Lie groups and Lie algebras

Eckhard Meinrenken

Lecture Notes, University of Toronto, Fall 2010

Contents

1. Terminology and notation

1.1. **Lie groups.**

Definition 1.1. A Lie group is a group G, equipped with a manifold structure such that the group operations

$$\mathrm{Mult}\colon G \times G \to G, \quad (g_1, g_2) \mapsto g_1 g_2$$
$$\mathrm{Inv}\colon G \to G, \quad g \mapsto g^{-1}$$

are smooth. A morphism of Lie groups G, G' is a morphism of groups $\phi\colon G \to G'$ that is smooth.

Remark 1.2. Using the implicit function theorem, one can show that smoothness of Inv is in fact automatic. (Exercise)

The first example of a Lie group is the *general linear group*

$$\mathrm{GL}(n,\mathbb{R}) = \{A \in \mathrm{Mat}_n(\mathbb{R}) |\ \det(A) \neq 0\}$$

of invertible $n \times n$ matrices. It is an open subset of $\mathrm{Mat}_n(\mathbb{R})$, hence a submanifold, and the smoothness of group multiplication follows since the product map for $\mathrm{Mat}_n(\mathbb{R})$ is obviously smooth.

Our next example is the orthogonal group

$$\mathrm{O}(n) = \{A \in \mathrm{Mat}_n(\mathbb{R}) |\ A^T A = I\}.$$

To see that it is a Lie group, it suffices to show that $\mathrm{O}(n)$ is an embedded submanifold of $\mathrm{Mat}_n(\mathbb{R})$. In order to construct submanifold charts, we use the exponential map of matrices

$$\exp\colon \mathrm{Mat}_n(\mathbb{R}) \to \mathrm{Mat}_n(\mathbb{R}), \quad B \mapsto \exp(B) = \sum_{n=0}^{\infty} \frac{1}{n!} B^n$$

(an absolutely convergent series). One has $\frac{d}{dt}|_{t=0} \exp(tB) = B$, hence the differential of exp at 0 is the identity $\mathrm{id}_{\mathrm{Mat}_n(\mathbb{R})}$. By the inverse function theorem, this means that there is $\epsilon > 0$ such that exp restricts to a diffeomorphism from the open neighborhood $U = \{B : \ ||B|| < \epsilon\}$ of 0 onto an open neighborhood $\exp(U)$ of I. Let

$$\mathfrak{o}(n) = \{B \in \mathrm{Mat}_n(\mathbb{R}) |\ B + B^T = 0\}.$$

We claim that

$$\exp(\mathfrak{o}(n) \cap U) = \mathrm{O}(n) \cap \exp(U),$$

so that exp gives a submanifold chart for $\mathrm{O}(n)$ over $\exp(U)$. To prove the claim, let $B \in U$. Then

$$\begin{aligned}
\exp(B) \in \mathrm{O}(n) &\Leftrightarrow \exp(B)^T = \exp(B)^{-1} \\
&\Leftrightarrow \exp(B^T) = \exp(-B) \\
&\Leftrightarrow B^T = -B \\
&\Leftrightarrow B \in \mathfrak{o}(n).
\end{aligned}$$

For a more general $A \in \mathrm{O}(n)$, we use that the map $\mathrm{Mat}_n(\mathbb{R}) \to \mathrm{Mat}_n(\mathbb{R})$ given by left multiplication is a diffeomorphism. Hence, $A\exp(U)$ is an open neighborhood of A, and we have

$$A\exp(U) \cap \mathrm{O}(n) = A(\exp(U) \cap \mathrm{O}(n)) = A\exp(U \cap \mathfrak{o}(n)).$$

Thus, we also get a submanifold chart near A. This proves that $\mathrm{O}(n)$ is a submanifold. Hence its group operations are induced from those of $\mathrm{GL}(n,\mathbb{R})$, they are smooth. Hence $\mathrm{O}(n)$ is a Lie group. Notice that $\mathrm{O}(n)$ is compact (the column vectors of an orthogonal matrix are an orthonormal basis of $\mathbb{R}^n$; hence $\mathrm{O}(n)$ is a subset of $S^{n-1} \times \cdots S^{n-1} \subset \mathbb{R}^n \times \cdots \mathbb{R}^n$).

A similar argument shows that the *special linear group*

$$\mathrm{SL}(n,\mathbb{R}) = \{A \in \mathrm{Mat}_n(\mathbb{R}) |\ \det(A) = 1\}$$

is an embedded submanifold of $\mathrm{GL}(n, \mathbb{R})$, and hence is a Lie group. The submanifold charts are obtained by exponentiating the subspace

$$\mathfrak{sl}(n, \mathbb{R}) = \{B \in \mathrm{Mat}_n(\mathbb{R}) |\ \mathrm{tr}(B) = 0\},$$

using the identity $\det(\exp(B)) = \exp(\mathrm{tr}(B))$.

Actually, we could have saved most of this work with $\mathrm{O}(n)$, $\mathrm{SL}(n, \mathbb{R})$ once we have the following beautiful result of E. Cartan:

> **Fact:** *Every closed subgroup of a Lie group is an embedded submanifold, hence is again a Lie group.*

We will prove this very soon, once we have developed some more basics of Lie group theory. A closed subgroup of $\mathrm{GL}(n, \mathbb{R})$ (for suitable n) is called a *matrix Lie group*. Let us now give a few more examples of Lie groups, without detailed justifications.

Examples 1.3. (a) Any finite-dimensional vector space V over $\mathbb{R}$ is a Lie group, with product Mult given by addition.

(b) Let $\mathcal{A}$ be a finite-dimensional associative algebra over $\mathbb{R}$, with unit $1_{\mathcal{A}}$. Then the group $\mathcal{A}^\times$ of invertible elements is a Lie group. More generally, if $n \in \mathbb{N}$ we can create the algebra $\mathrm{Mat}_n(\mathcal{A})$ of matrices with entries in $\mathcal{A}$, and the *general linear group*

$$\mathrm{GL}(n, \mathcal{A}) := \mathrm{Mat}_n(\mathcal{A})^\times$$

is a Lie group. If $\mathcal{A}$ is *commutative*, one has a determinant map $\det\colon \mathrm{Mat}_n(\mathcal{A}) \to \mathcal{A}$, and $\mathrm{GL}(n, \mathcal{A})$ is the pre-image of $\mathcal{A}^\times$. One may then define a *special linear group*

$$\mathrm{SL}(n, \mathcal{A}) = \{g \in \mathrm{GL}(n, \mathcal{A}) |\ \det(g) = 1\}.$$

(c) We mostly have in mind the cases $\mathcal{A} = \mathbb{R}, \mathbb{C}, \mathbb{H}$. Here $\mathbb{H}$ is the algebra of *quaternions* (due to Hamilton). Recall that $\mathbb{H} = \mathbb{R}^4$ as a vector space, with elements $(a, b, c, d) \in \mathbb{R}^4$ written as

$$x = a + ib + jc + kd$$

with imaginary units i, j, k. The algebra structure is determined by

$$i^2 = j^2 = k^2 = -1,\ ij = k,\ jk = i,\ ki = j.$$

Note that $\mathbb{H}$ is non-commutative (e.g. $ji = -ij$), hence $\mathrm{SL}(n, \mathbb{H})$ is *not* defined. On the other hand, one can define complex conjugates

$$\overline{x} = a - ib - jc - kd$$

and

$$|x|^2 := x\overline{x} = a^2 + b^2 + c^2 + d^2.$$

defines a norm $x \mapsto |x|$, with $|x_1 x_2| = |x_1||x_2|$ just as for complex or real numbers. The spaces $\mathbb{R}^n, \mathbb{C}^n, \mathbb{H}^n$ inherit norms, by putting

$$||x||^2 = \sum_{i=1}^n |x_i|^2, \quad x = (x_1, \ldots, x_n).$$

The subgroups of $\mathrm{GL}(n, \mathbb{R})$, $\mathrm{GL}(n, \mathbb{C})$, $\mathrm{GL}(n, \mathbb{H})$ preserving this norm (in the sense that $||Ax|| = ||x||$ for all x) are denoted

$$\mathrm{O}(n),\ \mathrm{U}(n),\ \mathrm{Sp}(n)$$

and are called the *orthogonal, unitary, and symplectic group*, respectively. Since the norms of $\mathbb{C}, \mathbb{H}$ coincide with those of $\mathbb{C} \cong \mathbb{R}^2$, $\mathbb{H} = \mathbb{C}^2 \cong \mathbb{R}^4$, we have

$$\mathrm{U}(n) = \mathrm{GL}(n, \mathbb{C}) \cap \mathrm{O}(2n), \quad \mathrm{Sp}(n) = \mathrm{GL}(n, \mathbb{H}) \cap \mathrm{O}(4n).$$

In particular, all of these groups are compact. One can also define

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}), \quad \mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n, \mathbb{C}),$$

these are called the *special orthogonal* and *special unitary* groups. The groups $\mathrm{SO}(n)$, $\mathrm{SU}(n), \mathrm{Sp}(n)$ are often called the *classical groups* (but this term is used a bit loosely).

(d) For any Lie group G, its universal cover $\widetilde{G}$ is again a Lie group. The universal cover $\widetilde{\mathrm{SL}(2, \mathbb{R})}$ is an example of a Lie group that is not isomorphic to a matrix Lie group.

1.2. **Lie algebras.**

Definition 1.4. A Lie algebra is a vector space $\mathfrak{g}$, together with a bilinear map $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying *anti-symmetry*

$$[\xi, \eta] = -[\eta, \xi] \text{ for all } \xi, \eta \in \mathfrak{g},$$

and the *Jacobi identity*,

$$[\xi, [\eta, \zeta]] + [\eta, [\zeta, \xi]] + [\zeta, [\xi, \eta]] = 0 \text{ for all } \xi, \eta, \zeta \in \mathfrak{g}.$$

The map $[\cdot, \cdot]$ is called the Lie bracket. A morphism of Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$ is a linear map $\phi \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ preserving brackets.

The space

$$\mathfrak{gl}(n, \mathbb{R}) = \mathrm{Mat}_n(\mathbb{R})$$

is a Lie algebra, with bracket the commutator of matrices. (The notation indicates that we think of $\mathrm{Mat}_n(\mathbb{R})$ as a Lie algebra, not as an algebra.)

A Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R})$, i.e. a subspace preserved under commutators, is called a *matrix Lie algebra*. For instance,

$$\mathfrak{sl}(n, \mathbb{R}) = \{B \in \mathrm{Mat}_n(\mathbb{R}) \colon \ \mathrm{tr}(B) = 0\}$$

and

$$\mathfrak{o}(n) = \{B \in \mathrm{Mat}_n(\mathbb{R}) \colon \ B^T = -B\}$$

are matrix Lie algebras (as one easily verifies). It turns out that every finite-dimensional real Lie algebra is isomorphic to a matrix Lie algebra (*Ado's theorem*), but the proof is not easy.

The following examples of finite-dimensional Lie algebras correspond to our examples for Lie groups. The origin of this correspondence will soon become clear.

Examples 1.5. (a) Any vector space V is a Lie algebra for the zero bracket.

