)$, $\mathrm{SU}(n)$, $\mathrm{SO}(n, \mathbb{C})$, similar argument can be used to compute the center of each of these groups. The answer is

$$Z(\mathrm{SL}(n, \mathbb{C})) = Z(\mathrm{SU}(n)) = \{\lambda\,\mathrm{id}, \quad \lambda^n = 1\} \qquad \mathfrak{z}(\mathfrak{sl}(n, \mathbb{C})) = \mathfrak{z}(\mathfrak{su}(n)) = 0$$
$$Z(\mathrm{U}(n)) = \{\lambda\,\mathrm{id}, \quad |\lambda| = 1\} \qquad\qquad\qquad \mathfrak{z}(\mathfrak{u}(n)) = \{\lambda\,\mathrm{id}, \quad \lambda \in i\mathbb{R}\}$$
$$Z(\mathrm{SO}(n, \mathbb{C})) = Z(\mathrm{SO}(n, \mathbb{R})) = \{\pm 1\} \qquad\qquad \mathfrak{z}(\mathfrak{so}(n, \mathbb{C})) = \mathfrak{z}(\mathfrak{so}(n, \mathbb{R})) = 0.$$

As an immediate corollary of Schur's lemma, we get the following result.

Corollary 4.24. *Let V be a completely reducible representation of Lie group G (respectively, Lie algebra $\mathfrak{g}$). Then*

(1) *If $V = \bigoplus V_i$, V_i — irreducible, pairwise non-isomorphic, then any intertwining operator $\Phi\colon V \to V$ is of the form $\Phi = \bigoplus \lambda_i\,\mathrm{id}_{V_i}$.*

(2) *If $V = \bigoplus N_i V_i = \bigoplus \mathbb{C}^{n_i} \otimes V_i$, V_i — irreducible, pairwise non-isomorphic, then any intertwining operator $\Phi\colon V \to V$ is of the form $\Phi = \bigoplus (A_i \otimes \mathrm{id}_{V_i})$, $A_i \in \mathrm{End}(\mathbb{C}^{n_i})$.*

Proof. For part (1), notice that any operator $V \to V$ can be written in a block form: $\Phi = \bigoplus \Phi_{ij}$, $\Phi_{ij}\colon V_i \to V_j$. By Schur lemma, $\Phi_{ij} = 0$ for $i \neq j$ and $\Phi_{ii} = \lambda_i\,\mathrm{id}_{V_i}$. Part (b) is proved similarly. $\square$

This result shows that indeed, if we can decompose a representation V into irreducible ones, this will give us a very effective tool for analysing intertwining operators. For example, if $V = \bigoplus V_i, V_i \not\simeq V_j$, then $\Phi = \bigoplus \lambda_i \operatorname{id}_{V_i}$, so one can find λ_i by computing $\Phi(v)$ for just one vector in V_i. It also shows that each eigenvalue λ_i will appear with multiplicity equal to $\dim V_i$. This is exactly what we did in the baby example in the introduction, where we had $G = \mathbb{Z}_n$.

Another useful corollary of Schur's lemma is the following result.

Proposition 4.25. *If G is a commutative group, then any irreducible complex representation of G is one-dimensional. Similarly, if $\mathfrak{g}$ is a commutative Lie algebra, then any irreducible representation of $\mathfrak{g}$ is one-dimensional.*

Indeed, since G is commutative, every $\rho(g)$ commutes with the action of G, so $\rho(g) = \lambda(g) \operatorname{id}$.

Example 4.26. Let $G = \mathbb{R}$. Then its irreducible representations are one-dimensional (this had already been dicussed before, see Example 4.18). In fact, it is easy to describe them: one-dimensional representations of the corresponding Lie algebra $\mathfrak{g} = \mathbb{R}$ are $a \mapsto \lambda a$, $\lambda \in \mathbb{C}$. Thus, irreducible representations of $\mathbb{R}$ are V_λ, $\lambda \in \mathbb{C}$, where each V_λ is a one-dimensional complex space with the action of $\mathbb{R}$ given by $\rho(a) = e^{\lambda a}$.

In a similar way, one can describe irreducible representations of $S^1 = \mathbb{R}/\mathbb{Z}$: they are exactly those representations of $\mathbb{R}$ which satisfy $\rho(a) = 1$ for $a \in \mathbb{Z}$. Thus, irreducible representations of S^1 are V_k, $k \in \mathbb{Z}$, where each V_k is a one-dimensional complex space with the action of S^1 given by $\rho(a) = e^{2\pi i k a}$. In the realization of $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ the formula is even simpler: in V_k, z acts by z^k.

4.5. Complete reducibility of unitary representations. Representations of finite groups

In this section, we will show that a large class of representations is completely reducible.

Definition 4.27. A complex representation V of a group G is called unitary if there is a G-invariant inner product: $(\rho(g)v, \rho(g)w) = (v, w)$, or equivalently, $\rho(g) \in \mathrm{U}(V)$ for any $g \in G$. (The word "inner product" means a positive definite Hermitian form.)

Similarly, a representation V of a Lie algebra $\mathfrak{g}$ is called unitary if there is an inner product which is $\mathfrak{g}$-invariant: $(\rho(x)v, w) + (v, \rho(x)w) = 0$, or equivalently, $\rho(x) \in \mathfrak{u}(V)$ for any $x \in \mathfrak{g}$

Example 4.28. Let $V = F(S)$ be the space of complex valued functions on a finite set S. Let G be a finite group acting by permutations on S; then it also acts on V by (2.1). Then $(f_1, f_2) = \sum_s f_1(s)\overline{f_2(s)}$ is an invariant inner product, so such a representation is unitary.

The following result explains why unitary representations are so important.

Theorem 4.29. *Each unitary representation is completely reducible.*

Proof. The proof goes by induction on dimension. Either V is irreducible, and we're done, or V has a subrepresentation W. Then $V = W \oplus W^\perp$, and $W^\perp$ is a subrepresentation as well. Indeed: if $w \in W^\perp$, then $(gw, v) = (w, g^{-1}v) = 0$ for any $v \in W$ (since $g^{-1}v \in W$), so $gw \in W^\perp$. $\square$

Theorem 4.30. *Any representation of a finite group is unitary.*

Proof. Let $B(v, w)$ be some positive definite inner product in V. Of course, it may not be G-invariant, so $B(gv, gw)$ may be different from $B(v, w)$. Let us "average" B by using group action:

$$\tilde{B}(v, w) = \frac{1}{|G|} \sum_{g \in G} B(gv, gw).$$

Then $\tilde{B}$ is a positive definite inner product (it is a sum of positive definite inner products), and it is G-invariant:

$$\tilde{B}(hv, hw) = \frac{1}{|G|} \sum_{g \in G} B(ghv, ghw) = \frac{1}{|G|} \sum_{g' \in G} B(g'v, g'w)$$

by making subsitution $gh = g'$ and noticing that as g runs over G, so does g'. $\qquad\square$

Combining this with Theorem 4.29, we immediately get the main result of this section.

Theorem 4.31. *Every representation of a finite group is completely reducible.*

Note that this theorem does not give an explicit recipe for decomposing a representation into direct sum of irreducibles. We will return to this problem later.

4.6. Haar measure on compact Lie groups

In the previous section we have proved complete reducibility of representations of a finite group G. The natural question is whether this proof can be generalized to Lie groups.

Analyzing the proof, we see that the key step was averaging a function over the group: $\tilde{f} = \frac{1}{|G|} \sum f(g)$ for a complex-valued function on a group. It seems reasonable to expect that in the case of Lie groups, we should replace the sum by suitably defined integral over G.

Definition 4.32. A right Haar measure on a real Lie group G is a measure dg such that $\int_G dg = 1$ and which is invariant under the right action of G on itself.

Right invariance implies (and, in fact, is equivalent to) the identity $\int f(gh) \, dg = \int f(g) \, dg$ for any $h \in G$ and integrable function f. In a similar way one defines left Haar measure on G.

In this definition we did not specify what exactly we mean by measure. For readers familiar with measure theory, we can make it more precise by saying that dg is supposed to be σ-additive Borel measure. This guarantees that any continuous function is measurable and thus, since $\int dg$ is finite, every continuous function is integrable.

However, for Lie groups there is no need to go deeply in the measure theory: we can restrict ourselves to measures which are given by some smooth density. Recall that a *density* (also called volume element) on a manifold M is a map $\mu \colon \Lambda^n TM \to C^\infty(M)$ such that $\mu(fs) = |f|\mu(s)$ for any smooth function f and a section s of $\Lambda^n TM$. We will always be considering positive densities: $d(s) > 0$ for any non-zero s. A typical example of a positive density is $d = |\omega|$ where ω is a non-vanishing top degree differential form. Moreover, locally every positive density can be written in such a form.

Densities are non-oriented analogs of top degree differential forms: they form an R_+-bundle over M and under change of coordinates, a density is multiplied by the absolute value of the Jacobian (compare with differential forms which are multiplied by Jacobian under change of coordinates). The standard results about integration on manifolds apply to densities too. In particular, for any smooth function f with a compact support and a density μ, the integral $\int_M f\mu$ is well-defined. Thus, every density gives a measure on the manifold. Unlike differential forms, there is no need to fix orientation of M or even require that M be orientable; however, if M is orientable, then every denisty can be written in the form $d = |\omega|$ and $\int_M |\omega| = \left|\int_M \omega\right|$.

Theorem 4.33. *Let G be a compact real Lie group. Then it has a unique right Haar measure which is given by some smooth positive density. In addition, this Haar measure is also left-invariant and invariant under $g \mapsto g^{-1}$.*

The Haar measure on G is usually denoted by dg.

Proof. Let us choose some non-zero element in $\Lambda^n \mathfrak{g}^*$, $n = \dim G$. Then it can be uniquely extended to a right-invariant differential form ω on G (see Theorem 2.25). Since G is compact, the integral $I = \int_G |\omega|$ is finite. Define the density dg by $dg = \frac{|\omega|}{I}$. Then dg is right-invariant and satisfies $\int_G dg = 1$, so it is a Haar measure. Uniqueness is obvious: right-invariant density is unique up to a constant (because $\Lambda^n \mathfrak{g}^*$ is one-dimensional), and constant is fixed by $\int_G dg = 1$.

To prove that dg is also left-invariant, it suffices to check that it is invariant under adjoint action (cf. Theorem 2.26). But $\Lambda^n \mathfrak{g}^*$ is a one-dimensional representation of G. Thus, this result immediately follows from the following lemma.

Lemma 4.34. *Let V be a one-dimensional real representation of a compact Lie group G. Then for any $g \in G$, $|\rho(g)| = 1$.*

Indeed, if $|\rho(g)| < 1$, then $\rho(g^n) \to 0$. But $\rho(G)$ is a compact subset in $\mathbb{R}^\times$, so it can not contain a sequence with limit 0. In a similar way, $|\rho(g)| > 1$ also leads to a contradiction.

To prove invariance under $i \colon g \mapsto g^{-1}$, notice that since dg is left-invariant, it is easy to see that $i^*(dg)$ is a right-invariant density on G; thus, it suffices to check that dg and $i^*(dg)$ coincide at $1 \in G$. But $i_* \colon \mathfrak{g} \to \mathfrak{g}$ is given by $x \mapsto -x$ (which follows from $i(\exp(tx)) = \exp(-tx)$). Thus, on $\Lambda^n \mathfrak{g}$, $i_* = (-1)^n$. $\qquad\square$

Remark 4.35. In fact, every compact topological group (with some technical restrictions) has a unique Haar measure — even if it is not a Lie group. However, in full generality this result is much harder to prove. ???Reference??? **?!**

Example 4.36. Let $G = S^1 = \mathbb{R}/\mathbb{Z}$. Then the Haar measure is the usual measure dx on $\mathbb{R}/\mathbb{Z}$.

Note that in general, writing the Haar measure on a group is not easy—for example, because in general there is no good choice of coordinates on G. Even in those cases when a coordinate system on G can be described explicitly, the Haar measure is usually given by rather complicated formulas. The only case where this measure can be written by a formula simple enough to be useful for practical computations is when we integrate conjugation-invariant functions (also called class functions).

Example 4.37. Let $G = U(n)$ and let f be a smooth function on G such that $f(ghg^{-1}) = f(h)$. Then

$$\int_{U(n)} f(g) dg = \frac{1}{n!} \int_T f \begin{pmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_n \end{pmatrix} \prod_{i<j} |t_i - t_j|^2 dt$$

where

$$T = \left\{ \begin{pmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_n \end{pmatrix}, t_k = e^{i\varphi_k} \right\}$$

is the subgroup of diagonal matrices and $dt = \frac{1}{(2\pi)^n} d\varphi_1 \ldots d\varphi_n$ is the Haar measure on T.

This is a special case of *Weyl Integration Formula*. The statement of this theorem in full generality and the proof can be found, for example, in [**3**]. The proof requires a fair amount of structure theory of compact Lie groups and will not be given here.

The main result of this section is the following theorem.

Theorem 4.38. *Any finite-dimensional representation of a compact Lie group is unitary and thus completely reducible.*

Proof. The proof is almost identical to the proof for the finite group: let $B(v, w)$ be some positive definite inner product in V and "average" it by using group action:

$$\tilde{B}(v, w) = \int_G B(gv, gw) \, dg$$

where dg is the Haar measure on G. Then $\tilde{B}(v, v) > 0$ (it is an integral of a positive function) and right invariance of Haar measure shows that $B(hv, hw) = B(v, w)$.

Complete reducibility now follows from Theorem 4.29.   □

4.7. Orthogonality of characters and Peter-Weyl theorem

In the previous section, we have established that any representation of a comapct Lie group is completely reducible: $V \simeq \bigoplus N_i V_i$, where $N_i \in \mathbb{Z}_+, V_i$ are pairwise non-isomorphic irreducible representations. However, we have not really discussed how one can explicitly decompose a given representation in a direct sum of irreducibles, or at least find the multiplicities N_i. This will be discussed in this section. Throughout this section, G is a compact real Lie group with Haar measure dg.

Let v_i be a basis in a representation V. Writing the operator $\rho(g) \colon V \to V$ in the basis v_i, we get a matrix-valued function on G. Equivalently, we can consider each entry $\rho_{ij}(g)$ as a scalar-valued function on G. Such functions are called *matrix coefficients* (of the representation V).

Theorem 4.39.

(1) *Let V, W be non-isomorphic irreducible representations of G. Choose bases $v_i \in V$, $i = 1 \ldots n$ and $w_a \in W$, $a = 1 \ldots m$. Then for any i, j, a, b, the matrix coefficients $\rho_{ij}^V(g), \rho_{ab}^W$ are orthogonal: $(\rho_{ij}^V(g), \rho_{ab}^W) = 0$, where $(\ ,\ )$ is the inner product on $C^\infty(G, \mathbb{C})$ given by*

(4.1)
$$(f_1, f_2) = \int_G f_1(g) \overline{f_2(g)} \, dg.$$

(2) *Let V be an irreducible representation of G and let $v_i \in V$ be an orthonormal basis with respect to a G-invariant inner product (which exists by Theorem 4.38). Then the matrix coefficients $\rho_{ij}^V(g)$ are pairwise orthogonal, and each has norm squared $\frac{1}{\dim V}$:*

(4.2)
$$(\rho_{ij}^V(g), \rho_{kl}^V) = \frac{\delta_{ik}\delta_{jl}}{\dim V}$$

Proof. The proof is based on the following easy lemma.

Lemma 4.40.

(1) *Let V, W be non-isomorphic irreducible representations of G and f a linear map $V \to W$. Then $\int_G gfg^{-1} \, dg = 0$.*

(2) *If f is a linear map $V \to V$, then $\int gfg^{-1} \, dg = \frac{\text{tr}(f)}{\dim V} \, \text{id}$.*

Indeed, let $\tilde{f} = \int_G gfg^{-1} \, dg$. Then $\tilde{f}$ commutes with action of g: $h\tilde{f}h^{-1} = \int_G (hg)f(hg)^{-1} \, dg = \tilde{f}$. By Schur lemma, $\tilde{f} = 0$ for $W \neq V$ and $\tilde{f} = \lambda \, \mathrm{id}$ for $W = V$. Since $\mathrm{tr}(gfg^{-1}) = \mathrm{tr}\, f$, we see that $\mathrm{tr}\, \tilde{f} = \mathrm{tr}\, f$, so $\lambda = \frac{\mathrm{tr}(f)}{\dim V} \, \mathrm{id}$. This proves the lemma.

Now let v_i, w_a be orthonormal bases in V, W. Choose a pair of indices i, a and apply this lemma to the map $E_{ai} \colon V \to W$ given by $E_{ai}(v_i) = w_a$, $E_{ai}v_j = 0$, $j \neq i$. Then we have

$$\int_G \rho^W(g) E_{ai} \rho^V(g^{-1}) \, dg = 0.$$

Rewriting this in the matrix form and using $\rho(g^{-1}) = \overline{\rho(g)^t}$ (which follows from unitarity of $\rho(g)$), we get that for any b, j,

$$\int \rho^W_{ba}(g) \overline{\rho^V_{ji}(g)} \, dg = 0$$

which proves the first part of the theorem in the case when the bases are orthonormal; general case immediately follows.

To prove the second part, apply the lemma to a matrix unit $E_{ki} \colon V \to V$ to get

$$\sum_{l,j} E_{lj} \int \rho^V_{lk}(g) \overline{\rho^V_{ji}(g)} \, dg = \frac{\mathrm{tr}\, E_{ki}}{\dim V} \, \mathrm{id}$$

which immediately yields the second part of the theorem. $\qquad\square$

So irreducible representations give us a way of constructing an orthonormal set of functions on the group. Unfortunately, they depend on the choice of basis. However, there is one particular combination of matrix coefficients that does not depend on the choice of basis.

Definition 4.41. A character of a representation V is the function on the group defined by

$$\chi_V(g) = \mathrm{tr}_V \, \rho(g) = \sum \rho^V_{ii}(g).$$

It is immediate from the definition that the character does not depend on the choice of basis in V. It also has a number of other properties, listed below; proof of them is left to the reader as an exercise.

Lemma 4.42.

    (1) *Let $V = \mathbb{C}$ be the trivial representation. Then $\chi_V = 1$.*

    (2) $\chi_{V \oplus W} = \chi_V + \chi_W$.

    (3) $\chi_{V \otimes W} = \chi_V \chi_W$

    (4) $\chi_V(ghg^{-1}) = \chi_V(h)$.

    (5) *Let V^* be the dual of representation V. Then $\chi_{V^*} = \overline{\chi_V}$.*

The orthogonality relation for matrix coefficients immediately implies the following result for the characters.

Theorem 4.43.

    (1) *Let V, W be non-isomorphic complex irreducible representations of a compact real Lie group G. Then the characters χ_V, χ_W are orthogonal with respect to inner product (4.1): $(\chi_V, \chi_W) = 0$.*

    (2) *For any irreducible representation V, $(\chi_V, \chi_V) = 1$.*

In other words, if we denote by $\widehat{G}$ the set of isomorphism classes of irreducible representations of G, then the set $\{\chi_V, V \in \widehat{G}\}$ is an orthonormal family of functions on G.

This immediately implies a number of corollaries.

Corollary 4.44. *Let V be a complex representation of a compact real Lie group G. Then*

(1) *V is irreducible iff $(\chi_V, \chi_V) = 1$.*

(2) *V can be uniquely written in the form $V \simeq \bigoplus N_i V_i$, V_i — pairwise non-isomorphic irreducible representations, and the multiplicities N_i are given by $N_i = (\chi_V, \chi_{V_i})$.*

In principle, this theorem gives a way of computing multiplicites N_i. In real life, it is only usable for finite groups and some special cases. Much more practical ways of finding multiplicities will be given later when we develop weight decomposition for representations of semisimple Lie algebras.

Finally, let us return to the matrix coefficients of representations. One might ask whether it is possible to give a formulation of of Theorem 4.39 in a way that does not require a choice of basis. The answer is "yes". Indeed, let $v \in V$, $v^* \in V^*$. Then we can define a function on the group $\rho_{v^*, v}(g)$ by

$$\rho_{v^*, v}(g) = \langle v^*, \rho(g)v \rangle.$$

This is a generalization of a matrix coefficient: if $v = v_j$, $v^* = v_i^*$, then we recover matrix coefficient $\rho_{ij}(g)$.

This shows that for any representation V, we have a map

$$m \colon V^* \otimes V \to C^\infty(G, \mathbb{C})$$
$$v^* \otimes v \mapsto \langle v^*, \rho(g)v \rangle.$$

The space $V^* \otimes V$ has additional structure. First, we have two commuting actions of G on it, given by action on the first factor and on the second one; in other words, $V^* \otimes V$ is a G-bimodule. In addition, if V is unitary, then the inner product defines an inner product on V^* (the simplest way to define it is to say that if v_i is an orthonormal basis in V, then the dual basis v_i^* is an orthonormal basis in V^*). Define an inner product on $V^* \otimes V$ by

(4.3) $$(v_1^* \otimes w_1, v_2^* \otimes w_2) = \frac{1}{\dim V}(v_1^*, v_2^*)(w_1, w_2).$$

Theorem 4.45. *Let $\widehat{G}$ be the set of isomorphism classes of irreducible representations of G. Define the map*

(4.4) $$m \colon \bigoplus_{V_i \in \widehat{G}} V_i^* \otimes V_i \to C^\infty(G, \mathbb{C})$$

by $m(v^* \otimes v)(g) = \langle v^*, \rho(g)v \rangle$. (Here $\bigoplus$ is the algebraic direct sum, i.e. the space of finite linear combinations.) Then

(1) *The map m is a morphism of G-bimodules:*

$$m((gv^*) \otimes v) = L_g(m(v^* \otimes v))$$
$$m(v^* \otimes gv) = R_g(m(v^* \otimes v)),$$

   where L_g, R_g are the left and right actions of G on $C^\infty(G, \mathbb{C})$: $(L_g f)(h) = f(g^{-1}h)$, $(R_g f)(h) = f(hg)$.

(2) *The map m preserves the inner product, if we define the inner product in $\bigoplus V_i^* \otimes V_i$ by (4.3) and inner product in $C^\infty(G)$ by (4.1).*

Proof. The first part is obtained by explicit computation:

$$(R_g m(v^* \otimes v))(h) = m(v^* \otimes v)(hg) = \langle v^*, \rho(hg)v \rangle = \langle v^*, \rho(h)\rho(g)v \rangle = m(v^* \otimes gv)(h)$$

$$(L_g m(v^* \otimes v))(h) = m(v^* \otimes v)(g^{-1}h) = \langle v^*, \rho(g^{-1})\rho(h)v \rangle = \langle gv^*, \rho(h)v \rangle = m(gv^* \otimes v)(h)$$

The second part immediately follows from Theorem 4.39.

$\square$

Corollary 4.46. *The map m is injective.*

It turns out that this map is also surjective if we replace the algebraic direct sum by an appropriate completion: every function on the group can be approximated by a linear combination of matrix coefficients. Precise statement is known as Peter–Weyl theorem.

Theorem 4.47. *The map* (4.4) *gives an isomorphism*

$$\widehat{\bigoplus}_{V_i \in \widehat{G}} V_i^* \otimes V_i \to L^2(G, dg)$$

where $\widehat{\bigoplus}$ *is the Hilbert space direct sum, i.e. the completion of the algebraic direct sum with respect to the metric given by inner product* (4.3).

The proof of this theorem requires some non-trivial analytic considerations and goes beyond the scope of this book. Interested reader can find it in [**22**] or [**15**].

Corollary 4.48. *The set of characters* $\{\chi_V, V \in \widehat{G}\}$ *is an orthonormal basis* (*in the sense of Hilbert spaces*) *of the space* $(L^2(G, dg))^G$ *of conjugation-invariant functions on G.*

Example 4.49. Let $G = S^1 = \mathbb{R}/\mathbb{Z}$. As we have already discussed, the Haar measure on G is given by dx and the irreducible representations are parametrized by $\mathbb{Z}$: for any $k \in \mathbb{Z}$, we have one-dimensional representation V_k with the action of S^1 given by $\rho(a) = e^{2\pi ika}$ (see Example 4.26). The corresponding matrix coefficient is the same as character and is given by $\chi_k(a) = e^{2\pi ika}$.

Then the orthogonality relation of Theorem 4.39 gives

$$\int_0^1 e^{2\pi ikx} \overline{e^{2\pi ilx}} \, dx = \delta_{kl},$$

which is the usual orthogonality relation for exponents. Peter-Weyl theorem in this case just says that the exponents $e^{2\pi ikx}, k \in \mathbb{Z}$, form an orthonormal basis of $L^2(S^1, dx)$ which is one of the main statements of the theory of Fourier series: every L^2 function on S^1 can be written as a series $f(x) = \sum_{k \in \mathbb{Z}} c_k e^{2\pi ikx}$ which converges in L^2 metric.

4.8. Universal enveloping algebra

In a Lie algebra $\mathfrak{g}$, in general there is no multiplication: the products of the form xy, $x, y \in \mathfrak{g}$ are not defined. However, if we consider a representation $\rho \colon \mathfrak{g} \to \mathfrak{gl}(V)$ of $\mathfrak{g}$, then the product $\rho(x)\rho(y)$ is well-defined in such a representation — and in fact, as we will see later, operators of this kind can be very useful in the study of representations. Moreover, commutation relation in $\mathfrak{g}$ imply some relations on the operators of this form. For example, commutation relation $[e, f] = h$ in $\mathfrak{sl}(2, \mathbb{C})$ implies that in any representation of $\mathfrak{sl}(2, \mathbb{C})$ we have $\rho(e)\rho(f) - \rho(f)\rho(e) = \rho(h)$, or equivalently, $\rho(e)\rho(f) = \rho(h) + \rho(f)\rho(e)$. These relations do not depend on the choice of representation ρ.

Motivated by this, we will define "universal" associative algebra generated by products of operators of the form $\rho(x)$, $x \in \mathfrak{g}$.

Definition 4.50. Let $\mathfrak{g}$ be a Lie algebra over a field $\mathbb{K}$. The universal enveloping algebra of $\mathfrak{g}$, denoted by $U\mathfrak{g}$, is the associative algebra with unit over $\mathbb{K}$ with generators $i(x)$, $x \in \mathfrak{g}$, subject to relations $i(x + y) = i(x) + i(y)$, $i(cx) = ci(x), c \in \mathbb{K}$, and

$$(4.5) \qquad\qquad i(x)i(y) - i(y)i(x) = i([x, y]).$$

To simplify the notation, we (and everyone else) will usually write simply $x \in U\mathfrak{g}$ instead of $i(x)$. This will be partly justified later (see Corollary 4.62) when we show that the map $i\colon \mathfrak{g} \to U\mathfrak{g}$ is injective and thus $\mathfrak{g}$ can be considered as a subspace in $U\mathfrak{g}$.

If we dropped relation (4.5), we would get the associative algebra generated by elements $x \in \mathfrak{g}$ with no relations other than linearity and associativity. By definition, this is exactly the tensor algebra of $\mathfrak{g}$:

$$(4.6) \qquad\qquad T\mathfrak{g} = \bigoplus_{n \geq 0} \mathfrak{g}^{\otimes n}.$$

Thus, one can alternatively describe the universal enveloping algebra as the quotient of the tensor algebra:

$$(4.7) \qquad\qquad U\mathfrak{g} = T\mathfrak{g}/(xy - yx - [x, y]), \qquad x, y \in \mathfrak{g}.$$

Example 4.51. Let $\mathfrak{g}$ be a commutative Lie algebra. Then $U\mathfrak{g}$ is generated by elements $x \in \mathfrak{g}$ with relations $xy = yx$. In other words, $U\mathfrak{g}$ is the symmetric alebra of $\mathfrak{g}$. Choosing a basis x_i in $\mathfrak{g}$ we see that $U\mathfrak{g} = \mathbb{K}[x_1, \ldots, x_n]$.

Note that in this example, the universal enveloping alegbra is infinite-dimensional. In fact, $U\mathfrak{g}$ is always infinite-dimensional (unless $\mathfrak{g} = 0$). We will return to the question of "how large $U\mathfrak{g}$ is" in the next section.

Example 4.52. The universal enveloping algebra of $\mathfrak{sl}(2, \mathbb{C})$ is the associative algebra over $\mathbb{C}$ generated by elements e, f, h with the relations $ef - ef = h$, $he - eh = 2e$, $hf = -2f$.

It should be noted that even when $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{K})$ is a matrix algebra, multiplication in $U\mathfrak{g}$ is different from multiplication of matrices. For example, let e be the standard generator of $\mathfrak{sl}(2, \mathbb{C})$. Then $e^2 = 0$ as a 2×2 matrix, but $e^2 \neq 0$ in $U\mathfrak{g}$ — and for a good reason: there are many representations of $\mathfrak{sl}(2, \mathbb{C})$ in which $\rho(e)^2 \neq 0$.

The following theorem shows that $U\mathfrak{g}$ is indeed universal in a certain sense, which justifies the name.

Theorem 4.53. *Let A be an associative algebra with unit over $\mathbb{K}$ and let $\rho\colon \mathfrak{g} \to A$ be a linear map such that $\rho(x)\rho(y) - \rho(y)\rho(x) = \rho([x, y])$. Then ρ can be uniquely extended to a morphism of associative algebras $U\mathfrak{g} \to A$.*

Corollary 4.54. *Any representation of $\mathfrak{g}$ (not necessailry finite-dimensional) has a canonical structure of a $U\mathfrak{g}$-module. Conversely, every $U\mathfrak{g}$-module has a canonical structure of a representation of $\mathfrak{g}$.*

In other words, categories of representations of $\mathfrak{g}$ and of $U\mathfrak{g}$-modules are equivalent.

As a useful application of this result, we can use $U\mathfrak{g}$ to construct various operators acting in representations of $\mathfrak{g}$ — in particular to construct intertwining operators.

Example 4.55. Let $C = ef + fe + \frac{1}{2}h^2 \in U\mathfrak{sl}(2, \mathbb{C})$. Then

$$eC = e^2 f + efe + \frac{1}{2}eh^2 = e(fe + h) + (fe + h)e + \frac{1}{2}(he - 2e)h$$

$$= efe + fe^2 + \frac{1}{2}heh + eh + he - eh = efe + fe^2 \frac{1}{2}h(he - 2e) + he = efe + fe^2 \frac{1}{2}h^2 e$$

$$= Ce.$$

The idea of this calculation is to move e to the right, using the relations $ef = fe + h$, $eh = he - 2e$ to interchange it with f, h. Similar calculations also show that $fC = Cf$, $hC = Ch$. Thus, C is central in $U\mathfrak{g}$.

In particular, this implies that in every representation V of $\mathfrak{sl}(2, \mathbb{C})$, the element $\rho(C) \colon V \to V$ commutes with the action of $\mathfrak{sl}(2, \mathbb{C})$ and thus is an intertwining operator. By Schur lemma, this shows that C acts by a constant in every irreducible representation. And if V is not irreducible, eigenspaces of V are subrepresentations, which could be used to decompose V into irreducible representations (see Lemma 4.20).

Element C is called the *Casimir operator* for $\mathfrak{sl}(2, \mathbb{C})$. We will discuss its generalization for other Lie algebras in Proposition 6.54.

Proposition 4.56.   (1) *The adjoint action of $\mathfrak{g}$ on $\mathfrak{g}$ can be uniquely extended to an action of $\mathfrak{g}$ on $U\mathfrak{g}$ which satisfies Leibnitz rule: $\operatorname{ad} x.(ab) = (\operatorname{ad} x.a)b + a(\operatorname{ad} x.b)$, $x \in \mathfrak{g}$, $a, b \in U\mathfrak{g}$. Moreover, $\operatorname{ad} x.a = xa - ax$.*

    (2) *Let $Z\mathfrak{g} = Z(U\mathfrak{g})$ be the center of universal enveloping algebra. Then $Z\mathfrak{g}$ coincides with the space of invariants on $U\mathfrak{g}$ with respect to the adjoint action of $\mathfrak{g}$:*

$$Z\mathfrak{g} = (U\mathfrak{g})^{\operatorname{ad} \mathfrak{g}}$$

Proof. Define adjoint action of $x \in \mathfrak{g}$ on $U\mathfrak{g}$ by $\operatorname{ad} x.a = xa - ax$. Clearly, for $a \in \mathfrak{g}$ this coincides with the usual definition of adjoint action. To see that it is indeed an action, we need to verify that $\operatorname{ad}[x, y].a = \operatorname{ad} x(\operatorname{ad} y.a) - \operatorname{ad} y(\operatorname{ad} x.a)$, or

$$[x, y]a - a[x, y] = \big(x(ya - ay) - (ya - ay)x\big) - \big(y(xa - ax) - (xa - ax)y\big)$$

which is given by explicit calculation.

Leibnitz rule follows from

$$xab - abx = (xa - ax)b + a(xb - bx).$$

This proves the first part. The second part follows immediately: $C \in U\mathfrak{g}$ is central iff it commutes with all generators, i.e. if $Cx = xC$ for any $x \in \mathfrak{g}$. The last condition is equivalent to $\operatorname{ad} x.C = 0$. $\qquad\square$

4.9. Poincare-Birkhoff-Witt theorem

In this section, $\mathfrak{g}$ is a finite-dimensional Lie algebra over the field $\mathbb{K}$ and $U\mathfrak{g}$ is the universal enveloping algebra of $\mathfrak{g}$.

We had already mentioned that $U\mathfrak{g}$ is infinite-dimensional. In this section, we will give more precise statement.

Unlike polynomial algebra, $U\mathfrak{g}$ is not graded: if we try to define degree by $\deg x_1 \ldots x_k = k, x_i \in \mathfrak{g}$, then we run into problem with defining relation (4.5): we would have $\deg(xy) = \deg(yx) = 2$, but

$\deg(xy - yx) = \deg([x, y]) = 1$. Instead, we have a weaker structure: we can define filtration on $U\mathfrak{g}$ by letting, for any $k \geq 0$,

$$(4.8) \qquad F^k U\mathfrak{g} = \text{Subspace in } U\mathfrak{g} \text{ generated by products } x_1 \ldots x_p, \quad p \leq k.$$

This defines a filtration on $U\mathfrak{g}$: we have

$$F^0 U\mathfrak{g} \subset F^1 U\mathfrak{g} \subset \ldots, \qquad U\mathfrak{g} = \bigcup F^p U\mathfrak{g}$$

The following proposition gives some properties of this filtration.

Proposition 4.57.

(1) $U\mathfrak{g}$ *is a filtered algebra: if* $x \in F^p U\mathfrak{g}$, $y \in F^q U\mathfrak{g}$, *then* $xy \in F^{p+q} U\mathfrak{g}$.

(2) *If* $x \in F^p U\mathfrak{g}$, $y \in F^q U\mathfrak{g}$, *then* $xy - yx \in F^{p+q-1} U\mathfrak{g}$.

(3) *Let* $x_1, \ldots, x_n$ *be an ordered basis in* $\mathfrak{g}$. *Then monomials*

$$(4.9) \qquad\qquad x_1^{k_1} \ldots x_n^{k_n}, \qquad \sum k_i \leq p$$

span $F^p U\mathfrak{g}$. *Note that we fix the order of basis elements.*

Proof. Part (1) is obvious. To prove the second part, note that for $p = 1$, we have

$$x(y_1 \ldots y_q) - (y_1 \ldots y_q)x = \sum_i y_1 \ldots [x, y_i] \ldots y_q \in F^q U\mathfrak{g}$$

Now we can argue by induction in p: if the statement is true for some p, then

$$x_1 \ldots x_{p+1} y \equiv x_1 \ldots x_p y x_{p+1} \equiv y x_1 \ldots x_p x_{p+1} \quad \mod F^{p+q} U\mathfrak{g}.$$

Part (3) is again proved by induction in p. Indeed, for $p = 1$ it is obvious. To establish the induction step, notice that $F^{p+1} U\mathfrak{g}$ is generated by elements of the form xy, $x \in \mathfrak{g}$, $y \in F^p U\mathfrak{g}$. By induction assumption, y can be written as linear combination of monomials of the form (4.9). But by part (2),

$$x_i(x_1^{k_1} \ldots x_n^{k_n}) - x_1^{k_1} \ldots x_i^{k_i+1} \ldots x_n^{k_n} \in F^p U\mathfrak{g}.$$

Using the induction assumption again, we see that $x_i(x_1^{k_1} \ldots x_n^{k_n})$ can again be written as linear combination of monomilas of the form (4.9), with $\sum k_i \leq p + 1$. $\qquad\square$

Corollary 4.58. *Each* $F^p U\mathfrak{g}$ *is finite-dimensional.*

Corollary 4.59. *The associated graded algebra*

$$(4.10) \qquad\qquad \operatorname{Gr} U\mathfrak{g} = \bigoplus_p (F^p U\mathfrak{g})/F^{p-1} U\mathfrak{g}$$

is commutative.

We can now formualte the main result of this section.

Theorem 4.60 (Poincaré–Birhoff–Witt)**.** *Let* $x_1, \ldots, x_n$ *be an ordered basis is* $\mathfrak{g}$. *Then monomials of the form* (4.9) *form a basis in* $F^p U\mathfrak{g}$.

The proof of this theorem is not given here; it can be found, for example, in [20],[6], [12]. Here is the main idea of the proof. Since we already know that monomials of the form (4.9) generate $F^p U\mathfrak{g}$ (see Proposition 4.57), it suffices to show that they are linearly independent. To show this, we construct a representation in which the operators corresponding to these monomials are linearly independent.

Namely, we consider (infinite-dimensional) vector space V with basis given by (4.9) (no restriction on $\sum k_i$). The action is uniquely defined by the requirement that $\rho(x_i).x_{j_1}\ldots x_{j_n} = x_i x_{j_1}\ldots x_{j_n}$ if $i \le j_1 \le j_2 \ldots$. For example, this forces $\rho(x_1).x_2 = x_1 x_2$.

This requirement also determines $\rho(x_i).x_{j_1}\ldots x_{j_n}$ if $i > j_1$. For example, to define $\rho(x_2).x_1$, we note that it must be equal to $\rho(x_2)\rho(x_1).1 = \rho(x_1)\rho(x_2).1 + \rho([x_2,x_1]).1 = x_1 x_2 + \sum a_i x_i$, where a_i are defined by $[x_1,x_2] = \sum a_i x_i$.

The difficult part is to check that it is indeed an action, i.e., satisfies $\rho(x)\rho(y) - \rho(y)\rho(x) = \rho[x,y]$, which is done by explicit calculation using Jacobi identity.

Note that this theorem would fail without Jacobi identity: if $[,] \colon \mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$ is an antisymmetric map not satisfying Jacobi identity, then the algebra defined by (4.5) can be trivial (i.e., all $i(x) = 0$).

This theorem can also be reformualted in a coordinate-independent way.

Theorem 4.61 (Poincaré–Birhoff–Witt)**.** *The graded algebra* $\operatorname{Gr} U\mathfrak{g}$ *is naturally isomorphic to the symmetric algebra* $S\mathfrak{g}$. *The isomorphism is given by*

(4.11)
$$S^p\mathfrak{g} \to \operatorname{Gr}^p U\mathfrak{g}$$
$$a_1 \ldots a_p \mapsto a_1 \ldots a_p \mod F^{p-1}U\mathfrak{g}$$

and the inverse isomorphism is given by

(4.12)
$$\operatorname{Gr}^p U\mathfrak{g} \to S^p\mathfrak{g}$$
$$a_1 \ldots a_p \mapsto a_1 \ldots a_p,$$
$$a_1 \ldots a_l \mapsto 0, \quad l < p$$

When written in this form, this theorem may seem trivial. The non-triviality is hidden in the statement that the maps (4.11), (4.12) are well-defined.