(b) Any associative algebra $\mathcal{A}$ can be viewed as a Lie algebra under commutator. Replacing $\mathcal{A}$ with matrix algebras over $\mathcal{A}$, it follows that $\mathfrak{gl}(n, \mathcal{A}) = \mathrm{Mat}_n(\mathcal{A})$, is a Lie algebra, with bracket the commutator. If $\mathcal{A}$ is commutative, then the subspace $\mathfrak{sl}(n, \mathcal{A}) \subset \mathfrak{gl}(n, \mathcal{A})$ of matrices of trace 0 is a Lie subalgebra.

(c) We are mainly interested in the cases $\mathcal{A} = \mathbb{R}, \mathbb{C}, \mathbb{H}$. Define an inner product on $\mathbb{R}^n, \mathbb{C}^n, \mathbb{H}^n$ by putting

$$\langle x, y\rangle = \sum_{i=1}^{n} \overline{x}_i y_i,$$

and define $\mathfrak{o}(n),\ \mathfrak{u}(n),\ \mathfrak{sp}(n)$ as the matrices in $\mathfrak{gl}(n,\mathbb{R}),\ \mathfrak{gl}(n,\mathbb{C}),\ \mathfrak{gl}(n,\mathbb{H})$ satisfying

$$\langle Bx, y\rangle = -\langle x, By\rangle$$

for all x, y. These are all Lie algebras called the (infinitesimal) orthogonal, unitary, and symplectic Lie algebras. For $\mathbb{R}, \mathbb{C}$ one can impose the additional condition $\mathrm{tr}(B) = 0$, thus defining the special orthogonal and special unitary Lie algebras $\mathfrak{so}(n),\ \mathfrak{su}(n)$. Actually,

$$\mathfrak{so}(n) = \mathfrak{o}(n)$$

since $B^T = -B$ already implies $\mathrm{tr}(B) = 0$.

Exercise 1.6. Show that $\mathrm{Sp}(n)$ can be characterized as follows. Let $J \in U(2n)$ be the unitary matrix

$$\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

where I_n is the $n \times n$ identity matrix. Then

$$\mathrm{Sp}(n) = \{A \in \mathrm{U}(2n) |\ \overline{A} = JAJ^{-1}\}.$$

Here $\overline{A}$ is the componentwise complex conjugate of A.

Exercise 1.7. Let $R(\theta)$ denote the 2×2 rotation matrix

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Show that for all $A \in \mathrm{SO}(2m)$ there exists $O \in \mathrm{SO}(2m)$ such that OAO^{-1} is of the block diagonal form

$$\begin{pmatrix} R(\theta_1) & 0 & 0 & \cdots & 0 \\ 0 & R(\theta_2) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & R(\theta_m) \end{pmatrix}.$$

For $A \in \mathrm{SO}(2m+1)$ one has a similar block diagonal presentation, with m 2×2 blocks $R(\theta_i)$ and an extra 1 in the lower right corner. Conclude that $\mathrm{SO}(n)$ is connected.

Exercise 1.8. Let G be a connected Lie group, and U an open neighborhood of the group unit e. Show that any $g \in G$ can be written as a product $g = g_1 \cdots g_N$ of elements $g_i \in U$.

Exercise 1.9. Let $\phi\colon G \to H$ be a morphism of connected Lie groups, and assume that the differential $\mathrm{d}_e\phi\colon T_eG \to T_eH$ is bijective (resp. surjective). Show that ϕ is a covering (resp. surjective). Hint: Use Exercise 1.8.

2. The covering SU(2) → SO(3)

The Lie group SO(3) consists of rotations in 3-dimensional space. Let $D \subset \mathbb{R}^3$ be the closed ball of radius π. Any element $x \in D$ represents a rotation by an angle $||x||$ in the direction of x. This is a 1-1 correspondence for points in the interior of D, but if $x \in \partial D$ is a boundary point then $x, -x$ represent the same rotation. Letting $\sim$ be the equivalence relation on D, given by antipodal identification on the boundary, we have $D^3/\sim = \mathbb{R}P(3)$. Thus

$$\mathrm{SO}(3) = \mathbb{R}P(3)$$

(at least, topologically). With a little extra effort (which we'll make below) one can make this into a diffeomorphism of manifolds.

By definition

$$\mathrm{SU}(2) = \{A \in \mathrm{Mat}_2(\mathbb{C})|\ A^\dagger = A^{-1},\ \det(A) = 1\}.$$

Using the formula for the inverse matrix, we see that SU(2) consists of matrices of the form

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix} \mid |w|^2 + |z|^2 = 1 \right\}.$$

That is, $\mathrm{SU}(2) = S^3$ as a manifold. Similarly,

$$\mathfrak{su}(2) = \left\{ \begin{pmatrix} it & -\overline{u} \\ u & -it \end{pmatrix} \mid t \in \mathbb{R},\ u \in \mathbb{C} \right\}$$

gives an identification $\mathfrak{su}(2) = \mathbb{R} \oplus \mathbb{C} = \mathbb{R}^3$. Note that for a matrix B of this form, $\det(B) = t^2 + |u|^2$, so that det corresponds to $|| \cdot ||^2$ under this identification.

The group SU(2) acts linearly on the vector space $\mathfrak{su}(2)$, by matrix conjugation: $B \mapsto ABA^{-1}$. Since the conjugation action preserves det, we obtain a linear action on $\mathbb{R}^3$, preserving the norm. This defines a Lie group morphism from SU(2) into O(3). Since SU(2) is connected, this must take values in the identity component:

$$\phi\colon\ \mathrm{SU}(2) \to \mathrm{SO}(3).$$

The kernel of this map consists of matrices $A \in \mathrm{SU}(2)$ such that $ABA^{-1} = B$ for all $B \in \mathfrak{su}(2)$. Thus, A commutes with all skew-adjoint matrices of trace 0. Since A commutes with multiples of the identity, it then commutes with all skew-adjoint matrices. But since $\mathrm{Mat}_n(\mathbb{C}) = \mathfrak{u}(n) \oplus i\mathfrak{u}(n)$ (the decomposition into skew-adjoint and self-adjoint parts), it then follows that A is a multiple of the identity matrix. Thus $\ker(\phi) = \{I, -I\}$ is discrete. Since $\mathrm{d}_e\phi$ is an isomorphism, it follows that the map ϕ is a double covering. This exhibits $\mathrm{SU}(2) = S^3$ as the double cover of SO(3).

Exercise 2.1. Give an explicit construction of a double covering of SO(4) by SU(2) × SU(2). Hint: Represent the quaternion algebra $\mathbb{H}$ as an algebra of matrices $\mathbb{H} \subset \mathrm{Mat}_2(\mathbb{C})$, by

$$x = a + ib + jc + kd \mapsto x = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}.$$

Note that $|x|^2 = \det(x)$, and that $\mathrm{SU}(2) = \{x \in \mathbb{H}|\ \det(x) = 1\}$. Use this to define an action of SU(2) × SU(2) on $\mathbb{H}$ preserving the norm.

3. The Lie algebra of a Lie group

3.1. **Review: Tangent vectors and vector fields.** We begin with a quick reminder of some manifold theory, partly just to set up our notational conventions.

Let M be a manifold, and $C^\infty(M)$ its algebra of smooth real-valued functions. For $m \in M$, we define the tangent space T_mM to be the space of directional derivatives:

$$T_mM = \{v \in \mathrm{Hom}(C^\infty(M), \mathbb{R})|\ v(fg) = v(f)g + v(g)f\}.$$

Here $v(f)$ is local, in the sense that $v(f) = v(f')$ if $f' - f$ vanishes on a neighborhood of m.

Example 3.1. If $\gamma\colon J \to M$, $J \subset \mathbb{R}$ is a smooth curve we obtain tangent vectors to the curve,

$$\dot\gamma(t) \in T_{\gamma(t)}M, \quad \dot\gamma(t)(f) = \frac{\partial}{\partial t}|_{t=0} f(\gamma(t)).$$

Example 3.2. We have $T_x\mathbb{R}^n = \mathbb{R}^n$, where the isomorphism takes $a \in \mathbb{R}^n$ to the corresponding velocity vector of the curve $x + ta$. That is,

$$v(f) = \frac{\partial}{\partial t}|_{t=0} f(x + ta) = \sum_{i=1}^n a_i \frac{\partial f}{\partial x_i}.$$

A smooth map of manifolds $\phi\colon M \to M'$ defines a *tangent map*:

$$\mathrm{d}_m\phi\colon T_mM \to T_{\phi(m)}M', \quad (\mathrm{d}_m\phi(v))(f) = v(f \circ \phi).$$

The locality property ensures that for an open neighborhood $U \subset M$, the inclusion identifies $T_mU = T_mM$. In particular, a coordinate chart $\phi\colon U \to \phi(U) \subset \mathbb{R}^n$ gives an isomorphism

$$\mathrm{d}_m\phi\colon T_mM = T_mU \to T_{\phi(m)}\phi(U) = T_{\phi(m)}\mathbb{R}^n = \mathbb{R}^n.$$

Hence T_mM is a vector space of dimension $n = \dim M$. The union $TM = \bigcup_{m\in M} T_mM$ is a vector bundle over M, called the tangent bundle. Coordinate charts for M give vector bundle charts for TM. For a smooth map of manifolds $\phi\colon M \to M'$, the entirety of all maps $\mathrm{d}_m\phi$ defines a smooth vector bundle map

$$\mathrm{d}\phi\colon TM \to TM'.$$

A *vector field* on M is a derivation $X\colon C^\infty(M) \to C^\infty(M)$. That is, it is a linear map satisfying

$$X(fg) = X(f)g + fX(g).$$

The space of vector fields is denoted $\mathfrak{X}(M) = \mathrm{Der}(C^\infty(M))$. Vector fields are local, in the sense that for any open subset U there is a well-defined restriction $X|_U \in \mathfrak{X}(U)$ such that $X|_U(f|_U) = (X(f))|_U$. For any vector field, one obtains tangent vectors $X_m \in T_mM$ by $X_m(f) = X(f)|_m$. One can think of a vector field as an assignment of tangent vectors, depending smoothly on m. More precisely, a vector field is a smooth section of the tangent bundle TM. In local coordinates, vector fields are of the form $\sum_i a_i \frac{\partial}{\partial x_i}$ where the a_i are smooth functions.

It is a general fact that the commutator of derivations of an algebra is again a derivation. Thus, $\mathfrak{X}(M)$ is a Lie algebra for the bracket

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

In general, smooth maps $\phi\colon M \to M'$ of manifolds do not induce maps of the Lie algebras of vector fields (unless ϕ is a diffeomorphism). One makes the following definition.

Definition 3.3. Let $\phi\colon M \to N$ be a smooth map. Vector fields X, Y on M, N are called ϕ-related, written $X \sim_\phi Y$, if

$$X(f \circ \phi) = Y(f) \circ \phi$$

for all $f \in C^\infty(M')$.

In short, $X \circ \phi^* = \phi^* \circ Y$ where $\phi^*\colon C^\infty(N) \to C^\infty(M),\ f \mapsto f \circ \phi$.

One has $X \sim_\phi Y$ if and only if $Y_{\phi(m)} = \mathrm{d}_m\phi(X_m)$. From the definitions, one checks

$$X_1 \sim_\phi Y_1,\ X_2 \sim_\phi Y_2 \ \Rightarrow\ [X_1, X_2] \sim_\phi [Y_1, Y_2].$$

Example 3.4. Let $j\colon S \hookrightarrow M$ be an embedded submanifold. We say that a vector field X is *tangent to* S if $X_m \in T_mS \subset T_mM$ for all $m \in S$. We claim that if two vector fields are tangent to S then so is their Lie bracket. That is, the vector fields on M that are tangent to S form a Lie subalgebra.