Poincaré–Birkhoff–Witt (or PBW for short) theorem has a number of useful corollaries. Here are some of them.

Corollary 4.62. *The natural map* $\mathfrak{g} \to Ug$ *is injective.*

Corollary 4.63. *Let* $\mathfrak{g}_1, \mathfrak{g}_2 \subset \mathfrak{g}$ *be subalgebras such that* $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ *as a vector space (we do not require that* $\mathfrak{g}_1, \mathfrak{g}_2$ *commute). Then the multiplication map*

$$U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 \to U\mathfrak{g}$$

is an isomorphism.

Corollary 4.64. *Algebra* $U\mathfrak{g}$ *has no zero divisors.*

Exercises

4.1. Let $\varphi \colon \mathrm{SU}(2) \to \mathrm{SO}(3,\mathbb{R})$ be the cover map constructed in Exercise 2.12.
  (1) Show that $\operatorname{Ker}\varphi = \{1, -1\} = \{1, e^{\pi i h}\}$, where h is defined by (3.20).
  (2) Using this, show that representations of $\mathrm{SO}(3,\mathbb{R})$ are the same as representations of $\mathfrak{sl}(2,\mathbb{C})$ satisfying $e^{\pi i \rho(h)} = \mathrm{id}$.

4.2. Let $V = \mathbb{C}^2$ be the standard 2-dimensional representation of the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$, and let $S^k V$ be the symmetric power of V.
  (1) Write explicitly the action of $e, f, h \in \mathfrak{sl}(2,\mathbb{C})$ (see Section 3.10) in the basis $e_1^i e_2^{k-i}$.
  (2) Show that $S^2 V$ is isomorphic to the adjoint representation of $\mathfrak{sl}(2,\mathbb{C})$.
  (3) By results of Section 4.1, each representation of $\mathfrak{sl}(2,\mathbb{C})$ can be considered as a representation of $\mathfrak{so}(3,\mathbb{R})$. Which of representations $S^k V$ can be lifted to a representation of $\mathrm{SO}(3,\mathbb{R})$?

4.3. Let $C = ef + fe + \frac{1}{2}h^2 \in U\,\mathfrak{sl}(2,\mathbb{C})$.
  (1) Show that C is central.
  (2) Find the eigenvalues of C in each of the representations $S^k V$ defined in Exercise 4.2.
  (3) Recall that we have an isomorphism $\mathfrak{so}(3,\mathbb{C}) \simeq \mathfrak{sl}(2,\mathbb{C})$ (see Section 3.10) which gives isomorphism of the corresponding enveloping algebras. Show that this isomorphism identifies element C above with a multiple of $J_x^2 + J_y^2 + J_z^2$.

4.4. (1) Let V, W be irreducible representations of a Lie group G. Show that $(V \otimes W^*)^G = 0$ if V is non-isomorphic to W, and that $(V \otimes V^*)^G$ is canonically isomorphic to $\mathbb{C}$.
  (2) Let V be an irreducible representation of a Lie algebra $\mathfrak{g}$. Show that V^* is also irreducible, and deduce from that the space of $\mathfrak{g}$-invariant bilinear forms on V is either zero or 1-dimensional.

4.5. For a representation V of a Lie algebra $\mathfrak{g}$, define the space of *coinvaraints* by $V_{\mathfrak{g}} = V/\mathfrak{g}V$, where $\mathfrak{g}V$ is the subspace spanned by xv, $x \in \mathfrak{g}, v \in V$.
  (1) Show that if V is completely reducible, then the composition $V^{\mathfrak{g}} \hookrightarrow V \to V_{\mathfrak{g}}$ is an isomorphism.
  (2) Show that in general, it is not so. (Hint: take $\mathfrak{g} = \mathbb{R}$ and an appropriate representation V.)

4.6. Let $\mathfrak{g}$ be a Lie algebra, and $(\, , \,)$ — a symmetric ad-invariant bilinear form on $\mathfrak{g}$. Show that the element $\omega \in (\mathfrak{g}^*)^{\otimes 3}$ given by

$$\omega(x, y, z) = ([x, y], z)$$

is skew-symmetric and ad-invariant.

4.7. Prove that if $A \colon \mathbb{C}^n \to \mathbb{C}^n$ is an operator of finite order: $A^k = I$ for some k, then A is diagonalizable. [Hint: use theorem about complete reducibility of representations of a finite group]

4.8. Let C be the standard cube in $\mathbb{R}^3$: $C = \{|x_i| \le 1\}$, and let S be the set of faces of C (thus, S consists of 6 coefficients). Consider the 6-dimensional complex vector V space of functions on S, and define $A \colon V \to V$ by

$$(Af)(\sigma) = \frac{1}{4} \sum_{\sigma'} f(\sigma')$$

where the sum is taken over all faces σ' which are neighbors of σ (i.e., have a common edge with σ). The goal of this problem is to diagonalize A.
  (1) Let $G = \{g \in \mathrm{O}(3, \mathbb{R}) \mid g(C) = C\}$ be the group of symmetries of C. Show that A commutes with the natural action of G on V.
  (2) Let $z = -I \in G$. Show that as a representation of G, V can be decomposed in the direct sum

$$V = V_+ \oplus V_-, \qquad V_\pm = \{f \in V \mid zf = \pm f\}.$$

  (3) Show that as a representation of G, V_+ can be decomposed in the direct sum

$$V_+ = V_+^0 \oplus V_+^1, \quad V_+^0 = \{f \in V_+ \mid \sum_\sigma f(\sigma) = 0\}, \quad V_+^1 = \mathbb{C} \cdot 1$$

where 1 denotes the constant function on S whose value at every $\sigma \in S$ is 1.
  (4) Find the eigenvalues of A on V_-, V_+^0, V_+^1.
  [Note: in fact, each of V_-, V_+^0, V_+^1 is an irreducible representation of G, but you do not need this fact.]

4.9. Let $G = \mathrm{SU}(2)$. Recall that we have a diffeomorphism $G \simeq S^3 \subset \mathbb{R}^4$ (see Example 2.3).
  (1) Show that the left action of G on $G \simeq S^3$ can be extended to an action of G by linear orthogonal transformations on $\mathbb{R}^4$.

(2) Let $\omega \in \Omega^3(G)$ be a left-invariant 3-form whose value at $1 \in G$ is defined by

$$\omega(x_1, x_2, x_3) = \operatorname{tr}([x_1, x_2]x_3), \qquad x_i \in \mathfrak{g}$$

(see Exercise 4.6). Show that $|\omega| = 4|dV|$ where $|dV|$ is the volume form on S^3 induced by the standard metric in $\mathbb{R}^4$ (hint: let x_1, x_2, x_3 be some orthonormal basis in $\mathfrak{su}(2)$ with respect to $\frac{1}{2}\operatorname{tr}(a\bar{b}^t)$).

(3) Show that $\frac{1}{8\pi^2}\omega$ is a bi-invariant form on G such that $\frac{1}{8\pi^2}\int_G |\omega| = 1$

Chapter 5

Representations of $\mathfrak{sl}(2,\mathbb{C})$ and Spherical Laplace Operator

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

We start by studying representations of the simplest possible Lie algebra, $\mathfrak{sl}(2,\mathbb{C})$. Recall that this Lie algebra has a basis e, f, h with the commutation relations

$$(5.1) \qquad [e, f] = h, \qquad [h, e] = 2e, \qquad [h, f] = -2f$$

(see Section 3.10). As was proved earlier, this Lie algebra is simple (Example 6.22).

The main idea of the study of representations of $\mathfrak{sl}(2,\mathbb{C})$ is to start by diagonalizing the operator h.

Definition 5.1. Let V be a representation of $\mathfrak{sl}(2,\mathbb{C})$. A vector $v \in V$ is called vector of weight λ, $\lambda \in \mathbb{C}$, if it is an eigenvector for h with eigenvalue λ:

$$hv = \lambda v.$$

We denote by $V[\lambda] \subset V$ the subspace of vectors of weight λ.

The following lemma plays the key role in the study of representations of $\mathfrak{sl}(2,\mathbb{C})$.

Lemma 5.2.
$$eV[\lambda] \subset V[\lambda + 2]$$
$$fV[\lambda] \subset V[\lambda - 2].$$

Proof. Let $v \in V[\lambda]$. Then

$$hev = [h, e]v + ehv = 2ev + \lambda ev = (\lambda + 2)ev$$

so $ev \in V[\lambda + 2]$. The proof for f is similar. $\qquad\square$

Theorem 5.3. *Every finite-dimensional representation V of $\mathfrak{sl}(2,\mathbb{C})$ can be written in the form*

$$V = \bigoplus_\lambda V[\lambda]$$

where $V[\lambda]$ is defined in Definition 5.1. This decomposition is called weight decomposition of V.

Proof. Since every representation of $\mathfrak{sl}(2,\mathbb{C})$ is completely reducible, it suffices to prove this for irreducible V. So assume that V is irreducible. Let $V' = \sum_\lambda V[\lambda]$ be the subspace spanned by eigenvectors of h. By well-known result of linear algebra, eigenvectors with different eigenvalues are linearly independent, so $V' = \bigoplus V[\lambda]$. By Lemma 5.2, V' is stable under action of e, f and h. Thus, V' is a subrepresentation. Since we assumed that V is irreducible, and $V' \neq 0$ (h has at least one eigenvector), we see that $V' = V$. $\qquad\square$

Our main goal will be classification of ireducible finite-dimensional representations. So from now on, let V be an irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$.

Let λ be a weight of V (i.e., $V[\lambda] \neq 0$) which is maximal in the following sense:

(5.2) $\qquad\qquad\qquad \operatorname{Re}\lambda \geq \operatorname{Re}\lambda' \qquad$ for every weight λ' of V.

Such a weight will be called "highest weight of V", and vectors $v \in V[\lambda]$ — highest weight vectors. It is obvious that every finite-dimensional representation has at least one non-zero highest weight vector.

Lemma 5.4. *Let $v \in V[\lambda]$ be a highest weight vector in V.*

(1) *$ev = 0$.*

(2) *Let*

$$v^k = \frac{f^k}{k!}v, \qquad k \geq 0$$

Then we have

$$hv^k = (\lambda - 2k)v^k,$$
(5.3) $\qquad\qquad fv^k = (k+1)v^{k+1},$
$$ev^k = (\lambda - k + 1)v^{k-1}, \quad k > 0$$

Proof. By Lemma 5.2, $ev \in V[\lambda + 2]$. But by definition of highest weight vector, $V[\lambda + 2] = 0$. This proves the first part.

To prove the second part, note that the formula for action of f is immediate from definition, and formula for action of h follows from Lemma 5.2. Thus, we need to prove the formula for the action of e.

The proof goes by induction. For $k = 1$ we have

$$ev^1 = efv = [e,f]v + fev = hv = \lambda v$$

(using $ev = 0$).

The induction step is proved by

$$ev^{k+1} = \frac{1}{k+1}efv^k = \frac{1}{k+1}(hv^k + fev^k) = \frac{1}{k+1}\left((\lambda - 2k)v^k + (\lambda - k + 1)fv^{k-1}\right)$$
$$= \frac{1}{k+1}(\lambda - 2k + (\lambda - k + 1)k)v^k = (\lambda - k)v^k.$$

$\qquad\square$

Of course, since V is finite-dimensional, only finitely many of v^k are non-zero. However, it is convenient to consider V as a quotient of infinite-dimensional vector space with basis v^k. This is done as follows.

Lemma 5.5. *Let $\lambda \in \mathbb{C}$. Define M_λ to be the infinite-dimensional vector space with basis $v^0, v^1, \ldots$.*

(1) *Formulas (5.3) and $ev^0 = 0$ define on M_λ a structure of (infinite-dimensional) represen-tation of $\mathfrak{sl}(2,\mathbb{C})$.*

(2) *If V is an irreducible finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ which contains a non-zero highest weight vector of highest weight λ, then $V = M_\lambda/W$ for some subrepresentation W.*

Proof. The first part is done by explicit calculation which is essentially equivalent to the calculation used in the proof of Lemma 5.4. The second part immediately follows from Lemma 5.4. □

Now we can prove the main theorem.

Theorem 5.6.

(1) *For any $n \geq 0$, let V_n be the finite-dimensional vector space with basis $v^0, v^1, \ldots, v^n$. Define the action of $\mathfrak{sl}(2,\mathbb{C})$ by*

(5.4)
$$hv^k = (n - 2k)v^k,$$
$$fv^k = (k+1)v^{k+1}, \quad k < n; \quad fv^n = 0$$
$$ev^k = (n + 1 - k)v^{k-1}, \quad k > 0; \quad ev^0 = 0.$$

Then V_n is an irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$; we will call it the irreducible represen-tation with highest weight n.

(2) *For $n \neq m$, representation V_n, V_m are non-isomorphic.*

(3) *Every finite-dimensional irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$ is isomorphic to one of rep-resentations V_n.*

Proof. Consider the infinite-dimensional representation M_λ defined in Lemma 5.5. If $\lambda = n$ is a non-negative integer, consider the subspace $M' \subset M_n$ spanned by vectors $v^{n+1}, v^{n+2}, \ldots$. Then this subspace is actually a subrepresentation. Indeed, it is obviously stable under action of h and f; the only non-trivial relation to check is that $ev^{n+1} \subset M'$. But $ev^{n+1} = (n + 1 - (n + 1))v^n = 0$.

Thus, the quotient space M_n/M' is a finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$. It is obvious that it has basis $v^0, \ldots, v^n$ and that the action of $\mathfrak{sl}(2,\mathbb{C})$ is given by (5.4). Irreducibility of this representation is also easy to prove: any subrepresentation must be spanned by some subset of $v, v^1, \ldots, v^n$, but it is easy to see that each of them generates (under the action of $\mathfrak{sl}(2,\mathbb{C})$) the whole representation V_n. Therefore, V_N is an irreduible finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$. Sonce $\dim V_n = n + 1$, it is obvious that V_N are pairwise non-iosmorphic.

To prove that every irreducible representation is of this form, let V be an irreducible represen-tation of $\mathfrak{sl}(2,\mathbb{C})$ and let $v \in V[\lambda]$ be a highest weight vector. By Lemma 5.5, V is a quotient of M_λ; in other words, it is spanned by vectors $v^k = \frac{f^k}{k!}v$.

Since v^k have different weights, if they are non-zero, then they must be linearly independent. On the other hand, V is finite-dimensional; thus, only finitely many of v^i are non-zero. Let n be maximal such that $v^n \neq 0$, so that $v^{n+1} = 0$. Obviously, in this case $v^0, \ldots, v^n$ are all non-zero and since they have different weight, they are linearly independent, so they form a basis in V.

Since $v^{n+1} = 0$, we must have $ev^{n+1} = 0$. On the other hand, by(5.3), we have $ev^{n+1} = (\lambda - n)v^n$. Since $v^n \neq 0$, this implies that $\lambda = n$ is a non-negative integer. Thus, V is a representation of the form discussed in part (1).

□

Irreducible representations V_n can also be described more explicitly, as symmetric powers of the usual two-dimensional representation (see Exercise 5.2).

As a corollary, we immediately get some useful information about any finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$.

Theorem 5.7. *Let V be a finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$.*

(1) *V admits a weight decomposition with integer weights:*

$$V = \bigoplus_{n\in\mathbb{Z}} V[n]$$

(2) *$\dim V[n] = \dim V[-n]$. Moreover, for $n \geq 0$ the maps*

$$e^n \colon V[n] \to V[-n]$$
$$f^n \colon V[-n] \to V[n]$$

are isomorphisms.

By results of Section 4.1, this also implies similar statements for representations of Lie algebra $\mathfrak{so}(3,\mathbb{R})$ and the group $\mathrm{SO}(3,\mathbb{R})$. These results are given in Exercise 5.3.

5.2. Spherical Laplace operator and hydrogen atom

In this section, we apply our knowledge of representation theory of Lie groups and Lie algebras to the study of Laplace operator on the sphere, thus answering the question raised in the introduction. The material of this section will not be used in the rest of the book, so it can be safely skipped. However, it is a very illustrative example of how one uses representation theory to the study of systems with a symmetry.

Let $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$ be the usual Laplace operator in $\mathbb{R}^3$. We would like to split it into "radial" and "spherical" part, which can be done as follows.

Notice that $\mathbb{R}^3 - \{0\}$ can be identified with the direct product

$$\mathbb{R}^3 - \{0\} \simeq S^2 \times \mathbb{R}_+$$
$$\vec{x} \mapsto (\mathbf{u}, r)$$

(5.5)

$$\mathbf{u} = \frac{\vec{x}}{|\vec{x}|} \in S^2, \quad r = |\vec{x}| \in \mathbb{R}_+.$$

The following well-known lemma shows how Δ can be rewritten in coordinates $\mathbf{u}, r$.

Lemma 5.8.

(1) *When rewritten in coordinates u, r, we have*

$$\Delta = \frac{1}{r^2}\Delta_{sph} + \Delta_{radial}$$

where Δ_{sph} is a differential operator on S^2 and Δ_{radial} is a differential operator on $\mathbb{R}_+$.

(2) *We have*

(5.6)

$$\Delta_{radial} = \partial_r^2 + \frac{2}{r}\partial_r$$
$$\Delta_{sph} = J_x^2 + J_y^2 + J_z^2$$

where

$$J_x = y\partial_z - z\partial y$$
$$J_y = z\partial_x - x\partial z$$
$$J_z = x\partial_y - y\partial x$$

are vector fields corresponding to generators of Lie algebra $\mathfrak{so}(3, \mathbb{R})$ (see Exercise 3.11).

Sketch of proof. Since for any $r > 0$, the vector fields J_x, J_y, J_z are tangent to the sphere of radius r, the operator Δ_{sph} defined by (5.6) is well defined as a differential operator on the sphere. Identity $\Delta = \frac{1}{r^2}\Delta_{sph} + \Delta_{radial}$ can be shown by explicit calculation (see Exercise 3.11). □

One can introduce usual coordinates on the sphere and write Δ_{sph} in these coordinates. Such an expression can be found in any book on multivariable calculus, but it is very messy; more importantly, it will be useless for our purposes. For this reason it is not given here.

The main question we want to answer is as follows:

(5.7)   Find eigenvalues of Δ_{sph} acting on functions on S^2

The motivation for this problem comes from physics. Namely, quantum mechanical description of a particle moving in a central force field (for example, an electron in the hydrogen atom) is given by Schrödinger equation

$$\dot{\psi} = iH\psi$$

where $\psi = \psi(t, \vec{x})$, $\vec{x} \in \mathbb{R}^3$, is the wave-function which describes the state of the system and

$$H = -\Delta + V(r)$$

is the Hamiltonian, or the energy operator. Solving the Schrödinger equation is essentially equivalent to diagonalizing the Hamiltonian. The usual approach to this problem is to use separation of variables, writing

(5.8) $$\psi(\vec{x}) = \sum f_i(r)g_i(\mathbf{u})$$

where $r \in \mathbb{R}_+$, $\mathbf{u} \in S^2$ are given by (5.5), and g_i are eigenfunctions for Δ_{sph}. Substituting this in the equation $H\psi = \lambda\psi$ gives a second-order differential equation on $f_i(r)$. For many potentials $V(r)$, one can explicitly solve this equation, thus giving eigenfunctions of the Hamiltonian — in particular, the energy levels for the hydrogen atom. Details can be found, for example, in [**16**].

Returning to question (5.7), we notice that the straightforward approach, based on introducing coordinates on the sphere and writing the corresponding partial differential equation, is rather complicated. Instead, we will use the symmetries of the sphere, much as we used the $\mathbb{Z}_n$ symmetry in the baby example in the Introduction. We have an obvious action of the group $G = \mathrm{SO}(3, \mathbb{R})$ on the sphere S^2 which therefore defines an action of G on the space of functions on S^2, by $g.f(x) = f(g^{-1}(x))$.

Lemma 5.9. $\Delta_{sph}\colon C^\infty(S^2) \to C^\infty(S^2)$ *commutes with the action of* $\mathrm{SO}(3, \mathbb{R})$.

Proof. This can be shown in several ways. The easiest way is to note that the action of $\mathrm{SO}(3, \mathbb{R})$ on $C^\infty(S^2)$ also defines an action of the corresponding Lie algebra $\mathfrak{g} = \mathfrak{so}(3, \mathbb{R})$ by first order differential operators on $C^\infty(S^2)$ and thus the action of the universal enveloping algebra $U\mathfrak{g}$ by differential operators. By definition, Δ_{sph} is exactly the differential operator corresponding to element $C = J_x^2 + J_y^2 + J_z^2 \in U\mathfrak{g}$. Thus, it suffices to show that C is invariant under adjoint action of G on $U\mathfrak{g}$.

By results of Section 4.2, it is equivalent to checking that for any $a \in \mathfrak{g}$, we have $[a, C] = 0$, i.e. that C is central element in $U\mathfrak{g}$. This can be easily shown by explicit calculation using commutation relations (3.17): for example, for $a = J_x$, we get

$$[J_x, C] = [J_x, J_y^2] + [J_x, J_z^2] = [J_x, J_y]J_y + J_y[J_x, J_y] + [J_x, J_z]J_z + J_z[J_x, J_z]$$
$$= J_z J_y + J_y J_z - J_y J_z - J_z J_y = 0.$$

An even easier way of checking that C is central is to notice that up to a factor, it is the Casimir element of $\mathfrak{so}(3, \mathbb{R})$ corresponding to the invariant bilinear form $\mathrm{tr}(ab)$ and thus, is central (see Proposition 6.54). $\qquad\square$

Therefore, by general results of Section 4.4, the best way to study Δ_{sph} would be to decompose the space of functions on S^2 into irreducible representations of $\mathrm{SO}(3, \mathbb{R})$. As usual, it is more convenient to work with complex representations, so we consider the space of complex-valued functions.

There are some obvious technical problems: the space of functions is infinite dimensional. To avoid dealing with convergence questions and other analytical difficulties, let us consider the space of polynomials

(5.9) $\qquad P_n = \begin{Bmatrix} \text{Complex-valued functions on } S^2 \text{ which can be written as} \\ \text{polynomials in } x, y, z \text{ of total degree} \leq n \end{Bmatrix}.$

One easily sees that each P_n is a finite-dimensional representation of $\mathrm{SO}(3, \mathbb{R})$ which is also Δ_{sph}-invariant. Thus, we can use the theory of finite-dimensional representations to decompose P_n into irreducible representations and then use this to find the eigenvalues of Δ_{sph} in P_n. Since $\bigcup P_n = P$ is the space of all polynomial functions on S^2, which is everywhere dense in $C^\infty(S^2)$, diagonalizing Δ_{sph} in P is essentially equivalent to diagonalizing Δ_{sph} in C^∞ (precise statement will be given below).

Thus, our immediate goal is to decompose P_n into direct sum of irreducible representations of $\mathrm{SO}(3, \mathbb{R})$. To do this, note that by results of Exercise 5.3, irreducible representations of $\mathrm{SO}(3, \mathbb{R})$ are of the form V_{2k}, $k \in \mathbb{Z}_+$. Thus, we can write

$$P_n = \bigoplus c_k V_{2k}.$$

To find coefficients c_k, we need to find the character of P_n, i.e., the dimensions of eigenspaces for J_z. We can do it by explicitly constructing an eigenbasis in P_n.

Lemma 5.10. *The following set of functions form a basis of P_n:*

$$f_{p,k} = z^p \left(\sqrt{1-z^2}\right)^{|k|} e^{ik\varphi}, \qquad p \in \mathbb{Z}_+, k \in \mathbb{Z}, p + |k| \leq n$$

where φ is defined by $x = \rho\cos\varphi, y = \rho\sin\varphi, \rho = \sqrt{x^2 + y^2}$.

Proof. Let $u = x + iy = \rho e^{i\varphi}$, $v = x - iy = \rho e^{-i\varphi}$. Then any polynomial in x, y, z can be written as a polynomial in z, u, v. Since on the sphere we have $1 - z^2 = x^2 + y^2 = uv$, every monomial $z^k u^l v^m$ can be written as a monomial which involves only u or v but not both. Thus, every element of P_n can be written as a linear combination of monomials

$$z^p,$$
$$z^p u^k = z^p \rho^k e^{ik\varphi} = f_{p,k},$$
$$z^p v^k = z^p \rho^k e^{-ik\varphi} = f_{p,-k}$$

with $p, k \in \mathbb{Z}_+$, $p + k \leq n$. Thus, elements $f_{p,k}$ span P_n.

To show that they are linearly independent, assume that

$$\sum a_{p,k} f_{p,k} = \sum_k a_k(z) e^{ik\varphi} = 0, \qquad a_k(z) = \sum_p a_{p,k} z^p \left(\sqrt{1-z^2}\right)^{|k|}.$$

By uniqueness of Fourier series, we see that for every $k \in \mathbb{Z}$, $z \in (-1, 1)$, we have $a_k(z) = 0$ which easily implies that for every p, k, $a_{p,k} = 0$. $\qquad\square$

We can now find the dimensions of the eigenspaces for J_z. Since J_z is generator of rotations around z axis, it is easy to see that in the cylindrical coordinates z, ρ, φ, $J_z = \frac{\partial}{\partial \varphi}$. Thus,

$$J_z f_{p,k} = ik f_{p,k}$$

so $P_n[2k] = \mathrm{Span}(f_{p,k})_{0 \le p \le n - |k|}$ and thus $\dim P_n[2k] = n + 1 - k$. Using the formula for multiplicities from Exercise 5.1, we see that

(5.10) $$P_n \simeq V_0 \oplus V_2 \oplus \cdots \oplus V_{2n}.$$

Now the computation of the eigenvalues of spherical Laplace operator is easy. Namely, by Exercise 4.3 $J_x^2 + J_y^2 + J_z^2$ acts in V_l by $-l(l+2)/4$. Thus, we get the following result.

Theorem 5.11. *The eigenvalues of spherical Laplace operator Δ_{sph} in the space P_n are*

(5.11) $$\lambda_k = -k(k+1), \qquad k = 0, \ldots, n$$

and multiplicity of λ_k is equal to $\dim V_{2k} = 2k + 1$.

Finally, we can formulate the final result about eigenfunctions in $C^\infty(S^2)$.

Theorem 5.12. *Each eigenfunction of Δ_{sph} is polynomial. The eigenvalues are given by (5.11), and multiplicity of λ_k is equal to $2k + 1$.*

Proof. Consider the space $L^2(S^2, \mathbb{C})$ of complex-valued L^2 functions on S^2. Since action of $SO(3)$ preserves the volume form, it also preserves the inner product in $L^2(S^2, \mathbb{C})$. It shows that operators J_x, J_y, J_z are skew-Hermitian, and thus, Δ_{sph} is Hermitian, or self-adjoint.

Let $E_n \subset P_n$ be the orthogonal complement to P_{n-1}. Then E_n is $SO(3)$-invariant, and it follows from (5.10) that as an $SO(3)$-module $E_n \simeq V_{2n}$, so Δ_{sph} acts on E_n by λ_n. On the other hand, since the space of polynomials is dense in L^2, we have

$$L^2(S^2, \mathbb{C}) = \bigoplus_{n \ge 0} E_n$$

(direct sum of Hilbert spaces). Thus, if $\Delta_{sph} f = \lambda f$ for some function $f \in C^\infty(S^2) \subset L^2(S^2)$, then either $\lambda \ne \lambda_n$ for all n, which forces $(f, E_n) = 0$ for all n, so $f = 0$, or $\lambda = \lambda_n$, so $(f, E_k) = 0$ for all $k \ne n$, so $f \in E_n$. $\qquad\square$

Exercises

5.1. Show that if V is a finite-dimensional representation of $\mathfrak{sl}(2, \mathbb{C})$, then $V \simeq \bigoplus n_k V_k$, and $n_k = \dim V[k] - \dim V[k+2]$. Show also that $\sum n_{2k} = \dim V[0]$, $\sum n_{2k+1} = \dim V[1]$.

5.2. Show that the symmetric power representation $S^k \mathbb{C}^2$, considered in Exercise 4.2, is isomorphic to the irreducible representation V_k with highest weight k.

5.3. Prove an analog of Theorem 5.7 for complex representations of $\mathfrak{so}(3, \mathbb{R})$, namely

(1) Every finite-dimensional representation of $\mathfrak{so}(3, \mathbb{R})$ admits a weight decomposition:

$$V = \bigoplus_{n \in \mathbb{Z}} V[n]$$

where $V[n] = \{v \in V \mid J_z v = \frac{in}{2} v\}$.

(2) A representation V of $\mathfrak{so}(3, \mathbb{R})$ can be lifted to a representation of $\mathrm{SO}(3, \mathbb{R})$ iff all weights are even: $V[k] = 0$ for all odd k (cf. with Exercise 4.1).

In physical literature, the number $j = \text{weight}/2$ is called the *spin*; thus, instead of talking say, of representation with highest weight 3, physicicts would talk about spin 3/2 representation. In this language, we see that a representation V of $\mathfrak{so}(3, \mathbb{R})$ can be lifted to a representation of $\mathrm{SO}(3, \mathbb{R})$ iff the spin is integer.

5.4. Complete the program sketched in Section 5.2 to find the eigenvalues and multiplicities of the operator

$$H = -\Delta - \frac{c}{r}, \qquad c > 0$$

in $L^2(\mathbb{R}^3, \mathbb{C})$ (this operator describes the hydrogen atom).

Chapter 6

Structure Theory of Lie Algebras

In this section, we will start developing the structure theory of Lie algebras, with the goal of getting eventually the full classification for semisimple Lie algebras and their representations.

In this chapter, $\mathfrak{g}$ will always stand for a finite-dimensional Lie algebra over the ground field $\mathbb{K}$ which can be either $\mathbb{R}$ or $\mathbb{C}$ (most results will apply equally in both cases and in fact for any field of characteristic zero). We will not be using the theory of Lie groups.

6.1. Ideals and commutant

Recall that a subalgebra of $\mathfrak{g}$ is a vector subspace closed under the commutator, and an ideal is a vector subspace $\mathfrak{h}$ such that $[x, y] \in \mathfrak{h}$ for any $x \in \mathfrak{g}, y \in \mathfrak{h}$. This definition is the natural analog of an ideal in an associative algebra. Note, however, that because of skew-symmetry of the commutator there is no difference between left and right ideals: every right ideal is also automatically a left ideal.

As in the theory of associative algebras, if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ then the quotient space $\mathfrak{g}/\mathfrak{h}$ has a canonical structure of a Lie algebra, and we have the following trivial result, proof of which is left to the reader as an exercise.

Lemma 6.1. *If $f\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is a morphism of Lie algebras, then $\operatorname{Ker} f$ is an ideal in $\mathfrak{g}_1$, $\operatorname{Im} f$ is a subalgebra in $\mathfrak{g}_2$, and f gives rise to an isomorphism of Lie algebras $\mathfrak{g}/\operatorname{Ker} f \simeq \operatorname{Im} f$.*

In addition, here is another important result about ideals.

Lemma 6.2. *Let I_1, I_2 be ideals in $\mathfrak{g}$. Define*

$$I_1 + I_2 = \{x + y \mid x \in I_1, y \in I_2\}$$
$$[I_1, I_2] = Subspace\ spanned\ by\ [x, y], \quad x \in I_1, y \in I_2.$$

Then $I_1 \cap I_2$, $I_1 + I_2$, $[I_1, I_2]$ are ideals in $\mathfrak{g}$.

One of the first ways to study Lie algebras is by analyzing how close the Lie algebra is to a commutative Lie algebra. There are several ways of making it precise.

First, we might look at how large the center $\mathfrak{z}(\mathfrak{g}) = \{x \in \mathfrak{g} \mid [x, y] = 0 \text{ for all } y \in \mathfrak{g}\}$ is. However, it turns out that it is more effective to study commutative quotients of $\mathfrak{g}$.

Definition 6.3. The commutant of a Lie algebra $\mathfrak{g}$ is the ideal $[\mathfrak{g}, \mathfrak{g}]$.

The following lemma explains the importance of the commutant.

Lemma 6.4. *The quotient $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is an abelian Lie algebra. Moreover, $[\mathfrak{g}, \mathfrak{g}]$ is the smallest ideal with this property: if $\mathfrak{g}/I$ is abelian for some ideal $I \subset \mathfrak{g}$, then $I \supset [\mathfrak{g}, \mathfrak{g}]$.*

Commutant gives us another way of measuring how far a Lie algebra is from being commutative: the smaller $[\mathfrak{g}, \mathfrak{g}]$ (and the larger $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$), the closer $\mathfrak{g}$ is to being commutative. For example, for commutative $\mathfrak{g}$, we have $[\mathfrak{g}, \mathfrak{g}] = 0$.

Example 6.5. The commutant $[\mathfrak{gl}(n, \mathbb{K}), \mathfrak{gl}(n, \mathbb{K})] = [\mathfrak{sl}(n, \mathbb{K}), \mathfrak{sl}(n, \mathbb{K})] = \mathfrak{sl}(n, \mathbb{K})$. Indeed, it is obvious that for any $z = [x, y]$ we have $\operatorname{tr} z = 0$. On the other hand, for $i \neq j$ we have $E_{ii} - E_{jj} = [E_{ij}, E_{ji}]$ and $2E_{ij} = [E_{ii} - E_{jj}, E_{ij}]$, which shows that $E_{ii} - E_{jj}, E_{ij} \in [\mathfrak{sl}(n, \mathbb{K}), \mathfrak{sl}(n, \mathbb{K})]$.

6.2. Solvable and nilpotent Lie algebras

We now can define an important class of Lie algebras.

Definition 6.6. For a Lie algebra $\mathfrak{g}$, define the series of ideals $D^i \mathfrak{g}$ (called *derived series*) by $D^0 \mathfrak{g} = \mathfrak{g}$ and
$$D^{i+1} \mathfrak{g} = [D^i \mathfrak{g}, D^i \mathfrak{g}].$$

It immediately follows from Lemma 6.2, Lemma 6.4 that each D^i is an ideal in $\mathfrak{g}$ and $D^i \mathfrak{g}/D^{i+1} \mathfrak{g}$ is abelian.

Proposition 6.7. *The following conditions are equivalent:*

(1) *$D^n \mathfrak{g} = 0$ for large enough n.*

(2) *There exists a sequence of subalgebras $\mathfrak{a}^0 = \mathfrak{g} \supset \mathfrak{a}^1 \supset \cdots \supset \mathfrak{a}^k = \{0\}$ such that $\mathfrak{a}^{i+1}$ is an ideal in $\mathfrak{a}^i$ and the quotient $\mathfrak{a}^i/\mathfrak{a}^{i+1}$ is abelian.*

(3) *For large enough n, every commutator of the form*
$$[\ldots [[x_1, x_2], [x_3, x_4]] \ldots]$$
(2^n terms, arranged in a binary tree of length n) is zero.

Proof. Equivalence of (1) and (3) is obvious. Implication (1) $\implies$ (2) is also clear: we can take $\mathfrak{a}^i = D^i \mathfrak{g}$. To prove (2) $\implies$ (1), note that if $\mathfrak{a}^i$ satisfies the conditions of the proposition, then by Lemma 6.4, we have $\mathfrak{a}^{i+1} \supset [\mathfrak{a}^i, \mathfrak{a}^i]$. Thus, reasoning by induction, we see that $\mathfrak{a}^i \supset D^i \mathfrak{g}$. $\qquad\square$

Definition 6.8. Lie algebra $\mathfrak{g}$ is called solvable if it satisfies any of the equivalent conditions of Proposition 6.7.

Informally, a solvable Lie algebra is an "almost commutative" Lie algebra: it is an algebra that can be obtained by successive extensions of commutative algebras.

This is not the only way of making the notion of "almost commutative" Lie algebra precise. Another class of Lie algebras can be defined as follows.

Definition 6.9. For a Lie algebra $\mathfrak{g}$, define a series of ideals $D_i \mathfrak{g} \subset \mathfrak{g}$ (called *lower central series*) by $D_0 \mathfrak{g} = \mathfrak{g}$ and
$$D_{i+1} \mathfrak{g} = [\mathfrak{g}, D_i \mathfrak{g}].$$

Proposition 6.10. *The following conditions are equivalent:*

(1) *$D_n \mathfrak{g} = 0$ for large enough n.*

(2) *There exists a sequence of ideals $\mathfrak{a}_0 = \mathfrak{g} \supset \mathfrak{a}_1 \supset \cdots \supset \mathfrak{a}_k = \{0\}$ such that $[\mathfrak{g}, \mathfrak{a}_i] \subset \mathfrak{a}_{i+1}$.*

(3) *For large enough n, every commutator of the form*

$$[\ldots[[x_1, x_2], x_3], x_4] \ldots x_n]$$

(*n terms*) *is zero.*

Proof. Equivalence of (1) and (3) is obvious. Implication (1) $\implies$ (2) is also clear: we can take $\mathfrak{a}_i = D_i\mathfrak{g}$. To prove (2) $\implies$ (1), note that if $\mathfrak{a}_i$ satisfies the conditions of the proposition, then by induction, we see that $\mathfrak{a}_i \supset D_i\mathfrak{g}$. $\qquad\square$

Definition 6.11. Lie algebra $\mathfrak{g}$ is called nilpotent if it satisfies any of the equivalent conditions of Proposition 6.10.

Example 6.12. Let $\mathfrak{b} \subset \mathfrak{gl}(n, \mathbb{K})$ be the subalgebra of upper triangular matrices, and $\mathfrak{n}$ be the subalgebra of all strictly upper triangular matrices. Then $\mathfrak{b}$ is solvable, and $\mathfrak{n}$ is nilpotent.

To prove it, let us first generalize it. Namely, if $\mathcal{F}$ is a flag in a finite-dimensional vector space V:

$$\mathcal{F} = (\{0\} \subset V_1 \subset V_2 \subset \ldots V_n = V)$$

with $\dim V_i < \dim V_{i+1}$ (we do not require that $\dim V_i = i$), then define

$$\mathfrak{b}(\mathcal{F}) = \{x \in \mathfrak{gl}(V) \mid xV_i \subset V_i \text{ for all } i\},$$
$$\mathfrak{n}(\mathcal{F}) = \{x \in \mathfrak{gl}(V) \mid xV_i \subset V_{i-1} \text{ for all } i\}.$$

By taking $\mathcal{F}$ to be the standard flag in $\mathbb{K}^n$ (see Example 2.23) we recover the Lie algebras $\mathfrak{b}, \mathfrak{n}$ defined above.

We claim that $\mathfrak{b}(\mathcal{F})$ is solvable and $\mathfrak{n}(\mathcal{F})$ is nilpotent. Indeed, define more general algebras

$$\mathfrak{a}_k(\mathcal{F}) = \{x \in \mathfrak{gl}(V) \mid xV_i \subset V_{i-k} \text{ for all } i\}$$

so that $\mathfrak{b}(\mathcal{F}) = \mathfrak{a}_0$, $\mathfrak{n}(\mathcal{F}) = \mathfrak{a}_1$. Then it is obvious that for $x \in \mathfrak{a}_k, y \in \mathfrak{a}_l$, we have $xy \in \mathfrak{a}_{k+l}$ (here xy is the usual product in $\text{End}(V)$); thus, $[\mathfrak{a}_k, \mathfrak{a}_l] \subset \mathfrak{a}_{k+l}$, so $D_i\mathfrak{n} \subset \mathfrak{a}_{i+1}$. This proves nilpotency of $\mathfrak{n}(\mathcal{F})$.

To show solvability of $\mathfrak{b}$, note that even though for $x, y \in \mathfrak{b}$ we can only say that $xy \in \mathfrak{b}$, for the commutator we have a stronger condition: $[x, y] \in \mathfrak{n} = \mathfrak{a}_1$. Indeed, diagonal entries of xy and yx coincide. Thus, $D^1\mathfrak{b} \subset \mathfrak{n} = \mathfrak{a}_1$. From here it easily follows by induction that $D^{i+1}\mathfrak{b} \subset \mathfrak{a}_{2^i}$.