Indeed, the definition means that there exists a vector field $X_S \in \mathfrak{X}(S)$ such that $X_S \sim_j X$. Hence, if X, Y are tangent to S, then $[X_S, Y_S] \sim_j [X, Y]$, so $[X_S, Y_S]$ is tangent.

Similarly, the vector fields vanishing on S are a Lie subalgebra.

Let $X \in \mathfrak{X}(M)$. A curve $\gamma(t),\ t \in J \subset \mathbb{R}$ is called an *integral curve* of X if for all $t \in J$,

$$\dot\gamma(t) = X_{\gamma(t)}.$$

In local coordinates, this is an ODE $\frac{\mathrm{d}x_i}{\mathrm{d}t} = a_i(x(t))$. The existence and uniqueness theorem for ODE's (applied in coordinate charts, and then patching the local solutions) shows that for any $m \in M$, there is a unique maximal integral curve $\gamma(t),\ t \in J_m$ with $\gamma(0) = m$.

Definition 3.5. A vector field X is *complete* if for all $m \in M$, the maximal integral curve with $\gamma(0) = m$ is defined for all $t \in \mathbb{R}$.

In this case, one obtains a *smooth* map

$$\Phi\colon \mathbb{R} \times M \to M,\ (t, m) \mapsto \Phi_t(m)$$

such that $\gamma(t) = \Phi_{-t}(m)$ is the integral curve through m. The uniqueness property gives

$$\Phi_0 = \mathrm{Id},\quad \Phi_{t_1+t_2} = \Phi_{t_1} \circ \Phi_{t_2}$$

i.e. $t \mapsto \Phi_t$ is a group homomorphism. Conversely, given such a group homomorphism such that the map Φ is smooth, one obtains a vector field X by setting

$$X = \frac{\partial}{\partial t}|_{t=0}\Phi^*_{-t},$$

as operators on functions. That is, $X(f)(m) = \frac{\partial}{\partial t}|_{t=0} f(\Phi_{-t}(m))$. [1]

[1] The minus sign is convention, but it is motivated as follows. Let $\mathrm{Diff}(M)$ be the infinite-dimensional group of diffeomorphisms of M. It acts on $C^\infty(M)$ by $\Phi.f = f \circ \Phi^{-1} = (\Phi^{-1})^* f$. Here, the inverse is needed so that $\Phi_1.\Phi_2.f = (\Phi_1\Phi_2).f$. We think of vector fields as 'infinitesimal flows', i.e. informally as the tangent space at id to $\mathrm{Diff}(M)$. Hence, given a curve $t \mapsto \Phi_t$ through $\Phi_0 = \mathrm{id}$, smooth in the sense that the map $\mathbb{R} \times M \to M,\ (t, m) \mapsto \Phi_t(m)$ is smooth, we define the corresponding vector field $X = \frac{\partial}{\partial t}|_{t=0}\Phi_t$ in terms of the action on functions: as

$$X.f = \frac{\partial}{\partial t}|_{t=0}\Phi_t.f = \frac{\partial}{\partial t}|_{t=0}(\Phi_t^{-1})^* f.$$

If Φ_t is a flow, we have $\Phi_t^{-1} = \Phi_{-t}$.

The Lie bracket of vector fields measure the non-commutativity of their flows. In particular, if X, Y are complete vector fields, with flows $\Phi_t^X,\ \Phi_s^Y$, then $[X, Y] = 0$ if and only if

$$\Phi_t^X \circ \Phi_s^Y = \Phi_s^Y \circ \Phi_t^X.$$

In this case, $X + Y$ is again a complete vector field with flow $\Phi_t^{X+Y} = \Phi_t^X \circ \Phi_t^Y$. (The right hand side defines a flow since the flows of X, Y commute, and the corresponding vector field is identified by taking a derivative at $t = 0$.)

3.2. **The Lie algebra of a Lie group.** Let G be a Lie group, and TG its tangent bundle. For all $a \in G$, the left,right translations

$$L_a \colon G \to G,\ g \mapsto ag$$

$$R_a \colon G \to G,\ g \mapsto ga$$

are smooth maps. Their differentials at e define isomorphisms $\mathrm{d}_g L_a \colon T_g G \to T_{ag} G$, and similarly for R_a. Let

$$\mathfrak{g} = T_e G$$

be the tangent space to the group unit.

A vector field $X \in \mathfrak{X}(G)$ is called left-invariant if

$$X \sim_{L_a} X$$

for all $a \in G$, i.e. if it commutes with L_a^*. The space $\mathfrak{X}^L(G)$ of left-invariant vector fields is thus a Lie subalgebra of $\mathfrak{X}(G)$. Similarly the space of right-invariant vector fields $\mathfrak{X}^R(G)$ is a Lie subalgebra.

Lemma 3.6. *The map*

$$\mathfrak{X}^L(G) \to \mathfrak{g},\ X \mapsto X_e$$

is an isomorphism of vector spaces. (Similarly for $\mathfrak{X}^R(G)$.)

Proof. For a left-invariant vector field, $X_a = (\mathrm{d}_e L_a) X_e$, hence the map is injective. To show that it is surjective, let $\xi \in \mathfrak{g}$, and put $X_a = (\mathrm{d}_e L_a)\xi \in T_a G$. We have to show that the map $G \to TG,\ a \mapsto X_a$ is smooth. It is the composition of the map $G \to G \times \mathfrak{g},\ g \mapsto (g, \xi)$ (which is obviously smooth) with the map $G \times \mathfrak{g} \to TG,\ (g, \xi) \mapsto \mathrm{d}_e L_g(\xi)$. The latter map is the restriction of $\mathrm{d}\,\mathrm{Mult} \colon TG \times TG \to TG$ to $G \times \mathfrak{g} \subset TG \times TG$, and hence is smooth. □

We denote by $\xi^L \in \mathfrak{X}^L(G),\ \xi^R \in \mathfrak{X}^R(G)$ the left,right invariant vector fields defined by $\xi \in \mathfrak{g}$. Thus

$$\xi^L|_e = \xi^R|_e = \xi$$

Definition 3.7. The Lie algebra of a Lie group G is the vector space $\mathfrak{g} = T_e G$, equipped with the unique bracket such that

$$[\xi, \eta]^L = [\xi^L, \eta^L], \quad \xi \in \mathfrak{g}.$$

Remark 3.8. If you use the right-invariant vector fields to define the bracket on $\mathfrak{g}$, we get a minus sign. Indeed, note that $\mathrm{Inv} \colon G \to G$ takes left translations to right translations. Thus, ξ^R is Inv-related to some left invariant vector field. Since $\mathrm{d}_e \mathrm{Inv} = -\mathrm{Id}$, we see $\xi^R \sim_{\mathrm{Inv}} -\xi^L$. Consequently,

$$[\xi^R, \eta^R] \sim_{\mathrm{Inv}} [-\xi^L, -\eta^L] = [\xi, \eta]^L.$$

But also $-[\xi,\eta]^R \sim_{\mathrm{Inv}} [\xi,\eta]^L$, hence we get

$$[\xi^R,\zeta^R] = -[\xi,\zeta]^R.$$

The construction of a Lie algebra is compatible with morphisms. That is, we have a *functor* from Lie groups to finite-dimensional Lie algebras.

Theorem 3.9. *For any morphism of Lie groups $\phi\colon G \to G'$, the tangent map $d_e\phi\colon \mathfrak{g} \to \mathfrak{g}'$ is a morphism of Lie algebras. For all $\xi \in \mathfrak{g}$, $\xi' = d_e\phi(\xi)$ one has*

$$\xi^L \sim_\phi (\xi')^L,\ \xi^R \sim_\phi (\xi')^R.$$

Proof. Suppose $\xi \in \mathfrak{g}$, and let $\xi' = \mathrm{d}_e\phi(\xi) \in \mathfrak{g}'$. The property $\phi(ab) = \phi(a)\phi(b)$ shows that $L_{\phi(a)} \circ \phi = \phi \circ L_a$. Taking the differential at e, and applying to ξ we find $(\mathrm{d}_e L_{\phi(a)})\xi' = (\mathrm{d}_a\phi)(\mathrm{d}_e L_a(\xi))$ hence $(\xi')^L_{\phi(a)} = (\mathrm{d}_a\phi)(\xi^L_a)$. That is $\xi^L \sim_\phi (\xi')^L$. The proof for right-invariant vector fields is similar. Since the Lie brackets of two pairs of ϕ-related vector fields are again ϕ-related, it follows that $\mathrm{d}_e\phi$ is a Lie algebra morphism. ☐

Remark 3.10. Two special cases are worth pointing out.

(a) Let V be a finite-dimensional (real) vector space. A representation of a Lie group G on V is a Lie group morphism $G \to \mathrm{GL}(V)$. A representation of a Lie algebra $\mathfrak{g}$ on V is a Lie algebra morphism $\mathfrak{g} \to \mathfrak{gl}(V)$. The Theorem shows that the differential of any Lie group representation is a representation of its a Lie algebra.

(b) An *automorphism of a Lie group* G is a Lie group morphism $\phi\colon G \to G$ from G to itself, with ϕ a diffeomorphism. An *automorphism of a Lie algebra* is an invertible morphism from $\mathfrak{g}$ to itself. By the Theorem, the differential of any Lie group automorphism is an automorphism of its Lie algebra. As an example, $\mathrm{SU}(n)$ has a Lie group automorphism given by complex conjugation of matrices; its differential is a Lie algebra automorphism of $\mathfrak{su}(n)$ given again by complex conjugation.

Exercise 3.11. Let $\phi\colon G \to G$ be a Lie group automorphism. Show that its fixed point set is a closed subgroup of G, hence a Lie subgroup. Similarly for Lie algebra automorphisms. What is the fixed point set for the complex conjugation automorphism of $\mathrm{SU}(n)$?

4. The exponential map

Theorem 4.1. *The left-invariant vector fields ξ^L are complete, i.e. they define a flow Φ^ξ_t such that*

$$\xi^L = \frac{\partial}{\partial t}|_{t=0}(\Phi^\xi_{-t})^*.$$

Letting $\phi^\xi(t)$ denote the unique integral curve with $\phi^\xi(0) = e$. It has the property

$$\phi^\xi(t_1+t_2) = \phi^\xi(t_1)\phi^\xi(t_2),$$

and the flow of ξ^L is given by right translations:

$$\Phi^\xi_t(g) = g\phi^\xi(-t).$$

Similarly, the right-invariant vector fields ξ^R are complete. $\phi^\xi(t)$ is an integral curve for ξ^R as well, and the flow of ξ^R is given by left translations, $g \mapsto \phi^\xi(-t)g$.

Proof. If $\gamma(t),\ t \in J \subset \mathbb{R}$ is an integral curve of a left-invariant vector field ξ^L, then its left translates $a\gamma(t)$ are again integral curves. In particular, for $t_0 \in J$ the curve $t \mapsto \gamma(t_0)\gamma(t)$ is again an integral curve. Hence it coincides with $\gamma(t_0+t)$ for all $t \in J \cap (J - t_0)$. In this way, an integral curve defined for small $|t|$ can be extended to an integral curve for all t, i.e. ξ^L is complete.