Note, finally, that $\mathfrak{b}$ is not nilpotent: $D_2\mathfrak{b} = [\mathfrak{b}, D_1\mathfrak{b}] = D_1\mathfrak{b} = \mathfrak{n}$, which can be easily deduced from $[x, E_{ij}] = (\lambda_i - \lambda_j)E_{ij}$ if x is a diagonal matrix with entries λ_i.

The following theorem summarizes some basic properties of solvable and nilpotent Lie algebras.

Theorem 6.13.

(1) *A real Lie algebra $\mathfrak{g}$ is solvable (respectively, nilpotent) iff its complexification $\mathfrak{g}_{\mathbb{C}}$ is solvable (respectively, nilpotent).*

(2) *If $\mathfrak{g}$ is solvable, then any subalgebra, quotient of $\mathfrak{g}$ is also solvable. If $\mathfrak{g}$ is nilpotent, then any subalgebra, quotient of $\mathfrak{g}$ is also nilpotent.*

(3) *If $\mathfrak{g}$ is nilpotent, then $\mathfrak{g}$ is solvable.*

(4) *If $I \subset \mathfrak{g}$ is an ideal such that both $I, \mathfrak{g}/I$ are solvable, then $\mathfrak{g}$ is solvable.*

Proof. Parts (1), (2) are obvious if we use definition of solvable algebra in the form "any commutator of the form ... is zero", and similarly for nilpotent. Part (3) follows from inclusion $D^i\mathfrak{g} \subset D_i\mathfrak{g}$, which can be easily proved by induction.

Finally, to prove part (4), denote by φ the canonical projection $\mathfrak{g} \to \mathfrak{g}/I$. Then $\varphi(D^n\mathfrak{g}) = D^n(\mathfrak{g}/I) = 0$ for some n. Thus, $D^n\mathfrak{g} \subset I$. Therefore, $D^{n+k}\mathfrak{g} \subset D^k I$, so $D^{n+k}\mathfrak{g} = 0$ for large enough k. $\qquad\square$

6.3. Lie and Engel theorems

The main result of this section is the following theorem.

Theorem 6.14 (Lie theorem about representations of solvable Lie algebra). *Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a complex representation of a solvable Lie algebra $\mathfrak{g}$. Then there exists a basis in V such that in this basis, all operators $\rho(x)$ are upper-triangular.*

This theorem is a generalization of a well-known result that any operator in a complex vector space can be brought to an upper-triangular form by a change of basis.

The key step in the proof of the theorem is the following result.

Proposition 6.15. *Let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a complex representation of a solvable Lie algebra $\mathfrak{g}$. Then there exists a vector $v \in V$ which is a common eigenvector of all $\rho(x)$, $x \in \mathfrak{g}$.*

Proof. The proof goes by induction in dimension of $\mathfrak{g}$. Since $\mathfrak{g}$ is solvable, $[\mathfrak{g}, \mathfrak{g}] \neq \mathfrak{g}$. Let $\mathfrak{g}' \subset \mathfrak{g}$ be a subspace which contains $[\mathfrak{g}, \mathfrak{g}]$ and has codimension 1 in $\mathfrak{g}$: $\mathfrak{g} = \mathfrak{g}' \oplus \mathbb{C}x$. Then $\mathfrak{g}'$ is an ideal in $\mathfrak{g}$; thus, $\mathfrak{g}'$ is solvable.

By induction assumption, there exists $v \in V$ which is a common eigenvector for all $\rho(h), h \in \mathfrak{g}'$: $\rho(h)v = \lambda(h)v$. Consider the vector space W spanned by $v, v^1 = \rho(x)v, v^2 = (\rho(x))^2v, \ldots$.

We claim that W is stable under action of any $h \in \mathfrak{g}'$; moreover,

$$hv^k = \lambda(h)v^k + \sum_{l<k} a_{kl}(h)v^l.$$

This is easily proved by induction: indeed,

(6.1) $hv^k = hxv^{k-1} = xhv^{k-1} + [h,x]v^{k-1} = \lambda(h)xv^{k-1} + \lambda([h,x])v^{k-1} + \ldots$

Thus, W is stable under the action of $\mathfrak{g}$. Since W is finite-dimensional, we can choose some n such that $v, v^1, \ldots v^n$ is a basis in W. Then we see that in this basis, any $\rho(h)$ is upper-triangular, with $\lambda(h)$ on the diagonal. In particular, this implies that $\operatorname{tr}_W \rho(h) = (n+1)\lambda(h)$.

Since $\operatorname{tr}_W[\rho(x), \rho(h)] = 0$, this implies that $\lambda([h,x]) = 0$ for any $h \in \mathfrak{g}'$. Going back to formula (6.1), we see that this implies $hv^k = \lambda(h)v^k$. Therefore, any vector $w \in W$ is a common eigenvector for all $h \in \mathfrak{g}'$. Choosing w to be an eigenvector for x, we get the statement of the proposition. $\quad\square$

This proposition immediately implies Lie theorem:

Proof of Theorem 6.14. Proof goes by induction in $\dim V$. By Proposition 6.15, there exists a common eigenvector v for all $x \in \mathfrak{g}$. Consider the space $V/\mathbb{C}v$. By induction assumption, there exists a basis $v_1, v_2, \ldots$ in $V/\mathbb{C}v$ such that the action of $\mathfrak{g}$ in this basis of $V/\mathbb{C}v$ is upper-triangular. For each of these vectors, choose a preimage $\tilde{v}_i \in V$. Then one immediately sees that the action of any $x \in \mathfrak{g}$ in the basis $v, \tilde{v}_1, \tilde{v}_2, \ldots$ is upper-triangular. $\quad\square$

This theorem has a number of useful corollaries.

Corollary 6.16.

(1) *Any irreducible representation of a solvable Lie algebra is 1-dimensional.*

(2) *If a complex Lie algebra $\mathfrak{g}$ is solvable, then there exists a sequence $0 \subset I_1 \subset \cdots \subset I_n = \mathfrak{g}$, where each I_k is an ideal in $\mathfrak{g}$ and I_{k+1}/I_k is one dimensional.*

(3) $\mathfrak{g}$ *is solvable if and only if $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent.*

Proof. Part (1) is obvious from Proposition 6.15; part (2) is immediately obtained if we apply Lie theorem to adjoint representation and note that a subrepresentation of the adjoint representation is the same as an ideal in $\mathfrak{g}$.

To prove part (3), note that implication in one direction is obvious. Indeed, if $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent, then it is also solvable; since $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is commutative (and thus solvable), by Theorem 6.13, $\mathfrak{g}$ itsefl is solvable.

Conversely, assume that $\mathfrak{g}$ is solvable. Without loss of generality, we may assume that $\mathfrak{g}$ is complex. Apply Lie theorem to the adjoint representation. Then $\operatorname{ad} \mathfrak{g} \subset \mathfrak{b}$ (algebra of upper-triangular matrices) in some basis of $\mathfrak{g}$; thus, by results of Example 6.12, the algebra $[\operatorname{ad} \mathfrak{g}, \operatorname{ad} \mathfrak{g}] = \operatorname{ad}[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{n}$ is nilpotent, so $\operatorname{ad}[x_1, [\ldots [x_{n-1}, x_n] \ldots] = 0$ for sufficiently large n and all $x_i \in [\mathfrak{g}, \mathfrak{g}]$. Thus, $[y, [x_1, [\ldots [x_{n-1}, x_n] \ldots] = 0$ for sufficiently large n and all $x_i, y \in [\mathfrak{g}, \mathfrak{g}]$. $\qquad\square$

One also might ask if there is an analog of Lie theorem for nilpotent Lie algebras. Of course, since every nilpotent Lie algebra is automatically solvable (Theorem 6.13), Lie theorem shows that in any representation of a nilpotent algebra, operators $\rho(x)$ are upper-triangular in a certain basis. One wonders whether one has a stronger result — for example, whether operators $\rho(x)$ can be made strictly upper–triangular. Here the answer is obviously negative: it suffices to take a commutative Lie algebra which acts diagonally in $\mathbb{C}^n$.

The proper analog of Lie theorem for nilpotent Lie algebras is given by the following result.

Theorem 6.17. *Let V be a finite-dimensional vector space, either real or complex, and let $\mathfrak{g} \subset \mathfrak{gl}(V)$ be a Lie subaglebra which consists of nilpotent operators. Then there exists a basis in V such that all operators $x \in \mathfrak{g}$ are strictly upper-triangular.*

The proof of this theorem will not be given here; interested reader can find it in [**20**], [**11**], or [**12**]. It is not very difficult and in fact is rather similar to the proof of Lie theorem; the only reason it is not given here is because it does not give any new insight.

As an immediate corollary, we get the following theorem.

Theorem 6.18 (Engel theorem). *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if for every $x \in \mathfrak{g}$, the operator $\operatorname{ad} x \in \operatorname{End}(\mathfrak{g})$ is nilpotent.*

Proof. One direction is obvious: if $\mathfrak{g}$ is nilpotent then by definition, $[x, [x, \ldots [x, y \ldots] = (\operatorname{ad} x)^n . y = 0$ for large enough n.

Conversely, if $\operatorname{ad} x$ is nilpotent for every x, then by the previous theorem, there exists a sequence of subspaces $0 \subset \mathfrak{g}_1 \subset \mathfrak{g}_2 \cdots \subset \mathfrak{g}_n = \mathfrak{g}$ such that $\operatorname{ad} x . \mathfrak{g}_i \subset \mathfrak{g}_{i-1}$. This shows that each $\mathfrak{g}_i$ is an ideal in $\mathfrak{g}$ and moreover, $[\mathfrak{g}, \mathfrak{g}_i] \subset \mathfrak{g}_{i-1}$. Thus, $\mathfrak{g}$ is nilpotent. $\qquad\square$

6.4. The radical. Semisimple and reductive algebras

So far, we have defined the notion of a solvable Lie algebra; informally, a solvable Lie algebra is the one which is close to being abelian. In this section, we will describe opposite extreme case, Lie algebras which are as far as possible from being abelian (they are called semisimple) and show that in a reasonable sense, any Lie algebra is built out of a solvable and semisimple one.

Definition 6.19. A Lie algebra $\mathfrak{g}$ is called semisimple if it contains no nonzero solvable ideals.

Note that this in particular implies that the center $\mathfrak{z}(\mathfrak{g}) = 0$.

A special case of semisimple Lie algebras is given by simple ones.

Definition 6.20. A Lie algebra $\mathfrak{g}$ is called simple if it is not abelian and contains no ideals other than 0 and $\mathfrak{g}$.

The condition that $\mathfrak{g}$ should not be abelian is included to rule out one-dimensional Lie algebra: there are many reasons not to include it in the class of simple Lie algebras. One of these reasons is the following lemma.

Lemma 6.21. *Any simple Lie algebra is semisimple.*

Proof. If $\mathfrak{g}$ is simple, then it contains no ideals other than 0 and $\mathfrak{g}$. Thus, if $\mathfrak{g}$ contains a nonzero solvable ideal, then it must coincide with $\mathfrak{g}$, so $\mathfrak{g}$ must be solvable. But then $[\mathfrak{g}, \mathfrak{g}]$ is an ideal which is strictly smaller than $\mathfrak{g}$ (because $\mathfrak{g}$ is solvable) and nonzero (because $\mathfrak{g}$ is not abelian). This gives a contradiction. $\square$

Example 6.22. The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is simple. Indeed, recall that $\operatorname{ad} h$ is diagonal in the basis e, f, h, with eigenvalues $2, -2, 0$ (see Section 3.10). Any ideal in $\mathfrak{g}$ must be stable under $\operatorname{ad} h$. Now we can use the following easy to prove result from linear algebra: if A is a diagonalizable operator in a finite-dimensional vector space, with distinct eigenvalues: $Av_i = \lambda_i v_i, \lambda_i \neq \lambda_j$, then the only subspaces invariant under A are those spanned by some of the eigenvectors v_i. Applying this to $\operatorname{ad} h$, we see that any ideal in $\mathfrak{sl}(2, \mathbb{C})$ must be spanned as a vector space by a subset of $\{e, f, h\}$.

But if an ideal I contains h, then $[h, e] = 2e \in I, [h, f] = -2f \in I$, so $I = \mathfrak{sl}(2, \mathbb{C})$. If I contains e, then $[e, f] = h \in I$, so again $I = \mathfrak{sl}(2, \mathbb{C})$. Similarly, if $f \in I$, then $I = \mathfrak{sl}(2, \mathbb{C})$. Thus, $\mathfrak{sl}(2, \mathbb{C})$ contains no non-trivial ideals.

In the next section, we will generalize this result and show that classical Lie algebras such as $\mathfrak{sl}(n, \mathbb{C}), \mathfrak{su}(n), \mathfrak{sp}(2n, \mathbb{C}), \mathfrak{so}(n, \mathbb{C})$ are semisimple.

For a general Lie algebra $\mathfrak{g}$, which is neither semisimple nor solvable, we can try to "separate" the solvable and semisimple parts.

Proposition 6.23. *In any Lie algebra $\mathfrak{g}$, there is a unique solvable ideal which contains any other solvable ideal. This solvable ideal is called the* radical *of $\mathfrak{g}$ and denoted by* $\operatorname{rad}(\mathfrak{g})$.

Proof. Uniqueness is obvious. To show existence, note that if I_1, I_2 are solvable ideals, then so is $I_1 + I_2$. Indeed, it contains solvable ideal I_1 and the quotient $(I_1 + I_2)/I_1 = I_2/(I_1 \cap I_2)$ is also solvable since it is a quotient of I_2. Thus, by Theorem 6.13, $I_1 + I_2$ is also solvable. By induction, this shows that any finite sum of solvable ideals is also solvable. Thus, we can let $\operatorname{rad}(\mathfrak{g}) = \sum I$, where the sum is taken over all solvable ideals (finite-dimensionality of $\mathfrak{g}$ shows that it suffices to take finite sum). $\square$

Using this definition, we can rewrite the definition of a semisimple Lie algebra as follows: $\mathfrak{g}$ is semisimple iff $\operatorname{rad}(\mathfrak{g}) = 0$.

Theorem 6.24. *For any Lie algebra $\mathfrak{g}$, the quotient $\mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ is semisimple. Conversely, if $\mathfrak{b}$ is a solvable ideal in $\mathfrak{g}$ such that $\mathfrak{g}/\mathfrak{b}$ is semisimple, then $\mathfrak{b} = \operatorname{rad}(\mathfrak{g})$.*

Proof. Assume that $\mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ contains a solvable ideal I. Consider the ideal $\tilde{I} = \pi^{-1}(I) \subset \mathfrak{g}$, where π is the canonical map $g \to \mathfrak{g}/\operatorname{rad}(\mathfrak{g})$. Then $\tilde{I} \supset \operatorname{rad}(\mathfrak{g})$ and $\tilde{I}/\operatorname{rad}(\mathfrak{g}) = I$ is solvable. Thus, by Theorem 6.13, $\tilde{I}$ is solvable, so $\tilde{I} = \operatorname{rad}(\mathfrak{g})$, $I = 0$.

Proof of the second statement is left to the reader as an exercise. $\qquad\square$

This theorem shows that any Lie algebra can be included in a short exact sequence $0 \to \mathfrak{b} \to \mathfrak{g} \to \mathfrak{g}_{ss} \to 0$, where $\mathfrak{b}$ is solvable and $\mathfrak{g}_{ss}$ is semisimple. In fact, one has a much stronger result.

Theorem 6.25 (Levi theorem). *Any Lie algebra can be written as a direct sum*

$$\text{(6.2)} \qquad \mathfrak{g} = \mathrm{rad}(\mathfrak{g}) \oplus \mathfrak{g}_{ss}$$

where $\mathfrak{g}_{ss}$ is a semisimple subalgebra (not an ideal!) in $\mathfrak{g}$. Such a decomposition is called Levi decomposition *for $\mathfrak{g}$.*

This theorem will not be proved here. A proof can be found in standard textbooks on Lie algebras, such as [**20**] or [**11**]. We only mention here that the proof essentially reduces to showing that cohomology $H^2(\mathfrak{g}, \mathbb{C}) = 0$ vanishes for semisimple $\mathfrak{g}$ and is done using methods similar to those we will use in Section 6.9 to show complete reducibility of representations.

Example 6.26. Let $G = \mathrm{SO}(3, \mathbb{R}) \ltimes \mathbb{R}^3$ be the Poincare group, i.e. the group of all maps $\mathbb{R}^3 \to \mathbb{R}^3$ which have the form $x \mapsto Ax + b, A \in \mathrm{SO}(3, \mathbb{R}), b \in \mathbb{R}^3$. The corresponding Lie algebra is $\mathfrak{g} = \mathfrak{so}(3, \mathbb{R}) \oplus \mathbb{R}^3$, where the commutator is given by $[(A_1, b_1), (A_2, b_2)] = ([A_1, A_2], A_1 b_2 - A_2 b_1)$. Thus, $\mathbb{R}^3$ is an ideal and $\mathfrak{so}(3, \mathbb{R})$ is a subalgebra. Since $\mathbb{R}^3$ is abelian and $\mathfrak{so}(3, \mathbb{R})$ is semisimple (which follows from semisimplicity of $\mathfrak{so}(3, \mathbb{R})_{\mathbb{C}} \simeq \mathfrak{sl}(2, \mathbb{C})$, see Example 6.22), we see that $\mathfrak{g} = \mathfrak{so}(3, \mathbb{R}) \oplus \mathbb{R}^3$ is exactly the Levi decomposition.

Another instructive example of Levi decomposition is the Levi decomposition for parabolic subalgebras; a special case is given in Exercise 6.3.

As in the theory of associative algebras, there is a relation between the radical of $\mathfrak{g}$ and kernels of irreducible representations.

Theorem 6.27. *Let V be an irreducible representation of $\mathfrak{g}$. Then any $h \in \mathrm{rad}(\mathfrak{g})$ acts in V by scalar operators: $\rho(h) = \lambda(h)\,\mathrm{id}$. Also, any $h \in [\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$ acts by zero.*

Proof. By Proposition 6.15, there is a common eigenvector in V for all $h \in \mathrm{rad}(\mathfrak{g})$: $\rho(h).v = \lambda(h)v$ for some $\lambda \colon \mathrm{rad}(\mathfrak{g}) \to \mathbb{C}$. Define $V_\lambda = \{w \in V \mid \rho(h)w = \lambda(h)w \text{ for all } h \in \mathrm{rad}(\mathfrak{g})\}$. Then the same argument as in the proof of Proposition 6.15 shows that for any $x \in \mathfrak{g}$, one has $\rho(x)(V_\lambda) \subset V_\lambda$. Thus, V_λ is a subrepresentation; since it is non-zero and V is irreducible, we must have $V = V_\lambda$, which proves the first statement of the theorem. The second statement immediately follows from the first one. $\qquad\square$

From the point of view of representation theory, having non-zero elements which act by zero in any irreducible representation significantly complicates the theory. Thus, it is natural to consider a class of algebras for which $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})] = 0$.

Definition 6.28. A Lie algebra is called *reductive* if $\mathrm{rad}(\mathfrak{g}) = \mathfrak{z}(\mathfrak{g})$, i.e. if $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is semisimple. (Recall that $\mathfrak{z}(\mathfrak{g})$ is the center of $\mathfrak{g}$.)

Of course, any semisimple Lie algebra is reductive (because then $\mathrm{rad}(\mathfrak{g}) = \mathfrak{z}(\mathfrak{g}) = 0$), but converse is not true: for example, any Lie algebra which is a direct sum of an abelian and semisimple

$$\text{(6.3)} \qquad \mathfrak{g} = \mathfrak{z} \oplus \mathfrak{g}_{ss}, \qquad [\mathfrak{z}, \mathfrak{g}_{ss}] = 0$$

is reductive. In fact, it follows from Levi theorem that any reductive Lie algebra must have such form. Later we will give an alternative proof of this result, which does not use Levi theorem (see Theorem 6.61).

In the next section we will show that many classical Lie algebras such as $\mathfrak{gl}(n, \mathbb{C})$ or $\mathfrak{u}(n)$ are reductive.

6.5. Invariant bilinear forms and semisimplicity of classical Lie algebras

So far, we have only one example of a semisimple Lie algebra, namely $\mathfrak{sl}(2, \mathbb{C})$ (see Example 6.22), and the proof of its semisimplicity was done by brute force, by analyzing all possibilities for an ideal. It is clear that such a proof would be difficult to generalize to higher-dimensional Lie algebras: we need better tools.

The standard approach to the study of semisimplicity is based on the notion of invariant bilinear form. Recall that a bilinear form B on $\mathfrak{g}$ is called invariant if

$$B(\operatorname{ad} x.y, z) + B(y, \operatorname{ad} x.z) = 0$$

for any $x, y, z \in \mathfrak{g}$ (see Example 4.14). The following lemma shows the importance of such forms.

Lemma 6.29. *Let B be an invariant bilinear form on $\mathfrak{g}$, and $I \subset \mathfrak{g}$ an ideal. Let $I^{\perp}$ be the orthogonal complement of I with respect to B: $I^{\perp} = \{x \in \mathfrak{g} \mid B(x, y) = 0 \text{ for all } y \in I\}$. Then $I^{\perp}$ is also an ideal in $\mathfrak{g}$. In particular, $\operatorname{Ker} B = \mathfrak{g}^{\perp}$ is an ideal in $\mathfrak{g}$.*

The proof of this lemma is trivial and left to the reader. Note, however, that in general we can not write $\mathfrak{g} = I \oplus I^{\perp}$, as it is quite possible that $I \cap I^{\perp} \neq 0$, even for a non-degenerate form B.

Example 6.30. Let $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$ and define the form by $B(x, y) = \operatorname{tr}(xy)$. Then it is a symmetric invariant bilinear form on $\mathfrak{g}$. Indeed, symmetry is well-known and invariance follows from the following identity

$$\operatorname{tr}([x, y]z + y[x, z]) = \operatorname{tr}(xyz - yxz + yxz - yzx) = \operatorname{tr}(xyz - yzx) = 0.$$

In fact, there is an even easier proof: since $\operatorname{tr}(gxg^{-1}gyg^{-1}) = \operatorname{tr}(gxyg^{-1}) = \operatorname{tr}(xy)$ for any $g \in \operatorname{GL}(n, \mathbb{C})$, we see that this form is invariant under adjoint action of $\operatorname{GL}(n, \mathbb{C})$ which is equivalent to the invariance under the action of $\mathfrak{gl}(n, \mathbb{C})$.

This example can be easily generalized.

Proposition 6.31. *Let V be a representation of $\mathfrak{g}$ and define bilinear form on $\mathfrak{g}$ by*

(6.4) $$B_V(x, y) = \operatorname{tr}_V(\rho(x)\rho(y)).$$

Then B_V is a symmetric invariant bilinear form on $\mathfrak{g}$.

The proof is identical to the proof in Example 6.30.

However, this form can be degenerate or even zero. It turns out, however, that there is a close relation between non-degeneracy of such forms and semisimplicity of $\mathfrak{g}$.

Theorem 6.32. *Let $\mathfrak{g}$ be a Lie algebra with a representation V such that the form B_V defined by (6.4) is non-degenerate. Then $\mathfrak{g}$ is reductive.*

Proof. It suffices to show that $[\mathfrak{g}, \operatorname{rad}(\mathfrak{g})] = 0$. Let $x \in [\mathfrak{g}, \operatorname{rad}(\mathfrak{g})]$; then, by Theorem 6.27, x acts by zero in any irreducible representation V_i and thus $x \in \operatorname{Ker} B_{V_i}$. But if we have a short exact sequence of representations $0 \to V_1 \to W \to V_2 \to 0$, then $B_W = B_{V_1} + B_{V_2}$ (see Exercise 6.1). Thus, arguing by induction it is easy to see that for any representation V, we would have $x \in \operatorname{Ker} B_V$. Since by assumption B_V is non-degenerate, this shows $x = 0$. $\qquad\square$

As an immediate corollary, we have the following important result.

Theorem 6.33. *All classical Lie algebras of Section* 2.5 *are reductive. Algebras* $\mathfrak{sl}(n,\mathbb{K})$, $\mathfrak{so}(n,\mathbb{K})$ $(n > 2)$, $\mathfrak{su}(n)$, $\mathfrak{sp}(2n,\mathbb{K})$ *are semisimple; algebras* $\mathfrak{gl}(n,\mathbb{K})$ *and* $\mathfrak{u}(n)$ *have one-dimensional center:* $\mathfrak{gl}(n,\mathbb{K}) = \mathbb{K} \cdot \mathrm{id} \oplus \mathfrak{sl}(n,\mathbb{K})$, $\mathfrak{u}(n) = i\mathbb{R} \cdot \mathrm{id} \oplus \mathfrak{su}(n)$. *(As before,* $\mathbb{K}$ *is either* $\mathbb{R}$ *or* $\mathbb{C}$.)

Proof. For each of these subalgebras, consider the trace form B_V where V is the defining representation ($\mathbb{K}^n$ for $\mathfrak{gl}(n,\mathbb{K})$, $\mathfrak{sl}(n,\mathbb{K})$, $\mathfrak{so}(n,\mathbb{K})$; $\mathbb{C}^n$ for $\mathfrak{su}(n)$, $\mathfrak{u}(n)$ and $\mathbb{K}^{2n}$ for $\mathfrak{sp}(2n,\mathbb{K})$). Then this form is non-degenerate. Indeed, for $\mathfrak{gl}(n)$ it follows because $B(x,y) = \sum x_{ij}y_{ji}$ which is obviously non-degenerate; for $\mathfrak{sl}(n)$ it follows from the result for $\mathfrak{gl}(n)$ and decomposition $\mathfrak{gl}(n) = \mathbb{K} \cdot \mathrm{id} \oplus \mathfrak{sl}(n)$, with the two summands being orthogonal with respect to the form B.

For $\mathfrak{so}(n)$, we have $B(x,y) = \sum x_{ij}y_{ji} = -2\sum_{i>j} x_{ij}y_{ij}$ so it is again non-degenerate. Similarly, for $\mathfrak{u}(n)$ we have $B(x,y) = -\operatorname{tr} x\overline{y}^t = -\sum x_{ij}\overline{y}_{ij}$; in particular, $B(x,x) = -\sum |x_{ij}|^2$, so this form is negative definite and in particular, non-degenerate. Therefore, its restriction to $\mathfrak{su}(n) \subset \mathfrak{u}(n)$ is also negative definite and thus non-degenerate.

The non-degeneracy of this form for $\mathfrak{sp}(2n,\mathbb{K})$ is left as an exercise (Exercise 6.4).

Thus, by Theorem 6.32 we see that each of these Lie algebras is reductive. Since the center of each of them is easy to compute (see Example 4.23), we get the statement of the theorem. $\square$

6.6. Killing form and Cartan criterion

In the previous section, we have shown that for any representation V of a Lie algebra $\mathfrak{g}$, the bilinear form $B_V(x,y) = \operatorname{tr}(\rho(x)\rho(y))$ is invariant and bilinear. An important special case is when we take V to be the adjoint representation.

Definition 6.34. The Killing form is the bilinear form on $\mathfrak{g}$ defined by $K(x,y) = \operatorname{tr}(\operatorname{ad} x \ \operatorname{ad} y)$.

The notation $K(x,y)$ can be ambiguous: if we have a subalgebra $\mathfrak{h} \subset \mathfrak{g}$, then $K(x,y), x,y \in \mathfrak{h}$, can mean either trace in $\mathfrak{g}$ or trace in $\mathfrak{h}$. In such cases we will write $K^{\mathfrak{h}}$ for Killing form of $\mathfrak{h}$ and $K^{\mathfrak{g}}$ for the restriction of Killing form of $\mathfrak{g}$ to $\mathfrak{h}$. Note, however, that if I is an ideal in $\mathfrak{g}$, then K^I coincides with the restriction of $K^{\mathfrak{g}}$ to I (see Exercise 6.1).

It follows from Proposition 6.31 that the Killing form is a symmetric invariant form on $\mathfrak{g}$.

Example 6.35. Let $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$. Then in the basis e, h, f, the operators $\operatorname{ad} e, \operatorname{ad} h, \operatorname{ad} f$ are given by

$$\operatorname{ad} e = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \operatorname{ad} h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \qquad \operatorname{ad} f = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

so explicit computation shows that the Killing form is given by $K(h,h) = 8$, $K(e,f) = K(f,e) = 4$, and $K(h,e) = K(h,f) = 0$. Thus, $K(x,y) = 4\operatorname{tr}(xy)$. This is not surprising: we already know that $\mathfrak{sl}(2,\mathbb{C})$ is simple, and by Exercise 4.4, this implies that the invariant bilinear form, if exists, is unique up to a factor.

The following two theorems show that non-degeneracy of Killing form is closely related to semisimplicity of $\mathfrak{g}$.

Theorem 6.36 (Cartan criterion of solvability). *Lie algebra* $\mathfrak{g}$ *is solvable iff* $K([\mathfrak{g},\mathfrak{g}],\mathfrak{g}) = 0$, *i.e.* $K(x,y) = 0$ *for any* $x \in [\mathfrak{g},\mathfrak{g}]$, $y \in \mathfrak{g}$.

Theorem 6.37 (Cartan criterion of semisimplicity). *Lie algebra is semisimple iff the Killing form is non-degenerate.*

The proof of these theorems will take the rest of this section. It is based on Jordan decomposition, i.e. the decomposition of a linear operator in a sum of diagonalizable (semisimple) and nilpotent one. We state here some results about this decomposition. Their proof, which is pure linear algebra, is given in Appendix B. In this theorem, we use the word "semisimple" as a synonym of "diagonalizable" (we are only discussing complex vector spaces).

Theorem 6.38. *Let V be a finite-dimensional complex vector space.*

(1) *Any linear operator A can be uniquely written as a sum of commuting semisimple and nilpotent ones:*

(6.5) $$A = A_s + A_n, \quad A_s A_n = A_n A_s, \quad A_n \text{ nilpotent}, \quad A_s \text{ semisimple}$$

(2) *For an operator $A\colon V \to V$, define $\operatorname{ad} A\colon \operatorname{End}(V) \to \operatorname{End}(V)$ by $\operatorname{ad} A.B = AB - BA$. Then*

$$(\operatorname{ad} A)_s = \operatorname{ad} A_s$$

and $\operatorname{ad} A_s$ can be written in the form $\operatorname{ad} A_s = P(\operatorname{ad} A)$ for some polynomial $P \in t\mathbb{C}[t]$ (depending on A).

(3) *Define $\overline{A}_s$ to be the operator which has the same eigenspaces as A_s but complex conjugate eigenvalues: if $A_s v = \lambda v$, then $\overline{A}_s v = \bar{\lambda} v$. Then $\operatorname{ad} \overline{A}_s$ can be written in the form $\operatorname{ad} \overline{A}_s = Q(\operatorname{ad} A)$ for some polynomial $Q \in t\mathbb{C}[t]$ (depending on A).*

Using this theorem, we can now give the proof of Cartan criterion.

Proof of Theorem 6.36. First, note that if $\mathfrak{g}$ is a real Lie algebra, then $\mathfrak{g}$ is solvable iff $\mathfrak{g}_\mathbb{C}$ is solvable (Theorem 6.13), and $K([\mathfrak{g}, \mathfrak{g}], \mathfrak{g}) = 0$ iff $K([\mathfrak{g}_\mathbb{C}, \mathfrak{g}_\mathbb{C}], \mathfrak{g}_\mathbb{C}) = 0$ (obvious). Thus, it suffices to prove the theorem for complex Lie algebras. So from now on we assume that $\mathfrak{g}$ is complex.

Assume that $\mathfrak{g}$ is solvable. Then by Lie theorem, there is a basis in $\mathfrak{g}$ such that all $\operatorname{ad} x$ are upper-triangular. Then in this basis, $\operatorname{ad} y, y \in [\mathfrak{g}, \mathfrak{g}]$ are strictly upper-triangular, so $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$.

To prove the opposite direction, we first prove the following lemma.

Lemma 6.39. *Let V be a complex vector space and $\mathfrak{g} \subset \mathfrak{gl}(V)$ — a Lie subalgebra such that for any $x \in [\mathfrak{g}, \mathfrak{g}], y \in \mathfrak{g}$ we have $\operatorname{tr}(xy) = 0$. Then $\mathfrak{g}$ is solvable.*

Proof. Let $x \in [\mathfrak{g}, \mathfrak{g}]$. By Theorem 6.38, it can be written in the form $x = x_s + x_n$. Consider now $\operatorname{tr}(x \overline{x}_s)$ where $\overline{x}_s$ is as in Theorem 6.38. On one hand, we see that $\operatorname{tr}(x \overline{x}_s) = \sum \lambda_i \bar{\lambda}_i = \sum |\lambda_i|^2$, where λ_i are eigenvalues of x. On the other hand, if $x = \sum [y_i, z_i]$, then

$$\operatorname{tr}(x \overline{x}_s) = \operatorname{tr}(\sum [y_i, z_i] \overline{x}_s) = \sum \operatorname{tr}(y_i [z_i, \overline{x}_s]) = -\sum \operatorname{tr}(y_i [\overline{x}_s, z_i])$$

By Theorem 6.38, $[\overline{x}_s, z_i] = \operatorname{ad} \overline{x}_s.z_i = Q(\operatorname{ad} x).z_i \in [\mathfrak{g}, \mathfrak{g}]$. Thus, $\operatorname{tr}(x \overline{x}_s) = \sum |\lambda_i|^2 = 0$, so all eigenvalues of x are zero and x is nilpotent. By one of the versions of Engel's theorem (Theorem 6.17), this implies that $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent, so $\mathfrak{g}$ is solvable. This completes the proof of Lemma 6.39. $\square$

Now the proof of Theorem 6.36 easily follows. Indeed, if $K(\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]) = 0$, then by Lemma 6.39, $\operatorname{ad}(\mathfrak{g}) \subset \mathfrak{gl}(\mathfrak{g})$ is solvable. Thus, both $\mathfrak{z}(\mathfrak{g})$, and $\mathfrak{g}/\mathfrak{z}(\mathfrak{g}) = \operatorname{ad}(\mathfrak{g})$ are solvable. By Theorem 6.13, this implies that $\mathfrak{g}$ is solvable. $\square$

Proof of Theorem 6.37. If K is non-degenerate, then by Theorem 6.32, $\mathfrak{g}$ is reductive. On the other hand, if $x \in \mathfrak{z}(\mathfrak{g})$, then $\operatorname{ad} x = 0$, so $x \in \operatorname{Ker} K$. Thus, $\mathfrak{z}(\mathfrak{g}) = 0$, so $\mathfrak{g}$ is semisimple.

Conversely, assume that $\mathfrak{g}$ is semisimple. Consider $I = \operatorname{Ker} K$; by Lemma 6.29, I is an ideal in $\mathfrak{g}$. Since restriction of K to I coincides with the Killing form of I (Exercise 6.1), the Killing form

of I is zero and thus, by previous theorem, I is solvable. But $\mathfrak{g}$ is semisimple, so $I = 0$. Thus, K is non-degenerate. $\qquad\square$

6.7. Properties of semisimple Lie algebras

Cartan criterion of semimplicity, proved in the previous section, is not very convenient for practical computations. However, it is extremely useful for theoretical considerations.

Proposition 6.40. *Let $\mathfrak{g}$ be a real Lie algebra. Then $\mathfrak{g}$ is semisimple iff $\mathfrak{g}_{\mathbb{C}}$ is semisimple.*

Proof. Immediately follows from Cartan criterion of semisimplicity. $\qquad\square$

Remark 6.41. This theorem fails if we replace the word "semisimple" by "simple": there exist simple real Lie algebras $\mathfrak{g}$ such that $\mathfrak{g}_{\mathbb{C}}$ is a direct sum of two simple algebras.

Theorem 6.42. *Let $\mathfrak{g}$ be a semisimple Lie algebra, and $I \subset \mathfrak{g}$ — an ideal. Then there is an ideal I' such that $\mathfrak{g} = I \oplus I'$.*

Proof. Let $I^{\perp}$ be the orthogonal complement with respect to the Killing form. By Lemma 6.29, $I^{\perp}$ is an ideal. Consider the intersection $I \cap I^{\perp}$. It is an ideal in $\mathfrak{g}$ with zero Killing form (by Exercise 6.1). Thus, by Cartan criterion, it is solvable. By definition of a semisimple Lie algebra, this means that $I \cap I^{\perp} = 0$, so $\mathfrak{g} = I \oplus I^{\perp}$. $\qquad\square$

Corollary 6.43. *A Lie algebra is semisimple iff it is a direct sum of simple Lie algebras.*

Proof. Any simple Lie algebra is semisimple by Lemma 6.21, and it is immediate from Cartan criterion that direct sum of semisimple Lie algebras is semisimple. This proves one direction.

Opposite direction — that each semisimple algebra is a direct sum of simple ones — easily follows by induction from the previous theorem. $\qquad\square$

Corollary 6.44. *If $\mathfrak{g}$ is a semisimple Lie algebra, then $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.*

Indeed, for a simple Lie algebra it is clear because $[\mathfrak{g}, \mathfrak{g}]$ is an ideal in $\mathfrak{g}$ which can not be zero (otherwise, $\mathfrak{g}$ would be abelian).

Proposition 6.45. *Let $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k$ be a semisimple Lie algebra, with $\mathfrak{g}_i$ being simple. Then any ideal I in $\mathfrak{g}$ is of the form $I = \bigoplus_{i \in I} \mathfrak{g}_i$ for some subset $I \subset \{1, \ldots, k\}$.*

Note that it is not an "up to isomorphism" statement: I is not just isomorphic to sum of some of $\mathfrak{g}_i$ but actually equal to such a sum as a subspace in $\mathfrak{g}$.

Proof. The proof goes by induction in k. Let $\pi_k \colon \mathfrak{g} \to \mathfrak{g}_k$ be the projection. Consider $\pi_k(I) \subset \mathfrak{g}_k$. Since $\mathfrak{g}_k$ is simple, either $\pi_k(I) = 0$, in which case $I \subset \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_{k-1}$ and we can use induction assumption, or $\pi_k(I) = \mathfrak{g}_k$. Then $[\mathfrak{g}_k, I] = [\mathfrak{g}_k, \pi_k(I)] = \mathfrak{g}_k$. Since I is an ideal, $I \supset \mathfrak{g}_k$, so $I = I' \oplus \mathfrak{g}_k$ for some $I' \subset \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_{k-1}$ and the result again follows from induction assumption. $\qquad\square$

Corollary 6.46. *Any ideal in a semisimple Lie algebra is semisimple. Also, any quotient of a semisimple Lie algebra is semisimple.*

Finally, recall that we have denoted by $\operatorname{Der} \mathfrak{g}$ the Lie algebra of all derivations of $\mathfrak{g}$ (see (3.13))

Proposition 6.47. *If $\mathfrak{g}$ is a semisimple Lie algebra, then $\operatorname{Der} \mathfrak{g} = \mathfrak{g}$, and $\operatorname{Aut} \mathfrak{g} / \operatorname{Ad} G$ is discrete.*

Proof. Recall that for any $x \in \mathfrak{g}$, $\operatorname{ad} x \colon \mathfrak{g} \to \mathfrak{g}$ is a derivation. This gives a natural morphism of Lie algebras $\mathfrak{g} \to \operatorname{Der} \mathfrak{g}$. Since $\mathfrak{z}(\mathfrak{g}) = 0$, this morphism is injective, so $\mathfrak{g}$ is a subalgebra in $\operatorname{Der} \mathfrak{g}$.