Since ξ^L is left-invariant, so is its flow Φ^ξ_t. Hence

$$\Phi^\xi_t(g) = \Phi^\xi_t \circ L_g(e) = L_g \circ \Phi^\xi_t(e) = g\Phi^\xi_t(e) = g\phi^\xi(-t).$$

The property $\Phi^\xi_{t_1+t_2} = \Phi^\xi_{t_1}\Phi^\xi_{t_2}$ shows that $\phi^\xi(t_1+t_2) = \phi^\xi(t_1)\phi^\xi(t_2)$. Finally, since $\xi^L \sim_{\mathrm{Inv}} -\xi^R$, the image

$$\mathrm{Inv}(\phi^\xi(t)) = \phi^\xi(t)^{-1} = \phi^\xi(-t)$$

is an integral curve of $-\xi^R$. Equivalently, $\phi^\xi(t)$ is an integral curve of ξ^R. □

Since left and right translations commute, it follows in particular that

$$[\xi^L, \eta^R] = 0.$$

Definition 4.2. A 1-parameter subgroup of G is a group homomorphism $\phi\colon \mathbb{R} \to G$.

We have seen that every $\xi \in \mathfrak{g}$ defines a 1-parameter group, by taking the integral curve through e of the left-invariant vector field ξ^L. Every 1-parameter group arises in this way:

Proposition 4.3. *If ϕ is a 1-parameter subgroup of G, then $\phi = \phi^\xi$ where $\xi = \dot\phi(0)$. One has*

$$\phi^{s\xi}(t) = \phi^\xi(st).$$

The map

$$\mathbb{R} \times \mathfrak{g} \to G,\ (t,\xi) \mapsto \phi^\xi(t)$$

is smooth.

Proof. Let $\phi(t)$ be a 1-parameter group. Then $\Phi_t(g) := g\phi(-t)$ defines a flow. Since this flow commutes with left translations, it is the flow of a left-invariant vector field, ξ^L. Here ξ is determined by taking the derivative of $\Phi_{-t}(e) = \phi(t)$ at $t = 0$: $\xi = \dot\phi(0)$. This shows $\phi = \phi^\xi$. As an application, since $\psi(t) = \phi^\xi(st)$ is a 1-parameter group with $\dot\psi^\xi(0) = s\dot\phi^\xi(0) = s\xi$, we have $\phi^\xi(st) = \phi^{s\xi}(t)$. Smoothness of the map $(t,\xi) \mapsto \phi^\xi(t)$ follows from the smooth dependence of solutions of ODE's on parameters. □

Definition 4.4. The *exponential map* for the Lie group G is the smooth map defined by

$$\exp\colon \mathfrak{g} \to G,\ \xi \mapsto \phi^\xi(1),$$

where $\phi^\xi(t)$ is the 1-parameter subgroup with $\dot\phi^\xi(0) = \xi$.

Proposition 4.5. *We have*

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

If $[\xi,\eta] = 0$ then

$$\exp(\xi+\eta) = \exp(\xi)\exp(\eta).$$

Proof. By the previous Proposition, $\phi^\xi(t) = \phi^{t\xi}(1) = \exp(t\xi)$. For the second claim, note that $[\xi, \eta] = 0$ implies that ξ^L, η^L commute. Hence their flows Φ_t^ξ, Φ_t^η, and $\Phi_t^\xi \circ \Phi_t^\eta$ is the flow of $\xi^L + \eta^L$. Hence it coincides with $\Phi_t^{\xi+\eta}$. Applying to e, we get $\phi^\xi(t)\phi^\eta(t) = \phi^{\xi+\eta}(t)$. Now put $t = 1$. ☐

In terms of the exponential map, we may now write the flow of ξ^L as $\Phi_t^\xi(g) = g\exp(-t\xi)$, and similarly for the flow of ξ^R. That is,

$$\xi^L = \frac{\partial}{\partial t}|_{t=0} R^*_{\exp(t\xi)}, \quad \xi^R = \frac{\partial}{\partial t}|_{t=0} L^*_{\exp(t\xi)}.$$

Proposition 4.6. *The exponential map is functorial with respect to Lie group homomorphisms $\phi\colon G \to H$. That is, we have a commutative diagram*

$$\begin{array}{ccc} G & \xrightarrow{\phi} & H \\ {\scriptstyle\exp}\uparrow & & \uparrow{\scriptstyle\exp} \\ \mathfrak{g} & \xrightarrow[d_e\phi]{} & \mathfrak{h} \end{array}$$

Proof. $t \mapsto \phi(\exp(t\xi))$ is a 1-parameter subgroup of H, with differential at e given by

$$\frac{d}{dt}\Big|_{t=0} \phi(\exp(t\xi)) = \mathrm{d}_e\phi(\xi).$$

Hence $\phi(\exp(t\xi)) = \exp(t\mathrm{d}_e\phi(\xi))$. Now put $t = 1$. ☐

Proposition 4.7. *Let $G \subset \mathrm{GL}(n, \mathbb{R})$ be a matrix Lie group, and $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{R})$ its Lie algebra. Then $\exp\colon \mathfrak{g} \to G$ is just the exponential map for matrices,*

$$\exp(\xi) = \sum_{n=0}^\infty \frac{1}{n!}\xi^n.$$

Furthermore, the Lie bracket on $\mathfrak{g}$ is just the commutator of matrices.

Proof. By the previous Proposition, applied to the inclusion of G in $\mathrm{GL}(n, \mathbb{R})$, the exponential map for G is just the restriction of that for $\mathrm{GL}(n, \mathbb{R})$. Hence it suffices to prove the claim for $G = \mathrm{GL}(n, \mathbb{R})$. The function $\sum_{n=0}^\infty \frac{t^n}{n!}\xi^n$ is a 1-parameter group in $\mathrm{GL}(n, \mathbb{R})$, with derivative at 0 equal to $\xi \in \mathfrak{gl}(n, \mathbb{R})$. Hence it coincides with $\exp(t\xi)$. Now put $t = 1$. ☐

Proposition 4.8. *For a matrix Lie group $G \subset \mathrm{GL}(n, \mathbb{R})$, the Lie bracket on $\mathfrak{g} = T_I G$ is just the commutator of matrices.*

Proof. It suffices to prove for $G = \mathrm{GL}(n, \mathbb{R})$. Using $\xi^L = \frac{\partial}{\partial t}\Big|_{t=0} R^*_{\exp(t\xi)}$ we have

$$\begin{aligned} &\frac{\partial}{\partial s}\Big|_{s=0} \frac{\partial}{\partial t}\Big|_{t=0} (R^*_{\exp(-t\xi)} R^*_{\exp(-s\eta)} R^*_{\exp(t\xi)} R^*_{\exp(s\eta)}) \\ &= \frac{\partial}{\partial s}\Big|_{s=0} (R^*_{\exp(-s\eta)} \xi^L R^*_{\exp(s\eta)} - \xi^L) \\ &= \xi^L \eta^L - \eta^L \xi^L \\ &= [\xi, \eta]^L. \end{aligned}$$

On the other hand, write

$$R^*_{\exp(-t\xi)}R^*_{\exp(-s\eta)}R^*_{\exp(t\xi)}R^*_{\exp(s\eta)} = R^*_{\exp(-t\xi)\exp(-s\eta)\exp(t\xi)\exp(s\eta)}.$$

Since the Lie group exponential map for $\mathrm{GL}(n,\mathbb{R})$ coincides with the exponential map for matrices, we may use Taylor's expansion,

$$\exp(-t\xi)\exp(-s\eta)\exp(t\xi)\exp(s\eta) = I + st(\xi\eta - \eta\xi) + \ldots = \exp(st(\xi\eta - \eta\xi)) + \ldots$$

where $\ldots$ denotes terms that are cubic or higher in s, t. Hence

$$R^*_{\exp(-t\xi)\exp(-s\eta)\exp(t\xi)\exp(s\eta)} = R^*_{\exp(st(\xi\eta-\eta\xi)} + \ldots$$

and consequently

$$\frac{\partial}{\partial s}\Big|_{s=0}\frac{\partial}{\partial t}\Big|_{t=0} R^*_{\exp(-t\xi)\exp(-s\eta)\exp(t\xi)\exp(s\eta)} = \frac{\partial}{\partial s}\Big|_{s=0}\frac{\partial}{\partial t}\Big|_{t=0} R^*_{\exp(st(\xi\eta-\eta\xi))} = (\xi\eta - \eta\xi)^L.$$

We conclude that $[\xi, \eta] = \xi\eta - \eta\xi$. □

Remark 4.9. Had we defined the Lie algebra using right-invariant vector fields, we would have obtained *minus* the commutator of matrices. Nonetheless, some authors use that convention.

The exponential map gives local coordinates for the group G on a neighborhood of e:

Proposition 4.10. *The differential of the exponential map at the origin is* $d_0 \exp = \mathrm{id}$. *As a consequence, there is an open neighborhood* U *of* $0 \in \mathfrak{g}$ *such that the exponential map restricts to a diffeomorphism* $U \to \exp(U)$.

Proof. Let $\gamma(t) = t\xi$. Then $\dot\gamma(0) = \xi$ since $\exp(\gamma(t)) = \exp(t\xi)$ is the 1-parameter group, we have

$$(\mathrm{d}_0 \exp)(\xi) = \frac{\partial}{\partial t}|_{t=0} \exp(t\xi) = \xi.$$

□

Exercise 4.11. Show hat the exponential map for $\mathrm{SU}(n)$, $\mathrm{SO}(n)$ $\mathrm{U}(n)$ are surjective. (We will soon see that the exponential map for any compact, connected Lie group is surjective.)

Exercise 4.12. A matrix Lie group $G \subset \mathrm{GL}(n,\mathbb{R})$ is called *unipotent* if for all $A \in G$, the matrix $A - I$ is nilpotent (i.e. $(A - I)^r = 0$ for some r). The prototype of such a group are the upper triangular matrices with 1's down the diagonal. Show that for a connected unipotent matrix Lie group, the exponential map is a diffeomorphism.

Exercise 4.13. Show that $\exp\colon \mathfrak{gl}(2,\mathbb{C}) \to \mathrm{GL}(2,\mathbb{C})$ is surjective. More generally, show that the exponential map for $\mathrm{GL}(n,\mathbb{C})$ is surjective. (Hint: First conjugate the given matrix into Jordan normal form).

Exercise 4.14. Show that $\exp\colon \mathfrak{sl}(2,\mathbb{R}) \to \mathrm{SL}(2,\mathbb{R})$ is not surjective, by proving that the matrices $\begin{pmatrix} -1 & \pm 1 \\ 0 & -1 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$ are not in the image. (Hint: Assuming these matrices are of the form $\exp(B)$, what would the eigenvalues of B have to be?) Show that these two matrices represent *all* conjugacy classes of elements that are not in the image of exp. (Hint: Find a classification of the conjugacy classes of $\mathrm{SL}(2,\mathbb{R})$, e.g. in terms of eigenvalues.)

5. Cartan's theorem on closed subgroups

Using the exponential map, we are now in position to prove Cartan's theorem on closed subgroups.

Theorem 5.1. *Let H be a closed subgroup of a Lie group G. Then H is an embedded submanifold, and hence is a Lie subgroup.*

We first need a Lemma. Let V be a Euclidean vector space, and $S(V)$ its unit sphere. For $v \in V\backslash\{0\}$, let $[v] = \frac{v}{||v||} \in S(V)$.

Lemma 5.2. *Let $v_n, v \in V\backslash\{0\}$ with $\lim_{n\to\infty} v_n = 0$. Then*

$$\lim_{n\to\infty}[v_n] = [v] \Leftrightarrow \exists a_n \in \mathbb{N}\colon \lim_{n\to\infty} a_n v_n = v.$$

Proof. The implication $\Leftarrow$ is obvious. For the opposite direction, suppose $\lim_{n\to\infty}[v_n] = [v]$. Let $a_n \in \mathbb{N}$ be defined by $a_n - 1 < \frac{||v||}{||v_n||} \le a_n$. Since $v_n \to 0$, we have $\lim_{n\to\infty} a_n \frac{||v_n||}{||v||} = 1$, and

$$a_n v_n = \left(a_n \frac{||v_n||}{||v||}\right)[v_n]\,||v|| \to [v]\,||v|| = v.$$

□

Proof of E. Cartan's theorem. It suffices to construct a submanifold chart near $e \in H$. (By left translation, one then obtains submanifold charts near arbitrary $a \in H$.) Choose an inner product on $\mathfrak{g}$.

We begin with a candidate for the Lie algebra of H. Let $W \subset \mathfrak{g}$ be the subset such that $\xi \in W$ if and only if either $\xi = 0$, or $\xi \neq 0$ and there exists $\xi_n \neq 0$ with

$$\exp(\xi_n) \in H, \quad \xi_n \to 0, \quad [\xi_n] \to [\xi].$$

We will now show the following:

(i) $\exp(W) \subset H$,

(ii) W is a subspace of $\mathfrak{g}$,

(iii) There is an open neighborhood U of 0 and a diffeomorphism $\phi\colon U \to \phi(U) \subset G$ with $\phi(0) = e$ such that

$$\phi(U \cap W) = \phi(U) \cap H.$$

(Thus ϕ defines a submanifold chart near e.)

Step (i). Let $\xi \in W\backslash\{0\}$, with sequence ξ_n as in the definition of W. By the Lemma, there are $a_n \in \mathbb{N}$ with $a_n\xi_n \to \xi$. Since $\exp(a_n\xi_n) = \exp(\xi_n)^{a_n} \in H$, and H is closed, it follows that

$$\exp(\xi) = \lim_{n\to\infty} \exp(a_n\xi_n) \in H.$$

Step (ii). Since the subset W is invariant under scalar multiplication, we just have to show that it is closed under addition. Suppose $\xi, \eta \in W$. To show that $\xi + \eta \in W$, we may assume that $\xi, \eta, \xi + \eta$ are all non-zero. For t sufficiently small, we have

$$\exp(t\xi)\exp(t\eta) = \exp(u(t))$$

for some smooth curve $t \mapsto u(t) \in \mathfrak{g}$ with $u(0) = 0$. Then $\exp(u(t)) \in H$ and

$$\lim_{n\to\infty} n\,u(\frac{1}{n}) = \lim_{h\to 0}\frac{u(h)}{h} = \dot{u}(0) = \xi + \eta.$$

hence $u(\frac{1}{n}) \to 0$, $\exp(u(\frac{1}{n}) \in H$, $[u(\frac{1}{n})] \to [\xi + \eta]$. This shows $[\xi + \eta] \in W$, proving (ii).

Step (iii). Let W' be a complement to W in $\mathfrak{g}$, and define

$$\phi\colon \mathfrak{g} \cong W \oplus W' \to G, \quad \phi(\xi + \xi') = \exp(\xi)\exp(\xi').$$

Since $\mathrm{d}_0\phi$ is the identity, there is an open neighborhood $U \subset \mathfrak{g}$ of 0 such that $\phi\colon U \to \phi(U)$ is a diffeomorphism. It is automatic that $\phi(W \cap U) \subset \phi(W) \cap \phi(U) \subset H \cap \phi(U)$. We want to show that we can take U sufficiently small so that we also have the opposite inclusion

$$H \cap \phi(U) \subset \phi(W \cap U).$$

Suppose not. Then, any neighborhood $U_n \subset \mathfrak{g} = W \oplus W'$ of 0 contains an element (η_n, η_n') such that

$$\phi(\eta_n, \eta_n') = \exp(\eta_n)\exp(\eta_n') \in H$$

(i.e. $\exp(\eta_n') \in H$) but $(\eta_n, \eta_n') \notin W$ (i.e. $\eta_n' \neq 0$). Thus, taking U_n to be a nested sequence of neighborhoods with intersection $\{0\}$, we could construct a sequence $\eta_n' \in W' - \{0\}$ with $\eta_n' \to 0$ and $\exp(\eta_n') \in H$. Passing to a subsequence we may assume that $[\eta_n'] \to [\eta]$ for some $\eta \in W' \backslash \{0\}$. On the other hand, such a convergence would mean $\eta \in W$, by definition of W. Contradiction. □

As remarked earlier, Cartan's theorem is very useful in practice. For a given Lie group G, the term 'closed subgroup' is often used as synonymous to 'embedded Lie subgroup'.

Examples 5.3. (a) The matrix groups $G = \mathrm{O}(n), \mathrm{Sp}(n), \mathrm{SL}(n,\mathbb{R}), \ldots$ are all closed subgroups of some $\mathrm{GL}(N,\mathbb{R})$, and hence are Lie groups.

(b) Suppose that $\phi\colon G \to H$ is a morphism of Lie groups. Then $\ker(\phi) = \phi^{-1}(e) \subset G$ is a closed subgroup. Hence it is an embedded Lie subgroup of G.

(c) The center $Z(G)$ of a Lie group G is the set of all $a \in G$ such that $ag = ga$ for all $a \in G$. It is a closed subgroup, and hence an embedded Lie subgroup.

(d) Suppose $H \subset G$ is a closed subgroup. Its *normalizer* $N_G(H) \subset G$ is the set of all $a \in G$ such that $aH = Ha$. (I.e. $h \in H$ implies $aha^{-1} \in H$.) This is a closed subgroup, hence a Lie subgroup. The *centralizer* $Z_G(H)$ is the set of all $a \in G$ such that $ah = ha$ for all $h \in H$, it too is a closed subgroup, hence a Lie subgroup.

The E. Cartan theorem is just one of many 'automatic smoothness' results in Lie theory. Here is another.

Theorem 5.4. *Let G, H be Lie groups, and $\phi\colon G \to H$ be a continuous group morphism. Then ϕ is smooth.*

As a corollary, a given topological group carries at most one smooth structure for which it is a Lie group. For profs of these (and stronger) statements, see the book by Duistermaat-Kolk.

6. The adjoint representation

6.1. **Automorphisms.** The group $\mathrm{Aut}(\mathfrak{g})$ of automorphisms of a Lie algebra $\mathfrak{g}$ is closed in the group $\mathrm{End}(\mathfrak{g})^\times$ of vector space automorphisms, hence it is a Lie group. To identify its Lie algebra, let $D \in \mathrm{End}(\mathfrak{g})$ be such that $\exp(tD) \in \mathrm{Aut}(\mathfrak{g})$ for $t \in \mathbb{R}$. Taking the derivative of the defining condition $\exp(tD)[\xi,\eta] = [\exp(tD)\xi, \exp(tD)\eta]$, we obtain the property

$$D[\xi,\eta] = [D\xi,\eta] + [\xi, D\eta]$$

saying that D is a *derivation* of the Lie algebra. Conversely, if D is a derivation then

$$D^n[\xi,\eta]=\sum_{k=0}^n \binom{n}{k}[D^k\xi,\, D^{n-k}\eta]$$

by induction, which then shows that $\exp(D)=\sum_n \frac{D^n}{n!}$ is an automorphism. Hence the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$ is the Lie algebra $\mathrm{Der}(\mathfrak{g})$ of derivations of $\mathfrak{g}$.

Exercise 6.1. Using similar arguments, verify that the Lie algebra of $\mathrm{SO}(n), \mathrm{SU}(n), \mathrm{Sp}(n), \ldots$ are $\mathfrak{so}(n),\ \mathfrak{su}(n),\ \mathfrak{sp}(n), \ldots$.

6.2. **The adjoint representation of G.** Recall that an automorphism of a Lie group G is an invertible morphism from G to itself. The automorphisms form a group $\mathrm{Aut}(G)$. Any $a\in G$ defines an 'inner' automorphism $\mathrm{Ad}_a\in\mathrm{Aut}(G)$ by conjugation:

$$\mathrm{Ad}_a(g)=aga^{-1}$$

Indeed, Ad_a is an automorphism since $\mathrm{Ad}_a^{-1}=\mathrm{Ad}_{a^{-1}}$ and

$$\mathrm{Ad}_a(g_1g_2)=ag_1g_2a^{-1}=ag_1a^{-1}ag_2a^{-1}=\mathrm{Ad}_a(g_1)\,\mathrm{Ad}_a(g_2).$$

Note also that $\mathrm{Ad}_{a_1a_2}=\mathrm{Ad}_{a_1}\mathrm{Ad}_{a_2}$, thus we have a group morphism

$$\mathrm{Ad}\colon G\to\mathrm{Aut}(G)$$

into the group of automorphisms. The kernel of this morphism is the center $Z(G)$, the image is (by definition) the subgroup $\mathrm{Inn}(G)$ of inner automorphisms. Note that for any $\phi\in\mathrm{Aut}(G)$, and any $a\in G$,

$$\phi\circ\mathrm{Ad}_a\circ\phi^{-1}=\mathrm{Ad}_{\phi(a)}\,.$$

That is, $\mathrm{Inn}(G)$ is a *normal* subgroup of $\mathrm{Aut}(G)$. (I.e. the conjugate of an inner automorphism by any automorphism is inner.) It follows that $\mathrm{Out}(G)=\mathrm{Aut}(G)/\mathrm{Inn}(G)$ inherits a group structure; it is called the *outer automorphism group*.

Example 6.2. If $G=\mathrm{SU}(2)$ the complex conjugation of matrices is an inner automorphism, but for $G=\mathrm{SU}(n)$ with $n\geq 3$ it cannot be inner (since an inner automorphism has to preserve the spectrum of a matrix). Indeed, one know that $\mathrm{Out}(\mathrm{SU}(n))=\mathbb{Z}_2$ for $n\geq 3$.

The differential of the automorphism $\mathrm{Ad}_a\colon G\to G$ is a Lie algebra automorphism, denoted by the same letter: $\mathrm{Ad}_a=\mathrm{d}_e\mathrm{Ad}_a\colon\mathfrak{g}\to\mathfrak{g}$. The resulting map

$$\mathrm{Ad}\colon G\to\mathrm{Aut}(\mathfrak{g})$$

is called the *adjoint representation of* G. Since the Ad_a are Lie algebra/group morphisms, they are compatible with the exponential map,

$$\exp(\mathrm{Ad}_a\xi)=\mathrm{Ad}_a\exp(\xi).$$

Remark 6.3. If $G\subset\mathrm{GL}(n,\mathbb{R})$ is a matrix Lie group, then $\mathrm{Ad}_a\in\mathrm{Aut}(\mathfrak{g})$ is the conjugation of matrices

$$\mathrm{Ad}_a(\xi)=a\xi a^{-1}.$$

This follows by taking the derivative of $\mathrm{Ad}_a(\exp(t\xi))=a\exp(t\xi)a^{-1}$, using that exp is just the exponential series for matrices.