Definition of derivation immediately shows that for any derivation δ and $x \in \mathfrak{g}$, we have $\operatorname{ad} \delta(x) = [\delta, \operatorname{ad} x]$. Thus, $\mathfrak{g} \subset \operatorname{Der} \mathfrak{g}$ is an ideal.

Let us now extend the Killing form of $\mathfrak{g}$ to $\operatorname{Der} \mathfrak{g}$ by letting $K(\delta_1, \delta_2) = \operatorname{tr}_{\mathfrak{g}}(\delta_1 \delta_2)$ and consider the orthogonal complement $I = \mathfrak{g}^\perp \subset \operatorname{Der} \mathfrak{g}$. Since K is invariant, I is an ideal; since restriction of K to $\mathfrak{g}$ is non-degenerate, $I \cap \mathfrak{g}^\perp = 0$. Thus, $\operatorname{Der} \mathfrak{g} = \mathfrak{g} \oplus I$; since both $\mathfrak{g}, I$ are ideals, we have $[I, \mathfrak{g}] = 0$, which implies that for every $\delta \in I, x \in \mathfrak{g}$, we have $\operatorname{ad} \delta(x) = [\delta, \operatorname{ad} x] = 0$, so $\delta(x) = 0$. Thus, $I = 0$.

Since $\operatorname{Aut} \mathfrak{g}$ is a Lie group with Lie algebra $\operatorname{Der} \mathfrak{g}$ (see Example 3.30), the second statement of the theorem immediately follows from the first one. $\qquad\square$

6.8. Relation with compact groups

In the previous section, we have shown that the Killing form on $\mathfrak{g}$ is non-degenerate if and only if $\mathfrak{g}$ is semisimple. However, in the case of *real* $\mathfrak{g}$, one might also ask whether the Killing form is positive definite, negative definite, or neither. More generally, the same question can be asked about the trace form in any representation: $B_V(x, y) = \operatorname{tr}_V(xy)$.

It turns out that the answer to this question is closely related with the question of compactness of the corresponding Lie group.

Example 6.48. Let $\mathfrak{g} = \mathfrak{u}(n)$. Then the form $(x, y) = \operatorname{tr}(xy)$ is negative definite.

Indeed, $\operatorname{tr}(xy) = -\operatorname{tr}(x\overline{y}^t)$, and $\operatorname{tr}(x^2) = -\operatorname{tr}(x\overline{x}^t) = -\sum |x_{ij}|^2 \leq 0$, with equality only for $x = 0$.

Theorem 6.49. *Let G be a compact real Lie group. Then $\mathfrak{g} = \operatorname{Lie}(G)$ is reductive, and the Killing form on $\mathfrak{g}$ is negative semidefinite, with $\operatorname{Ker} K = \mathfrak{z}(\mathfrak{g})$; the Killing form of the semisimple part $\mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ is negative definite.*

Conversely, let $\mathfrak{g}$ be a semisimple real Lie algebra with negative definite Killing form. Then $\mathfrak{g}$ is a Lie algebra of a compact real Lie group.

Proof. If G is compact, then by Theorem 4.38, every complex representation of G is unitary, so $\rho(G) \subset U(V)$, $\rho(\mathfrak{g}) \subset \mathfrak{u}(V)$. By Example 6.48, this implies that the trace form $B_V(x, y)$ is negative semidefinite, with $\operatorname{Ker} B_V = \operatorname{Ker} \rho$.

Applying this to the complexified adjoint representation $V = \mathfrak{g}_{\mathbb{C}}$, we see that the Killing form is negative semidefinite, with $\operatorname{Ker} K = \mathfrak{z}(\mathfrak{g})$.

Conversely, assume that $\mathfrak{g}$ is a real Lie algebra with negative definite Killing form. Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Then $B(x, y) = -K(x, y)$ is positive definite and $\operatorname{Ad} G$ invariant. This shows that $\operatorname{Ad}(G) \subset \operatorname{SO}(\mathfrak{g})$ (the orthogonal group). Since $\operatorname{Ad}(G)$ is the connected component of unity of the group $\operatorname{Aut} \mathfrak{g}$ (see Proposition 6.47), and $\operatorname{Aut} \mathfrak{g} \subset \operatorname{GL}(\mathfrak{g})$ is a Lie subgroup (see Example 3.30), $\operatorname{Ad}(G)$ is a Lie subgroup in the compact group $\operatorname{SO}(\mathfrak{g})$. Thus, $\operatorname{Ad}(G)$ is a compact Lie group. Since $\operatorname{Ad}(G) = G/Z(G)$, we have $\operatorname{Lie}(\operatorname{Ad}(G)) = \mathfrak{g}/\mathfrak{z}(\mathfrak{g}) = \mathfrak{g}$, which proves the theorem. $\qquad\square$

Remark 6.50. In fact, one can prove a stronger result: if $\mathfrak{g}$ is a real Lie algebra with negative definite Killing form, then any connected Lie group with Lie algebra $\mathfrak{g}$ is compact. In particular, the simply-connected Lie group with Lie algebra $\mathfrak{g}$ is compact.

One might also ask for which real Lie algebras the Killing form is *positive definite*. Unfortunately, it turns out that there are not many such algebras.

Lemma 6.51. *If $\mathfrak{g}$ is a real Lie algebra with a positive definite Killing form, then $\mathfrak{g} = 0$.*

The proof of this lemma is given as an exercise (Exercise 6.5).

Finally, let us discuss the relation between complex semisimple Lie groups and algebras with compact groups. We already know that there are no compact complex Lie groups. Instead, we could take real compact Lie groups and corresponding Lie algebras, then consider their complexifications. By Theorem 6.49, such complex Lie algebras will be reductive. A natural question is whether any reductive complex Lie algebra can be obtained in this way. The following theorem (which for simplicity is stated only for semisimple Lie algebras) provides the answer.

?!

Theorem 6.52. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra. Then there exists a real subalgebra $\mathfrak{k}$ such that $g = \mathfrak{k} \otimes \mathbb{C}$ and $\mathfrak{k}$ is a Lie algebra of a compact Lie group K. The Lie algebra $\mathfrak{k}$ is called the compact real form of $\mathfrak{g}$; it is unique up to conjugation.*

If G is a connected complex Lie group with Lie algebra $\mathfrak{g}$, then K can be chosen so that $K \subset G$. In this case, K is called the compact real form of the Lie group G.

The proof of this theorem will not be given here. Interested readers can find a discussion of this theorem in [**8**, Section I.7].

Example 6.53. For $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$, $G = \mathrm{SL}(n, \mathbb{C})$, the compact form is $\mathfrak{k} = \mathfrak{su}(n)$, $K = \mathrm{SU}(n)$.

6.9. Complete reducibility of representations

In this section, we will show one of fundamental results of the theory of semisimple Lie algebras: every representation of a semisimple Lie algebra is completely reducible. Throughout this section, $\mathfrak{g}$ is a semisimple complex Lie algebra and V — a finite-dimensional complex representation of $\mathfrak{g}$.

This result can be proved in several ways. Historically, the first proof of this result was given by H. Weyl using the theory of compact groups. Namely, if $\mathfrak{g}$ is a semisimple complex Lie algebra, then by Theorem 6.52 $\mathfrak{g}$ can be written as a complexification of a real Lie algebra $\mathfrak{k} = \mathrm{Lie}(K)$ for some compact connected, simply connected group K. Then complex representations of $\mathfrak{g}, \mathfrak{k}$ and K are the same, and by Theorem 4.38, every representation of K is completely reducible. This argument is commonly called "Weyl's unitary trick".

However, there is a completely algebraic proof of complete reducibility. It uses some basic homological algebra: obstruction to complete reducibility is described by a certain type of cohomology, and we will show that cohomology vanishes. To do it, we will use a special central element in the universal enveloping algebra, called the *Casimir element*.

Proposition 6.54. *Let $\mathfrak{g}$ be a Lie algebra, and B — a non-degenerate invariant symmetric bilinear form on $\mathfrak{g}$. Let x_i be a basis of $\mathfrak{g}$, and x^i — the dual basis with respect to B. Then the element*

$$C_B = \sum x_i x^i \in U\mathfrak{g}$$

does not depend on the choice of basis x_i and is central. It is called the Casimir element determined by form B.

Proof. Independence of choice of basis follows from the fact that the element $\sum x_i \otimes x^i \in \mathfrak{g} \otimes \mathfrak{g}$ is independent of the choice of basis: under the identification $\mathfrak{g} \otimes \mathfrak{g} \simeq \mathfrak{g} \otimes \mathfrak{g}^* = \mathrm{End}(\mathfrak{g})$ given by the form B, this element becomes the identity operator in $\mathrm{End}(\mathfrak{g})$.

To show that it is central, choose x_i to be an orthonormal basis with respect to B (for real $\mathfrak{g}$, this may require replacing $\mathfrak{g}$ by $\mathfrak{g}_{\mathbb{C}}$). Then for any $y \in \mathfrak{g}$, the matrix of $\operatorname{ad} y$ in the basis x_i is skew-symmetric (since $\operatorname{ad} y \in \mathfrak{so}(\mathfrak{g})$) and thus we have

$$[y, C_B] = \operatorname{ad} y. C_B = \sum (\operatorname{ad} y.x_i) x_i + x_i (\operatorname{ad} y.x_i) = \sum_{i,j} a_{ij}(x_i x_j + x_j x_i) = 0$$

where a_{ij} is the matrix of $\operatorname{ad} y$ in basis x_i. $\qquad\square$

Note that if $\mathfrak{g}$ is simple, then by Exercise 4.4, the invariant bilinear form is unique up to a constant. Thus, in this case the Casimir element is also unique up to a constant; in this case, it is usually referred to as the Casimir element, without specifying the form.

Proposition 6.55. *Let V be a non-trivial irreducible representation of $\mathfrak{g}$. Then there exists a central element $C_V \in Z(U\mathfrak{g})$ which acts by a non-zero constant in V and which acts by zero in the trivial representation.*

Proof. Let $B_V(x, y) = \operatorname{tr}(\rho(x)\rho(y))$; by Proposition 6.31, this form is an invariant bilinear form. If B_V is non-degenerate, then let $C_V = C_{B_V}$ be the Casimir element of $\mathfrak{g}$ defined by form B_V. Obviously, C_V acts by zero in $\mathbb{C}$. Since V is non-degenerate, by Schur lemma C_V acts in V by a constant: $C_V = \lambda \operatorname{id}_V$. On the other hand, $\operatorname{tr}(C_V) = \sum \operatorname{tr}(x_i x^i) = \dim \mathfrak{g}$, because by definition of B, $\operatorname{tr}(x_i x^i) = B(x_i, x^i) = 1$. Thus, $\lambda = \frac{\dim \mathfrak{g}}{\dim V} \neq 0$, which proves the proposition in this special case.

In general, let $I = \operatorname{Ker} B_V \subset \mathfrak{g}$. Then it is an ideal in $\mathfrak{g}$, and $I \neq \mathfrak{g}$ (otherwise, by Lemma 6.39, $\rho(\mathfrak{g}) \subset \mathfrak{gl}(V)$ is solvable, which is impossible as it is a quotient of a semisimple Lie algebra and thus itself semisimple). By results of Theorem 6.42, $\mathfrak{g} = I \oplus \mathfrak{g}'$ for some non-zero ideal $\mathfrak{g}' \subset \mathfrak{g}$. By Proposition 6.45, $\mathfrak{g}'$ is semisimple, and restriction of B_V to $\mathfrak{g}'$ is non-degenerate. Let C_V be the Casimir element of $\mathfrak{g}'$ corresponding to the form B_V. Since $I, \mathfrak{g}'$ commute, C_V will be central in $U\mathfrak{g}$, and the same argument as before shows that it acts in V by $\frac{\dim \mathfrak{g}'}{\dim V} \neq 0$, which completes the proof. $\qquad\square$

Remark 6.56. In fact, a stronger result is known: if we let C be the Casimir element defined by the Killing form, then C acts by a non-zero constant in any nontrivial irreducible representation. However, this is slightly more difficult to prove.

Now we are ready to prove the main result of this section.

Theorem 6.57. *Any representation of a semisimple Lie algebra $\mathfrak{g}$ is completely reducible.*

Proof. The proof assumes some familiarity with basic homological algebra, such as the notion of functors $\operatorname{Ext}^i(V, W)$. They are defined for representations of Lie algebras in the same way as for modules over an associative algebra. In fact, the same definition works for any abelian category, i.e. a category where morphisms form abelian groups and where we have the notion of image and kernel of a morphism satisfying the usual properties.

In particular, the standard argument of homological agebra shows that for fixed V_1, V_2 equivalence classes of extensions $0 \to V_1 \to W \to V_2 \to 0$ are in bijection with $\operatorname{Ext}^1(V_1, V_2)$. Thus, our goal is to show that $\operatorname{Ext}^1(V_1, V_2) = 0$ for any two representations V_1, V_2. This will be done in several steps. For convenience, we introduce the notation $H^1(\mathfrak{g}, V) = \operatorname{Ext}^1(\mathbb{C}, V)$.

Lemma 6.58. *For any irreducible representation V, one has $H^1(\mathfrak{g}, V) = 0$.*

Proof. To prove that $\operatorname{Ext}^1(\mathbb{C}, V) = 0$ it suffices to show that every short exact sequence of the form $0 \to \mathbb{C} \to W \to V \to 0$ splits. So let us assume that we have such an exact sequence.

Let us consider separately two cases: V is a non-trivial irreducible representaton and $V = \mathbb{C}$.

If V is a non-trivial irreducible representation, consider the Casimir element C_V as defined in Proposition 6.55. Since it acts in $\mathbb{C}$ by zero and in V by a non-zero constant λ, its eigenvalues in W are 0 with multiplicity 1 and λ with multiplicity $\dim V$. Thus, we can split W in a direct sum of generalized eigenspaces: $W = \mathbb{C} \oplus W_\lambda$, where W_λ is the generalized eigenspace for C_V with egenvalue λ. Since C_V commutes with the action of $\mathfrak{g}$, this decomposition is preserved by the action of $\mathfrak{g}$, i.e. W_λ is a subrepresentation. Since kernel of the projection $W \to V$ is $\mathbb{C}$, this projection gives an isomorphism $W_\lambda \simeq V$; thus, W splits as direct sum $\mathbb{C} \oplus V$.

If $V = \mathbb{C}$ is a trivial representation, so we have an exact sequence $0 \to \mathbb{C} \to W \to \mathbb{C} \to 0$, then W is a two-dimensional representation such that the action of $\rho(x)$ is strictly upper triangular for all $x \in \mathfrak{g}$. Thus, $\rho(\mathfrak{g})$ is nilpotent, so by Corollary 6.46, $\rho(\mathfrak{g}) = 0$. Thus, $W \simeq \mathbb{C} \oplus \mathbb{C}$ as a representation. $\qquad\square$

This lemma provides the crucial step; the rest is simple homological algebra.

Lemma 6.59. $H^1(\mathfrak{g}, V) = 0$ *for any representation* V.

Proof. If we have a short exact sequence of representations $0 \to V_1 \to V \to V_2 \to 0$, then we have a long exact sequence of Ext groups; in particular,

$$\cdots \to H^1(\mathfrak{g}, V_1) \to H^1(\mathfrak{g}, V) \to H^1(\mathfrak{g}, V_2) \to \ldots$$

Thus, if $H^1(\mathfrak{g}, V_1) = H^1(\mathfrak{g}, V_2) = 0$, then $H^1(\mathfrak{g}, V) = 0$. Since for irreducible representations we have proved that $H^1(\mathfrak{g}, V) = 0$, it is easy to show by induction in dimension that for any representation, $H^1(\mathfrak{g}, V) = 0$. $\qquad\square$

We are now ready to prove Theorem 6.57. Let $0 \to V_1 \to W \to V_2 \to 0$ be a short exact sequence of $\mathfrak{g}$-modules. We need to show that it splits.

Let us apply to this sequence the functor $X \mapsto \operatorname{Hom}_{\mathbb{C}}(V_2, X) = V_2^* \otimes X$ (considered as a $\mathfrak{g}$-module, see Example 4.11). Obviously, this gives a short exact sequence of $\mathfrak{g}$-modules

$$0 \to \operatorname{Hom}_{\mathbb{C}}(V_2, V_1) \to \operatorname{Hom}_{\mathbb{C}}(V_2, W) \to \operatorname{Hom}_{\mathbb{C}}(V_2, V_2) \to 0$$

Now, let us apply to this sequence the functor of $\mathfrak{g}$-invariants: $X \mapsto X^{\mathfrak{g}} = \operatorname{Hom}_{\mathfrak{g}}(\mathbb{C}, X)$. Applying this functor to $\operatorname{Hom}_{\mathbb{C}}(A, B)$ gives $(\operatorname{Hom}_{\mathbb{C}}(A, B))^{\mathfrak{g}} = \operatorname{Hom}_{\mathfrak{g}}(A, B)$ (see Example 4.13). This functor is left exact but in general not exact, so we get a long exact sequence

$$0 \to \operatorname{Hom}_{\mathfrak{g}}(V_2, V_1) \to \operatorname{Hom}_{\mathfrak{g}}(V_2, W) \to \operatorname{Hom}_{\mathfrak{g}}(V_2, V_2) \to \operatorname{Ext}^1(\mathbb{C}, V_2^* \otimes V_1) = H^1(\mathfrak{g}, V_2^* \otimes V_1) \to \ldots$$

But since we have already proved that $H^1(\mathfrak{g}, V) = 0$ for any module V, we see that in fact we do have a short exact sequence

$$0 \to \operatorname{Hom}_{\mathfrak{g}}(V_2, V_1) \to \operatorname{Hom}_{\mathfrak{g}}(V_2, W) \to \operatorname{Hom}_{\mathfrak{g}}(V_2, V_2) \to 0$$

In particular, this shows that there exists a morphism $f \colon V_2 \to W$ which, when composed with projection $W \to V_2$, gives identity morphism $V_2 \to V_2$. This gives a splitting of exact sequence $0 \to V_1 \to W \to V_2 \to 0$. This completes the proof of Theorem 6.57. $\qquad\square$

Remark 6.60. The same proof can be rewritten without using the language of Ext groups; see, for example, [**20**]. This would make it formally accessible to readers with no knowledge of homological algebra. However, this does not change the fact that all arguments are essentially homological in nature; in fact, such a rewriting would obscure the ideas of the proof rather than make them clearer.

In fact, the groups $\operatorname{Ext}^1(V, W)$ and in particular, $H^1(\mathfrak{g}, V) = \operatorname{Ext}^1(\mathbb{C}, V)$ used in this proof are just the beginning of a well-developed cohomology theory of Lie algebras. In particular, one can define higher cohomology groups $H^i(\mathfrak{g}, V)$ in a very similar way. The same argument with

the Casimir element can be used to show that for non-trivial irreducible representation V, one has $H^i(\mathfrak{g}, V) = 0$ for $i > 0$. However, it is not true for the trivial representation: $H^i(\mathfrak{g}, \mathbb{C})$ can be non-zero. For example, it can be shown that if G is a connected, simply connected compact real Lie group, and $\mathfrak{g} = \mathrm{Lie}(G)$, then $H^i(\mathfrak{g}, \mathbb{R}) = H^i(G, \mathbb{R})$, where $H^i(G, \mathbb{R})$ are the usual topological cohomology (which can be defined, for example, as De Rham cohomology). This and much more can be found, for example, in [**7**].

Complete reducibility has a number of useful corollaries. One of them is the following result, announced in Section 6.4.

Theorem 6.61. *Any reductive Lie algebra can be written as a direct sum (as Lie algebra) of semisimple and commutative:*

$$\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{g}_{ss}, \qquad \mathfrak{z} \text{ commutative}, \quad \mathfrak{g}_{ss} \text{ semisimple}.$$

Proof. Consider adjoint representation of $\mathfrak{g}$. Since $\mathfrak{z}(\mathfrak{g})$ acts by zero in adjoint representation, the adjoint action descends to an action of $\mathfrak{g}' = \mathfrak{g}/\mathfrak{z}(\mathfrak{g})$. By definition of reductive algebra, $\mathfrak{g}'$ is semisimple. Thus, $\mathfrak{g}$ considered as a representation of $\mathfrak{g}'$ is completely reducible. Since $\mathfrak{z} \subset \mathfrak{g}$ is stable under adjoint action, it is a subrepresentation. By complete reducibility, we can write $\mathfrak{g} = \mathfrak{z} \oplus I$ for some $I \subset \mathfrak{g}$ such that $\mathrm{ad}\, x. I \subset I$ for any $x \in I$. Thus, I is an ideal in $\mathfrak{g}$, so $\mathfrak{g} = \mathfrak{z} \oplus I$ as Lie algebras. Obviously, $I \simeq \mathfrak{g}/\mathfrak{z} = \mathfrak{g}'$ is semisimple. $\qquad\square$

In a similar way one can prove Levi theorem (Theorem 6.25). We do not give this proof here, referring the reader to [**20**],[**18**], [**11**]. Instead, we just mention that in the language of homological algebra, Levi theorem follows from vanishing of cohomology $H^2(\mathfrak{g}, \mathbb{C})$.

Exercises

6.1.
  (1) Let V be a representation of $\mathfrak{g}$ and $W \subset V$ be a subrepresentation. Then $B_V = B_W + B_{V/W}$, where B_V is defined by (6.4).
  (2) Let $I \subset \mathfrak{g}$ be an ideal. Then the restriction of Killing form of $\mathfrak{g}$ to I coincides with the Killing form of I.

6.2. Show that for $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$, the Killing form is given by $K(x, y) = 2n \,\mathrm{tr}(xy)$.

6.3. Let $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{C})$ be the subspace consisting of block-triangular matrices:

$$\mathfrak{g} = \left\{ \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \right\}$$

where A is a $k \times k$ matrix, B is a $k \times (n - k)$ matrix, and D is a $(n - k) \times (n - k)$ matrix.
  (1) Show that $\mathfrak{g}$ is a Lie subalgebra (this is a special case of so-called *parabolic subalgebras*).
  (2) Show that radical of $\mathfrak{g}$ consists of matrices of the form $\begin{pmatrix} \lambda \cdot I & B \\ 0 & \mu \cdot I \end{pmatrix}$, and describe $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$.

6.4. Show that the bilinear form $\mathrm{tr}(xy)$ on $\mathfrak{sp}(2n, \mathbb{K})$ is non-degenerate.

6.5. Let $\mathfrak{g}$ be a real Lie algebra with a positive definite Killing form. Show that then $\mathfrak{g} = 0$.

Chapter 7

Complex Semisimple Lie Algebras

In this chapter, we begin the study of semisimple Lie algebras and their representations. This is one of the most beautiful areas in all of mathematics.

?!

Throughout this chapter, $\mathfrak{g}$ is a complex semisimple Lie algebra.

7.1. Semisimple elements and toroidal subalgebras

Recall that the main tool used in the study of representations of $\mathfrak{sl}(2, \mathbb{C})$ in Section 5.1 was the weight decomposition, i.e. decomposing a representation of $\mathfrak{sl}(2, \mathbb{C})$ as a direct sum of eigenspaces for h. In order to generalize this idea to other Lie algebras, we need to find a proper analog of h.

Looking closely at the proofs of Section 5.1, we see that the key property of h were the commutation relations $[h, e] = 2e$, $[h, f] = -2f$ which were used to to show that e, f shift the weight. This justifies the following definition.

Definition 7.1. An element $x \in \mathfrak{g}$ is called *semisimple* if $\operatorname{ad} x$ is a semisimple operator $\mathfrak{g} \to \mathfrak{g}$.

An element $x \in \mathfrak{g}$ is called *nilpotent* if $\operatorname{ad} x$ is a nilpotent operator $\mathfrak{g} \to \mathfrak{g}$.

It is easy to show (see Exercise 7.1) that for $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$ this definition coincides with the usual definition of a semisimple operator.

Of course, we do not yet know if such elements exist for any $\mathfrak{g}$. The following theorem, which generalizes Jordan decomposition theorem (Theorem B.2), answers this question.

Theorem 7.2. *If $\mathfrak{g}$ is a semisimple complex Lie algebra, then any $x \in \mathfrak{g}$ can be uniquely written in the form*

$$x = x_s + x_n$$

where x_s is semisimple, x_n is nilpotent, and $[x_s, x_n] = 0$. Moreover, $\operatorname{ad} x_s = P(\operatorname{ad} x)$ for some polynomial $P \in t\mathbb{C}[t]$ depending on x.

Proof. Uniqueness immediately follows from uniqueness of Jacobi decomposition for $\operatorname{ad} x$ (Theorem B.2): if $x = x_s + x_n = x'_s + x'_n$, then $(\operatorname{ad} x)_s = \operatorname{ad} x_s = \operatorname{ad} x'_s$, so $\operatorname{ad}(x_s - x'_s) = 0$. But by definition, a semisimple Lie algebra has zero center, so this implies $x_s - x'_s = 0$.

To prove existence, let us write $\mathfrak{g}$ as direct sum of generalized eigenspaces for $\operatorname{ad} x$: $\mathfrak{g} = \bigoplus \mathfrak{g}_\lambda$, $(\operatorname{ad} x - \lambda \operatorname{id})^n|_{\mathfrak{g}_\lambda} = 0$ for $n \gg 0$.

Lemma 7.3. $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subset \mathfrak{g}_{\lambda+\mu}$.

Proof. By Jacobi identity, $(\operatorname{ad} x - \lambda - \mu)[y, z] = [(\operatorname{ad} x - \lambda)y, z] + [y, (\operatorname{ad} x - \mu)z]$. Thus, if $y \in \mathfrak{g}_\lambda, z \in \mathfrak{g}_\mu$, then

$$(\operatorname{ad} x - \lambda - \mu)^n [y, z] = \sum \binom{n}{k} [(\operatorname{ad} x - \lambda)^k y, (\operatorname{ad} x - \mu)^{n-k} z] = 0$$

for $n \gg 0$. $\qquad \square$

Let $\operatorname{ad} x = (\operatorname{ad} x)_s + (\operatorname{ad} x)_n$ be the Jacobi decomposition of operator $\operatorname{ad} x$ (see Theorem B.2), so that $(\operatorname{ad} x)_s|_{\mathfrak{g}_\lambda} = \lambda$. Then the lemma implies that $(\operatorname{ad} x)_s$ is a derivation of $\mathfrak{g}$. By Proposition 6.47, any derivation is inner, so $(\operatorname{ad} x)_s = \operatorname{ad} x_s$ for some $x_s \in \mathfrak{g}$; thus, $(\operatorname{ad} x)_n = \operatorname{ad}(x - x_s)$. This proves existence of Jordan decomposition for x.

Finally, by Theorem B.2, we have $(\operatorname{ad} x)_s = P(\operatorname{ad} x)$ for some polynomial P. Applying both sides to x itself, we see that this polynomial has zero constant term. $\qquad \square$

Corollary 7.4. *In any semisimple complex Lie algebra, there exist non-zero semisimple elements.*

Proof. If any semisimple element is zero, then, by Theorem 7.2, any $x \in \mathfrak{g}$ is nilpotent. By Engel theorem, this implies that $\mathfrak{g}$ is nilpotent, which contradicts semisimplicity of $\mathfrak{g}$. $\qquad \square$

Theorem 7.5. *Let $\mathfrak{g}$ be a semisimple Lie algebra, and $x \in \mathfrak{g}$ — a semisimple element. Let V be a finite-dimensional complex representation of $\mathfrak{g}$. Then $\rho(x) \in \mathfrak{gl}(V)$ is semisimple.*

Proof. The proof essentially repeats the proof in $\mathfrak{sl}(2, \mathbb{C})$ case (see Theorem 5.3) with necessary changes.

By complete reducibility theorem, it suffices to consider the case when V is irreducible. Let $V[\lambda] = \{v \in V \mid \rho(x).v = \lambda v\}$, and let $V' = \bigoplus_\lambda V[\lambda]$. Our goal is to show that $V = V'$.

Since x is semisimple, we can write $\mathfrak{g} = \bigoplus_{\alpha \in \mathbb{C}} \mathfrak{g}_\alpha$, where $\mathfrak{g}_\alpha = \{y \in \mathfrak{g} \mid [x, y] = \alpha y\}$. Easy explicit calculation shows that if $y \in \mathfrak{g}_\alpha$, $v \in V[\lambda]$, then $\rho(y).v \in V[\lambda + \alpha]$. Thus, V' is a subrepresentation of V. Since V is irreducible, and $V' \neq 0$, we have $V' = V$. $\qquad \square$

Our next step would be considering not just one semisimple element but a family of commuting semisimple elements.

Definition 7.6. A subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is called *toroidal* if it is commutative and consists of semisimple elements.

Theorem 7.7. *Let $\mathfrak{h} \subset \mathfrak{g}$ be a toroidal subalgebra. Then*

(1) $\mathfrak{g} = \bigoplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}_\alpha$, *where $\mathfrak{g}_\alpha$ is a common eigenspace for all operators $\operatorname{ad} h, h \in \mathfrak{h}$, with eigenvalue α:*

$$\operatorname{ad} h.x = \langle \alpha, h \rangle x, \quad h \in \mathfrak{h}, x \in \mathfrak{g}_\alpha.$$

In particular, $\mathfrak{h} \subset \mathfrak{g}_0$.

(2) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.

(3) *If $\alpha + \beta \neq 0$, then $\mathfrak{g}_\alpha, \mathfrak{g}_\beta$ are orthogonal with respect to the Killing form K.*

(4) *For any α, the Killing form gives a non-degenerate pairing $\mathfrak{g}_\alpha \otimes \mathfrak{g}_{-\alpha} \to \mathbb{C}$. In particular, restriction of K to $\mathfrak{g}_0$ is non-degenerate.*

Proof. By definition, for each $h \in \mathfrak{h}$, the operator $\operatorname{ad} h$ is diagonalizable. Since all operators $\operatorname{ad} h$ commute, they can be simultaneously diagonalized, which is exactly the statement of the first part of the theorem.

The second part is in fact a very special case of Lemma 7.3. However, in this case it can be proved much easier: if $y \in \mathfrak{g}_\alpha$, $z \in \mathfrak{g}_\beta$, then

$$\operatorname{ad} h.[y, z] = [\operatorname{ad} h.y, z] + [y, \operatorname{ad} h.z] = \langle \alpha, h \rangle [y, z] + \langle \beta, h \rangle [y, z] = \langle \alpha + \beta, h \rangle [y, z].$$

For the third part, notice that if $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_\beta$, then $\operatorname{ad} x \operatorname{ad} y(\mathfrak{g}_\gamma) \subset \mathfrak{g}_{\gamma + \alpha + \beta}$. Thus, if $\alpha + \beta \neq 0$, then all diagonal entries of $\operatorname{ad} x \operatorname{ad} y$ are zero.

The final part immediately follows from the previous part and non-degeneracy of Killing form (Theorem 6.37). $\qquad \square$

Of course, since $\mathfrak{g}$ is finite-dimensional, $\mathfrak{g}_\alpha = 0$ for all but finitely many $\alpha \subset \mathfrak{h}^*$.

7.2. Cartan subalgebra

Our next goal is to produce as large a toroidal subalgebra in $\mathfrak{g}$ as possible. The standard way of formalizing this is as follows.

Definition 7.8. Let $\mathfrak{g}$ be a semisimple Lie algebra. A Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is a toroidal subalgebra which coincides with its centralizer: $C(\mathfrak{h}) = \{x \mid [x, \mathfrak{h}] = 0\} = \mathfrak{h}$.

Remark 7.9. This definition should only be used for semisimple Lie algebras: for general Lie algebras, Cartan subalgebras are defined in a slightly different way (see, e.g., [**21**]). However, it can be shown that for semisimple algebras our definition is equivalent to the usual one.

Example 7.10. Let $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$ and $\mathfrak{h} = \{$diagonal matrices with trace $0\}$. Then $\mathfrak{h}$ is a Cartan subalgebra. Indeed, it is obviously commutative, and every diagonal element is semisimple (see Exercise 7.1), so it is a toroidal subalgebra. On the other hand, choose $h \in \mathfrak{h}$ to be a diagonal matrix with distinct eigenvalues. By a well-known result of linear algebra, if $[x, h] = 0$, and h has distinct eigenvalues, then any eigenvector of h is also an eigenvector of x; thus, x must also be diagonal. Thus, $C(h) = \mathfrak{h}$.

We still need to prove existence of Cartan subalgebras.

Theorem 7.11. *Let $\mathfrak{h} \subset \mathfrak{g}$ be a maximal toroidal subalgebra, i.e. a toroidal subalgebra which is not properly contained in any other toroidal subalgebra. Then $\mathfrak{h}$ is a Cartan subalgebra.*

Proof. Let $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$ be the decomposition of $\mathfrak{g}$ into eigenspaces for $\operatorname{ad} h$ as in Theorem 7.7. Then $C(\mathfrak{h}) = \mathfrak{g}_0$. It is a subalgebra in $\mathfrak{g}$, and by Theorem 7.7, the restriction of the Killing form of $\mathfrak{g}$ to $C(\mathfrak{h})$ is non-degenerate.

Step 1. Let $x \in C(\mathfrak{h})$ and let $x = x_s + x_n$ be the Jordan decomposition of x (see Theorem 7.2). Then $x_s, x_n \in C(\mathfrak{h})$.

Indeed, if $\operatorname{ad} x.h = 0$ for any $h \in \mathfrak{h}$, then by Theorem 7.2, $\operatorname{ad} x_s = P(\operatorname{ad} x)$, so $\operatorname{ad} x_s.h = 0$. Thus, $x_s \in C(\mathfrak{h})$.

Step 2. $C(\mathfrak{h})$ is reductive.

Consider $\mathfrak{g}$ as a representation of $C(\mathfrak{h})$. Then the corresponding trace form $(h_1, h_2) = \operatorname{tr}_\mathfrak{g}(\operatorname{ad} h_1 \operatorname{ad} h_2)$ on $C(\mathfrak{h})$ is exactly the restriction of Killing form $K^\mathfrak{g}$ to $C(\mathfrak{h})$ and by Theorem 7.7 is non-degenerate. But by one of the forms of Cartan criterion (Theorem 6.32), this implies that $C(\mathfrak{h})$ is reductive.

Step 3. $C(\mathfrak{h})$ is commutative.

Indeed, otherwise $C(\mathfrak{h}) = \mathfrak{z} \oplus \mathfrak{g}'$ for some semisimple $\mathfrak{g}'$. Let x be a non-zero semisimple element in $\mathfrak{g}'$ (which exists by Corollary 7.4). Applying Theorem 7.5 to $\mathfrak{g}$ considered as a representation of $\mathfrak{g}'$, we see that x is also semisimple as an element of $\mathfrak{g}$.

Then $[x, \mathfrak{h}] = 0$, so $\mathfrak{h} \oplus \mathbb{C} \cdot x$ would be a toroidal subalgebra, which contradicts maximality of $\mathfrak{h}$.

Step 4. $C(\mathfrak{h})$ contains no non-zero nilpotent elements.

Indeed, if $x \in C(\mathfrak{h})$ is nilpotent, then $\operatorname{ad} x \colon \mathfrak{g} \to \mathfrak{g}$ is nilpotent. Since $C(\mathfrak{h})$ is commutative, for any $y \in C(\mathfrak{h})$, $\operatorname{ad} x \operatorname{ad} y$ is also nilpotent, so $\operatorname{tr}_{\mathfrak{g}}(\operatorname{ad} x \operatorname{ad} y) = 0$. This contradicts to non-degeneracy of Killing form on $C(\mathfrak{h})$.

Now combining the results of step 1 with step 4, we see that every element $x \in C(\mathfrak{h})$ is semisimple. Thus, $C(\mathfrak{h})$ is a toroidal subalgebra which contains $\mathfrak{h}$. Since $\mathfrak{h}$ was chosen to be maximal, $C(\mathfrak{h}) = \mathfrak{h}$. $\qquad\square$

Corollary 7.12. *In every complex semisimple Lie algebra $\mathfrak{g}$, there exists a Cartan subalgebra.*

Finally, we quote without proof the following important result, proof of which can be found in [**21**], [**11**].

Theorem 7.13. *Any two Cartan subalgebras in a semisimple Lie algebra are conjugate: if $\mathfrak{h}_1, \mathfrak{h}_2 \subset \mathfrak{g}$ are Cartan subalgebras, then there exists an element g in the Lie group G corresponding to $\mathfrak{g}$ such that $\mathfrak{h}_2 = \operatorname{Ad} g(\mathfrak{h}_1)$.*

Corollary 7.14. *Any two Cartan subalgebras in $\mathfrak{g}$ have the same dimension. This dimension is called* rank *of $\mathfrak{g}$:*

$$\operatorname{rank}(\mathfrak{g}) = \dim \mathfrak{h}.$$

Example 7.15. Rank of $\mathfrak{sl}(n, \mathbb{C})$ is equal to $n - 1$.

7.3. Root decomposition and root systems

From now on, we fix a semisimple Lie algebra $\mathfrak{g}$ and a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$.

Theorem 7.16.

(1) *We have the following decomposition for $\mathfrak{g}$, called the* root decomposition

$$(7.1) \qquad\qquad \mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha,$$

where

$$(7.2) \qquad \begin{aligned} \mathfrak{g}_\alpha &= \{x \mid [h, x] = \langle \alpha, h \rangle x \text{ for all } h \in \mathfrak{h}\} \\ R &= \{\alpha \in \mathfrak{h}^* - \{0\} \mid \mathfrak{g}_\alpha \neq 0\} \end{aligned}$$

The set R is called the root system *of $\mathfrak{g}$, and subspaces $\mathfrak{g}_\alpha$ are called the* root subspaces.

(2) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha + \beta}$ *(here and below, we let $\mathfrak{g}_0 = \mathfrak{h}$).*

(3) *If $\alpha + \beta \neq 0$, then $\mathfrak{g}_\alpha, \mathfrak{g}_\beta$ are orthogonal with respect to the Killing form K.*

(4) *For any α, the Killing form gives a non-degenerate pairing $\mathfrak{g}_\alpha \otimes \mathfrak{g}_{-\alpha} \to \mathbb{C}$. In particular, restriction of K to $\mathfrak{h}$ is non-degenerate.*

Proof. This immediately follows from Theorem 7.7 and $\mathfrak{g}_0 = \mathfrak{h}$, which is the definition of Cartan subalgebra. $\qquad\square$

This decomposition is the most important result one should know about semisimple Lie algebras — much more important than the definition of semisimple algebras (in fact, this could be taken as the definition, see Exercise 7.2). Our goal is to use this decomposition to get as much information as possible about the structure of semisimple Lie algebras, eventually getting full classification theorem for them.

From now on, we will denote the Killing form on $\mathfrak{g}$ by $(\,,\,)$. Since its restriction to $\mathfrak{h}$ is non-degenerate, it defines an isomorphism $\mathfrak{h} \xrightarrow{\sim} \mathfrak{h}^*$ and a non-degenerate bilinear form on $\mathfrak{h}^*$, which we will also denote by $(\,,\,)$. It can be explicitly defined as follows: if we denote for $\alpha \in \mathfrak{h}^*$ by H_α the corresponding element of $\mathfrak{h}$, then

$$(7.3) \qquad\qquad (\alpha, \beta) = (H_\alpha, H_\beta) = \langle H_\alpha, \beta \rangle$$

for any $\alpha, \beta \in \mathfrak{h}^*$.