6.3. **The adjoint representation of $\mathfrak{g}$.** Let $\mathrm{Der}(\mathfrak{g})$ be the Lie algebra of derivations of the Lie algebra $\mathfrak{g}$. There is a Lie algebra morphism,

$$\mathrm{ad}\colon \mathfrak{g} \to \mathrm{Der}(\mathfrak{g}), \quad \xi \mapsto [\xi, \cdot].$$

The fact that ad_ξ is a derivation follows from the Jacobi identity; the fact that $\xi \mapsto \mathrm{ad}_\xi$ it is a Lie algebra morphism is again the Jacobi identity. The kernel of ad is the center of the Lie algebra $\mathfrak{g}$, i.e. elements having zero bracket with all elements of $\mathfrak{g}$, while the image is the Lie subalgebra $\mathrm{Inn}(\mathfrak{g}) \subset \mathrm{Der}(\mathfrak{g})$ of *inner* derivations. It is a normal Lie subalgebra, i.e $[\mathrm{Der}(\mathfrak{g}), \mathrm{Inn}(\mathfrak{g})] \subset \mathrm{Inn}(\mathfrak{g})$, and the quotient Lie algebra $\mathrm{Out}(\mathfrak{g})$ are the *outer automorphims.*

Suppose now that G is a Lie group, with Lie algebra $\mathfrak{g}$. We have remarked above that the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$ is $\mathrm{Der}(\mathfrak{g})$. Recall that the differential of any G-representation is a $\mathfrak{g}$-representation. In particular, we can consider the differential of $G \to \mathrm{Aut}(\mathfrak{g})$.

Theorem 6.4. *If $\mathfrak{g}$ is the Lie algebra of G, then the adjoint representation $\mathrm{ad}\colon \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$ is the differential of the adjoint representation $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathfrak{g})$. One has the equality of operators*

$$\exp(\mathrm{ad}_\xi) = \mathrm{Ad}(\exp \xi)$$

for all $\xi \in \mathfrak{g}$.

Proof. For the first part we have to show $\frac{\partial}{\partial t}\big|_{t=0} \mathrm{Ad}_{\exp(t\xi)}\, \eta = \mathrm{ad}_\xi\, \eta$. This is easy if G is a matrix Lie group:

$$\frac{\partial}{\partial t}\Big|_{t=0} \mathrm{Ad}_{\exp(t\xi)}\, \eta = \frac{\partial}{\partial t}\Big|_{t=0} \exp(t\xi)\eta\exp(-t\xi) = \xi\eta - \eta\xi = [\xi, \eta].$$

For general Lie groups we compute, using

$$\exp(s\,\mathrm{Ad}_{\exp(t\xi)}\, \eta) = \mathrm{Ad}_{\exp(t\xi)} \exp(s\eta) = \exp(t\xi)\exp(s\eta)\exp(-t\xi),$$

$$\begin{aligned}
\frac{\partial}{\partial t}\Big|_{t=0} (\mathrm{Ad}_{\exp(t\xi)}\, \eta)^L &= \frac{\partial}{\partial t}\Big|_{t=0} \frac{\partial}{\partial s}\Big|_{s=0} R^*_{\exp(s\,\mathrm{Ad}_{\exp(t\xi)}\, \eta)} \\
&= \frac{\partial}{\partial t}\Big|_{t=0} \frac{\partial}{\partial s}\Big|_{s=0} R^*_{\exp(t\xi)\exp(s\eta)\exp(-t\xi)} \\
&= \frac{\partial}{\partial t}\Big|_{t=0} \frac{\partial}{\partial s}\Big|_{s=0} R^*_{\exp(t\xi)} R^*_{\exp(s\eta)} R^*_{\exp(-t\xi)} \\
&= \frac{\partial}{\partial t}\Big|_{t=0} R^*_{\exp(t\xi)}\ \eta^L\ R^*_{\exp(-t\xi)} \\
&= [\xi^L, \eta^L] \\
&= [\xi, \eta]^L = (\mathrm{ad}_\xi\, \eta)^L.
\end{aligned}$$

This proves the first part. The second part is the commutativity of the diagram

$$\begin{array}{ccc}
G & \xrightarrow{\mathrm{Ad}} & \mathrm{Aut}(\mathfrak{g}) \\
{\scriptstyle \exp}\uparrow & & \uparrow{\scriptstyle \exp} \\
\mathfrak{g} & \xrightarrow[\mathrm{ad}]{} & \mathrm{Der}(\mathfrak{g})
\end{array}$$

which is just a special case of the functoriality property of exp with respect to Lie group morphisms. ∎

Remark 6.5. As a special case, this formula holds for matrices. That is, for $B, C \in \mathrm{Mat}_n(\mathbb{R})$,

$$e^B\, C\, e^{-B} = \sum_{n=0}^{\infty} \frac{1}{n!}[B, [B, \cdots [B, C] \cdots]].$$

The formula also holds in some other contexts, e.g. if B, C are elements of an algebra with B nilpotent (i.e. $B^N = 0$ for some N). In this case, both the exponential series for e^B and the series on the right hand side are finite. (Indeed, $[B, [B, \cdots [B, C] \cdots]]$ with n B's is a sum of terms $B^j C B^{n-j}$, and hence must vanish if $n \geq 2N$.)

7. The differential of the exponential map

We had seen that $\mathrm{d}_0 \exp = \mathrm{id}$. More generally, one can derive a formula for the differential of the exponential map at arbitrary points $\xi \in \mathfrak{g}$,

$$\mathrm{d}_\xi \exp\colon \mathfrak{g} = T_\xi \mathfrak{g} \to T_{\exp \xi} G.$$

Using left translation, we can move $T_{\exp \xi} G$ back to $\mathfrak{g}$, and obtain an endomorphism of $\mathfrak{g}$.

Theorem 7.1. *The differential of the exponential map* $\exp\colon \mathfrak{g} \to G$ *at* $\xi \in \mathfrak{g}$ *is the linear operator* $d_\xi \exp\colon \mathfrak{g} \to T_{\exp(\xi)}\mathfrak{g}$ *given by the formula,*

$$d_\xi \exp = (d_e L_{\exp \xi}) \circ \frac{1 - \exp(-\operatorname{ad}_\xi)}{\operatorname{ad}_\xi}.$$

Here the operator on the right hand side is defined to be the result of substituting ad_ξ in the entire holomorphic function $\frac{1-e^{-z}}{z}$. Equivalently, it may be written as an integral

$$\frac{1 - \exp(-\operatorname{ad}_\xi)}{\operatorname{ad}_\xi} = \int_0^1 \mathrm{d}s\ \exp(-s \operatorname{ad}_\xi).$$

Proof. We have to show that for all $\xi, \eta \in \mathfrak{g}$,

$$(\mathrm{d}_\xi \exp)(\eta) \circ L^*_{\exp(-\xi)} = \int_0^1 \mathrm{d}s\ (\exp(-s \operatorname{ad}_\xi)\eta)$$

as operators on functions $f \in C^\infty(G)$. To compute the left had side, write

$$(\mathrm{d}_\xi \exp)(\eta) \circ L^*_{\exp(-\xi)}(f) = \frac{\partial}{\partial t}\Big|_{t=0} (L^*_{\exp(-\xi)}(f))(\exp(\xi + t\eta)) = \frac{\partial}{\partial t}\Big|_{t=0} f(\exp(-\xi)\exp(\xi + t\eta)).$$

We think of this as the value of $\frac{\partial}{\partial t}\Big|_{t=0} R^*_{\exp(-\xi)} R^*_{\exp(\xi+t\eta)} f$ at e, and compute as follows: [2]

$$\begin{aligned}\frac{\partial}{\partial t}\Big|_{t=0} R^*_{\exp(-\xi)} R^*_{\exp(\xi+t\eta)} &= \int_0^1 \mathrm{d}s\ \frac{\partial}{\partial t}\Big|_{t=0} \frac{\partial}{\partial s} R^*_{\exp(-s\xi)} R^*_{\exp(s(\xi+t\eta)}\\ &= \int_0^1 \mathrm{d}s\ \frac{\partial}{\partial t}\Big|_{t=0} R^*_{\exp(-s\xi)} (t\eta)^L R^*_{\exp(s(\xi+t\eta)}\\ &= \int_0^1 \mathrm{d}s\ R^*_{\exp(-s\xi)}\ \eta^L\ R^*_{\exp(s(\xi))}\\ &= \int_0^1 \mathrm{d}s\ (\mathrm{Ad}_{\exp(-s\xi)}\,\eta)^L\\ &= \int_0^1 \mathrm{d}s\ (\exp(-s\,\mathrm{ad}_\xi)\eta)^L.\end{aligned}$$

Applying this result to f at e, we obtain $\int_0^1 \mathrm{d}s\ (\exp(-s\,\mathrm{ad}_\xi)\eta)(f)$ as desired. □

Corollary 7.2. *The exponential map is a local diffeomorphism near $\xi \in \mathfrak{g}$ if and only if ad_ξ has no eigenvalue in the set $2\pi i\mathbb{Z}\backslash\{0\}$.*

Proof. $\mathrm{d}_\xi \exp$ is an isomorphism if and only if $\frac{1-\exp(-\mathrm{ad}_\xi)}{\mathrm{ad}_\xi}$ is invertible, i.e. has non-zero determinant. The determinant is given in terms of the eigenvalues of ad_ξ as a product, $\prod_\lambda \frac{1-e^{-\lambda}}{\lambda}$. This vanishes if and only if there is a non-zero eigenvalue λ with $e^\lambda = 1$. □

As an application, one obtains a version of the *Baker-Campbell-Hausdorff formula*. Let $g \mapsto \log(g)$ be the inverse function to exp, defined for g close to e. For $\xi, \eta \in \mathfrak{g}$ close to 0, the function

$$\log(\exp(\xi)\exp(\eta))$$

The BCH formula gives the Taylor series expansion of this function. The series starts out with

$$\log(\exp(\xi)\exp(\eta)) = \xi + \eta + \tfrac{1}{2}[\xi,\eta] + \cdots$$

but gets rather complicated. To derive the formula, introduce a t-dependence, and let $f(t,\xi,\eta)$ be defined by $\exp(\xi)\exp(t\eta) = \exp(f(t,\xi,\eta))$ (for ξ,η sufficiently small). Thus

$$\exp(f) = \exp(\xi)\exp(t\eta)$$

We have, on the one hand,

$$(\mathrm{d}_e L_{\exp(f)})^{-1}\frac{\partial}{\partial t}\exp(f) = \mathrm{d}_e L^{-1}_{\exp(t\eta)}\frac{\partial}{\partial t}\exp(t\eta) = \eta.$$

On the other hand, by the formula for the differential of exp,

$$(\mathrm{d}_e L_{\exp(f)})^{-1}\frac{\partial}{\partial t}\exp(f) = (\mathrm{d}_e L_{\exp(f)})^{-1}(\mathrm{d}_f \exp)(\frac{\partial f}{\partial t}) = \frac{1-e^{-\mathrm{ad}_f}}{\mathrm{ad}_f}(\frac{\partial f}{\partial t}).$$

Hence

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\mathrm{ad}_f}{1-e^{-\mathrm{ad}_f}}\eta.$$

[2]We will use the identities $\frac{\partial}{\partial s} R^*_{\exp(s\zeta)} = R^*_{\exp(s\zeta)}\ \zeta^L = \zeta^L\ R^*_{\exp(s\zeta)}$ for all $\zeta \in \mathfrak{g}$. Proof: $\frac{\partial}{\partial s} R^*_{\exp(s\zeta)} = \frac{\partial}{\partial u}|_{u=0} R^*_{\exp((s+u)\zeta)} = \frac{\partial}{\partial u}|_{u=0} R^*_{\exp(u\zeta)} R^*_{\exp(s\zeta)} = \zeta^L R^*_{\exp(s\zeta)}$.

Letting χ be the function, holomorphic near $w = 1$,

$$\chi(w) = \frac{\log(w)}{1 - w^{-1}} = 1 + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n(n+1)}(w-1)^n,$$

we may write the right hand side as $\chi(e^{\mathrm{ad}_f})\eta$. By Applying Ad to the defining equation for f we obtain $e^{\mathrm{ad}_f} = e^{\mathrm{ad}_\xi} e^{t\,\mathrm{ad}_\eta}$. Hence

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \chi(e^{\mathrm{ad}_\xi} e^{t\,\mathrm{ad}_\eta})\eta.$$

Finally, integrating from 0 to 1 and using $f(0) = \xi,\ f(1) = \log(\exp(\xi)\exp(\eta))$, we find:

$$\log(\exp(\xi)\exp(\eta)) = \xi + \Big(\int_0^1 \chi(e^{\mathrm{ad}_\xi} e^{t\,\mathrm{ad}_\eta})\mathrm{d}t\Big)\eta.$$

To work out the terms of the series, one puts

$$w - 1 = e^{\mathrm{ad}_\xi} e^{t\,\mathrm{ad}_\eta} - 1 = \sum_{i+j\geq 1} \frac{t^j}{i!j!} \mathrm{ad}_\xi^i \mathrm{ad}_\eta^j$$

in the power series expansion of χ, and integrates the resulting series in t. We arrive at:

Theorem 7.3 (Baker-Campbell-Hausdorff series)**.** *Let G be a Lie group, with exponential map* $\exp\colon \mathfrak{g} \to G$*. For $\xi, \eta \in \mathfrak{g}$ sufficiently small we have the following formula*

$$\log(\exp(\xi)\exp(\eta)) = \xi + \eta + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n(n+1)} \Big(\int_0^1 dt\, \Big(\sum_{i+j\geq 1} \frac{t^j}{i!j!} \mathrm{ad}_\xi^i \mathrm{ad}_\eta^j \Big)^n \Big)\eta.$$

An important point is that the resulting Taylor series in ξ, η is a *Lie series*: all terms of the series are of the form of a constant times $\mathrm{ad}_\xi^{n_1} \mathrm{ad}_\eta^{m_2} \cdots \mathrm{ad}_\xi^{n_r} \eta$. The first few terms read,

$$\log(\exp(\xi)\exp(\eta)) = \xi + \eta + \tfrac{1}{2}[\xi,\eta] + \frac{1}{12}[\xi,[\xi,\eta]] - \frac{1}{12}[\eta,[\xi,\eta]] + \frac{1}{24}[\eta,[\xi,[\eta,\xi]]] + \dots.$$

Exercise 7.4. Work out these terms from the formula.

There is a somewhat better version of the BCH formula, due to Dynkin. A good discussion can be found in the book by Onishchik-Vinberg, Chapter I.3.2.

8. Actions of Lie groups and Lie algebras

8.1. **Lie group actions.**

Definition 8.1. An *action of a Lie group G on a manifold M* is a group homomorphism

$$\mathcal{A}\colon G \to \mathrm{Diff}(M),\ g \mapsto \mathcal{A}_g$$

into the group of diffeomorphisms on M, such that the *action map*

$$G \times M \to M, \quad (g,m) \mapsto \mathcal{A}_g(m)$$

is smooth.

We will often write $g.m$ rather than $\mathcal{A}_g(m)$. With this notation, $g_1.(g_2.m) = (g_1g_2).m$ and $e.m = m$. A map $\Phi\colon M_1 \to M_2$ between G-manifolds is called G-equivariant if $g.\Phi(m) = \Phi(g.m)$ for all $m \in M$, i.e. the following diagram commutes:

$$\begin{array}{ccc} G\times M_1 & \longrightarrow & M_1 \\ \Big\downarrow{\scriptstyle \mathrm{id}\times\Phi} & & \Big\downarrow{\scriptstyle \Phi} \\ G\times M_2 & \longrightarrow & M_2 \end{array}$$

where the horizontal maps are the action maps.

Examples 8.2. (a) An $\mathbb{R}$-action on M is the same thing as a global flow.

(b) The group G acts $M = G$ by right multiplication, $\mathcal{A}_g = R_{g^{-1}}$, left multiplication, $\mathcal{A}_g = L_g$, and by conjugation, $\mathcal{A}_g = \mathrm{Ad}_g = L_g \circ R_{g^{-1}}$. The left and right action commute, hence they define an action of $G\times G$. The conjugation action can be regarded as the action of the diagonal subgroup $G \subset G \times G$.

(c) Any G-representation $G \to \mathrm{End}(V)$ can be regarded as a G-action, by viewing V as a manifold.

(d) For any closed subgroup $H \subset G$, the space of right cosets $G/H = \{gH|\ g \in G\}$ has a unique manifold structure such that the quotient map $G \to G/H$ is a smooth submersion, and the action of G by left multiplication on G descends to a smooth G-action on G/H. (Some ideas of the proof will be explained below.)

(e) The defining representation of the orthogonal group $\mathrm{O}(n)$ on $\mathbb{R}^n$ restricts to an action on the unit sphere S^{n-1}, which in turn descends to an action on the projective space $\mathbb{R}P(n-1)$. One also has actions on the Grassmann manifold $\mathrm{Gr}_{\mathbb{R}}(k,n)$ of k-planes in $\mathbb{R}^n$, on the flag manifold $\mathrm{Fl}(n) \subset \mathrm{Gr}_{\mathbb{R}}(1,n) \times \cdots \mathrm{Gr}_{\mathbb{R}}(n-1,n)$ (consisting of sequences of subspaces $V_1 \subset \cdots V_{n-1} \subset \mathbb{R}^n$ with $\dim V_i = i$), and various types of 'partial' flag manifolds. These examples are all of the form $\mathrm{O}(n)/H$ for various choices of H. (E.g, for $\mathrm{Gr}(k,n)$ one takes H to be the subgroup preserving $\mathbb{R}^k \subset \mathbb{R}^n$.)

8.2. **Lie algebra actions.**

Definition 8.3. An *action of a finite-dimensional Lie algebra* $\mathfrak{g}$ on M is a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{X}(M)$, $\xi \mapsto \mathcal{A}_\xi$ such that the action map

$$\mathfrak{g} \times M \to TM, \quad (\xi, m) \mapsto \mathcal{A}_\xi|_m$$

is smooth.

We will often write $\xi_M =: \mathcal{A}_\xi$ for the vector field corresponding to ξ. Thus, $[\xi_M, \eta_M] = [\xi,\eta]_M$ for all $\xi, \eta \in \mathfrak{g}$. A smooth map $\Phi\colon M_1 \to M_2$ between $\mathfrak{g}$-manifolds is called equivariant if $\xi_{M_1} \sim_\Phi \xi_{M_2}$ for all $\xi \in \mathfrak{g}$, i.e. if the following diagram commutes

$$\begin{array}{ccc} \mathfrak{g}\times M_1 & \longrightarrow & TM_1 \\ \Big\downarrow{\scriptstyle \mathrm{id}\times\Phi} & & \Big\downarrow{\scriptstyle \mathrm{d}\Phi} \\ \mathfrak{g}\times M_2 & \longrightarrow & TM_2 \end{array}$$

where the horizontal maps are the action maps.

Examples 8.4. (a) Any vector field X defines an action of the Abelian Lie algebra $\mathbb{R}$, by $\lambda \mapsto \lambda X$.

(b) Any Lie algebra representation $\phi\colon \mathfrak{g} \to \mathrm{gl}(V)$ may be viewed as a Lie algebra action
$$(\mathcal{A}_\xi f)(v) = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} f(v - t\phi(\xi)v) = -\langle \mathrm{d}_v f,\ \phi(\xi)v\rangle, \quad f \in C^\infty(V)$$
defines a $\mathfrak{g}$-action. Here $\mathrm{d}_v f\colon T_vV \to \mathbb{R}$ is viewed as an element of V^*. Using a basis e_a of V to identify $V = \mathbb{R}^n$, and introducing the components of $\xi \in \mathfrak{g}$ in the representation as $\phi(\xi).e_a = \sum_b \phi(\xi)^b_a e_b$ the generating vector fields are
$$\xi_V = -\sum_{ab} \phi(\xi)^b_a x^a \frac{\partial}{\partial x^b}.$$
Note that the components of the generating vector fields are homogeneous linear functions in x. Any $\mathfrak{g}$-action on V with this property comes from a linear $\mathfrak{g}$-representation.

(c) For any Lie group G, we have actions of its Lie algebra $\mathfrak{g}$ by $\mathcal{A}_\xi = \xi^L$, $\mathcal{A}_\xi = -\xi^R$ and $\mathcal{A}_\xi = \xi^L - \xi^R$.

(d) Given a closed subgroup $H \subset G$, the vector fields $-\xi^R \in \mathfrak{X}(G)$, $\xi \in \mathfrak{g}$ are invariant under the right multiplication, hence they are related under the quotient map to vector fields on G/H. That is, there is a unique $\mathfrak{g}$-action on G/H such that the quotient map $G \to G/H$ is equivariant.

Definition 8.5. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Given a G-action $g \mapsto \mathcal{A}_g$ on M, one defines its *generating vector fields* by
$$\mathcal{A}_\xi = \frac{d}{dt}\Big|_{t=0} \mathcal{A}^*_{\exp(-t\xi)}.$$

Example 8.6. The generating vector field for the action by right multiplication $\mathcal{A}_a = R_{a^{-1}}$ are the left-invariant vector fields,
$$\mathcal{A}_\xi = \frac{\partial}{\partial t}|_{t=0} R^*_{\exp(t\xi)} = \xi^L.$$
Similarly, the generating vector fields for the action by left multiplication $\mathcal{A}_a = L_a$ are $-\xi^R$, and those for the conjugation action $\mathrm{Ad}_a = L_a \circ R_{a^{-1}}$ are $\xi^L - \xi^R$.

Observe that if $\Phi\colon M_1 \to M_2$ is an equivariant map of G-manifolds, then the generating vector fields for the action are Φ-related.