Example 7.17. Let $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{h} = $ diagonal matrices with trace 0 (see Example 7.10). Denote by $e_i \colon \mathfrak{h} \to \mathbb{C}$ the functional which computes i^{th} diagonal entry of h:

$$e_i \colon \begin{bmatrix} h_1 & 0 & \dots \\ 0 & h_2 & \dots \\ & \dots & \\ 0 & \dots & h_n \end{bmatrix} \mapsto h_i$$

Then one easily sees that $\sum e_i = 0$, so

$$\mathfrak{h}^* = \bigoplus \mathbb{C}e_i / \mathbb{C}(e_1 + \dots + e_n).$$

It is easy to see that matrix units E_{ij} are eigenvectors for $\operatorname{ad} h$, $h \in \mathfrak{h}$: $[h, E_{ij}] = (h_i - h_j)E_{ij} = (e_i - e_j)(h)E_{ij}$. Thus, the root decomposition is given by

$$R = \{e_i - e_j \mid i \neq j\} \subset \bigoplus \mathbb{C}e_i / \mathbb{C}(e_1 + \dots + e_n)$$
$$\mathfrak{g}_{e_i - e_j} = \mathbb{C}E_{ij}.$$

The Killing form on $\mathfrak{h}$ is given by

$$(h, h') = \sum_{i \neq j}(h_i - h_j)(h'_i - h'_j) = 2n \sum_i h_i h'_i = 2n \operatorname{tr}(hh').$$

From this, it is easy to show that if $\lambda = \sum \lambda_i e_i$, $\mu = \sum \mu_i e_i \in \mathfrak{h}^*$, and λ_i, μ_i are chosen so that $\sum \lambda_i = \sum \mu_i = 0$ (which is always possible), then the corresponding form on $\mathfrak{h}^*$ is given by

$$(\alpha, \mu) = \frac{1}{2n} \sum_i \lambda_i \mu_i.$$

Lemma 7.18. *Let $e \in \mathfrak{g}_\alpha, f \in \mathfrak{g}_{-\alpha}$. Then*

$$[e, f] = (e, f)H_\alpha.$$

Proof. Let us compute the inner product $([e, f], h)$ for some $h \in \mathfrak{h}$. Since Killing form is invariant, we have

$$([e, f], h) = (e, [f, h]) = -(e, [h, f]) = \langle h, \alpha \rangle (e, f) = (e, f)(h, H_\alpha)$$

Since $(\,,\,)$ is a non-degenerate form on $\mathfrak{h}$, this implies that $[e, f] = (e, f)H_\alpha$. $\qquad\square$

Lemma 7.19.   (1) *Let $\alpha \in R$. Then $(\alpha, \alpha) = (H_\alpha, H_\alpha) \neq 0$.*

    (2) *Let $e \in \mathfrak{g}_\alpha$, $f \in \mathfrak{g}_{-\alpha}$ be such that $(e, f) = \frac{2}{(\alpha, \alpha)}$, and let*

$$(7.4) \qquad\qquad h_\alpha = \frac{2H_\alpha}{(\alpha, \alpha)}.$$

Then $\langle h_\alpha, \alpha \rangle = 2$ and the elements e, f, h_α satisfy the relations of Lie algebra $\mathfrak{sl}(2, \mathbb{C})$. We will denote such a subalgebra by $\mathfrak{sl}(2, \mathbb{C})_\alpha \subset \mathfrak{g}$.

Proof. Assume that $(\alpha, \alpha) = 0$; then $\langle H_\alpha, \alpha \rangle = 0$. Choose $e \in \mathfrak{g}_\alpha, f \in \mathfrak{g}_{-\alpha}$ such that $(e, f) \neq 0$ (possible by Theorem 7.16). Let $h = [e, f] = (e, f)H_\alpha$ and consider the algebra $\mathfrak{a}$ generated by e, f, h. Then we see that $[h, e] = \langle h, \alpha \rangle e = 0$, $[h, f] = -\langle h, \alpha \rangle f = 0$, so $\mathfrak{a}$ is solvable Lie algebra. By Lie theorem (Theorem 6.14), we can choose a basis in $\mathfrak{g}$ such that operators $\operatorname{ad} e, \operatorname{ad} f, \operatorname{ad} h$ are upper tringular. Since $h = [e, f]$, $\operatorname{ad} h$ will be strictly upper-triangular and thus nilpotent. But since $h \in \mathfrak{h}$, it is also semisimple. Thus, $h = 0$. On the other hand, $h = (e, f)H_\alpha \neq 0$. This contradiction proves the first part of the theorem.

The second part is immediate from definitions and Lemma 7.18. $\square$

This lemma gives us a very powerful tool for study of $\mathfrak{g}$: we can consider $\mathfrak{g}$ as a module over the subalgebra $\mathfrak{sl}(2, \mathbb{C})_\alpha$ and then use results about representations of $\mathfrak{sl}(2, \mathbb{C})$ proved in Section 5.1.

Lemma 7.20. *Let α be a root, and let $\mathfrak{sl}(2, \mathbb{C})_\alpha$ be the Lie subalgebra generated by $e \in \mathfrak{g}_\alpha, f \in \mathfrak{g}_{-\alpha}$ and h_α as in Lemma 7.19.*

Consider the subspace

$$V = \mathbb{C}h_\alpha \oplus \bigoplus_{k \in \mathbb{Z}, k \neq 0} \mathfrak{g}_{k\alpha} \subset \mathfrak{g}.$$

Then V is an irreducible representation of $\mathfrak{sl}(2, \mathbb{C})_\alpha$.

Proof. Since $\operatorname{ad} e.\mathfrak{g}_{k\alpha} \subset \mathfrak{g}_{(k+1)\alpha}$, and by Lemma 7.18, $\operatorname{ad} e.\mathfrak{g}_{-\alpha} \subset \mathbb{C}h_\alpha$, and similarly for f, V is a representation of $\mathfrak{sl}(2, \mathbb{C})_\alpha$. Since $\langle h_\alpha, \alpha \rangle = 2$, we see that the weight decomposition of V is given by $V[k] = 0$ for odd k and $V[2k] = \mathfrak{g}_{k\alpha}$, $V[0] = \mathbb{C}h_\alpha$. In particular, zero weight space $V[0]$ is one-dimensional. By Exercise 5.1, this implies that V is irreducible. $\square$

Now we can prove the main theorem about the structure of semisimple Lie algebras.

Theorem 7.21. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra with Cartan subalgebra $\mathfrak{h}$ and root decomposition $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha$.*

(1) *R spans $\mathfrak{h}^*$ as a vector space, and elements $h_\alpha, \alpha \in R$, span $\mathfrak{h}$ as a vector space*

(2) *For each $\alpha \in R$, the root subspace $\mathfrak{g}_\alpha$ is one-dimensional.*

(3) *For any two roots α, β, the number*

$$\langle h_\alpha, \beta \rangle = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}$$

is integer.

(4) *For $\alpha \in R$, define the reflection operator $s_\alpha \colon \mathfrak{h}^* \to \mathfrak{h}^*$ by*

$$s_\alpha(\lambda) = \lambda - \langle h_\alpha, \lambda \rangle \alpha = \lambda - \frac{2(\alpha, \lambda)}{(\alpha, \alpha)} \alpha.$$

Then for any roots α, β, $s_\alpha(\beta)$ is also a root. In particular, if $\alpha \in R$, then $-\alpha = s_\alpha(\alpha) \in R$.

(5) *For any root α, the only multiples of α which are also roots are $\pm\alpha$.*

(6) *For roots $\alpha, \beta \neq \pm\alpha$, the subspace*

$$V = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{\beta + k\alpha}$$

is an irreducible representation of $\mathfrak{sl}(2, \mathbb{C})_\alpha$.

(7) *If α, β are roots such that $\alpha + \beta$ is also a root, then $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.*

Proof. (1) Assume that R does not generate $\mathfrak{h}^*$; then there exists a non-zero $h \in \mathfrak{h}$ such that $\langle h, \alpha \rangle = 0$ for all $\alpha \in R$. But then root decomposition (7.1) implies that $\operatorname{ad} h = 0$. However, by definition in a semisimple Lie algebra, the center is trivial: $\mathfrak{z}(\mathfrak{g}) = 0$.

The fact that h_α span $\mathfrak{h}$ now immediately follows: using identification of $\mathfrak{h}$ with $\mathfrak{h}^*$ given by the Killing form, elements h_α are identified with non-zero multiples of α.

(2) Immediate from Lemma 7.20 and the fact that in any irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$, weight subspaces are one-dimensional.

(3) Consider $\mathfrak{g}$ as a representation of $\mathfrak{sl}(2, \mathbb{C})_\alpha$. Then elements of $\mathfrak{g}_\beta$ have weight equal to $\langle h_\alpha, \beta \rangle$. But by Theorem 5.7, weights of any finite-dimensional representation of $\mathfrak{sl}(2, \mathbb{C})$ are integer.

(4) Assume that $\langle h_\alpha, \beta \rangle = n \geq 0$. Then elements of $\mathfrak{g}_\beta$ have weight n with respect to action of $\mathfrak{sl}(2, \mathbb{C})_\alpha$. By Theorem 5.7, operator f_α^n is an isomorphism of the space of vectors of weight n with the space of vectors of weight $-n$. In particular, it means that if $v \in \mathfrak{g}_\beta$ is non-zero vector, then $f_\alpha^n v \in \mathfrak{g}_{\beta - n\alpha}$ is also non-zero. Thus, $\beta - n\alpha = s_\alpha(\beta) \in R$.

For $n \leq 0$, the proof is similar, using e^{-n} insead of f^n.

(5) Assume that α and $\beta = c\alpha$, $c \in \mathbb{C}$ are both roots. By part (3), $\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = 2c$ is integer, so c is a half-integer. Same argument shows that $1/c$ is also a half-integer. It is easy to see that this implies that $c = \pm 1, \pm 2, \pm 1/2$. Interchanging the roots if necessary and possibly replacing α by $-\alpha$, we have $c = 1$ or $c = 2$.

Now let us consider the subspace

$$V = \mathbb{C} h_\alpha \oplus \bigoplus_{k \in \mathbb{Z}, k \neq 0} \mathfrak{g}_{k\alpha} \subset \mathfrak{g}.$$

By Lemma 7.20, V is an irreducible representation of $\mathfrak{sl}(2, \mathbb{C})_\alpha$, and by part (2), $V[2] = \mathfrak{g}_\alpha = \mathbb{C} e_\alpha$. Thus, the map $\operatorname{ad} e_\alpha \colon \mathfrak{g}_\alpha \to \mathfrak{g}_{2\alpha}$ is zero. But the results of Section 5.1 show that in an irreducible representation, kernel of e is exactly the highest weight subspace. Thus, we see that V has highest weight 2: $V[4] = V[6] = \cdots = 0$. This means that $V = \mathfrak{g}_{-\alpha} \oplus \mathbb{C} h_\alpha \oplus \mathfrak{g}_\alpha$, so the only integer multiples of α which are roots are $\pm\alpha$. In particular, 2α is not a root.

Combining these two results, we see that if $\alpha, c\alpha$ are both roots, then $c = \pm 1$.

(6) Proof is immediate from $\dim \mathfrak{g}_{\beta + k\alpha} = 1$.

(7) We already know that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha + \beta}$. Since $\dim \mathfrak{g}_{\alpha + \beta} = 1$, we need to show that for non-zero $e_\alpha \in \mathfrak{g}_\alpha, e_\beta \in \mathfrak{g}_\beta$, we have $[e_\alpha, e_\beta] \neq 0$. This follows from the previous part and the fact that in an irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$, if $v \in V[k]$ is non-zero and $V[k+2] \neq 0$, then $e.v \neq 0$.

$\square$

In the next chapter, we will study in detail the set of roots R. As we will see, it gives us a key to the classification of semisimple Lie algebras.

Theorem 7.22.

(1) *Let $\mathfrak{h}_\mathbb{R} \subset \mathfrak{h}$ be the real vector space generated by $h_\alpha, \alpha \in R$. Then $\mathfrak{h} = \mathfrak{h}_\mathbb{R} \oplus i\mathfrak{h}_\mathbb{R}$, and the restriction of Killing form to $\mathfrak{h}_\mathbb{R}$ is positive definite.*

(2) *Let $\mathfrak{h}_\mathbb{R}^* \subset \mathfrak{h}^*$ be the real vector space generated by $\alpha \in R$. Then $\mathfrak{h}^* = \mathfrak{h}_\mathbb{R}^* \oplus i\mathfrak{h}_\mathbb{R}^*$. Also, $\mathfrak{h}_\mathbb{R}^* = \{\lambda \in \mathfrak{h}^* \mid \langle \lambda, h \rangle \in \mathbb{R} \text{ for all } h \in \mathfrak{h}_\mathbb{R}\} = (\mathfrak{h}_\mathbb{R})^*$.*

Proof. Let us first prove that the restriction of the Killing form to $\mathfrak{h}_\mathbb{R}$ is real and positive definite. Indeed,

$$(h_\alpha, h_\beta) = \operatorname{tr}(\operatorname{ad} h_\alpha \operatorname{ad} h_\beta) = \sum_{\gamma \in R} \langle h_\alpha, \gamma \rangle \langle h_\beta, \gamma \rangle.$$

But by Theorem 7.21, $\langle h_\alpha, \gamma \rangle$, $\langle h_\beta, \gamma \rangle \in \mathbb{Z}$, so $(h_\alpha, h_\beta) \in \mathbb{Z}$.

Now let $h = \sum c_\alpha h_\alpha \in \mathfrak{h}_\mathbb{R}$. Then $\langle h, \gamma \rangle = \sum c_\alpha \langle h_\alpha, \gamma \rangle \in \mathbb{R}$ for any root γ, so

$$(h, h) = \operatorname{tr}(\operatorname{ad} h)^2 = \sum_\gamma \langle h, \gamma \rangle^2 \geq 0$$

which proves that the Killing form is positive definite.

By a well known result of linear algebra, this shows that $\dim_\mathbb{R} \mathfrak{h}_\mathbb{R} \leq \frac{1}{2} \dim_\mathbb{R} \mathfrak{h} = r$, where $r = \dim_\mathbb{C} \mathfrak{h}$ is the rank of $\mathfrak{g}$. On the other hand, since h_α generate $\mathfrak{h}$ over $\mathbb{C}$, we see that $\dim_\mathbb{R} \mathfrak{h}_\mathbb{R} \geq r$. Thus, $\dim_\mathbb{R} \mathfrak{h}_\mathbb{R} = r$, so $\mathfrak{h} = \mathfrak{h}_\mathbb{R} \oplus i\mathfrak{h}_\mathbb{R}$.

The second part easily follows from the first one. $\square$

Exercises

7.1. Show that for $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$, Definition 7.1 is equivalent to the usual definition of a semisimple operator (hint: use results of Appendix B).

7.2. Let $\mathfrak{g}$ be a Lie algebra which has a root decomposition:

$$g = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha$$

where R is a finite subset in $\mathfrak{h}^* - \{0\}$, $\mathfrak{h}$ is commutative and for $h \in \mathfrak{h}, x \in \mathfrak{g}_\alpha$ we have $[h, x] = \langle \mathfrak{h}, \alpha \rangle x$. Show that then $\mathfrak{g}$ is semisimple, and $\mathfrak{h}$ is a Cartan subalgebra.

7.3. Let $\mathfrak{h} \subset \mathfrak{so}(4, \mathbb{C})$ be the subalgebra consisisting of matrices of the form

$$\begin{bmatrix} & a & & \\ -a & & & \\ & & & b \\ & & -b & \end{bmatrix}$$

(entries not shows are zeros). Show that then $\mathfrak{h}$ is a Cartan subalgebra abd find the corresponding root decomposition.

Chapter 8

Root Systems

8.1. Abstract root systems

The results of Section 7.3 show that the set of roots R of a semisimple complex Lie algebra has a number of remarkable properties. It turns out that the sets with similar properties also appear in many other areas of mathematics. Thus, we will introduce the notion of abstract root system and study such objects, leaving for some time the theory of Lie algebras.

Definition 8.1. An abstract root system is a finite set of elements $R \subset E \setminus \{0\}$, where E is a real vector space with a positive definite inner product, such that the following properties hold:

(R1) R generates E as a vector space.

(R2) For any two roots α, β, the number

$$(8.1) \qquad n_{\alpha\beta} = \frac{2(\alpha, \beta)}{(\beta, \beta)}$$

is integer.

(R3) Let $s_\alpha \colon E \to E$ be defined by

$$(8.2) \qquad s_\alpha(\lambda) = \lambda - \frac{2(\alpha, \lambda)}{(\alpha, \alpha)} \alpha.$$

Then for any roots α, β, $s_\alpha(\beta) \in R$.

The number $r = \dim E$ is called the *rank* of R.

If, in addition, R satisfies the following property

(R4) If $\alpha, c\alpha$ are both roots, then $c = \pm 1$.

then R is called a *reduced* root system.

Remark 8.2. It is easy to deduce from (R1)—(R3) that if $\alpha, c\alpha$ are both roots, then $c = \pm 1, \pm 2, \pm \frac{1}{2}$ (see proof of Theorem 7.21). However, there are indeed examples of non-reduced root systems, which contain α and 2α as roots — see Exercise 8.1. Thus, condition (R4) does not follow from (R1)—(R3).

Note that conditions (R2), (R3) have a very simple geometric meaning. Namely, s_α is the reflection around the hyperplane

$$(8.3) \qquad L_\alpha = \{\lambda \in E \mid (\alpha, \lambda) = 0\}$$

It can be defined by $s_\alpha(\lambda) = \lambda$ if $(\alpha, \lambda) = 0$ and $s_\alpha(\alpha) = -\alpha$.

Similarly, the number $n_{\alpha\beta}$ also has a simple geometric meaning: if we denote by p_α the operator of orthogonal projection onto the line containing α, then $p_\alpha(\beta) = \frac{n_{\beta\alpha}}{2}\alpha$. Thus, (R2) says that the projection of β onto α is a half-integer multiple of α.

Using the notion of a root system, much of the results of the previous chapter can be reformulated as follows.

Theorem 8.3. *Let $\mathfrak{g}$ be a semisimple complex Lie algebra, with root decomposition* (7.1). *Then the set of roots $R \subset \mathfrak{h}_\mathbb{R}^* \setminus \{0\}$ is a reduced root system.*

Finally, for future use it is convenient to introduce, for every root $\alpha \in R$, the corresponding *coroot* $\alpha^\vee \in E^*$ defined by

$$(8.4) \qquad \langle \alpha^\vee, \lambda \rangle = \frac{2(\alpha, \lambda)}{(\alpha, \alpha)}.$$

Note that for the root system of a semisimple Lie algebra, this coincides with the definition of $h_\alpha \in \mathfrak{h}$ defined by (7.4): $\alpha^\vee = h_\alpha$.

Then one easily sees that $\langle \alpha^\vee, \alpha \rangle = 2$ and that

$$(8.5) \qquad \begin{aligned} n_{\alpha\beta} &= \langle \alpha, \beta^\vee \rangle \\ s_\alpha(\lambda) &= \lambda - \langle \lambda, \alpha^\vee \rangle \alpha. \end{aligned}$$

Example 8.4. Let e_i be the standard basis of $\mathbb{R}^n$, with usual inner product: $(e_i, e_j) = \delta_{ij}$. Let $E = \{(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n \mid \sum \lambda_i = 0\}$, and $R = \{e_i - e_j \mid 1 \leq i, j \leq n, i \neq j\} \subset E$. Then R is a reduced root system. Indeed, one easily sees that for $\alpha = e_i - e_j$, the corresponding reflection $s_\alpha \colon E \to E$ is transposition of i, j entries:

$$s_{e_i - e_j}(\ldots, \lambda_i, \ldots, \lambda_j, \ldots) = (\ldots, \lambda_j, \ldots, \lambda_i, \ldots)$$

Clearly, R is stable under such transpositions (and, more generally, under all permutations). Thus, condition (R3) is satisfied.

Since $(\alpha, \alpha) = 2$ for any $\alpha \in R$, condition (R2) is equivalent to $(\alpha, \beta) \in \mathbb{Z}$ for any $\alpha, \beta \in R$ which is immediate.

Finally, condition (R1) is obvious. Thus, R is a root system of rank $n-1$. For historical reasons, this root system is usually referred to as "root system of type A_{n-1}" (subscript is chosen to match the rank of the root system).

Alternatively, one can also define E as a quotient of $\mathbb{R}^n$:

$$E = \mathbb{R}^n / \mathbb{R}(1, \ldots, 1).$$

In this description, we see that this root system is exactly the root system of Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ (see Example 7.17).

8.2. Automorphisms and Weyl group

Most important information about the root system is contained in the numbers $n_{\alpha\beta}$ rather than in inner product themselves. Thism motivates the following definition.

Definition 8.5. Let $R_1 \subset E_1$, $R_2 \subset E_2$ be two root systems. An isomorphism $\varphi \colon R_1 \to R_2$ is a vector space isomorphism $\varphi \colon E_1 \to E_2$ which also gives a bijection R_1 $simeq R_2$ and such that $n_{\varphi(\alpha)\varphi(\beta)} = n_{\alpha\beta}$ for any $\alpha, \beta \in R_1$.

Note that condition $n_{\varphi(\alpha)\varphi(\beta)} = n_{\alpha\beta}$ will be automatically satisfied if φ preserves inner product. However, not every isomorphism of root systems preserves inner products. For example, for any $c \in R_+$, the root systems R and $cR = \{c\alpha, \alpha \in R\}$ are isomorphic. The isomorphism is given by $v \mapsto cv$, which does not preserve the inner product.

A special class of automorphisms of a root system R are those generated by reflections s_α.

Definition 8.6. The Weyl group W of a root system R is the subgroup of $\mathrm{GL}(E)$ generated by reflections s_α, $\alpha \in R$.

Lemma 8.7.

    (1) *The Weyl group W is a finite subgroup in the orthogonal group $\mathrm{O}(E)$, and the root system R is invariant under the action of W.*

    (2) *For any $w \in W$, we have $s_{w(\alpha)} = w s_\alpha w^{-1}$.*

Proof. Since every reflection s_α is an orthogonal tranformation, $W \subset \mathrm{O}(E)$. Since $s_\alpha(R) = R$ (by the axioms of a root system), we have $w(R) = R$ for any $w \in W$. Moreover, if some $w \in W$ leaves every root invariant, then $w = \mathrm{id}$ (because R generates E). Thus, W is a subgroup of the group $\mathrm{Aut}(R)$ of all automorphisms of R. Since R is a finite set, $\mathrm{Aut}(R)$ is finite; thus W is also finite.

The second identity is obvious: indeed, $w s_\alpha w^{-1}$ acts as identity on the hyperplane $w L_\alpha = L_{w(\alpha)}$, and $w s_\alpha w^{-1}(w(\alpha)) = -w(\alpha)$, so it is a reflection corresponding to root $w(\alpha)$. $\qquad\square$

Example 8.8. Let R be the root system of type A_{n-1} (see Example 8.4). Then W is the group generated by transpositions s_{ij}. It is easy to see that these transpositions generate the symmetric group S_n; thus, for this root system $W = S_n$.

In particular, for root system A_1 (i.e., root system of $\mathfrak{sl}(2,\mathbb{C})$), we have $W = S_2 = \mathbb{Z}_2 = \{1, \sigma\}$ where σ acts on $E \simeq \mathbb{R}$ by $\lambda \mapsto -\lambda$.

It should be noted, however, that not all automorphisms of a root system are given by elements of Weyl group. For example, for A_n, $n > 2$, the automorphism $\alpha \mapsto -\alpha$ is not in the Weyl group.

8.3. Pairs of roots and rank two root systems

Our main goal is to give a full classification of all possible reduced root systems, which in turn will be used to get a classification of all semisimple Lie algebras. The first step is considering the rank two case.

Throughout this section, R is a reduced root system.

The first observation is that conditions (R2), (R3) impose very strong restrictions on relative position of two roots.

Theorem 8.9. *Let $\alpha, \beta \in R$ roots which are not multiples of one another, with $|\alpha| \geq |\beta|$, and let φ be the angle between them. Then we must have one of the following possibilities:*

    (1) $\varphi = \pi/2$ *(i.e., α, β are orthogonal)*, $n_{\alpha\beta} = n_{\beta\alpha} = 0$

    (2a) $\varphi = 2\pi/3$, $|\alpha| = |\beta|$, $n_{\alpha\beta} = n_{\beta\alpha} = -1$

    (2b) $\varphi = \pi/3$, $|\alpha| = |\beta|$, $n_{\alpha\beta} = n_{\beta\alpha} = 1$.

    (3a) $\varphi = 3\pi/4$, $|\alpha| = \sqrt{2}|\beta|$, $n_{\alpha\beta} = -2, n_{\beta\alpha} = -1$.

    (3b) $\varphi = \pi/4$, $|\alpha| = \sqrt{2}|\beta|$, $n_{\alpha\beta} = 2, n_{\beta\alpha} = 1$.

    (4a) $\varphi = 5\pi/6$, $|\alpha| = \sqrt{3}|\beta|$, $n_{\alpha\beta} = 3, n_{\beta\alpha} = 1$.

    (4b) $\varphi = \pi/6$, $|\alpha| = \sqrt{3}|\beta|$, $n_{\alpha\beta} = -3, n_{\beta\alpha} = -1$.

Proof. Recall $n_{\alpha\beta}$ defined by (8.1). Since $(\alpha, \beta) = |\alpha||\beta|\cos\varphi$, we see that $n_{\alpha\beta} = 2\dfrac{\sqrt{|\alpha|}}{\sqrt{|\beta|}}\cos\varphi$. Thus, $n_{\alpha\beta}n_{\beta\alpha} = 4\cos^2\varphi$. Since $n_{\alpha\beta}n_{\beta\alpha} \in \mathbb{Z}$, this means that $n_{\alpha\beta}n_{\beta\alpha}$ must be one of $0, 1, 2, 3$. Analyzing each of these possibilities and using $\frac{n_{\alpha\beta}}{n_{\beta\alpha}} = \frac{|\alpha|}{|\beta|}$ if $\cos\varphi \neq 0$, we get the statement of the theorem. $\square$

It turns out that each of the possibilities listed in this theorem is indeed realized.

Theorem 8.10.

(1) *Let $A_1 \times A_1$, A_2, B_2, G_2 be the sets of vectors in $\mathbb{R}^2$ shown in Figure 8.1. Then each of them is a rank two root system.*

(2) *Any rank two reduced root system is isomorphic to one of root systems $A_1 \times A_1$, A_2, B_2, G_2.*
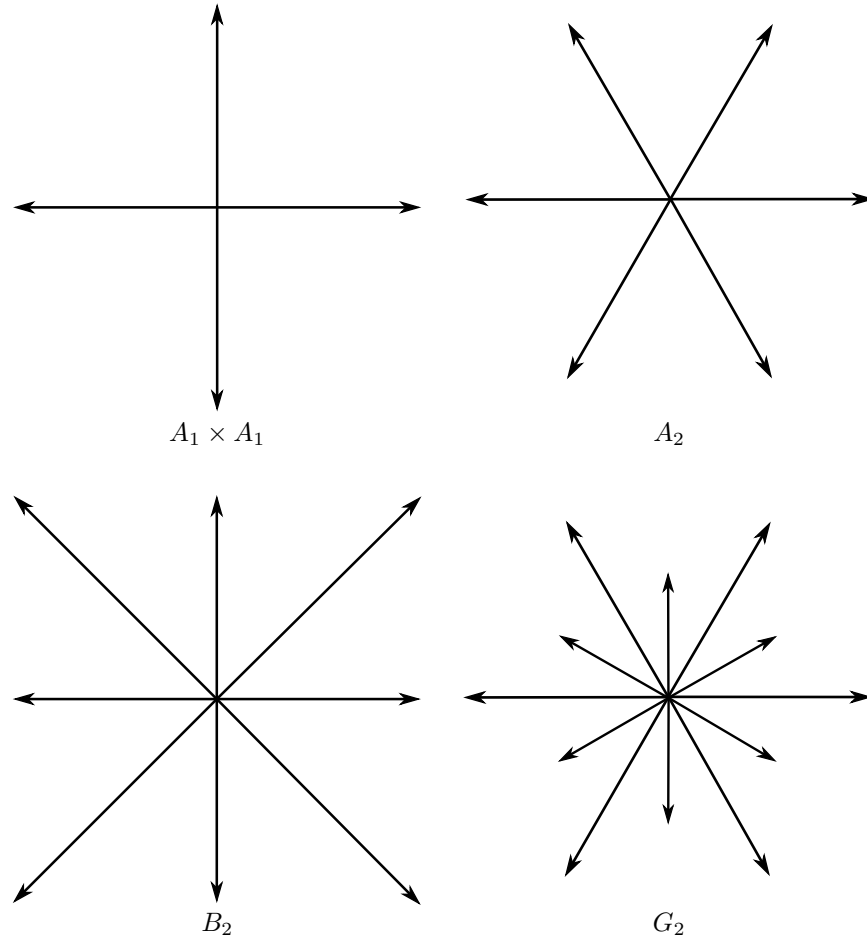
Figure 8.1. Rank two root systems

Proof. Proof of part (1) is given by explicit analysis. Since for any pair of vectors in these systems, the angle and ratio of lengths is among one of the possibilities listed in Theorem 8.9, condition (R2) is satisfed. It is also easy to see that condition (R3) is satisfied.

To prove the second part, assume that R is a reduced rank 2 root system. Let us choose α, β to be two roots such that the angle φ between them is as large as possible and $|\alpha| \geq |\beta|$. Then $\varphi \geq \pi/2$ (otherwise, we could take the pair $\alpha, s_\alpha(\beta)$ and get a larger angle). Thus, we must be in one of situations (1), (2a), (3a), (4a) of Theorem 8.9.

Consider, for example, case (2a): $|\alpha| = |\beta|$, $\varphi = 2\pi/3$. By definitions of root system, R is stable under reflections s_α, s_β. But succesively applying these two reflections to α, β we get exactly the root system of type A_2. Thus, in this case R contains as a subset the root system A_2 generated by α, β.

To show that in this case $R = A_2$, note that if we have another root γ which is not in A_2, then γ must be between some of the roots of A_2 (since R is reduced). Thus, the angle between γ and some root δ is less than $\pi/3$, and the angle between γ and $-\delta$ is greater than $2\pi/3$, which is impossible because angle between α, β is the maximal possible. Thus, $R = A_2$.

Similar analysis shows that in cases (1), (3a), (4a) of Theorem 8.9, we will get $R = A_1 \times A_1$, B_2, G_2 respectively. $\qquad\square$

For future use, we also give the following result.

Lemma 8.11. *Let $\alpha, \beta \in R$ be two roots such that $(\alpha, \beta) < 0$, $\alpha \neq c\beta$. Then $\alpha + \beta \in R$.*

Proof. It suffices to prove this for each of rank two root systems described in Theorem 8.10. For each of them, it is easy to check directly. $\qquad\square$

8.4. Positive roots and simple roots

In order to proceed with classification of root systems, we would like to find for each root system some small set of "generating roots", similar to what was done in the previous section for rank 2 root systems. In general it can be done as follows.

Let $t \in E$ be such that for any root α, $(t, \alpha) \neq 0$ (such elements t are called *regular*). Then we can write

(8.6)
$$R = R_+ \sqcup R_-$$
$$R_+ = \{\alpha \in R \mid (\alpha, t) > 0\}, \quad R_- = \{\alpha \in R \mid (\alpha, t) < 0\}$$

Such a decomposition will be called a *polarization* of R. Note that polarization depends on the choice of t. The roots $\alpha \in R_+$ will be called *positive*, and the roots $\alpha \in R_-$ will be called *negative*.

From now on, let us assume that we have fixed a polarization (8.6) of the root system R.

Definition 8.12. A root $\alpha \in R_+$ is called *simple* if it can not be written as a sum of two positive roots.

We will denote the set of simple roots by $\Pi \subset R_+$.

We have the following easy lemma.

Lemma 8.13. *Every positive root can be written as a sum of simple roots.*

Proof. If a positive root α is not simple, it can be written in the form $\alpha = \alpha' + \alpha''$, with $\alpha', \alpha'' \in R_+$, and $(\alpha', t) < (\alpha, t)$, $(\alpha'', t) < (\alpha, t)$. If α', α'' are not simple, we can apply the same argument to them to write them as a sum of positive roots. Since (α, t) can only take finitely many values, the process will terminate after finitely many steps. $\qquad\square$
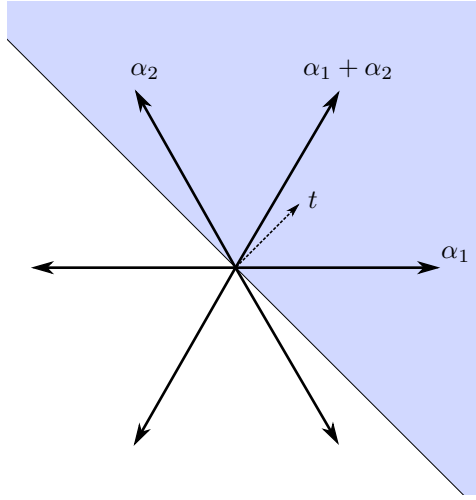
Figure 8.2. Positive and simple roots for A_2

Example 8.14. Let us consider the root system A_2 and let t be as shown in the figure below. Then there are three positive roots: two of them are denoted by α_1, α_2, and the third one is $\alpha_1 + \alpha_2$. Thus, one easily sees that α_1, α_2 are simple roots, and $\alpha_1 + \alpha_2$ is not simple.

Lemma 8.15. *If $\alpha, \beta \in R_+$ are simple, then $(\alpha, \beta) \leq 0$.*

Proof. Assume that $(\alpha, \beta) > 0$. Then, applying Lemma 8.11 to $-\alpha, \beta$, we see that $\beta' = \beta - \alpha \in R$. If $\beta' \in R_+$, then $\beta = \beta' + \alpha$ can not be simple. If $\beta' \in R_-$, then $-\beta' \in R_+$, so $\alpha = -\beta' + \beta$ can not be simple. This contradiction shows that $(\alpha, \beta) > 0$ is impossible. $\square$

Theorem 8.16. *Let $R = R_+ \sqcup R_- \subset E$ be a root system. Then the simple roots form a basis of the vector space E.*

Proof. By Lemma 8.13, every positive root can be written as linear combination of simple roots. Since R spans E, this implies that the set of simple roots spans E.

Linear independence of simple roots follows from the results of Lemma 8.15 and the following linear algebra lemma proof of which is given in the exercises (Exercise 8.3).

Lemma 8.17. *Let $v_1, \ldots v_k$ be a collection of non-zero vectors in a Euclidean space E such that for $i \neq j$, $(v_i, v_j) \leq 0$. Then $\{v_1, \ldots, v_k\}$ are linearly independent.*

$\square$

Corollary 8.18. *Every $\alpha \in R$ can be uniquely written as a linear combination of simple roots with integer coefficients:*

$$(8.7) \qquad \alpha = \sum_{i=1}^{r} n_i \alpha_i, \qquad n_i \in \mathbb{Z}$$

where $\{\alpha_1, \ldots, \alpha_r\} = \Pi$ is the set of simple roots. If $\alpha \in R_+$, then all $n_i \geq 0$; if $\alpha \in R_-$, then all $n_i \leq 0$.

For a positive root $\alpha \in R_+$, we define its *height* by

$$(8.8) \qquad \mathrm{ht}\left(\sum n_i \alpha_i\right) = \sum n_i \in \mathbb{Z}_+$$

so that $\mathrm{ht}(\alpha_i) = 1$. In many cases, statements about positive roots can be proved by induction in height.

Example 8.19. Let R be the root system of type A_{n-1} or equivalently, the root system of $\mathfrak{sl}(n, \mathbb{C})$ (see Example 7.17, Example 8.4). Choose the polarization as follows:

$$R_+ = \{e_i - e_j \mid i < j\}$$

(the corresponding root subspaces $E_{ij}, i < j$, generate the Lie subalgebra $\mathfrak{n}$ of strictly upper-triangular matrices in $\mathfrak{sl}(n, \mathbb{C})$).

Then it is easy to show that the simple roots are

$$\alpha_1 = e_1 - e_2, \qquad \alpha_2 = e_2 - e_3, \quad \ldots, \quad \alpha_{n-1} = e_{n-1} - e_n$$

and indeed, any positive root can be written as a sum of simple roots with non-negative integer coefficients. For example, $e_2 - e_4 = (e_2 - e_3) + (e_3 - e_4) = \alpha_2 + \alpha_3$. The height is given by $\mathrm{ht}(e_i - e_j) = j - i$.

8.5. Weight and root lattices

In the study of root systems of simple Lie algebras, we will frequently use the following lattices. Recall that a lattice in a real vector space E is an abelian group generated by a basis in E. Of course, by a suitable change of basis any lattice $L \subset E$ can be identified with $\mathbb{Z}^n \subset \mathbb{R}^n$.

Every root system $R \subset E$ gives rise to the following lattices:

$$
\begin{aligned}
Q &= \{\text{abelian group generated by } \alpha \in R\} \subset E \\
Q^\vee &= \{\text{abelian group generated by } \alpha^\vee, \alpha \in R\} \subset E^*
\end{aligned}
$$
(8.9)

Lattice Q is called the *root lattice* of R, and $Q^\vee$ is the *coroot lattice*.

To justify the use of the word lattice, we need to show that $Q, Q^\vee$ are indeed generated by a basis in E (respectively E^*). This can be done as follows. Fix a polarization of R and let $\Pi = \{\alpha_1, \ldots, \alpha_r\}$ be the corresponding system of simple roots. Since every root can be written as a linear combination of simple roots with integer coefficients (Corollary 8.18), one has

$$(8.10) \qquad\qquad Q = \bigoplus \mathbb{Z}\alpha_i,$$

which shows that Q is indeed a lattice. Similarly, it follows from Exercise 8.2 that

$$(8.11) \qquad\qquad Q^\vee = \bigoplus \mathbb{Z}\alpha_i^\vee.$$

Even more important in the applications to representation theory of semisimple Lie algebras is the *weight lattice* $P \subset E$ defined as follows:

$$(8.12) \qquad P = \{\lambda \in E \mid \langle \lambda, \alpha^\vee \rangle \in \mathbb{Z} \text{ for all } \alpha \in R\} = \{\lambda \in E \mid \langle \lambda, \alpha^\vee \rangle \in \mathbb{Z} \text{ for all } \alpha^\vee \in Q^\vee\}.$$

In other words, $P \subset E$ is exactly the dual lattice of $Q^\vee \subset E^*$. Elements of P are frequently called *integral* weights. Their role in representation theory will be discussed in Chapter 9.

Since $Q^\vee$ is generated by $\alpha_i^\vee$, the weight lattice can also be defined by

$$(8.13) \qquad\qquad P = \{\lambda \in E \mid \langle \lambda, \alpha_i^\vee \rangle \in \mathbb{Z} \text{ for all simple roots } \alpha_i\}.$$

One can easily define a basis in P. Namely, define *fundamental weights* $\omega_i \in E$ by

$$(8.14) \qquad\qquad \langle \omega_i, \alpha_j^\vee \rangle = \delta_{ij}.$$

Then one easily sees that so defined ω_i form a basis in E and that

$$P = \bigoplus_i \mathbb{Z}\omega_i.$$

Finally, note that by the axioms of a root system, we have $n_{\alpha\beta} = \langle \alpha, \beta^\vee \rangle \in \mathbb{Z}$ for any roots α, β. Thus, $R \subset P$ which implies that

$$Q \subset P$$

However, in general $P \neq Q$, as the examples below show. Since both P, Q are free abelian groups of rank r, general theory of finitely generated abelian groups implies that the quotient group P/Q is a finite abelian group. It is also possible to describe the order $|P/Q|$ in terms of the matrix $a_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle$ (see Exercise 8.4).

Example 8.20. Consider the root system A_1. It has the unique positive root α, so $Q = \mathbb{Z}\alpha$, $Q^\vee = \mathbb{Z}\alpha^\vee$. If we define the inner product $(\, ,\,)$ by $(\alpha, \alpha) = 2$, and use this product to identify $E^* \simeq E$, then under this identification $\alpha^\vee \mapsto \alpha$, $Q^\vee \xrightarrow{\sim} Q$.

Since $\langle \alpha, \alpha^\vee \rangle = 2$, we see that the fundamental weight is $\omega = \frac{\alpha}{2}$, and $P = \mathbb{Z}\frac{\alpha}{2}$. Thus, in this case $P/Q = \mathbb{Z}_2$.

Example 8.21. For the root system A_2, the root and weight lattices are shown in Figure 8.3. This figure also shows simple roots α_1, α_2 and fundamental weights ω_1, ω_2.
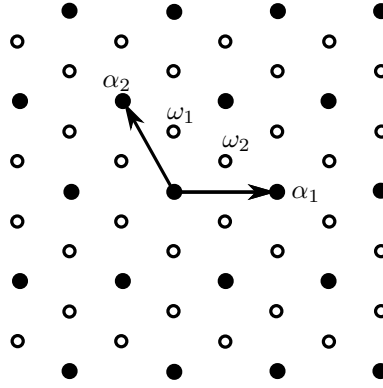
Figure 8.3. Weight and root lattices for A_2. Bold dots show $\alpha \in Q$, empty dots $\alpha \in P - Q$.

It is easy to see from the figure (and also easy to prove algebraically) that one can take α_1, ω_2 as a basis of P, and that $\alpha_1, 3\omega_2 = \alpha_2 + 2\alpha_1$ is a basis of Q. Thus, $P/Q = \mathbb{Z}_3$.

8.6. Weyl chambers

In the previous sections, we have constructed, starting with a root system R, first the set of positive roots R_+ and then a smaller set of simple roots $\Pi = \{\alpha_1, \ldots, \alpha_r\}$ which in a suitable sense generates R. Schematically this can be shown as follows:

$$R \longrightarrow R_+ \longrightarrow \Pi = \{\alpha_1, \ldots, \alpha_r\}$$

The first step (passage from R to R_+) requires a choice of polarization, which is determined by a regular element $t \in E$; the second step is independent of any choices.

Our next goal is to use this information to get a classification of reduced root systems, by classifying possible sets of simple roots. However, before doing this we need to answer the following two questions:

(1) Is it possible to recover R from Π?

(2) Do different choices of polarization give rise to equivalent in a suitable sense sets of simple roots Π, Π'?

We will start with the second question as it is easier to answer.

Recall that a polarization is defined by an element $t \in E$ which does not lie on any of the hyperplanes orthogonal to roots:

(8.15)
$$t \in E \setminus \bigcup_{\alpha \in R} L_\alpha$$
$$L_\alpha = \{\lambda \in E \mid (\alpha, \lambda) = 0\}$$

Moreover, the polarization actually depends not on t itself but only on the signs of (t, α); thus, polarization is unchanged if we change t as long as we do not cross any of the hyperplanes. This justifies the following definition.

Definition 8.22. A Weyl chamber is a connected component of the complement to the hyperplanes:

$$C = \text{connected component of } \left(E \setminus \bigcup_{\alpha \in R} L_\alpha \right)$$

For example, for root system A_2 there are 6 Weyl chambers; one of them is shaded in the figure below.
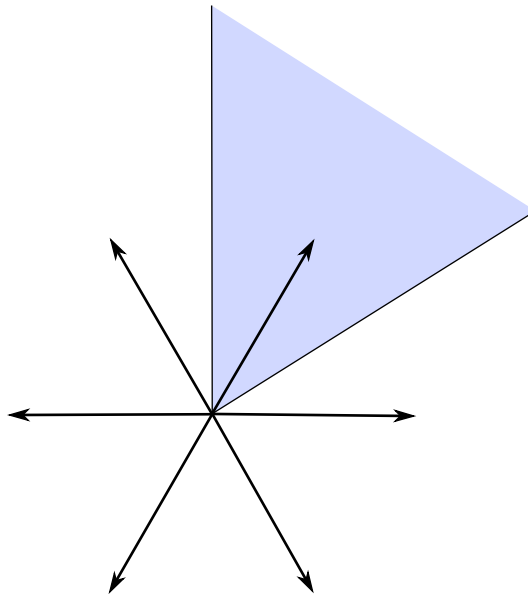
Figure 8.4. A Weyl chamber for A_2

Clearly, to specify a Weyl chamber we need to specify, for each hyperplane L_α, on which side of the hyperplane the Weyl chamber lies. Thus, a Weyl chamber is defined by a system of inequalities of the form

$$\pm(\alpha, \lambda) > 0$$

(one inequality for each pair of roots $\alpha, -\alpha$). Any such system of inequalities defines either an empty set or a Weyl chamber.

For future use, we state here some results about geometry of Weyl chambers.

Lemma 8.23.

(1) *The closure $\overline{C}$ of a Weyl chamber C is an unbounded convex cone.*

(2) *The boundary $\partial\overline{C}$ is a union of finite number of codimension one faces: $\partial\overline{C} = \bigcup F_i$. Each F_i is a closed convex unbounded subset in one of the hyperplanes L_α, given by a system of inequalities. The hyperplanes containing F_i are called* walls *of C.*

This lemma is geometrically obvious (in fact, it equally applies to any subset in a Euclidean space defined by a finite system of strict inequalities) and we omit the proof.

We can now return to the polarizations. Note that any Weyl chamber C defines a polarization given by

$$(8.16) \qquad\qquad R_+ = \{\alpha \in R \mid \alpha(t) > 0\}, \qquad t \in C$$

(this does not depend on the choice of $t \in C$). Conversely, given a polarization $R = R_+ \sqcup R_-$, define the corresponding *positive Weyl chamber C_+* by

$$(8.17) \qquad C_+ = \{\lambda \in E \mid (\lambda, \alpha) > 0 \text{ for all } \alpha \in R_+\} = \{\lambda \in E \mid (\lambda, \alpha_i) > 0 \text{ for all } \alpha_i \in \Pi\}$$

(to prove the last equality, note that if $(\lambda, \alpha_i) > 0$ for all $\alpha_i \in \Pi$, then by Lemma 8.13, for any $\alpha = \sum n_i \alpha_i$, we have $(\lambda, \alpha) > 0$). This system of inequalities does have solutions (because the element t used to define the polarization satisfies these inequalities) and thus defines a Weyl chamber.

Lemma 8.24. *Formulas (8.16), (8.17) define a bijection between the set of all polarizations of R and the set of Weyl chambers.*

Proof is left as an exercise to the reader.

In order to relate polarizations defined by different Weyl chambers, recall the Weyl group W defined in Section 8.2. Since action of W maps root hyperplanes to root hyperplanes, we have a well-defined action of W on the set of Weyl chambers.

Theorem 8.25. *The Weyl group acts transitively on the set of Weyl chambers.*

Proof. The proof is based on several facts which are of significant interest in their own right. Namely, let us say that two Weyl chambers C, C' are *adjacent* if they have a common codimension one face F (obviously, they have to be on different sides of F). If L_α is the hyperplane containing this common face F, then we will say that C, C' are adjacent chambers separated by L_α.

Then we have the following two lemmas, proof of which as an exercise to the reader.

Lemma 8.26. *Any two Weyl chambers C, C' can be connected by a sequence of chambers $C_0 = C$, $C_1, \ldots, C_l = C'$ such that C_i is adjacent to C_{i+1}.*

Lemma 8.27. *If C, C' are adjacent Weyl chambers separated by hyperplane L_α then $s_\alpha(C) = C'$.*

The statement of theorem now easily follows from these two lemmas. Indeed, let C, C' be two Weyl chambers. By Lemma 8.26, they can be connected by a sequence of Weyl chambers $C_0 = C$, $C_1, \ldots, C_l = C'$. Let L_{β_i} be the hyperplane separating C_{i-1} and C_i. Then, by Lemma 8.27,

$$(8.18) \qquad \begin{aligned} C_l &= s_{\beta_l}(C_{l-1}) = s_{\beta_l} s_{\beta_{l-1}}(C_{l-2}) = \ldots \\ &= s_{\beta_l} \ldots s_{\beta_1}(C_0) \end{aligned}$$

so $C' = w(C)$, with $w = s_{\beta_l} \ldots s_{\beta_1}$. This completes the proof of Theorem 8.25. $\qquad\square$

Corollary 8.28. *Every Weyl chamber has exactly $r = \mathrm{rank}(R)$ walls. Walls of positive Weyl chamber C_+ are L_{α_i}, $\alpha_i \in \Pi$.*

Proof. For positive Weyl chamber C_+, this follows from (8.17). Since every Weyl chamber can be written in the form $C = w(C_+)$ for some $w \in W$, all Weyl chambers have the same number of walls. $\square$

Corollary 8.29. *Let $R = R_+ \sqcup R_- = R'_+ \sqcup R'_-$ be two polarizations of the same root system, and Π, Π' the corresponding sets of simple roots. Then there exists an element $w \in W$ such that $\Pi = w(\Pi')$.*

Proof. By Lemma 8.24, each polarization is defined by a Weyl chamber. Since W acts transitvely on the set of Weyl chambers, it also acts transitively on the set of all polarizations. $\square$

This last corollary provides an answer to the question asked in the beginning of this section: sets of simple roots obtained from different polarizations can be related by an orthogonal transformation of E.

8.7. Simple reflections

We can now return to the first question asked in the beginning of the previous section: is it possible to recover R from the set of simple roots Π? The answer is again based on the use of Weyl group.

Theorem 8.30. *Let R be a reduced root system, with fixed polarization $R = R_+ \sqcup R_-$. Let $\Pi = \{\alpha_1, \ldots, \alpha_r\}$ be the set of simple roots. Consider reflections corresponding to simple roots $s_i = s_{\alpha_i}$ (they are called simple reflections).*

(1) *The simple refelections s_i generate W.*

(2) *$W(\Pi) = R$: every $\alpha \in R$ can be written in the form $w(\alpha_i)$ for some $w \in W$ and $\alpha_i \in \Pi$.*

Proof. We start by proving the following result

Lemma 8.31. *Any Weyl chamber can be written as*
$$C = s_{i_1} \ldots s_{i_l}(C_+)$$
for some sequence of indices $i_1, \ldots, i_l \in \{1, \ldots, r\}$.

Proof. By the construction given in the proof of Theorem 8.25, we can connect C_+, C by a chain of adjacent Weyl chambers $C_0 = C_+, C_1, \ldots, C_l = C$. Then $C = s_{\beta_l} \ldots s_{\beta_1}(C_+)$, where L_{β_i} is the hyperplane separating C_{i-1} and C_i.

Since β_1 separates $C_0 = C_+$ from C_1, it means that L_{β_1} is one of the walls of C_+. Since the walls of C_+ are exactly hyperplanes L_{α_i} corresponding to simple roots (see Corollary 8.28), we see that $\beta_1 = \pm\alpha_{i_1}$ for some index $i_1 \in \{1, \ldots, r\}$, so $s_{\beta_1} = s_{i_1}$ and $C_1 = s_{i_1}(C_+)$.

Consider now the hyperplane L_{β_2} separating C_1 from C_2. It is a wall of $C_1 = s_{i_1}(C_+)$; thus, it must be of the form $L_{\beta_2} = s_{i_1}(L)$ for some hyperplane L which is a wall of C_+. Thus, we get that $\beta_2 = \pm s_{i_1}(\alpha_{i_2})$ for some index i_2. By Lemma 8.7, we therefore have $s_{\beta_2} = s_{i_1} s_{i_2} s_{i_1}$ and thus

$$s_{\beta_2} s_{\beta_1} = s_{i_1} s_{i_2} s_{i_1} \cdot s_{i_1} = s_{i_1} s_{i_2}$$
$$C_2 = s_{i_1} s_{i_2}(C_+)$$

Repeating the same argument, we finally get that
$$C = s_{i_1} \ldots s_{i_l}(C_+)$$
and the indices i_k are computed inductively, by

(8.19) $$\beta_k = s_{i_1} \ldots s_{i_{k-1}}(\alpha_{i_k})$$

which completes the proof of the lemma. $\square$

Now the theorem easily follows. Indeed, every hyperplane L_α is a wall of some Weyl chamber C. Using the lemma, we can write $C = w(C_+)$ for some $w = s_{i_1} \ldots s_{i_l}$. Thus, $L_\alpha = w(L_{\alpha_j})$ for some index j, so $\alpha = \pm w(\alpha_j)$ and $s_\alpha = w s_j w^{-1}$, which proves both statements of the theorem.

$\square$

It is also possible to write the full set of defining relations for W (see Exercise 8.8).

Example 8.32. Let R be the root system of type A_{n-1}. Then the Weyl group is $W = S_n$ (see Example 8.8) and simple reflections are transpositions $s_i = (i\ i+1)$. And indeed, it is well known that these transpositions generate the symmetric group.

We can also describe the Weyl chambers in this case. Namely, the positive Weyl chamber is

$$C_+ = \{(\lambda_1, \ldots, \lambda_n) \in E \mid \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n\}$$

and all other Weyl chambers are obtained by applying to C_+ permutations $\sigma \in S_n$. Thus, they are of the form

$$C_\sigma = \{(\lambda_1, \ldots, \lambda_n) \in E \mid \lambda_{\sigma(1)} \geq \alpha_{\sigma(2)} \geq \cdots \geq \lambda_{\sigma(n)}\}, \qquad \sigma \in S_n.$$

Corollary 8.33. *The root system R can be recovered from the set of simple roots Π.*

Proof. Given Π, we can recover W as the group generated by s_i and then recover $R = W(\Pi)$. $\square$

Let us say that a root hyperplane L_α *separates* two Weyl chambers C, C' if these two chambers are on different sides of H_α, i.e. $\alpha(C), \alpha(C')$ have different signs.

Definition 8.34. Let R be a reduced root system, with set of simple roots Π. Then we define, for an element $w \in W$, its *length* by

(8.20) $l(w) =$ number of root hyperplanes separating C_+ and $w(C_+) = |\{\alpha \in R_+ \mid w(\alpha) \in R_-\}|.$

It should be noted that $l(w)$ depends not only on w itself but also on the choice of polarization $R = R_+ \sqcup R_-$ or equivalently, the set of simple roots.

Example 8.35. Let $w = s_i$ be a simple reflection. Then the Weyl chambers C_+ and $s_i(C_+)$ are separated by exactly one hyperplane, namely L_{α_i}. Therefore, $l(s_i) = 1$, and

(8.21) $\{\alpha \in R_+ \mid w(\alpha) \in R_-\} = \{\alpha_i\}.$

In other words, $s_i(\alpha_i) = -\alpha_i \in R_-$ and s_i permutes elements of $R_+ \setminus \{\alpha_i\}$.

This example is very useful in many arguments involving Weyl group, such as the following lemma.

Lemma 8.36. *Let*

(8.22)
$$\rho = \frac{1}{2} \sum_{\alpha \in R_+} \alpha.$$

Then $\langle \rho, \alpha_i^\vee \rangle = \frac{2(\rho, \alpha_i)}{(\alpha_i, \alpha_i)} = 1$.

Proof. Writing $\rho = (\alpha_i + \sum_{\alpha \in R_+ \setminus \{\alpha_i\}} \alpha)/2$ and using results of Example 8.35, we see that $s_i(\rho) = \rho - \alpha_i$. On the other hand, by definition $s_i(\lambda) = \lambda - \langle \alpha_i^\vee, \lambda \rangle \alpha_i$. $\square$

Theorem 8.37. *Let*

$$w = s_{i_1} \ldots s_{i_l}$$

be an expression for w as product of simple reflections which has minimal possible length (such expressions are called reduced*). Then $l = l(w)$.*

Proof. We can connect C and C_+ by a chain of Weyl chambers $C_0 = C_+, C_1, \ldots, C_l = w(C_+)$, where $C_k = s_{i_1} \ldots s_{i_k}(C_+)$. The same argument as in the proof of Theorem 8.31 shows that then C_k and C_{k-1} are adjacent Weyl chambers separated by root hypeplane L_{β_k}, with $\beta_k = s_{i_1} \ldots s_{i_{k-1}}(\alpha_{i_k})$. This shows that we can connect C_+ and $w(C_+)$ by a path crossing exactly l hyperplanes. In particular, this means that C_+ and $w(C_+)$ are separated by at most l hyperplanes, so $l(w) \leq l$.

Note, however, that we can not yet conclude that $l(w) = l$: it is possible that the path we had constructed crosses some hyperplane more than once. For example, we can write $1 = s_i s_i$, which gives us a path connecting C_+ with itself but crossing hypeplane L_{α_i} twice. So to show that that $l(w) = l$, we need to show that if $w = s_{i_1} \ldots s_{i_l}$ is a reduced expression, then all hyperplanes $L_{\beta_1}, \ldots, L_{\beta_l}$ are distinct: we never cross any hyperplane more than once. The proof of this fact is given as an exercise (see Exercise 8.5). $\square$

Corollary 8.38. *The action of W on the set of Weyl chambers is simply transitive.*

Proof. Otherwise, there exists $w \in W$ such that $w(C_+) = C_+$. But then, by definition, $l(w) = 0$. $\square$

Lemma 8.39. *Let C_- be the negative Weyl chamber: $C_- = -C_+$ and let $w_0 \in W$ be such that $w_0(C_+) = C_-$ (by Corollary 8.38, such an element exists and is unique). Then $l(w_0) = |R_+|$ and for any $w \in W$, $w \neq w_0$ we have $l(w) < l(w_0)$. For this reason w_0 is called the* longest element *in W.*

The proof of this lemma is left to the reader as an exercise.

8.8. Dynkin diagrams and classification of root systems

In the previous sections, we have discussed that given a reduced root system R, we can choose a polarization $R = R_+ \sqcup R_-$ and then define the set of simple roots $\Pi = \{\alpha_1, \ldots, \alpha_r\}$. We have shown that R can be recovered from Π (Corollary 8.33) and that different choices of polarization give rise to sets of simple roots which are related by the action of the Weyl group (Corollary 8.29). Thus, classifying root systems is equivalent to classifying possible sets of simple roots Π.

The main goal of this section will be to give a complete solution of this problem, i.e. give a classification of all root systems.

The first step is to note that there is an obvious construction which allows one to construct larger root systems from smaller ones. Namely, if $R_1 \subset E_1$ and $R_2 \subset E_2$ are two root systems, then we can define $R = R_1 \sqcup R_2 \subset E_1 \oplus E_2$, with the inner product on $E_1 \oplus E_2$ defined so that $E_1 \perp E_2$. It is easy to see that so defined R is again a root system.

Definition 8.40. A root system R is called *reducible* if it can be written in the form $R = R_1 \sqcup R_2$, with $R_1 \perp R_2$. Otherwise, R is called *irreducible*.

For example, root system $A_1 \times A_1$ discussed in Section 8.3 is reducible; all other root systems discussed in that section are irreducible.

Remark 8.41. It should be noted that a root system being reducible or irreducible is completely unrelated to whether the root system is reduced or not. It would be best if a different terminology were used, to avoid confusion; however, both of these terms are so widely used that changing them is not feasible.

There is an analogous notion for the set of simple roots.

Lemma 8.42. *Let R be a reduced root system, with given polarization, and let Π be the set of simple roots.*

    (1) *If R is reducible: $R = R_1 \sqcup R_2$, then $\Pi = \Pi_1 \sqcup \Pi_2$, where $\Pi_i = \Pi \cap R_i$ is the set of simple roots for R_i.*

    (2) *Conversely, if $\Pi = \Pi_1 \sqcup \Pi_2$, with $\Pi_1 \perp \Pi_2$, then $R = R_1 \sqcup R_2$, where R_i is the root system generated by Π_i.*

Proof. The first part is obvious. To prove the second part, notice that if $\alpha \in \Pi_1$, $\beta \in \Pi_2$, then $s_\alpha(\beta) = \beta$ and s_α, s_β commute. Thus, if we denote by W_i the group generated by simple reflections $s_\alpha, \alpha \in \Pi_i$, then $W = W_1 \times W_2$, and W_1 acts trivially on Π_2, W_2 acts trivially on Π_1. Thus, $R = W(\Pi_1 \sqcup \Pi_2) = W_1(\Pi_1) \sqcup W_2(\Pi_2)$. $\qquad\qquad\square$

It can be shown that every reducible root system can be uniquely written in the form $R_1 \sqcup R_2 \cdots \sqcup R_n$, where R_i are mutually orthogonal irreducible root systems. Thus, in order to classify all root systems, it suffices to classify all irreducible root systems. Fro this reason, from now on R is an irreducible root system and Π is the corresponding set of simple roots. We assume that we have chosen an order on the set of simple roots: $\Pi = \{\alpha_1 \ldots, \alpha_r\}$.

The compact way of describing relative position of roots $\alpha_i \in \Pi$ is by writing all inner products between these roots. However, this is not invariant under isomorphisms of root systems. A better way of describing relative position of simple roots is given by Cartan matrix.

Definition 8.43. The Cartan matrix A of a set of simple roots $\Pi \subset R$ is the $r \times r$ matrix with entries

$$(8.23) \qquad\qquad a_{ij} = n_{\alpha_j \alpha_i} = \langle \alpha_i^\vee, \alpha_j \rangle = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}.$$

The following properties of Cartan matrix immediately follow from definitions and from known properties of simple roots.

Lemma 8.44.

    (1) For any i, $a_{ii} = 2$.

    (2) For any $i \neq j$, a_{ij} is a non-positive integer: $a_{ij} \in \mathbb{Z}$, $a_{ij} \leq 0$.

    (3) For any $i \neq j$, $a_{ij} a_{ji} = 4\cos^2 \varphi$, where φ is angle between α_i, α_j. If $\varphi \neq \pi/2$, then

$$\frac{|\alpha_i|^2}{|\alpha_j|^2} = \frac{a_{ji}}{a_{ij}}.$$

Example 8.45. For the root system A_n, the Cartan matrix is

$$A = \begin{bmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & & \vdots & & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}$$

(entries which are not shown are zeroes).

The same information contained in the Cartan matrix can also be presented in a graphical way.

Definition 8.46. Let Π be a set of simple roots of a root sytem R. The *Dynkin diagram* of Π is the graph constructed as follows:

- For each simple root α_i, we construct a vertex v_i of the Dynkin diagram (traditionally, vertices are drawn as small circles rather than as dots)

- For each pair of simple roots $\alpha_i \neq \alpha_j$, we connect the corresponding vertices by n edges, where n depends on the angle φ between α_i, α_j:
  For $\varphi = \pi/2$, $n = 0$ (vertices are not connnected)
  For $\varphi = 2\pi/3$, $n = 1$ (case of A_2 system)
  For $\varphi = 3\pi/4$, $n = 2$ (case of B_2 system)
  For $\varphi = 5\pi/6$, $n = 3$ (case of G_2 system)

- Finally, for every pair of distinct simple roots $\alpha_i \neq \alpha_j$, if $|\alpha_i| \neq |\alpha_j|$ and they are not orthogonal, we orient the corresponding (multiple) edge by putting on it an arrow pointing towards the shorter root.

Example 8.47. The Dynkin diagrams for rank two root systems are shown in Figure 8.5.
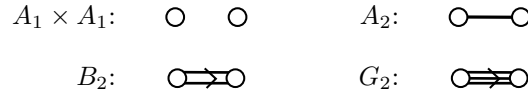
Figure 8.5. Dynkin diagrams of rank two root systems

Theorem 8.48. *Let Π be a set of simple roots of a reduced root system R.*

(1) *The Dynkin diagram is connected if and only if R is irreducible.*

(2) *The Dynkin diagram determines the Cartan matrix A.*

(3) *R is determined by the Dynkin diagram uniquely up to an isomorphism: if R, R' are two reduced root systems with the same Dynkin diagram, then they are isomorphic.*

Proof.   (1) Assume that R is reducible; then, by Lemma 8.42, we have $\Pi = Pi_1 \sqcup \Pi_2$, with $\Pi_1 \perp \Pi_2$. Thus, by construction of Dynkin diagram, it will be a disjoint union of the Dynkin diagram of Π_1 and the Dynkin diagram of Π_2. Proof in the oppsoite direction is similar.

(2) Dynkin diagram determines, for each pair of simple roots α_i, α_j, the angle between them and shows which of them is longer. Since all possible configurations of two roots are listed in Theorem 8.9, one easily sees that this information, together with $(\alpha_i, \alpha_j) \leq 0$, uniquely determines $n_{\alpha_i \alpha_j}$, $n_{\alpha_j \alpha_i}$.

(3) By part (2), the Dynkin diagrma determines Π uniquely up to an isomorphism. By Corollary 8.33, Π determines R uniquely up to an isomorphism.

$\square$

Thus, the problem of classifying all irreducible root systems reduces to the following problem: which graphs can appear as Dynkin diagrams of an irreducible root systems? The answer is given by the following theorem.

Theorem 8.49. *Let R be a reduced irreducible root system. Then its Dynkin diagram is isomorphic to one of the diagrams below (in each diagram, the number of vertices is equal to the subscript, so A_n has exactly n vertices):*

- A_n $(n \geq 1)$:
- B_n $(n \geq 2)$:
- C_n $(n \geq 2)$:

- $D_n\ (n \geq 4):$

- $E_6:$

- $E_7:$

- $E_8:$

- $F_4:$
- $G_2:$
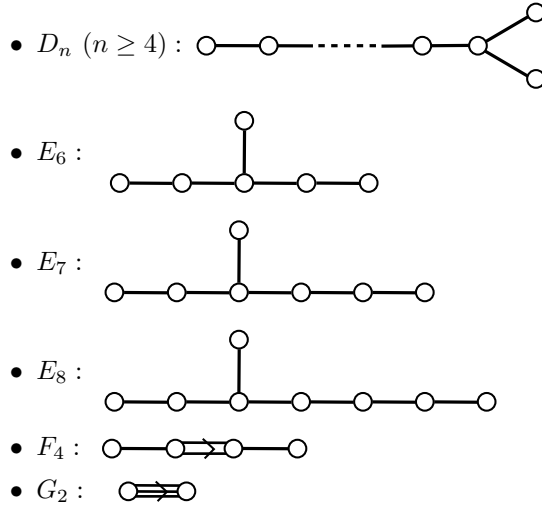
Conversely, each of these diagrams does appear as a Dynkin diagram of a reduced irreducible root system.

The proof of this theorem is rather long. A sketch of proof is given in ??; full proof can be found in [**12**], [**11**], or [**2**] and will not be repeated here.

Explicit construction of the root systems corresponding to each of these diagrams is given in Appendix C, along with useful information such as a description of the Weyl group, and much more.

The letters $A, B, \ldots, G$ do not have any deep meaning: these are just the first 7 letters of the alphabet. However, this notation has become standard. Since the Dynkin diagram determines the root system up to isomorphism, it is also common to use the same notation $A_n, \ldots, G_2$ for the corresponding root system.

Remark 8.50. In the list above, we have imposed restrictions $n \geq 2$ for B_n, C_n and $n \geq 4$ for D_n, which is natural from the point of view of diagrams. In fact, constructions of root systems given in Appendix C also work for smaller values of n. Namely, constructions of root systems B_n, C_n also make sense for $n = 1$; however, these root systems coincide with A_1: $B_1 = C_1 = A_1$, so they do not give any new diagrams (note also that for $n = 2$, $B_2 = C_2$). Similarly, construction of the root system D_n also makes sense for $n = 2, 3$, in which case it gives $D_2 = A_2$, $D_3 = A_3$. Other then the equalities listed above, all root systems $A_n, \ldots, G_2$ are distinct.

Corollary 8.51. *If R is a reduced irreducible root system, then (α, α) can take at most two different values. The number*

$$(8.24) \qquad\qquad m = \frac{\max(\alpha, \alpha)}{\min(\alpha, \alpha)}$$

is equal to the maximal multiplicity of an edge in the Dynkin diagram; thus, $m = 1$ for root systems of types ADE (these are called simply-laced *diagrams), $m = 2$ for types BCF, and $m = 3$ for G_2.*

For non-simply laced systems, the roots with (α, α) being the larger of two possible values are called the long roots, *and the remaining roots are called* short.

8.9. Serre relations and classification of semisimple Lie algebras

We can now return to the question of classification of complex semisimple Lie algebras. Since every semisimple algebra is a direct sum of simple ones, it suffices to classify simple algebras.

According to the results of Section 7.3, every semisimple Lie algebra defines a reduced root system; if the algebra is not simple but only semisimple, then the root system is reducible. The one question we have not yet answered is whether one can go back and recover the Lie algebra from the root system. If the answer is positive, then the isomorphism classes of simple Lie algebras are in bijection with the isomorphism classes of reduced irreducible root systems, and thus we could use classification results of Section 8.8 to classify simple Lie algebras.

Theorem 8.52. *Let $\mathfrak{g}$ be a semisimple complex Lie algebra with root system $R \subset \mathfrak{h}^*$. Let $R = R_+ \sqcup R_-$ be a polarization of R, and $\Pi = \{\alpha_1, \ldots, \alpha_r\}$ — the corresponding system of simple roots.*

(1) *The subspaces*

$$(8.25) \qquad \mathfrak{n}_\pm = \bigoplus_{\alpha \in R_\pm} \mathfrak{g}_\alpha$$

are subalgebras in $\mathfrak{g}$, and

$$(8.26) \qquad \mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$$

as a vector space.

(2) *Let $e_i \in \mathfrak{g}_{\alpha_i}$, $f_i \in \mathfrak{g}_{-\alpha_i}$ be chosen so that $(e_i, f_i) = \frac{2}{(\alpha_i, \alpha_i)}$, and let $h_i = h_{\alpha_i} \in \mathfrak{h}$ be defined by (7.4). Then $e_1, \ldots, e_r$ generate $\mathfrak{n}_+$, $f_1, \ldots, f_r$ generate $\mathfrak{n}_-$, and $h_1, \ldots, h_r$ form a basis of $\mathfrak{h}$. In particular, $\{e_i, f_i, h_i\}_{i=1\ldots r}$ generate $\mathfrak{g}$.*

(3) *The elements e_i, f_i, h_i satisfy the following relations, called the* Serre relations:

$$(8.27) \qquad [h_i, h_j] = 0$$

$$(8.28) \qquad [h_i, e_j] = a_{ij} e_j, \qquad [h_i, f_j] = -a_{ij} f_j$$

$$(8.29) \qquad [e_i, f_j] = \delta_{ij} h_i$$

$$(8.30) \qquad (\operatorname{ad} e_i)^{1-a_{ij}} e_j = 0$$

$$(8.31) \qquad (\operatorname{ad} f_i)^{1-a_{ij}} f_j = 0$$

where $a_{ij} = n_{\alpha_j, \alpha_i} = \langle \alpha_i^\vee, \alpha_j \rangle$ are the entries of the Cartan matrix.

Proof.

(1) The fact that $\mathfrak{n}_+$ is a subalgebra follows from $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$ (see Theorem 7.16) and the fact that sum of positive roots is positive. Equation (8.26) is obvious.

(2) The fact that h_i form a basis of $\mathfrak{h}$ follows from Theorem 8.16. To prove that e_i generate $\mathfrak{n}_+$, we first prove the following lemma.

Lemma 8.53. *Let $R = R_+ \sqcup R_-$ be a reduced root system,, with set of simple roots $\{\alpha_1, \ldots, \alpha_r\}$. Let α be a positive root which is not simple. Then $\alpha = \beta + \alpha_i$ for some positive root β and simple root α_i.*

Proof. Let us consider the inner products (α, α_i). If all of them are non-positive, then, by Lemma 8.17, $\{\alpha, \alpha_1, \ldots, \alpha_r\}$ are linearly independent, which is impossible since $\{\alpha_i\}$ is a basis. Thus, there exists i such that $(\alpha, \alpha_i) > 0$. Thus, $(\alpha, -\alpha_i) < 0$. By Lemma 8.11, this implies that $\beta = \alpha - \alpha_i$ is a root, so $\alpha = \beta + \alpha_i$. We leave it to the reader to check that β must be a positive root. This completes the proof of the lemma. $\square$

By Theorem 7.21, under the assumption of the lemma we have $\mathfrak{g}_\alpha = [\mathfrak{g}_\beta, e_i]$. Using induction in height $\operatorname{ht}(\alpha)$ (see equation (8.8)), it is now easy to show that e_i generate $\mathfrak{n}_+$. Similar argument shows that f_i generate $\mathfrak{n}_-$.

(3) Relations (8.27), (8.28) are an immediate corollary of the definition of Cartan subalgebra and root subspace. Commutation relation $[e_i, f_i] = h_i$ is part of Lemma 7.19 (about $\mathfrak{sl}$-triple determined by a root). Commutation relation $[e_i, f_j] = 0$ for $i \neq j$ follows from the fact that $[e_i, f_j] \in \mathfrak{g}_{\alpha_i - \alpha_j}$. But $\alpha_i - \alpha_j$ is not a root (it can not be a positive root because the coefficient of α_j is negative, and it can not be a negative root because the coefficient of α_i is positive). Thus, $[e_i, f_j] = 0$.

To prove relations (8.31), consider the subspace $\bigoplus \mathfrak{g}_{\alpha_j + k\alpha_i} \subset \mathfrak{g}$ as a module over $\mathfrak{sl}(2, \mathbb{C})$ triple generated by e_i, f_i, h_i. Since $\operatorname{ad} e_i . f_j = 0$, f_j is a highest-weight vector; by (8.29), its weight is equal to $-a_{ij}$. Results of Section 5.1 about representation theory of $\mathfrak{sl}(2, \mathbb{C})$, imply that if v is a vector of weight λ in a finite-dimensional representation, with $e.v = 0$, then $f^{\lambda+1}.v = 0$. Applying it to f_j, we get (8.31). Equality (8.30) is proved similarly.

This completes the proof of Theorem 8.52.

$\square$

A natural question is whether (8.27)–(8.31) is a full set of defining relations for $\mathfrak{g}$. The answer is given by the following theorem.

Theorem 8.54. *Let R be a reduced irreducible root system, with a polarization $R = R_+ \sqcup R_-$ and system of simple roots $\Pi = \{\alpha_1, \ldots, \alpha_r\}$. Let $\mathfrak{g}(R)$ be the complex Lie algebra with generators e_i, f_i, h_i, $i = 1 \ldots, r$ and relations (8.27)-(8.31). Then $\mathfrak{g}$ is a finite-dimensional semisimple Lie algebra with root system R.*

The proof of this theorem is not given here; interested reader can find it in [**21**], [**12**], or [**11**]. We note only that it is highly non-trivial that $\mathfrak{g}(R)$ is finite-dimensional (in fact, this is the key step of the proof), which in turn is based on the use of Weyl group.

Corollary 8.55.

(1) *If $\mathfrak{g}$ is a semisimple Lie algebra with root system R, then there is a natural isomorphism $\mathfrak{g} \simeq \mathfrak{g}(R)$.*

(2) *There is a natural bijection between the set of isomorphism classes of reduced root systems and the set of isomorphism classes of finite-dimensional complex semisimple Lie algebras. The Lie algebra is simple iff the root system is irreducible.*

Combining this corollary with the classification given in Theorem 8.49, we get the following famous result.

Theorem 8.56. *Simple finite-dimensional complex Lie algebras are classified by Dynkin diagrams $A_n \ldots G_2$ listed in Theorem 8.49.*

It is common to refer to the simple Lie algebra corresponding to the Dynkin diagram, say, E_6, as "simple Lie algebra of type E_6".

It is possible to give an explicit construction of the simple Lie algebra corresponding to each of the Dynkin diagrams of Theorem 8.49. For example, Lie algebra of type A_n is nothing but $\mathfrak{sl}(n + 1, \mathbb{C})$. Series B_n, C_n, D_n correspond to classical Lie algebras $\mathfrak{so}$ and $\mathfrak{sp}$. These root systems and Lie algebras are described in detail in Appendix C.

Exercises

8.1. Let $R \subset \mathbb{R}^n$ be given by

$$R = \{\pm e_i, \pm 2e_i \mid 1 \leq i \leq n\} \cup \{\pm e_i \pm e_j \mid 1 \leq i, j \leq n, i \neq j\}$$

where e_i is the standard basis in $\mathbb{R}^n$. Show that R is a non-reduced root system. (This root system is usually denoted BC_n.)

8.2. (1) Let $R \subset E$ be a root system. Show that the set

$$R^\vee = \{\alpha^\vee \mid \alpha \in R\} \subset E^*$$

is also a root system. It is usually called *dual root system* of R.

(2) Let $\Pi = \{\alpha_1, \ldots, \alpha_r\} \subset R$ be the set of simple roots. Show that the set $\Pi^\vee = \{\alpha_1^\vee, \ldots, \alpha_r^\vee\} \subset R^\vee$ is the set of simple roots of $R^\vee$. [Note: this is not completely trivial, as $\alpha \mapsto \alpha^\vee$ is not a linear map. Try using equation (8.17).]

8.3. Prove Lemma 8.17. (Hint: any linear dependence can be written in the form

$$\sum_{i \in I} c_i v_i = \sum_j c_j v_j$$

where $I \cap J = \varnothing$, $c_i, c_j \geq 0$. Show that if one denotes $v = \sum_{i \in I} c_i v_i$, then $(v, v) \leq 0$.)

8.4. Show that $|P/Q| = |\det A|$, where A is the Cartan matrix: $a_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle$.

8.5. Complete the gap in the proof of Theorem 8.37. Namely, assume that $w = s_{i_1} \ldots s_{i_l}$. Let $\beta_k = s_{i_1} \ldots s_{i_{k-1}}(\alpha_{i_k})$. Show that if we have $\beta_k = \pm\beta_j$ for some $j < k$ (thus, the path constructed in the proof of Theorem 8.37 crosses the hyperplane L_{β_j} twice), then $w = s_{i_1} \ldots \widehat{s_{i_j}} \ldots \widehat{s_{i_k}} \ldots s_{i_l}$ (hat means that the corresponding factor should be skipped) and thus, original expression was not reduced.

8.6. Let $w = s_{i_1} \ldots s_{i_l}$ be a reduced expression. Show that then

$$\{\alpha \in R_+ \mid w(\alpha) \in R_-\} = \{\beta_1, \ldots, \beta_l\}$$

where $\beta_k = s_{i_1} \ldots s_{i_{k-1}}(\alpha_{i_k})$ (cf. proof of Lemma 8.31).

8.7. Let $W = S_n$ be the Weyl group of root system A_{n-1}. Show that the longest element $w_0 \in W$ is the permutation $w_0 = (n \ n-1 \ \ldots 1)$.

8.8.

(1) Let R be a reduced root system of rank 2, with simple roots α_1, α_2. Show that the longest element in the corresponding Weyl group is

$$w_0 = s_1 s_2 s_1 \cdots = s_2 s_1 s_2 \ldots \qquad (m \text{ factors in each of the products})$$

where m depends on the angle φ between α_1, α_2: $\varphi = \pi - \frac{\pi}{m}$ (so $m = 2$ for $A_1 \times A_1$, $m = 3$ for A_2, $m = 4$ for B_2, $m = 6$ for G_2). If you can not think of any other proof, give a case-by-case proof.

(2) Show that the following relations hold in W (these are called *Coxeter relations*):

(8.32)
$$s_i^2 = 1$$
$$(s_i s_j)^{m_{ij}} = 1,$$

where m_{ij} is determined by the angle between α_i, α_j in the same way as in the previous part. (It can be shown that the Coxeter relations is a defining set of relations for the Weyl group: W could be defined as the group generated by elements s_i subject to Coxeter relations. A proof of this fact can be found in [**13**] or in [**2**].)

8.9. Let $\varphi \colon R_1 \xrightarrow{\sim} R_2$ be an isomorphism between irreducible root systems. Show that then φ is a composition of an isometry and a scalar operator: $(\varphi(v), \varphi(w)) = c(v, w)$ for any $v, w \in E_1$.

8.10. Let $R \subset E$ be an irreducible root system. Show that then E is an irreducible representation of the Weyl group W.

Chapter 9

Representations of Semisimple Lie Algebras

In this chapter, we study representations of complex semisimple Lie algebras. Recall that by results of Section 6.9, every finite-dimensional representation is completely reducible and thus can be written in the form $V = \bigoplus n_i V_i$, where V_i are irreducible representations, and $n_i \in \mathbb{Z}_+$ are the multiplicities. Thus, the study of representations reduces to classification of irreducible representations and finding a way to determine, for a given representation V, the multiplicities n_i. Both of these questions have a complete answer, which will be given below.

Throughout this section, $\mathfrak{g}$ is a complex finite-dimensional semisimple Lie algebra. We fix a choice of a Cartan subalgebra and thus the root decomposition $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_R \mathfrak{g}_\alpha$ (see Section 7.3).

All representations considered in this section are complex and unless specified otherwise, finite-dimensional.

9.1. Weight decomposition

As in the study of representations of $\mathfrak{sl}(2, \mathbb{C})$ (see Section 5.1), the key to the study of representations of $\mathfrak{g}$ is decomposing the representation into eigenspaces for the Cartan subalgebra.

Definition 9.1. Let V be a representation of $\mathfrak{g}$. A vector $v \in V$ is called a vector of weight $\lambda \in \mathfrak{h}^*$ if, for any $h \in \mathfrak{h}$, one has $hv = \langle \lambda, h \rangle v$. The space of all vectors of weight λ is called the *weight space* and denoted $V[\lambda]$:

$$(9.1) \qquad V[\lambda] = \{v \in V \mid hv = \langle \lambda, h \rangle v \; \forall h \in \mathfrak{h}\}.$$

If $V[\lambda] \neq 0$, then λ is called a weight of V. The set of all weights of V is denoted by $P(V)$:

$$(9.2) \qquad P(V) = \{\lambda \in \mathfrak{h}^* \mid V[\lambda] \neq 0\}$$

Note that it easily follows from standard linear algebra results that vectors of different weights are linearly independent. This, in particular, implies that $P(V)$ is finite for a finite-dimensional representation.

Theorem 9.2. *Every finite-dimensional representation of $\mathfrak{g}$ admits a weight decomposition:*

$$(9.3) \qquad V = \bigoplus_{\lambda \in P(V)} V[\lambda].$$

Moreover, all weights of V are integral: $P(V) \subset P$, where P is the weight lattice defined in Section 8.5.

Proof. Let $\alpha \in R$ be a root. Consider the corresponding $\mathfrak{sl}(2, \mathbb{C})$ subalgebra in $\mathfrak{g}$ generated by $e_\alpha, f_\alpha, h_\alpha$ as in Lemma 7.19. Considering V is a module over this $\mathfrak{sl}(2, \mathbb{C})$ and using results of Section 5.1, we see that h_α is a diagonalizable operator in V. Since elements $h_\alpha, \alpha \in R$ span $\mathfrak{h}$, and sum of commuting diagonalizable operators is diagonalizable, we see that any $h \in \mathfrak{h}$ is diagonalizable. Since $\mathfrak{h}$ is commutative, all of them can be diagonalized simultaneously, which gives the weight decomposition.

Since weights of $\mathfrak{sl}(2, \mathbb{C})$ must be integer, we see that for any weight λ of V, we must have $\langle \lambda, h_\alpha \rangle \in \mathbb{Z}$, which by definition implies that $\lambda \in P$. $\square$

As in the $\mathfrak{sl}(2, \mathbb{C})$ case, this weight decomposition agrees nicely with the root decomposition of $\mathfrak{g}$.

Lemma 9.3. *If $x \in \mathfrak{g}_\alpha$, then $x.V[\lambda] \subset V[\lambda + \alpha]$.*

Proof of this lemma is almost identical to the proof in $\mathfrak{sl}(2, \mathbb{C})$ case (see Lemma 5.2). Details are left to the reader.

9.2. Highest-weight representations and Verma modules

To study irreducible representations, we introduce a class of representatins which are generted by a single vector. As we will later show, all finite-dimensional irreducible representations fall into this class. However, it turns out that to study finite-dimensional representations, we need to consider infinite-dimensional representations as an auxiliary tool. Thus, in the definition below we allow V to be infinite-dimensional.

Definition 9.4. A non-zero representation V (not necessarily finite-dimensional) of $\mathfrak{g}$ is called a *highest weight representation* if it is generated by a vector $v \in V[\lambda]$ such that $e_\alpha.v = 0$ for all $\alpha \in R_+$. In this case, v is called the *highest weight vector*, and λ is the highest weight of V.

The importance of such representations is explained by the following theorem.

Theorem 9.5. *Every irreducible finite-dimensional representation of $\mathfrak{g}$ is a highest-weight representation.*

Proof. Let $P(V)$ be the set of weights of V. Let $\lambda \in P(V)$ be such that for all $\alpha \in R_+$, $\lambda + \alpha \notin P(V)$. Such a λ exists: for example, we can take $h \in \mathfrak{h}$ such that $\langle h, \alpha \rangle 0$ for all $\alpha \in R_+$, and then consider λ such that $\langle h, \lambda \rangle$ is maximal possible.

Now let $v \in V[\lambda]$ be a non-zero vector. Since $\lambda + \alpha \notin P(V)$, we have $e_\alpha v = 0$ for any $\alpha \in R_+$. Consider the subrepresentation $V' \subset V$ generated by v. By definition V' is a highest weight representation. On the other hand, since V is irreducible, one has $V' = V$. $\square$

Note that there can be many non-isomorphic highest weight representations with the same highest weight. However, in any highest weight representation with highest weight vector $v_\lambda \in V[\lambda]$, the following conditions hold:

$$
(9.4) \qquad
\begin{aligned}
h v_\lambda &= \langle h, \lambda \rangle v_\lambda \quad \forall h \in \mathfrak{h} \\
x v_\lambda &= 0 \quad \forall x \in \mathfrak{n}_+ = \bigoplus_{R_+} \mathfrak{g}_\alpha
\end{aligned}
$$

Let us define the universal highest weight representation M_λ as a representation generated by a vector v_λ satisfying conditions (9.4) and no other relations. More formally, define

(9.5) $$M_\lambda = U\mathfrak{g}/I_\lambda$$

where I_λ is the left ideal in $U\mathfrak{g}$ generated by $(h - \langle h, \lambda \rangle), \mathfrak{n}_+$. This module is called *Verma module* and plays an important role in the future.

Remark 9.6. For readers familiar with the notion of induced representation, we note that M_λ can be naturally described as an induced representation:

$$M_\lambda = \mathrm{Ind}_{U\mathfrak{b}}^{U\mathfrak{g}} \mathbb{C}_\lambda = U\mathfrak{g} \otimes_{U\mathfrak{b}} \mathbb{C}_\lambda$$

where $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}_+$ is the Borel subalgebra, and $\mathbb{C}_\lambda$ is one-dimensional representation of $\mathfrak{b}$ defined by equations (9.4).

Example 9.7. Let $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$ and identify $\mathfrak{h}^* \simeq \mathbb{C}$ by $\lambda \mapsto \langle h, \lambda \rangle$. Then Verma module M_λ, $\lambda \in \mathbb{C}$, is the module described in Lemma 5.5.

The following lemma shows that Verma modules are indeed universal in a suitable sense.

Lemma 9.8. *If V is a highest weight module with highest weight λ, then*

$$V \simeq M_\lambda/W$$

for some submodule $W \subset M_\lambda$.

Thus, the study of highest weight representations essentially reduces to the study of submodules in Verma modules.

Theorem 9.9. *Let $\lambda \in \mathfrak{h}_*$ and let M_λ be the Verma module with highest weight λ.*

(1) *Every vector $v \in M_\lambda$ can be uniquely written in the form $v = uv_\lambda$, $u \in U\mathfrak{n}_-$, where $\mathfrak{n}_- = \bigoplus_{R_-} \mathfrak{g}_\alpha$. In other words, the map*

$$U\mathfrak{n}_- \to M_\lambda$$
$$u \mapsto uv_\lambda$$

is an isomorphism of vector spaces.

(2) *M_λ admits a weight decomposition: $M_\lambda = \bigoplus_\mu M_\lambda[\mu]$, with finite-dimensional weight spaces. The set of weights of M_λ is*

$$P(M_\lambda) = \lambda - Q_+, \qquad Q_+ = \left\{ \sum n_i \alpha_i, n_i \in \mathbb{Z}_+ \right\}$$

(3) $\dim M_\lambda[\lambda] = 1$.

Proof. To prove part (1), note that since v_λ generates V, any vector $v \in v_\lambda$ can be written in the form $v = xv_\lambda, x \in U\mathfrak{g}$. But by one of the corollaries of PBW theorem (Corollary 4.63), we have an isomorphism $U\mathfrak{g} = U\mathfrak{n}_- \otimes U\mathfrak{h} \otimes U\mathfrak{n}_+$. Thus, we can write.... $\square$ **?!**

Since every highest-weight representation is a quotient of a Verma module, the above theorem can be generalized to arbitrary highest-weight module.

Theorem 9.10. *Let V be a highest weight representation with highest weight λ (not necessarily finite-dimensional).*

(1) *Every vector $v \in M_\lambda$ can be written in the form $v = uv_\lambda$, $u \in U\mathfrak{n}_-$. In other words, the map*

$$U\mathfrak{n}_- \to M_\lambda$$
$$u \mapsto uv_\lambda$$

is surjective.

(2) *V admits a weight decomposition: $V = \bigoplus_\mu V[\mu]$, with finite-dimensional weight subspaces, and*

(9.6) $$P(V) \subset \lambda - Q_+, \qquad Q_+ = \left\{ \sum n_i \alpha_i, n_i \in \mathbb{Z}_+ \right\}$$

(3) $\dim M_\lambda[\lambda] = 1$.

Proof. Part (1) immediately follows from the similar statement for Verma modules. Part (2) also follows from weight decomposition for Verma modules and the following linear algebra lemma, the proof of which is left as an exercise (see Exercise 9.1).

Lemma 9.11. *Let $\mathfrak{h}$ be a commutative finite-dimensional Lie algebra and M a module over $\mathfrak{h}$ (not necessarily finite-dimensional) which admits weight decomposition with finite-dimensional weight spaces:*

$$M = \bigoplus M[\lambda], \qquad M[\lambda] = \{v \mid hv = \langle h, \lambda \rangle v\}$$

Then any submodule, quotient of M also admits a weight decomposition.

Similarly, to prove part (3), note that $\dim V[\lambda] \le \dim M_\lambda[\lambda] = 1$. On hte other hand, by definition of a highets-weight module, V does have a non-zero highest-weight vectro $v \in V[\lambda]$.

$\square$

Corollary 9.12. *In any highest weight module, there is a unique highest weight and unique up to a scalar highest weight vector.*

Proof. Indeed, if λ, μ are highest weights, then by (9.6), $\lambda - \mu \in Q_+$ and also $\mu - \lambda \in Q_+$, which is impossible unless $\lambda = \mu$. $\square$

9.3. Classification of irreducible finite-dimensional representations

Our next goal is to classify all irreducible finite-dimensional representations. Since, by Theorem 9.5, every such representation is a highest weight representation, this question can be reformulated as follows: classify all highest weight representations which are finite-dimensional and irreducible.

The first step is the following easy result.

Theorem 9.13. *For any $\lambda \in \mathfrak{h}^*$, there exists a unique up to isomorphism irreducible highest weight representation with highest weight λ. This representation is denoted L_λ.*

Proof. All highest weight representations with highest weight λ are of the form M_λ/W for some $W \subset M_\lambda$. It is easy to see that M_λ/W is irreducible iff W is a maximal proper subrepresentation (that is, not properly contained in any other proper subrespresentation). Thus, it suffices to prove that M_λ has a unique proper submodule.

Note that every proper submodule $W \subset M_\lambda$ has the property $W[\lambda] = 0$ (otherwise, we would have $W[\lambda] = M_\lambda[\lambda]$, which would force $W = M_\lambda$). Let J_λ be the sum of all submodules $W \subset M_\lambda$ such that $W[\lambda] = 0$. Then $J_\lambda \subset M_\lambda$ is proper (because $J_\lambda[\lambda] = 0$). Since it contains every other

proper submodule of M_λ, it is the unique maximal proper submodule of M_λ. Thus, $L_\lambda = M_\lambda / J_\lambda$ is the unique irreducible highest-weight module with highest weight λ. $\qquad\square$

Example 9.14. For $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$, results of Section 5.1 show that if $\lambda \in \mathbb{Z}_+$, then $L_\lambda = V_\lambda$ is the finite-dimensional irreducible module of dimension $\lambda + 1$, and $L_\lambda = M_\lambda$ for $\lambda \notin \mathbb{Z}_+$.

As we will later see, the situation is similar for other Lie algebras. In particular, for "generic" λ, Verma module is irreducible, so $M_\lambda = L_\lambda$.

Corollary 9.15. *Every irreducible finite-dimensionl representation V must be isomorphic to one of L_λ.*

Thus, to classify all irreducible finite-dimensional representations of $\mathfrak{g}$, we need to find which of L_λ are finite-dimensional.

To give an answer to this question, we need to introduce some notation. Recall the weight lattice $P \subset \mathfrak{h}^*$ defined by (8.12).

Definition 9.16. A weight $\lambda \in \mathfrak{h}^*$ is called *dominant integral* the following condition holds

$$(9.7) \qquad\qquad \langle \lambda, \alpha^\vee \rangle \in \mathbb{Z}_+ \qquad \text{for all } \alpha \in R_+.$$

The set of all integral dominant weights is denoted by P_+.

Note that it is immediate form the definitions that $P_+ \subset P$. It also follows from results of Exercise 8.2 that condition (9.7) is equivalent to

$$(9.8) \qquad\qquad \langle \lambda, \alpha_i^\vee \rangle \in \mathbb{Z}_+ \qquad \text{for all } \alpha_i \in \Pi.$$

Theorem 9.17. *Irreducible highest-weight representation L_λ is finite-dimensional iff $\lambda \in P_+$.*

Before proving this theorem, note that it immediately implies the following corollary.

Corollary 9.18. *For every $\lambda \in P_+$, representation L_λ is an irreducible finite-dimensional representation. These representations are pairwise non-isomorphic, and every irreducible finite-dimensional representation is isomorphic to one of them.*

Proof. First, let us prove that if L_λ is finite-dimensional, then $\lambda \in P_+$. Indeed, let α_i be a simple root and let $\mathfrak{sl}(2, \mathbb{C})_i$ be the subalgebra in $\mathfrak{g}$ generated by $e_{\alpha_i}, h_{\alpha_i}, f_{\alpha_i}$ (see). Consider L_λ as $\mathfrak{sl}(2, \mathbb{C})_i$ module. Then v_λ satisfies relations $e_i v_\lambda = 0, h_i v_\lambda = \langle h_i, \lambda \rangle = \langle \alpha_i^\vee, \lambda \rangle$. It generates a highest weight $\mathfrak{sl}(2, \mathbb{C})_i$ submodule, which is finite-dimensional (since L_λ is finite-dimensional). By classification theorem for irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$ (Theorem 5.6), this implies that the highest weight $\langle h_i, \lambda \rangle \in \mathbb{Z}_+$. Repeating the argument for each simple root, we see that $\lambda \in P_+$.

Now let us prove that if $\lambda \in P_+$, then L_λ is finite-dimensional. This is a more difficult result; we break the proof in several steps.

Step 1. Let $n_i = \langle \alpha_i^\vee, \lambda \rangle \in \mathbb{Z}_+$. Consider the vector

$$(9.9) \qquad\qquad v_{s_i.\lambda} = f_i^{n_i+1} v_\lambda \in M_\lambda[s_i.\lambda],$$

where $v_\lambda \in M_\lambda$ is the highest-weight vector and

$$(9.10) \qquad\qquad s_i.\lambda = \lambda - (n_i + 1)\alpha_i.$$

(see ?? for more general definition). Then we have

$$e_j v_{s_i.\lambda} = 0 \qquad \text{for all } i, j$$

Indeed, for $i \neq j$ we have $[e_j, f_i] = 0$ (see equation (8.29)), so $e_j f_i^{n_i+1} v_\lambda = f_i^{n_i+1} e_j v_\lambda = 0$. For $i = j$, this follows from the results of Section 5.1: if v is a vector of weight n in a representation of $\mathfrak{sl}(2, \mathbb{C})$ such that $ev = 0$, then $e f^{n+1} v = 0$.

Step 2. Let $M_i \subset M_\lambda$ be the subrepresentation generated by vector $v_{s_i\lambda}$. By (9.10), M_i is a highest weight representation. In particular, all weights of M_i are of the form $s_i.\lambda - Q_+ = \lambda - (n_i + 1)\alpha_i - Q_+$. Thus, λ is not a weight of M_i; therefore, each M_i is a proper submodule in M_λ.

Consider now the quotient

$$(9.11) \qquad\qquad\qquad \tilde{L}_\lambda = M_\lambda / \sum M_i.$$

Since each M_i is a proper subrepresentation, so is $\sum M_I$ (see the proof of Theorem 9.13); thus, $\tilde{L}_\lambda$ is a non-zero highest-weight module.

Step 3. The key step of the proof is the following theorem.

Theorem 9.19. *Let $\lambda \in P_+$, and let $\tilde{L}_\lambda$ be defined by (9.11). Then L_λ is finite-dimensional.*

The proof of this theorem is rather long. It is given in a separate section (Section 9.7) at the end of this chapter.

Now we can complete the proof of Theorem 9.17. Since L_λ is the quotient of M_λ by the maximal proper ideal, we see that L_λ is a quotient of $\tilde{L}_\lambda$. Since $\tilde{L}_\lambda$ is finite-dimensional, so is L_λ. This completes the proof. □

Note that this theorem does not exactly give an explicit construction of the modules L_λ. To some extent, this will be adddressed in the next section, where we study the structure of the modules $M_i \subset M_\lambda$ and L_λ in more detail. However, for all classical algebras there also exist very explicit constructions of the irreducible finite-dimensional representations which are usually not based on realizing L_λ as quotients of Verma modules. We will give an example of this for $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$ in Section 9.6.

9.4. Bernstein–Gelfand–Gelfand resolution

In the previous section, we have shown that for $\lambda \in P_+$, the irreducible highest-weight mdoule L_λ is irreducible. Our next goal is to study the structure of these representations.

Recall that the proof of Theorem 9.17 defined, for each $\lambda \in P_+$, a collection of submodules $M_i \subset M_\lambda$. We have shown that each of them is a highest weight module. In fact, we can make a more precise statement.

Lemma 9.20. *Let $v \in M_\lambda[\mu]$ be a vector such that $e_\alpha v = 0$ for all $\alpha \in R_+$, and let $M' \subset M$ be the submodule generated by v. Then M' is a Verma module with highest weight μ.*

Proof. Since $e_\alpha v = 0$, by definition M' is a highest weight module with highest weight μ and thus is isomorphic to a quotient of the Verma module M_μ. To show that $M' = M_\mu$, it suffices to show that the map $U\mathfrak{n}_- \to M' : u \mapsto uv$ is injective.

Indeed, assume that $uv = 0$ for some $u \in U\mathfrak{n}_-$. On the other hand, by Theorem 9.9, we can write $v = u'v_\lambda$ for some $u' \in U\mathfrak{n}_-$. Thus, we get $uu'v_\lambda = 0$. By Theorem 9.9, this implies that $uu' = 0$ as eleemnt of $U\mathfrak{n}_-$. But by Corollary 4.64, $U\mathfrak{n}_-$ has no zero divisors. □

9.5. Characters and Weyl character formula

9.6. Representations of $\mathfrak{sl}(n, \mathbb{C})$

9.7. Proof of Theorem 9.19

Exercises

9.1. Prove Lemma 9.11. You can do it by breaking it into several steps as shown below.
(1) Show that given any finite set of distinct weights $\lambda_1, \ldots, \lambda_n \in P(V)$, there exists an element $p \in U\mathfrak{h}$ such that $p(\lambda_1) = 1, p(\lambda_i) = 0$ for $i \neq 1$ (considering elements of $U\mathfrak{h} = S\mathfrak{h}$ as polynomial functions on $\mathfrak{h}^*$).
(2) Let $V \subset M$ be an $\mathfrak{h}$-submodule, and $v \in V$. Write $v = \sum v_i, v_i \in M[\lambda_i]$. Show that then each of $v_i \in V$.
(3) Deduce Lemma 9.11.

Appendix A

Manifolds and immersions

Appendix B

Jordan Decomposition

In this section, we give the proof of the Jordan decomposition for linear operators, which was used in Section 6.6, and several related results.

Throughout this appendix, V is a finite-dimensional complex vector space. Recall that an operator $A\colon V \to V$ is called nilpotent if $A^n = 0$ for sufficiently large n. We will call A *semisimple* if any A-invariant subspace has an A-invariant complement: if $W \subset V$, $AW \subset W$, then there exists $W' \subset V$ such that $V = W \oplus W'$, $AW' \subset W'$. It is easy to see that for operators in a finite-dimensional complex vector space, this is equivalent to requiring that A be diagonalizable.

Lemma B.1.

    (1) *Sum of two commuting semisimple operators is semisimple. Sum of two commuting nilpotent operators is nilpotent.*

    (2) *Let $A\colon V \to V$ be semisimple, and $W \subset V$ stable under A: $AW \subset W$. Then restrictions of A to W and to V/W are semisimple operators.*

Proof. Replacing in the first part the "semisimple" by "diagonalizable", we arrive at a well-known result of linear algebra. The second follows from the following result: A is semisimple iff there exists a polynomial p without multiple roots such that $p(A) = 0$. $\qquad\square$

Theorem B.2. *Any linear operator A can be uniquely written as a sum of commuting semisimple and nilpotent ones:*

(B.1) $A = A_s + A_n, \quad A_s A_n = A_n A_s, \quad A_n$ *nilpotent,* $\quad A_s$ *semisimple*

Moreover, A_s, A_n can be written as polynomials of A: $A_s = p(A)$, $A_n = A - p(A)$ for some polynomial $p \in \mathbb{C}[t]$ depending on A.

Decomposition (B.1) is called the Jordan decomposition of A.

Proof. It is well-known from linear algebra that one can decompose V in the direct sum of generalized eigenvalues: $V = \bigoplus V_{(\lambda)}$, where λ runs over the set of distinct eigenvalues of A and $V_{(\lambda)}$ is generalized eigenspace with eigenvalue λ: restriction of $A - \lambda\,\mathrm{id}$ to $V_{(\lambda)}$ is nilpotent.

Define A_s by $A_s|_{V_{(\lambda)}} = \lambda\,\mathrm{id}$, and A_n by $A_n = A - A_s$. Then it is immediate from definition that A_s is semisimple and A_n is nilpotent. It is also easy to see that they commute: in fact, A_s commutes with any operator $V_{(\lambda)} \to V_{(\lambda)}$. This shows existence of Jordan decomposition.

Let us also show that so defined A_s, A_n can be written as polynomials in A. Indeed, let $p \in \mathbb{C}[t]$ be defined by system of congruences

$$p(t) \equiv \lambda_i \mod (t - \lambda_i)^{n_i}$$

where λ_i are distinct eigenvalues of A and $n_i = \dim V_{(\lambda_i)}$. By Chinese remainder theorem, such a polynomial exists. Since $(A - \lambda_i)^{n_i} = 0$ on $V_{(\lambda_i)}$, we see that $p(A)|_{V_{(\lambda)}} = \lambda = A_s|_{V_{(\lambda)}}$. Thus, $A_s = p(A)$.

Finally, let us prove uniqueness. Let A_s, A_n be as defined above. Assume that $A = A'_s + A'_n$ be another Jordan decomposition. Then $A_s + A_n = A'_s + A'_n$. Since A'_s, A'_n commute with each other, they commute with A; since $A_s = p(A)$, we see that A_s, A_n commute with A'_s, A'_n. Consider now the operator $A_s - A'_s = A'_n - A_n$. On one hand, it is semisimple as difference of two commuting semisimple operators. on the other hand, it is nilpotent as difference of two commuting nilpotent operators (see Lemma B.1). Thus, all its eigenvalues are zero; since it is semisimple, it shows that it is a zero operator, so $A_s = A'_s, A_n = A'_n$. $\qquad\square$

The proof also shows that it is possible to choose a basis in V such that in this basis, A_s is diagonal and A_n is strictly upper-triangular, and that if 0 is an eigenvalue of A, then the polynomial p can be chosen so that $p(0) = 0$.

Theorem B.3. *Let A be an operator $V \to V$. Define $\operatorname{ad} A \colon \operatorname{End}(V) \to \operatorname{End}(V)$ by $\operatorname{ad} A.b = AB - BA$. Then $(\operatorname{ad} A)_s = \operatorname{ad} A_s$, and $\operatorname{ad} A_s$ can be written in the form $\operatorname{ad} A_s = Q(\operatorname{ad} A)$ for some polynomial $Q \in \mathbb{C}[t]$ such that $Q(0) = 0$.*

Proof. Let $A = A_s + A_n$ be the Jordan decomposition for A. Then $\operatorname{ad} A = \operatorname{ad} A_s + \operatorname{ad} A_n$, and it is immediate to check that $\operatorname{ad} A_s$, $\operatorname{ad} A_n$ commute.

Choose a basis in V such that in this basis, A_s is diagonal, A_n is strictly upper-triangular. Then it also gives a basis of matrix units E_{ij} in $\operatorname{End}(V)$. In this basis, action of $\operatorname{ad} A_s$ is diagonal: $\operatorname{ad} A_s.E_{ij} = (\lambda_i - \lambda_j)E_{ij}$, as is easily verified by direct computation. Using this basis, it is also easy to check that $\operatorname{ad} A_n$ is nilpotent (see Exercise B.1). Thus, $\operatorname{ad} A = \operatorname{ad} A_s + \operatorname{ad} A_n$ is the Jordan decomposition for $\operatorname{ad} A$, so $(\operatorname{ad} A)_s = \operatorname{ad} A_s$.

By Theorem B.2 applied to operator $\operatorname{ad} A$, we see that $(\operatorname{ad} A)_s$ can be written in the form $(\operatorname{ad} A)_s = Q(\operatorname{ad} A)$ for some polynomial $Q \in \mathbb{C}[t]$; moreover, since 0 is an eigenvalue of $\operatorname{ad} A$ (e.g., $\operatorname{ad} A.A = 0$), we see that $Q(0) = 0$. $\qquad\square$

Theorem B.4. *Let A be an operator $V \to V$. Define $\overline{A}_s$ to be the operator which has the same eigenspaces as A_s but complex conjugate eigenvalues: if $A_s v = \lambda v$, then $\overline{A}_s v = \bar{\lambda} v$. Then $\operatorname{ad} \overline{A}_s$ can be written in the form $\operatorname{ad} \overline{A}_s = Q(\operatorname{ad} A)$ for some polynomial $Q \in t\mathbb{C}[t]$ (depending on A).*

Proof. Let $\{v_i\}$ be a basis of eigenvectors for A_s: $A_s v_i = \lambda_i v_i$ so that $\overline{A}_s v_i = \bar{\lambda}_i v_i$. Let E_{ij} be the corresponding basis in $\operatorname{End}(V)$; then, as discussed in the proof of Theorem B.3, in this basis $\operatorname{ad} A_s$ is given by $\operatorname{ad} A_s.E_{ij} = (\lambda_i - \lambda_j)E_{ij}$, and $\operatorname{ad} \overline{A}_s.E_{ij} = (\bar{\lambda}_i - \bar{\lambda}_j)E_{ij}$.

Choose a polynomial $f \in \mathbb{C}[t]$ such that $f(\lambda_i - \lambda_j) = \bar{\lambda}_i - \bar{\lambda}_j$ (in particular, $f(0) = 0$); such a polynomial exists by interpolation theorem. Then $\operatorname{ad} \overline{A}_s = f(\operatorname{ad} A_s) = f(Q(\operatorname{ad} A))$. $\qquad\square$

Exercises

B.1. Let $A \colon V \to V$ be an upper-triangular operator. Let $F^k \subset \operatorname{End}(V)$ be the subspace spanned by matrix units E_{ij} with $i - j = k$. Show that then $\operatorname{ad} A.F^k \subset F^{k-1}$ and thus, $\operatorname{ad} A \colon \operatorname{End}(V) \to \operatorname{End}(V)$ is nilpotent.

Appendix C

Root Systems and Simple Lie Algebras

In this appendix, for each of the Dynkin diagrams of types $A_n, \ldots D_n$, we give an explicit description of the corresponding root system and simple Lie algebra, along with some relevant information such as the description of the Weyl group. This section contains no proofs; we refer the reader to ?? for proofs and descriptions of exceptional root systems $E_6, \ldots, G_2$.

In this appendix, we use the following notation:

$\mathfrak{g}$ — a complex simple Lie algebra, with fixed Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$

$R \subset \mathfrak{h}^*$ — the root system

$E = \mathfrak{h}^*_{\mathbb{R}}$ — the real vector space spanned by roots

$(\,,\,)$ — the symmetric invariant bilinear form on $\mathfrak{h}^*$ normalized so that $(\alpha, \alpha) = 2$ for long roots

R_+ — set of positive roots (of course, there are many ways to choose it; we will only give the most common and widely used choice)

$\Pi = \{\alpha_1, \ldots, \alpha_r\}$, $r = \mathrm{rank}(R)$ — set of simple roots (see Definition 8.12)

W — the Weyl group (see Section 8.2)

$P \subset E$ — the weight lattice (see Section 8.5)

$Q \subset E$ — the root lattice (see Section 8.5)

θ — highest root

$\rho = \frac{1}{2} \sum_{R_+} \alpha$ (see (8.22))

$h = \mathrm{ht}(\theta) + 1$, $h^\vee = \mathrm{ht}(\theta^\vee) + 1 = \langle \rho, \theta^\vee \rangle + 1$ — Coxeter number and dual Coxeter number

C.1. $A_n = \mathfrak{sl}(n+1, \mathbb{C})$

Lie algebra: $\mathfrak{g} = \mathfrak{sl}(n+1, \mathbb{C})$, with Cartan subalgebra $\mathfrak{h} = \{\text{diagonal matrices with trace } 0\}$. We denote by $e_i \in \mathfrak{h}^*$ the functional defined by

$$e_i \colon \begin{bmatrix} h_1 & 0 \ldots & 0 \\ & \ldots & \\ 0 & \ldots & h_{n+1} \end{bmatrix} \mapsto h_i$$

Then $\mathfrak{h}^* = \bigoplus \mathbb{C}e_i / \mathbb{C}(e_1 + \cdots + e_{n+1})$, and

$$E = \mathfrak{h}^*_{\mathbb{R}} = \bigoplus \mathbb{R}e_i / \mathbb{R}(e_1 + \cdots + e_{n+1})$$

with the inner product defined by defined by $(\lambda, \mu) = \sum \lambda_i \mu_i$ if representatives λ, μ are chosen so that $\sum \lambda_i = \sum \mu_i = 0$.

Root system: $R = \{e_i - e_j \mid i \neq j\}$

Root subspace corresponding to root $\alpha = e_i - e_j$ is $\mathfrak{g}_\alpha = \mathbb{C}E_{ij}$, and the corresponding coroot $h_\alpha = \alpha^\vee \in \mathfrak{h}$ is $h_\alpha = E_{ii} - E_{jj}$.

Positive and simple roots: $R_+ = \{e_i - e_j \mid i < j\}$, $|R_+| = \frac{n(n+1)}{2}$

$\Pi = \{\alpha_1, \ldots, \alpha_n\}$, $\alpha_i = e_i - e_{i+1}$.

Dynkin diagram: ○—○— · · · · · · · · —○—○

Cartan matrix:

$$A = \begin{bmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}$$

Weyl group: $W = S_{n+1}$, acting on E by permutations. Simple reflections are $s_i = (i\ i+1)$.

Weight and root lattices:

$P = \{(\lambda_1, \ldots, \lambda_{n+1}) \mid \lambda_i - \lambda_j \in \mathbb{Z}\} / \mathbb{R}(1, \ldots, 1) =$
$\ = \{(\lambda_1, \ldots, \lambda_n, 0) \mid \lambda_i \in \mathbb{Z}\}$.
$Q = \{(\lambda_1, \ldots, \lambda_{n+1}) \mid \lambda_i \in \mathbb{Z}, \sum \lambda_i = 0\}$
$P/Q \simeq \mathbb{Z}/(n+1)\mathbb{Z}$

Dominant weights and positive Weyl chamber:

$C_+ = \{\lambda_1, \ldots, \lambda_{n+1}) \mid \lambda_1 > \lambda_2 > \cdots > \lambda_{n+1}\} / \mathbb{R}(1, \ldots, 1)$
$\ = \{\lambda_1, \ldots, \lambda_n, 0) \mid \lambda_1 > \lambda_2 > \cdots > \lambda_n > 0\}$.
$P_+ = \{(\lambda_1, \ldots, \lambda_{n+1}) \mid \lambda_i - \lambda_{i+1} \in \mathbb{Z}_+\} / \mathbb{R}(1, \ldots, 1)$
$\ = \{(\lambda_1, \ldots, \lambda_n, 0) \mid \lambda_i \in \mathbb{Z}, \lambda_1 \geq \lambda_2 \cdots \geq \lambda_n \geq 0\}$.

Maximal root, ρ, and the Coxeter number:

$\theta = e_1 - e_{n+1} = (1, 0, \ldots, 0, -1)$
$\rho = (n, n-1, \ldots, 1, 0) = \left(\frac{n}{2}, \frac{n-2}{2}, \ldots, \frac{-n}{2}\right)$
$h = h^\vee = n+1$

C.2. $B_n = \mathfrak{so}(2n+1, \mathbb{C})$

Lie algebra: $\mathfrak{g} = \mathfrak{so}(2n+1, \mathbb{C})$, with Cartan subalgebra consisiting of block-diagonal matrices

$$\mathfrak{h} = \left\{ \begin{bmatrix} A_1 & & & & \\ & A_2 & & & \\ & & \ddots & & \\ & & & A_n & \\ & & & & 0 \end{bmatrix} \right\}, \quad A_i = \begin{bmatrix} 0 & a_i \\ -a_i & 0 \end{bmatrix}$$

Lie algebra (alternative description): $\mathfrak{g} = \mathfrak{so}(B) = \{a \in \mathfrak{gl}(2n+1, \mathbb{C}) \mid a + Ba^t B^{-1} = 0\}$,
where B is the symmetric non-degenerate bilinear form on $\mathbb{C}^{2n+1}$ with the matrix

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

This Lie algebra is isomorphic to the usual $\mathfrak{so}(2n+1, \mathbb{C})$; the isomorphism is given by $a \mapsto aB$.

In this description, the Cartan subalgebra is

$$\mathfrak{h} = \mathfrak{g} \cap \{\text{diagonal matrices}\} = \{\text{diag}(x_1, \ldots, x_n, -x_1, \ldots, -x_n, 0)\}$$

Define $e_i \in \mathfrak{h}^*$ by

$$e_i \colon \text{diag}(x_1, \ldots, x_n, -x_1, \ldots, -x_n, 0) \mapsto x_i.$$

Then $e_i, i = 1 \ldots n$ form a basis in $\mathfrak{h}^*$. The bilinear form is defined by $(e_i, e_j) = \delta_{ij}$.

Root system: $R = \{\pm e_i \pm e_j \ (i \neq j), \pm e_i\}$ (signs are chosen independently)

The corresponding root subspaces and coroots in $\mathfrak{g}$ (using the alternative description) are given by

- For $\alpha = e_i - e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{ij} - E_{j+n,i+n})$, $h_\alpha = H_i - H_j$.
- For $\alpha = e_i + e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{i,j+n} - E_{j,i+n})$, $h_\alpha = H_i + H_j$.
- For $\alpha = -e_i - e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{i+n,j} - E_{j+n,i})$, $h_\alpha = -H_i - H_j$.
- For $\alpha = e_i$, $\mathfrak{g}_\alpha = \mathbb{C}(E_{i,2n+1} - E_{2n+1,n+i})$, $h_\alpha = 2H_i$.
- For $\alpha = -e_i$, $\mathfrak{g}_\alpha = \mathbb{C}(E_{n+i,2n+1} - E_{2n+1,i})$, $h_\alpha = -2H_i$

where $H_i = E_{ii} - E_{i+n,i+n}$.

Positive and simple roots: $R_+ = \{e_i \pm e_j \ (i < j), e_i\}$, $|R_+| = n^2$

$\Pi = \{\alpha_1, \ldots, \alpha_n\}$, $\alpha_1 = e_1 - e_2, \ldots, \alpha_{n-1} = e_{n-1} - e_n, \alpha_n = e_n$.

Dynkin diagram:

Cartan matrix:

$$A = \begin{bmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -2 & 2 \end{bmatrix}$$

Weyl group: $W = S_n \ltimes (\mathbb{Z}_2)^n$, acting on E by permutations and sign changes of coordinates.
Simple reflections are $s_i = (i \ i+1) \ (i = 1 \ldots n-1)$, $s_n \colon (\lambda_1, \ldots, \lambda_n) \mapsto (\lambda_1, \ldots, -\lambda_n)$.

Weight and root lattices:
$P = \{(\lambda_1, \ldots, \lambda_n) \mid \lambda_i \in \frac{1}{2}\mathbb{Z}, \lambda_i - \lambda_j \in \mathbb{Z}\}$
$Q = \mathbb{Z}^n$

$P/Q \simeq \mathbb{Z}_2$

Dominant weights and positive Weyl chamber:

$C_+ = \{\lambda_1, \ldots, \lambda_n) \mid \lambda_1 > \lambda_2 > \cdots > \lambda_n > 0\}.$

$P_+ = \{(\lambda_1, \ldots, \lambda_n) \mid \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0, \lambda_i \in \frac{1}{2}\mathbb{Z}, \lambda_i - \lambda_j \in \mathbb{Z}\}.$

Maximal root, ρ, and the Coxeter number:

$\theta = e_1 + e_2 = (1, 1, 0, \ldots, 0)$

$\rho = (n - \frac{1}{2}, n - \frac{3}{2}, \ldots, \frac{1}{2})$

$h = 2n,\ h^\vee = 2n - 1$

C.3. $C_n = \mathfrak{sp}(n, \mathbb{C})$

Lie algebra: $\mathfrak{g} = \mathfrak{sp}(n, \mathbb{C}) = \{a \in \mathfrak{gl}(2n, \mathbb{C}) \mid a + Ja^t J^{-1} = 0\}$, where J is the skew-symmetric nondegenerate matrix

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

The Cartan subalgebra is given by

$$\mathfrak{h} = \mathfrak{g} \cap \{\text{diagonal matrices}\} = \{\operatorname{diag}(x_1, \ldots, x_n, -x_1, \ldots, -x_n)\}$$

Define $e_i \in \mathfrak{h}^*$ by

$$e_i \colon \operatorname{diag}(x_1, \ldots, x_n, -x_1, \ldots, -x_n) \mapsto x_i.$$

Then $e_i, i = 1 \ldots n$, form a basis in $\mathfrak{h}^*$. The bilinear form is defined by $(e_i, e_j) = \frac{1}{2}\delta_{ij}$.

Root system: $R = \{\pm e_i \pm e_j \ (i \neq j), \pm 2e_i\}$ (signs are chosen independently)
 The corresponding root subspaces and coroots are given by
 - For $\alpha = e_i - e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{ij} - E_{j+n,i+n})$, $h_\alpha = H_i - H_j$.
 - For $\alpha = e_i + e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{i,j+n} + E_{j,i+n})$, $h_\alpha = H_i + H_j$.
 - For $\alpha = -e_i - e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{i+n,j} + E_{j+n,i})$, $h_\alpha = -H_i - H_j$.
 - For $\alpha = 2e_i$, $\mathfrak{g}_\alpha = \mathbb{C}E_{i,i+n}$, $h_\alpha = H_i$
 - For $\alpha = -2e_i$, $\mathfrak{g}_\alpha = \mathbb{C}E_{i+n,i}$, $h_\alpha = -H_i$
 where $H_i = E_{ii} - E_{i+n,i+n}$.

Positive and simple roots: $R_+ = \{e_i \pm e_j \ (i < j), 2e_i\}$, $|R_+| = n^2$
 $\Pi = \{\alpha_1, \ldots, \alpha_n\}$, $\alpha_1 = e_1 - e_2, \ldots, \alpha_{n-1} = e_{n-1} - e_n, \alpha_n = 2e_n$.

Dynkin diagram: ○—○— · · · · · —○—○⇐○

Cartan matrix:

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 & -1 \\ & -1 & 2 & -1 \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 2 & -2 \\ & & & & -1 & 2 \end{bmatrix}$$

Weyl group: $W = S_n \ltimes (\mathbb{Z}_2)^n$, acting on E by permutations and sign changes of coordinates.
 Simple reflections are $s_i = (i \ i+1) \ (i = 1 \ldots n-1)$, $s_n \colon (\lambda_1, \ldots, \lambda_n) \mapsto (\lambda_1, \ldots, -\lambda_n)$.

Weight and root lattices: $P = \mathbb{Z}^n$
 $Q = \{(\lambda_1, \ldots, \lambda_n) \mid \lambda_i \in \mathbb{Z}, \sum \lambda_i \in 2\mathbb{Z}\}$
 $P/Q \simeq \mathbb{Z}_2$

Dominant weights and positive Weyl chamber:
 $C_+ = \{\lambda_1, \ldots, \lambda_n) \mid \lambda_1 > \lambda_2 > \cdots > \lambda_n > 0\}$.
 $P_+ = \{(\lambda_1, \ldots, \lambda_n) \mid \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0, \lambda_i \in \mathbb{Z}\}$.

Maximal root, ρ, and the Coxeter number:
 $\theta = 2e_1 = (2, 0, \ldots, 0)$
 $\rho = (n, n-1, \ldots, 1)$
 $h = 2n, \ h^\vee = n + 1$

C.4. $D_n = \mathfrak{so}(2n, \mathbb{C})$

Lie algebra: $\mathfrak{g} = \mathfrak{so}(2n, \mathbb{C})$, with Cartan subalgebra consisting of block-diagonal matrices

$$\mathfrak{h} = \left\{ \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{bmatrix} \right\}, \quad A_i = \begin{bmatrix} 0 & h_i \\ -h_i & 0 \end{bmatrix}$$

Lie algebra (alternative description): $\mathfrak{g} = \mathfrak{so}(B) = \{a \in \mathfrak{gl}(2n, \mathbb{C}) \mid a + Ba^t B^{-1} = 0\}$, where B is the symmetric non-degenerate bilinear form on $\mathbb{C}^{2n}$ with the matrix

$$B = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}$$

This Lie algebra is isomorphic to the usual $\mathfrak{so}(2n + 1, \mathbb{C})$; the isomorphism is given by $a \mapsto aB$.

In this description, the Cartan subalgebra is

$$\mathfrak{h} = \mathfrak{g} \cap \{\text{diagonal matrices}\} = \{\text{diag}(x_1, \ldots, x_n, -x_1, \ldots, -x_n)\}$$

Define $e_i \in \mathfrak{h}^*$ by

$$e_i \colon \text{diag}(x_1, \ldots, x_n, -x_1, \ldots, -x_n) \mapsto x_i.$$

Then $e_i, i = 1 \ldots n$ form a basis in $\mathfrak{h}^*$. The bilinear form is given by $(e_i, e_j) = \delta_{ij}$.

Root system: $R = \{\pm e_i \pm e_j \ (i \neq j)\}$ (signs are chosen independently)
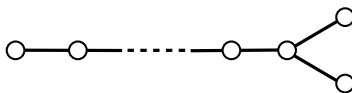
The corresponding root subspaces and coroots in $\mathfrak{g}$ (using the alternative description) are given by

- For $\alpha = e_i - e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{ij} - E_{j+n, i+n})$, $h_\alpha = H_i - H_j$.
- For $\alpha = e_i + e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{i, j+n} - E_{j, i+n})$, $h_\alpha = H_i + H_j$.
- For $\alpha = -e_i - e_j$: $\mathfrak{g}_\alpha = \mathbb{C}(E_{i+n, j} - E_{j+n, i})$, $h_\alpha = -H_i - H_j$

where $H_i = E_{ii} - E_{i+n, i+n}$.

Positive and simple roots: $R_+ = \{e_i \pm e_j \ (i < j)\}$, $|R_+| = n(n-1)$

$\Pi = \{\alpha_1, \ldots, \alpha_n\}$, $\alpha_1 = e_1 - e_2, \ldots, \alpha_{n-1} = e_{n-1} - e_n, \alpha_n = e_{n-1} + e_n$.

Dynkin diagram:

Cartan matrix:

$$A = \begin{bmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 & -1 \\ & & & & -1 & 2 & \\ & & & & -1 & & 2 \end{bmatrix}$$

Weyl group: $W = \{\text{permutations and even number of sign changes}\}$. Simple reflections are $s_i = (i \ i+1)$, $i = 1 \ldots n-1$, $s_n \colon (\lambda_1, \ldots, \lambda_{n-1}, \lambda_n) \mapsto (\lambda_1, \ldots, -\lambda_n, -\lambda_{n-1})$.

Weight and root lattices:

$P = \{(\lambda_1, \ldots, \lambda_n) \mid \lambda_i \in \frac{1}{2}\mathbb{Z}, \lambda_i - \lambda_j \in \mathbb{Z}\}$

$Q = \{(\lambda_1, \ldots, \lambda_n) \mid \lambda_i \in \mathbb{Z}, \sum \lambda_i \in 2\mathbb{Z}\}$

$P/Q \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$ for even n, $P/Q \simeq \mathbb{Z}_4$ for odd n

Dominant weights and positive Weyl chamber:

$$C_+ = \{(\lambda_1, \ldots, \lambda_n) \mid \lambda_1 > \lambda_2 > \cdots > \lambda_n, \, \lambda_{n-1} + \lambda_n > 0\}.$$
$$P_+ = \{(\lambda_1, \ldots, \lambda_n) \mid \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n, \lambda_{n-1} + \lambda_n \geq 0, \lambda_i \in \tfrac{1}{2}\mathbb{Z}, \lambda_i - \lambda_j \in \mathbb{Z}\}.$$

Maximal root, ρ, and the Coxeter number:

$$\theta = e_1 + e_2 = (1, 1, 0, \ldots, 0)$$
$$\rho = (n - 1, n - 2, \ldots, 0)$$
$$h = h^\vee = 2n - 2$$

List of Notation

$\mathbb{K}$ — either $\mathbb{R}$ or $\mathbb{C}$. This notation is used when a result holds for both $\mathbb{R}$ and $\mathbb{C}$.

Linear algebra

V^* — dual vector space

$\langle\,,\,\rangle\colon V \otimes V^* \to \mathbb{C}$ — canonical pairing of V with V^*.

$\mathrm{Hom}(V, W)$ — space of linear maps $V \to W$

$\mathrm{End}(V) = \mathrm{Hom}(V, V)$ — space of linear maps $V \to V$ considered as an associative algebra

$\mathfrak{gl}(V) = \mathrm{Hom}(V, V)$ — space of linear maps $V \to V$ considered as a Lie algebra, see Example 3.13

$\mathrm{tr}\, A$ — trace of a linear operator

$\mathrm{Ker}\, B = \{v \mid B(V, w) = 0 \text{ for all } w\}$, for a symmetric bilinear form B — kernel, or radical of B

A^t — adjoint operator: if $A\colon V \to W$ is a linear operator, then $A^t\colon W^* \to V^*$.

$A = A_s + A_n$ — Jordan decomposition of an operator A, see Theorem B.2

Differential geometry

$T_p M$ — tangent space to manifold M at point p

$\mathrm{Vect}(M)$ — space of C^∞ vector fields on M

$\mathrm{Diff}(M)$ — group of diffeomorphisms of a manifold M

Φ_ξ^t — time t flow of vector field ξ, see Section 3.5

Lie groups and Lie algebras

$\mathfrak{g} = \mathrm{Lie}(G)$ — Lie algebra of group G, see Theorem 3.17

$\mathfrak{z}(\mathfrak{g})$ — center of $\mathfrak{g}$, see Definition 3.31

$\mathrm{Der}(\mathfrak{g})$ — Lie algebra of derivations of $\mathfrak{g}$, see (3.13)

$[\mathfrak{g}, \mathfrak{g}]$ — commutant of $\mathfrak{g}$, see Definition 6.3

$\mathrm{rad}(\mathfrak{g})$ — radical of Lie algebra $\mathfrak{g}$, see Proposition 6.23

$K(x, y)$ — Killing form, see Definition 6.34

$\mathrm{Ad}\, g$ — adjoint action of G on $\mathfrak{g}$, see

$\mathrm{ad}\, x.y = [x, y]$, see (2.3)

$\mathrm{Stab}(m)$ — stabilizer of point m, see Lemma 2.18

$\exp\colon \mathfrak{g} \to G$ — exponential map, see Definition 3.2

Representations

χ_V — character of representation V, see Definition 4.41

V^G, $V^{\mathfrak{g}}$ — spaces of invariants, see Definition 4.12

$\mathrm{Hom}_G(V, W)$, $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ — spaces of intertwining operators, see Definition 4.1

Semisimple Lie algebras and root systems

$\mathfrak{h}$ — Cartan subalgebra, see Definition 7.8

$\mathfrak{g}_\alpha$ — root subspace, see Theorem 7.16

$R \subset \mathfrak{h}^* \setminus \{0\})$ — root system

$h_\alpha = \alpha^\vee = \frac{2H_\alpha}{(\alpha, \alpha)} \in \mathfrak{h}$ — dual root, see (7.4), (7.3) (for root system of a Lie algebra) and (8.4) for an abstract root system

$\mathrm{rank}(\mathfrak{g}) = \dim \mathfrak{h}$ — rank of a semisimple Lie algebra, see Corollary 7.14

s_α — reflection defined by a root α, see Definition 8.1

$V[\lambda]$ — weight subspace, see Definition 5.1

$R_\pm$ — positive and negative roots, see (8.6)

$\Pi = \{\alpha_1, \ldots, \alpha_r\} \subset R_+$ — simple roots, see Definition 8.12

$\mathrm{ht}(\alpha)$ — height of a positive root, see (8.8)

$L_\alpha = \{\lambda \in E \mid (\lambda, \alpha) = 0\}$ — root hyperplane, see (8.15)

C_+ — positive Weyl chamber, see (8.17)

W — Weyl group, see Definition 8.6

$s_i = s_{\alpha_i}$ — simple reflection, see Theorem 8.30

$l(w)$ — length of element $w \in W$, see Definition 8.34, Theorem 8.37

$\rho = \frac{1}{2}\sum_{R_+} \alpha$, see Lemma 8.36

P — weight lattice, see

Q — root lattice, see

Index

Bibliography

[1] V. Arnold, *Mathematical methods of classical mechanics*,

[2] N. Bourbaki, *Lie groups and Lie algebras. Chapters 1–3*, Springer, Berlin, 1998; *Lie groups and Lie algebras. Chapters 4–6*, Springer, Berlin, 2002;

[3] D. Bump, *Lie Groups*, Graduate Texts in Mathematics, vol. 225, Springer, New York, 2004

[4] T. Bröcker, T. tom Dieck, *Representations of Compact Lie Groups*, Graduate Texts in Mathematics, vol. 98, Springer, New York, 1985

[5] R. Carter, G. Segal, I. Macdonald, *Lectures on Lie groups and Lie algebras*, London Mathematical Society Student Texts, 32. Cambridge University Press, Cambridge, 1995.

[6] J. Dixmier, *Enveloping algebras*, Amer. Math. Soc., Providence, RI, 1996. [This is a reprint of 1977 translation of 1974 French original]

[7] B. L. Feigin, D. B. Fuchs, *Cohomologies of Lie groups and Lie algebras*, in *Lie groups and Lie algebras II*, Encyclopaedia Math. Sci., vol. 21, Springer, Berlin, 2000.

[8] V. V. Gorbatsevich, A. L. Onischik, E. B. Vinberg, *Structure of Lie Groups and Lie Algebras*, in *Lie groups and Lie algebras III*, Encyclopaedia Math. Sci., vol. 41, Springer, Berlin, 1994.

[9] B. Hall, *Lie Groups, Lie Algebras, and Representations. An Elementary Introduction*, Graduate Texts in Mathematics, Vol. 222, Springer, 2003. Corr. 2nd printing, 2004

[10] J. J. Duistermaat, J. A. C. Kolk, *Lie groups*, Springer-Verlag, Berlin, 2000.

[11] N. Jacobson, *Lie algebras*, Republication of the 1962 original. Dover Publications, Inc., New York, 1979.

[12] J. Humphreys, *Introduction to Lie Algebras and Representation Theory*

[13] J. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Univ. Press, Cambridge, 1990

[14] A. A. Kirillov, *Lectures on the orbit method*, Graduate Studies in Mathematics, 64. Amer. Math. Soc., Providence, RI, 2004

[15] A. Knapp, *Lie groups beyond an introduction*, Second edition. Progress in Mathematics, 140. Birkhäuser Boston, Inc., Boston, MA, 2002.

[16] Landau, Lifshits

[17] D. Mumford, J. Fogarty, F. Kirwan, *Geometric invariant theory*, third edition, Springer-Verlag, Berlin, 1994.

[18] A. L. Onishchik, E. B. Vinberg, *Foundations of Lie Theory*, in *Lie Groups and Lie Algebras II*, Enc. Math. Sci., vol. 20, Springer-Verlag, Berlin, 1990

[19] J.-P. Serre, *Linear representations of finite groups*, Graduate Texts in Mathematics, Vol. 42. Springer-Verlag, New York-Heidelberg, 1977

[20] J.-P. Serre, *Lie algebras and Lie groups*, 1964 lectures given at Harvard University, Second edition. Lecture Notes in Mathematics, 1500. Springer-Verlag, Berlin, 1992

[21] J.-P. Serre, *Complex semisimple Lie algebras*, Springer-Verlag, Berlin, 2001.

[22] B. Simon, *Representations of finite and compact groups*, Amer. Math. Soc., Providence, RI, 1996.

[23] M. Spivak, *A comprehensive introduction to differential geometry*, Vol. I. Second edition. Publish or Perish, Inc., Wilmington, Del., 1979.

[24] T. A. Springer, *Linear algebraic groups*, Second edition. Progress in Mathematics, 9. Birkhäuser Boston, Inc., Boston, MA, 1998.

[25] F. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Springer-Verlag, Berlin, 1983.

[26] D. P. Želobenko, *Compact Lie groups and their representations*, Translations of Mathematical Monographs, Vol. 40. Amer. Math. Soc., Providence, R.I., 1973

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.
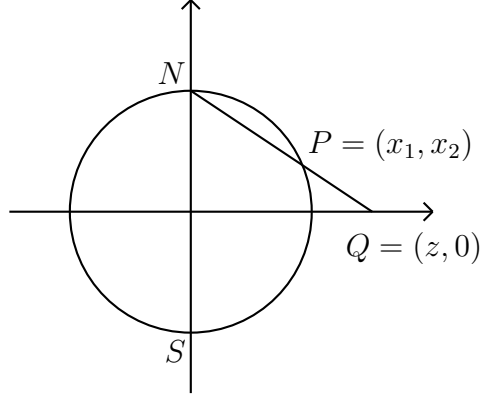
A line through N and P meets the x_1-axis in a unique point. This defines a map $S^1 \smallsetminus \{N\} \to \mathbb{R}$. To get an explicit formula for this, the line through N and P is described by $\{t(x_1, x_2) + (1 - t)(0, 1) \mid t \in \mathbb{R}\}$. This implies that $z = \frac{x_1}{1 - x_2}$. Therefore we define $U_1 = S^1 \smallsetminus \{N\}$, $V_1 = \mathbb{R}$ 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 $U_2 = S^1 \smallsetminus \{S\}$ 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 $U_1 \cap U_1 = S^1 \smallsetminus \{N, S\}$, $\phi_1(U_1 \cap U_2) = \phi_2(U_1 \cap U_2) = \mathbb{R} \smallsetminus \{0\}$. Hence $\phi_2 \circ \phi_1^{-1} : \mathbb{R} \smallsetminus \{0\} \to \mathbb{R} \smallsetminus \{0\}$. 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 $(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} \smallsetminus \{0\}$ and $[x_0 : x_1 : \cdots : x_n]$ represents the line through 0 and $(x_0, x_1, \ldots, x_n)$. Thus, for each $\lambda \in \mathbb{R}^\times$,

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

3

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 $\phi_i : U_i \xrightarrow{\sim} \mathbb{R}^n$, $[x_0 : x_1 : \cdots : x_n] \mapsto (\frac{x_0}{x_i}, \ldots, \widehat{x_i}, \ldots, \frac{x_n}{x_i})$ 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 $(x_0, x_1, \ldots, x_n)$. 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 $f : \mathbb{R}^n \to \mathbb{R}^m$, $f = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$, 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*

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

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 $\nu : C^\infty(M) \to \mathbb{R}$ 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 $f, g \in C^\infty(M)$.

2. It satisfies the derivation rule:

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

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 $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, we define the point derivation v_a at u by

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

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 $a = (a_1, \ldots, a_n)$, 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

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

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 $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}$ 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,

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

$\square$

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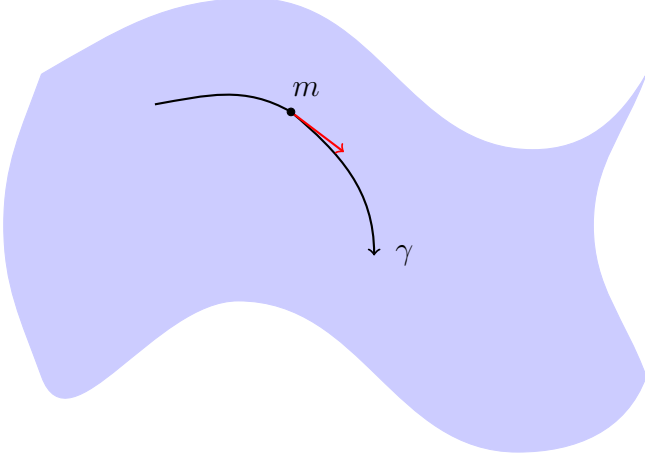
Figure 1: A curve through m and its point derivation.

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

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

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

7

have

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

Thus, it suffices to assume that $M = \mathbb{R}^n$ and $m = (m_1, \ldots, m_n) \in \mathbb{R}^n$. 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 $\gamma(t) = (a_1 t + m_1, \ldots, a_n t + m_n)$. Then $\gamma(0) = m$ and equation (3) shows that $\bar{\gamma} = \nu$. $\qquad\square$

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 $(1, 0, \ldots, 0)$, 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 $\mathbb{C}^\infty(M) \to \mathbb{C}^\infty(M)$ satisfying the product rule*

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

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

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

Hence $X(fg) = X(f)g + fX(g)$. $\qquad\square$

A linear map $\mathbb{C}^\infty(M) \to \mathbb{C}^\infty(M)$ 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

$$\frac{d}{dt} \gamma(t) \big|_{t=x} = X_{\gamma(x)}$$

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 $\mathcal{C}^{\mathrm{hol}}(M)$, 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 $C = G \setminus H$. 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 $f \in C' \setminus H$ then, in particular, $f \in (f^{-1})^{-1}(G \setminus H)$ i.e. there exists $g \in G \setminus H$ such that $f = (f^{-1})^{-1}g = fg$. But this implies that $g = e$ belongs to $G \setminus H$; a contradiction. $\square$

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 $G = \mathbb{R}^\times (=: \mathbb{R} \setminus \{0\})$, 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 $d_e \phi : T_e G \to T_e G$.

12

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

$$
\begin{CD}
G @>f>> H \\
@V\mathrm{Ad}(g)VV @VV\mathrm{Ad}(f(g))V \\
G @>>f> H
\end{CD}
\tag{5}
$$

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

$$
\begin{array}{ccc}
T_eG & \xrightarrow{\;d_ef\;} & T_eH \\
{\scriptstyle d_e\,\mathrm{Ad}(g)}\big\downarrow & & \big\downarrow{\scriptstyle d_e\,\mathrm{Ad}(f(g))} \\
T_eG & \xrightarrow[\;d_ef\;]{} & T_eH
\end{array}
\tag{6}
$$

However, if we want to check that d_ef really is the differential of some homomorphism $f : G \to H$ then (6) doesn't really help because the right vertical arrow is $d_e\,\mathrm{Ad}(f(g))$ 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

$$
\begin{array}{ccc}
G \times G & \xrightarrow{\;f \times f\;} & H \times H \\
{\scriptstyle \mathrm{Ad}}\big\downarrow & & \big\downarrow{\scriptstyle \mathrm{Ad}} \\
G & \xrightarrow[\;f\;]{} & H.
\end{array}
\tag{7}
$$

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 $T_mM \times T_nN \to T_{f(m,n)}K$. 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 $d_nf_m : T_nN \to T_{f(m,n)}K$. Then we fix $w \in T_nN$ and define $f'_w : M \to TK$ by $f'(m) = d_nf_m(w)$. Differentiating again we get $d_mf'_w : T_mM \to T_{f(m,n)}K$ and hence $b_{(m,n)}f : T_mM \times T_nN \to T_{f(m,n)}K$ given by $(v, w) \mapsto (d_mf'_w)(v)$. If one differentiates along M first and then along N then we get the same bilinear map. This follows from the fact that $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$

14

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

Thus, we get a commutative diagram of bilinear maps

$$
\begin{array}{ccc}
T_eG \times T_eG & \xrightarrow{\;b_{(e,e)}f \times f\;} & T_eH \times T_eH \\
{\scriptstyle b_{(e,e)}\,\mathrm{Ad}}\Big\downarrow & & \Big\downarrow{\scriptstyle b_{(e,e)}\,\mathrm{Ad}} \\
T_eG & \xrightarrow[\;d_ef\;]{} & T_eH.
\end{array}
\tag{8}
$$

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 $\text{ad}(Y) : \text{End}(V) \to \text{End}(V)$. If $X, Y \in \text{End}(V)$, then $(1 + \epsilon X)^{-1} = 1 - \epsilon X + \epsilon^2 X^2 - \cdots$. So,

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

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

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

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

The map $\text{Ad} : G \to GL(\mathfrak{g})$ is a morphism of Lie groups. Its differential at the identity is denoted $\text{ad} : \mathfrak{g} \to \text{End}(\mathfrak{g})$. 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.* $\text{ad}([X, Y]_G) = [\text{ad}(X), \text{ad}(Y)]_E$, *where* $[A, B]_E := AB - BA$ *is the bracket on* $\text{End}(\mathfrak{g})$.

Remark 3.19. One can check that $\text{ad}(Y)(X) = \text{ad}(X)(Y) = [X, Y]_G$ 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 $\text{Vect}(M)$ 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 $\mathfrak{gl}(V) = \text{End}(V)$ 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 $g \in G$, $\text{Ad}(\gamma(t))(g) = \gamma(t)g\gamma(t)^{-1}$. Then, taking $g = \rho(s)$ and differentiating ρ to get Y,

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

Thus,

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

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. $\qquad\square$

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, $d_h(g \cdot f) = d_{g^{-1}h}f \circ d_hL_{g^{-1}}$. Then,

$$[(g \cdot X)(f)](h) = [g \cdot (X(g^{-1} \cdot f))](h) = [X(g^{-1} \cdot f)](g^{-1}h)$$
$$= [d_{g^{-1}h}(g^{-1} \cdot f)](X_{g^{-1}h})$$
$$= (d_hf) \circ (d_{g^{-1}h}L_g)(X_{g^{-1}h}).$$

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

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. $\qquad\square$

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 $\lim_{n \to \infty} [x_n] = [x]$.

$\square$

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

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

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

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

24

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 $v \in V \setminus \{0\}$ 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. $\qquad\square$

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

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

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.

$\square$

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. $\qquad\square$

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 $K \simeq G \simeq G'$. $\qquad\square$

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. $\qquad \square$

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 $d_e f : T_e G \to T_e H$.*

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 $d_e u = \mathsf{id}_\mathfrak{g}$. Thus, $d_e(f \circ u) = d_e f$. By Theorem 4.18 and remark 4.19, there is a unique homomorphism $h : \widetilde{G} \to H$ such that $d_e h = d_e(f \circ u)$. This implies that $h = f \circ u$. If $g : G \to H$ was another map such that $d_e g = d_e f$, then $u \circ f = u \circ g$. But u is surjective, therefore this implies that $f = g$. $\qquad \square$

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

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

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 $A^* = \overline{A^T} = (\overline{A})^T$ 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

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

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

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

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

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

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

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

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

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:

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

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.

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

Their Lie algebras are

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

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

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

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 $u\bar{u} = a^2 + b^2 + c^2 + d^2$ 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 $\overline{z + w\mathbf{j}} = \bar{z} - w\mathbf{j}$ 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 $v\bar{v} = 1$ 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 $\mathbf{j}z = \bar{z}\mathbf{j}$.

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 $\operatorname{Hom}_{\mathfrak{g}}(M, N)$. When $M = N$ we write $\operatorname{End}_{\mathfrak{g}}(M)$ for this space. Notice that $\operatorname{Hom}_{\mathfrak{g}}(M, N)$ is a subspace of $\operatorname{Hom}_{\mathbb{C}}(M, N)$ and $\operatorname{End}_{\mathfrak{g}}(M)$ is a subspace of $\operatorname{End}_{\mathbb{C}}(M)$.

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. $\operatorname{End}_{\mathfrak{g}}(V) = \mathbb{C}$.*

Exercise 6.7. 1. Prove Schur's lemma.

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

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 $U = V \oplus V$. $\qquad\square$

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 $v' \otimes w' \in U$. $\qquad\square$

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 $v_1, \ldots, v_k$ 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 $v_1, \ldots, v_k$ are linearly independent, we can define an injective $\mathfrak{l}$-module homomorphism $\psi : W \oplus \cdots \oplus W \to V \otimes W$, $(w'_1, \ldots, w'_k) \mapsto \sum_{i=1}^{k} v_i \otimes w'_i$. Then u is the image of $\mathbf{w} := (w_1, \ldots, w_k)$ 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 $\mathfrak{k}$-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. $\qquad\square$

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 $\text{ad} : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{gl}(\mathfrak{sl}(2, \mathbb{C}))$. 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 $I_1, \ldots, I_k$ 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. $\qquad\square$

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 $Y(w) = 0$. $\qquad\square$

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. $\qquad\square$

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 $\{e_1, \ldots, e_n\}$ of V then the standard complete flag is $V_\bullet$, where $V_i = \mathrm{Span}\{e_1, \ldots, e_i\}$. 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 $i = 2, \ldots, n$. Then $\mathfrak{l}(V_i) \subset V_{i-1}$ as required. $\qquad\square$

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 $w, Y(w), Y^2(w), \ldots$. 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, $\operatorname{Tr}(X|_U) = \lambda(X) \dim U$. This applies to $[X, Y]$ too, $\operatorname{Tr}([X, Y]|_U) = \lambda([X, Y]) \dim U$. But the trace of the commutator of two endomorphisms is zero. Hence[1] $\lambda([X, Y]) = 0$. $\square$

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. $\qquad\square$

7.4 The Killing form

Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Recall that the adjoint representation defines a homomorphism 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 $\kappa(X, Y) = \text{Tr}(\text{ad}(X) \circ \text{ad}(Y))$.

Since $\text{Tr}(AB) = \text{Tr}(BA)$ 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 $\text{Tr}(X, Y) = 0$ 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 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 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. $\qquad\square$

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 $[\mathfrak{g}_i, \mathfrak{h}] = \mathfrak{g}_i$. $\qquad\square$

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. $\qquad \square$

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 $v \in V$. $\qquad \square$

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 $\beta_V(X_i, Y_i) = 1$. $\qquad\square$

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. $\qquad\square$

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. $\qquad \square$

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

Since $Y \in \mathfrak{h}$ too, we may diagonalize $\operatorname{ad}(Y)$. That is, there exists a basis $X_1, \ldots, X_n$ of $\mathfrak{h}$ and $\alpha_i \in \mathbb{C}$ such that $\operatorname{ad}(Y)(X_i) = \alpha_i X_i$. 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,

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

Since $Y, X_2, \ldots, X_n$ 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 $\mathfrak{g}_{-\alpha}$. $\square$

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 $\zeta(\mathfrak{g}) = 0$. $\qquad\square$

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. $\quad\square$

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. $\quad\square$

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. $\quad\square$

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 $V(n) = \mathfrak{g}_{\beta+r\alpha} \oplus \cdots \oplus \mathfrak{g}_{\beta-s\alpha}$. 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 $\gamma = \beta - \beta(H_\alpha)\alpha$. $\quad\square$

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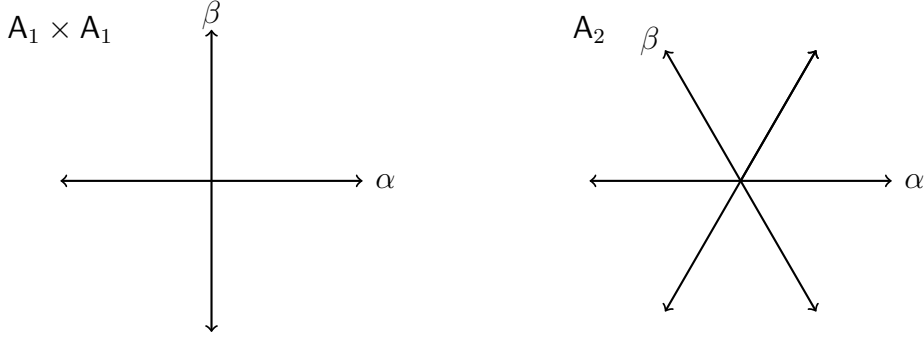
Proposition 9.16. *Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\mathfrak{h}$ a Cartan subalgebra. The decomposition of $\mathfrak{g}$ into $\mathfrak{h}$ weight spaces defines a root system R, in the abstract sense of Definition 9.15.*

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,

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

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 $0 \leq \cos^2 \theta \leq 1$, we have $0 \leq \langle \beta, \alpha \rangle \langle \alpha, \beta \rangle \leq 4$. Hence the only possible values of $\langle \alpha, \beta \rangle$ are:

$$
\begin{array}{cc|c}
\langle \beta, \alpha \rangle & \langle \alpha, \beta \rangle & \theta \\
\hline
0 & 0 & \frac{\pi}{2} \\
1 & 1 & \frac{\pi}{3} \\
-1 & -1 & \frac{2\pi}{3} \\
1 & 2 & \star \\
-1 & -2 & \frac{3\pi}{4} \\
1 & 3 & \frac{\pi}{6} \\
-1 & -3 & \star\star \\
1 & 4 & 0 \\
-1 & -4 & \pi
\end{array}
\tag{19}
$$

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

B$_2$ G$_2$
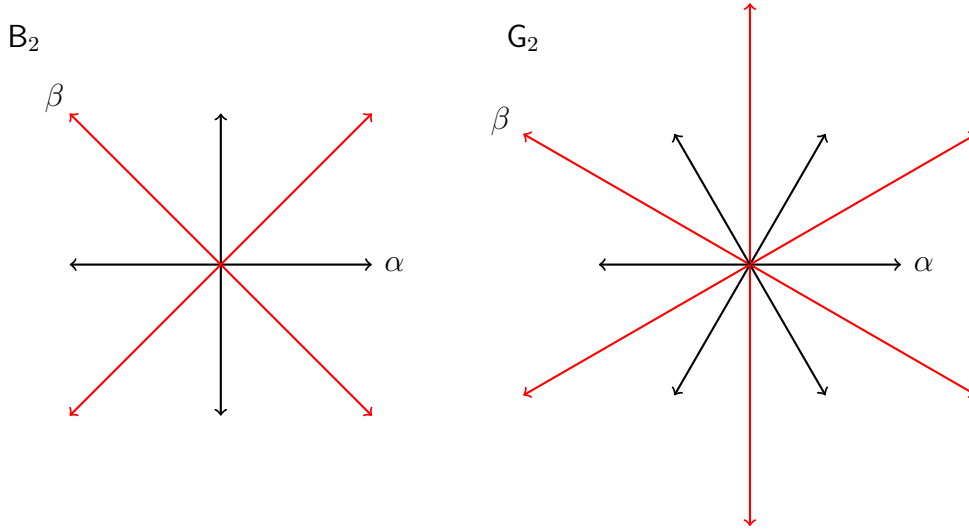
its action on R is trivial then, in particular, w fixes a basis of E. Hence w acts trivially on E i.e. $w = 1$. $\qquad\square$

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 $R = \{\epsilon_i - \epsilon_j \mid 1 \le i \ne j \le n+1\}$.

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 $E \setminus \bigcup_{\alpha \in R} H_\alpha$ 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
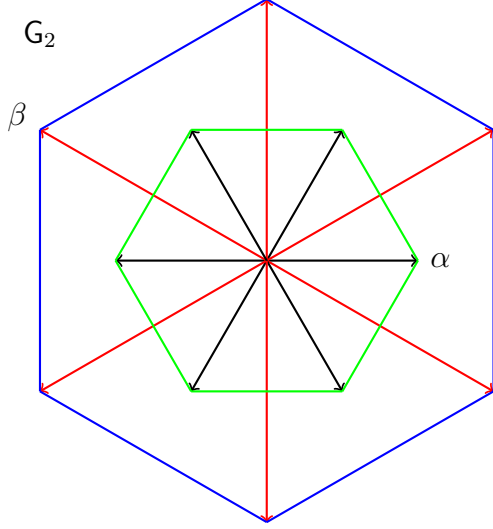
needed to recover the Lie algebra $\mathfrak{g}$. It is this fact that allows us to completely classify the simple Lie algebras.

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:

$$\mathsf{A_1 \times A_1} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}; \quad \mathsf{A_2} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}; \quad \mathsf{B_2} \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}; \quad \mathsf{G_2} \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}.$$

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, $\phi(R) = R'$. $\qquad\square$

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:

$$\mathsf{A}_1 \times \mathsf{A}_1 \quad \text{O} \quad \text{O}; \qquad \mathsf{A}_2 \quad \text{O—O}; \qquad \mathsf{B}_2 \quad \text{O}\!\!\Rightarrow\!\!\text{O}; \qquad \mathsf{G}_2 \quad \text{O}\!\!\Lleftarrow\!\!\text{O};$$

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

$$\mathsf{F}_4 \quad \text{O—O}\!\!\Rightarrow\!\!\text{O—O}$$

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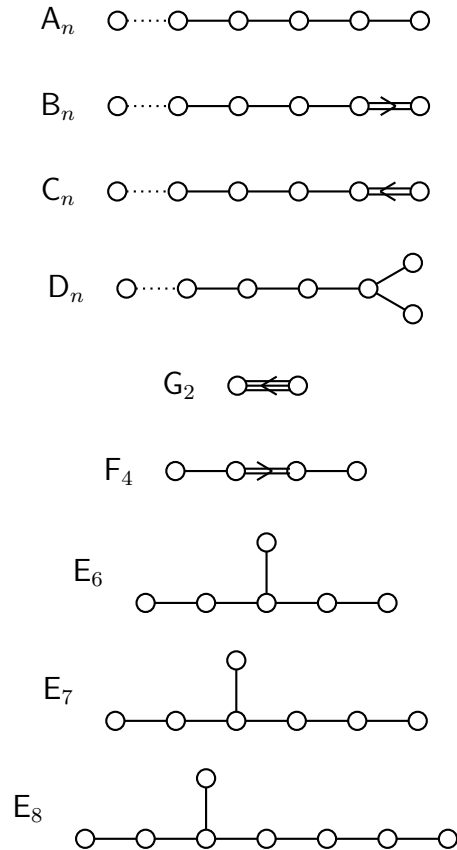
The Dynkin diagrams of type A_n, B_n, C_n and D_n correspond to the classical complex Lie algebras $\mathfrak{sl}(n+1, \mathbb{C})$, $\mathfrak{so}(2n+1, \mathbb{C})$, $\mathfrak{sp}(2n, \mathbb{C})$ and $\mathfrak{so}(2n, \mathbb{C})$ respectively.

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, $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$. Then $\mathfrak{g}$ will be generated by elements $H_1, \ldots, H_\ell, E_1, \ldots, E_\ell$ and $F_1, \ldots, F_\ell$. Now we need to give all nessecary relations amongst the generators $H_1, \ldots, F_\ell$, 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 $H_1, \ldots, H_\ell, E_1, \ldots, E_\ell$ and $F_1, \ldots, F_\ell$ and satisfying the relations*

1. *$[H_i, H_j] = 0$, for all $1 \leq i, j \leq \ell$;*

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 $1 \leq i, j \leq \ell$;*

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 $H_1, \ldots, H_\ell$ 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. $\qquad\square$

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 $m \in M_\lambda$.

70

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'. $\qquad\square$

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 $i_1, \ldots, i_k$ 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 $\beta_1, \ldots, \beta_\ell$. 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, $\{\beta_1, \ldots, \beta_\ell\}$ 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.

$\square$

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 $\text{Ch}(V)$ 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 $\text{Ch}(V)$ 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

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

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 $\text{sgn}(\sigma) := (-1)^{\ell(\sigma)}$, 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*

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

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 $\text{Id} \in \mathfrak{gl}(n, \mathbb{C})$ act by a fixed scalar (which we are free to choose).

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

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

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 $x_1^2 x_2 + x_1^2 x_3 + x_2^2 x_1 + x_2^2 x_3 + x_3^2 x_1 + x_3^2 x_2 + 2x_1 x_2 x_3$. To show this, we note that

$$\sum_{w \in \mathfrak{S}_3} \text{sgn}(w) e^{w(3,1,-1)} = x_1^3 x_2 x_3^{-1} - x_2^3 x_1 x_3^{-1} - x_3^3 x_2 x_1^{-1} - x_1^3 x_3 x_2^{-1} + x_2^3 x_3 x_1^{-1} + x_3^3 x_1 x_2^{-1}.$$

Then, one checks, by explicitly multiplying, that

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

74

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

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

First, we show spanning. If $v \in V$, then the fact that $[v_1], \ldots, [v_k]$ is a basis of V/W implies that there exists $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ 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 $w_1, \ldots, w_\ell$ are a basis of W, we can find $\beta_1, \ldots, \beta_\ell \in \mathbb{F}$ 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 $v_1, \ldots, v_k, w_1, \ldots, w_\ell$ span V.

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

Notice that we didn't really have to prove linear independence in the above proof. If we know that $v_1, \ldots, v_k, w_1, \ldots, w_\ell$ 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 $v_1, \ldots, v_k \in V$ such that $[v_1], \ldots, [v_k]$ is a basis of V/W, then $W' = \mathbb{C}\{v_1, \ldots, v_k\}$ is a compliment to W in V.*

Proof. Since $[v_1], \ldots, [v_k]$ 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 $V = W + W'$. $\qquad\square$

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:

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

$\square$

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. $\square$

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. $\square$

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 $W \subset \mathrm{Ker}\,\lambda$ 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 $v_1, \ldots, v_\ell, w_1, \ldots, w_k$ 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. $\qquad\square$

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.