Theorem 8.7. *The generating vector fields of any G-action $g \to \mathcal{A}_g$ define a $\mathfrak{g}$-action $\xi \to \mathcal{A}_\xi$.*

Proof. Write $\xi_M := \mathcal{A}_\xi$ for the generating vector fields of a G-action on M. We have to show that $\xi \mapsto \xi_M$ is a Lie algebra morphism. Note that the action map
$$\Phi\colon G \times M \to M,\ (a, m) \mapsto a.m$$
is G-equivariant, relative to the given G-action on M and the action $g.(a, m) = (ga, m)$ on $G \times M$. Hence $\xi_{G\times M} \sim_\Phi \xi_M$. But $\xi_{G\times M} = -\xi^R$ (viewed as vector fields on the product $G \times M$), hence $\xi \mapsto \xi_{G\times M}$ is a Lie algebra morphism. It follows that
$$0 = [(\xi_1)_{G\times M}, (\xi_1)_{G\times M}] - [\xi_1, \xi_2]_{G\times M} \sim_\Phi [(\xi_1)_M, (\xi_2)_M] - [\xi_1, \xi_2]_M.$$
Since Φ is a surjective submersion (i.e. the differential $\mathrm{d}\Phi\colon T(G \times M) \to TM$ is surjective), this shows that $[(\xi_1)_M, (\xi_2)_M] - [\xi_1, \xi_2]_M = 0$. ☐

8.3. **Integrating Lie algebra actions.** Let us now consider the inverse problem: For a Lie group G with Lie algebra $\mathfrak{g}$, integrating a given $\mathfrak{g}$-action to a G-action. The construction will use some facts about *foliations*.

Let M be a manifold. A *rank k distribution* on M is a $C^\infty(M)$-linear subspace $\mathfrak{R} \subset \mathfrak{X}(M)$ of the space of vector fields, such that at any point $m \in M$, the subspace

$$E_m = \{X_m |\ X \in \mathfrak{R}\}$$

is of dimension k. The subspaces E_m define a rank k vector bundle $E \subset TM$ with $\mathfrak{R} = \Gamma(E)$, hence a distribution is equivalently given by this subbundle E. An *integral submanifold* of the distribution $\mathfrak{R}$ is a k-dimensional submanifold S such that all $X \in \mathfrak{R}$ are tangent to S. In terms of E, this means that $T_m S = E_m$ for all $m \in S$. The distribution is called *integrable* if for all $m \in M$ there exists an integral submanifold containing m. In this case, there exists a maximal such submanifold, $\mathcal{L}_m$. The decomposition of M into maximal integral submanifolds is called a k-dimensional foliation of M, the maximal integral submanifolds themselves are called the *leaves* of the foliation.

Not every distribution is integrable. Recall that if two vector fields are tangent to a submanifold, then so is their Lie bracket. Hence, a *necessary* condition for integrability of a distribution is that $\mathfrak{R}$ is a Lie subalgebra. Frobenius' theorem gives the converse:

Theorem 8.8 (Frobenius theorem)**.** *A rank k distribution $\mathfrak{R} \subset \mathfrak{X}(M)$ is integrable if and only if $\mathfrak{R}$ is a Lie subalgebra.*

The idea of proof is to show that if $\mathfrak{R}$ is a Lie subalgebra, then the $C^\infty(M)$-module $\mathfrak{R}$ is spanned, near any $m \in M$, by k *commuting* vector fields. One then uses the flow of these vector fields to construct integral submanifold.

Exercise 8.9. Prove Frobenius' theorem for distributions $\mathfrak{R}$ of rank $k = 2$. (Hint: If $X \in \mathfrak{R}$ with $X_m \neq 0$, one can choose local coordinates such that $X = \frac{\partial}{\partial x_1}$. Given a second vector field $Y \in \mathfrak{R}$, such that $[X, Y] \in \mathfrak{R}$ and X_m, Y_m are linearly independent, show that one can replace Y by some $Z = aX + bY \in \mathfrak{R}$ such that $b_m \neq 0$ and $[X, Z] = 0$ on a neighborhood of m.)

Exercise 8.10. Give an example of a non-integrable rank 2 distribution on $\mathbb{R}^3$.

Given a Lie algebra of dimension k and a *free* $\mathfrak{g}$-action on M (i.e. $\xi_M|_m = 0$ implies $\xi = 0$), one obtains an integrable rank k distribution $\mathfrak{R}$ as the span (over $C^\infty(M)$) of the ξ_M's. We use this to prove:

Theorem 8.11. *Let G be a connected, simply connected Lie group with Lie algebra $\mathfrak{g}$. A Lie algebra action $\mathfrak{g} \to \mathfrak{X}(M)$, $\xi \mapsto \xi_M$ integrates to an action of G if and only if the vector fields ξ_M are all complete.*

Proof of the theorem. The idea of proof is to express the G-action in terms of a foliation. Given a G-action on M, consider the diagonal G-action on $G \times M$, where G acts on itself by left multiplication. The orbits of this action define a foliation of $G \times M$, with leaves indexed by the elements of m:

$$\mathcal{L}_m = \{(g, g.m) |\, g \in G\}.$$

Let $\mathrm{pr}_1, \mathrm{pr}_2$ the projections from $G \times M$ to the two factors. Then pr_1 restricts to diffeomorphisms $\pi_m \colon \mathcal{L}_m \to G$, and we recover the action as

$$g.m = \mathrm{pr}_2(\pi_m^{-1}(g)).$$

Given a $\mathfrak{g}$-action, our plan is to construct the foliation from an integrable distribution.

Let $\xi \mapsto \xi_M$ be a given $\mathfrak{g}$-action. Consider the diagonal $\mathfrak{g}$ action on $G \times M$,

$$\xi_{G\times M} = (-\xi^R, \xi_M) \in \mathfrak{X}(G \times M).$$

Note that the vector fields $\xi_{\widehat{M}}$ are complete, since it is the sum of commuting vector fields, both of which are complete. If Φ_t^ξ is the flow of ξ_M, the flow of $\xi_{\widehat{M}} = (-\xi^R, \xi_M)$ is given by

$$\widehat{\Phi}_t^\xi = (L_{\exp(t\xi)}, \Phi_t^\xi) \in \mathrm{Diff}(G \times M).$$

The action $\xi \mapsto \xi_{G\times M}$ is free, hence it defines an integrable $\dim G$-dimensional distribution $\mathfrak{R} \subset \mathfrak{X}(G \times M)$. Let $\mathcal{L}_m \hookrightarrow G \times M$ be the unique leaf containing the point (e, m). Projection to the first factor induces a smooth map $\pi_m : \mathcal{L}_m \to G$.

We claim that π_m is *surjective*. To see this, recall that any $g \in G$ can be written in the form $g = \exp(\xi_r) \cdots \exp(\xi_1)$ with $\xi_i \in \mathfrak{g}$. Define $g_0 = e,\ m_0 = m$, and

$$g_i = \exp(\xi_i) \cdots \exp(\xi_1), \quad m_i = (\Phi_1^{\xi_i} \circ \cdots \circ \Phi_1^{\xi_1})(m)$$

for $i = 1, \ldots, r$. Each path

$$\widehat{\Phi}_t^{\xi_i}\big(g_{i-1},\, m_{i-1}\big) = (\exp(t\xi_i) g_{i-1},\ \Phi_t^{\xi_i}(m_{i-1})), \quad t \in [0, 1]$$

connects (g_{i-1}, m_{i-1}) to (g_i, m_i), and stays within a leaf of the foliation (since it is given by the flow). Hence, by concatenation we obtain a (piecewise smooth) path in $\mathcal{L}_m$ connecting (e, m) to $(g_r, m_r) = (g, m_r)$. In particular, $\pi_m^{-1}(g) \neq \emptyset$.

For any $(g, x) \in \mathcal{L}_m$ the tangent map $\mathrm{d}_{(g,x)}\pi_m$ is an isomorphism. Hence $\pi_m \colon \mathcal{L}_m \to G$ is a (surjective) covering map. Since G is simply connected by assumption, we conclude that $\pi_m \colon \mathcal{L}_m \to G$ is a diffeomorphism. We now define $\mathcal{A}_g(m) = \mathrm{pr}_2(\pi_m^{-1}(g))$. Concretely, the construction above shows that if $g = \exp(\xi_r) \cdots \exp(\xi_1)$ then

$$\mathcal{A}_g(m) = (\Phi_1^{\xi_r} \circ \cdots \circ \Phi_1^{\xi_1})(m).$$

From this description it is clear that $\mathcal{A}_{gh} = \mathcal{A}_g \circ \mathcal{A}_h$. □

Let us remark that, in general, one cannot drop the assumption that G is simply connected. Consider for example $G = \mathrm{SU}(2)$, with $\mathfrak{su}(2)$-action $\xi \mapsto -\xi^R$. This exponentiates to an action of $\mathrm{SU}(2)$ by left multiplication. But $\mathfrak{su}(2) \cong \mathfrak{so}(3)$ as Lie algebras, and the $\mathfrak{so}(3)$-action does not exponentiate to an action of the group $\mathrm{SO}(3)$.

As an important special case, we obtain:

Theorem 8.12. *Let H, G be Lie groups, with Lie algebras $\mathfrak{h}, \mathfrak{g}$. If H is connected and simply connected, then any Lie algebra morphism $\phi \colon \mathfrak{h} \to \mathfrak{g}$ integrates uniquely to a Lie group morphism $\psi \colon H \to G$.*

Proof. Define an $\mathfrak{h}$-action on G by $\xi \mapsto -\phi(\xi)^R$. Since the right-invariant vector fields are complete, this action integrates to a Lie group action $\mathcal{A} \colon H \to \mathrm{Diff}(G)$. This action commutes with the action of G by right multiplication. Hence, $\mathcal{A}_h(g) = \psi(h)g$ where $\psi(h) = \mathcal{A}_h(e)$. The action property now shows $\psi(h_1)\psi(h_2) = \psi(h_1 h_2)$, so that $\psi \colon H \to G$ is a Lie group morphism integrating ϕ. □

Corollary 8.13. *Let G be a connected, simply connected Lie group, with Lie algebra $\mathfrak{g}$. Then any $\mathfrak{g}$-representation on a finite-dimensional vector space V integrates to a G-representation on V.*

Proof. A $\mathfrak{g}$-representation on V is a Lie algebra morphism $\mathfrak{g} \to \mathfrak{gl}(V)$, hence it integrates to a Lie group morphism $G \to \mathrm{GL}(V)$. □

Definition 8.14. A *Lie subgroup* of a Lie group G is a subgroup $H \subset G$, equipped with a Lie group structure such that the inclusion is a morphism of Lie groups (i.e. , is smooth).

Note that a Lie subgroup need not be closed in G, since the inclusion map need not be an embedding. Also, the one-parameter subgroups $\phi\colon \mathbb{R} \to G$ need not be subgroups (strictly speaking) since ϕ need not be injective.

Proposition 8.15. *Let G be a Lie group, with Lie algebra $\mathfrak{g}$. For any Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ there is a unique connected Lie subgroup H of G such that the differential of the inclusion $H \hookrightarrow G$ is the inclusion $\mathfrak{h} \hookrightarrow \mathfrak{g}$.*

Proof. Consider the distribution on G spanned by the vector fields $-\xi^R$, $\xi \in \mathfrak{g}$. It is integrable, hence it defines a foliation of G. The leaves of any foliation carry a unique manifold structure such that the inclusion map is smooth. Take $H \subset G$ to be the leaf through $e \in H$, with this manifold structure. Explicitly,

$$H = \{g \in G |\ g = \exp(\xi_r) \cdots \exp(\xi_1), \xi_i \in \mathfrak{h}\}.$$

From this description it follows that H is a Lie group. □

By Ado's theorem, any finite-dimensional Lie algebra $\mathfrak{g}$ is isomorphic to a matrix Lie algebra. We will skip the proof of this important (but relatively deep) result, since it involves a considerable amount of structure theory of Lie algebras. Given such a presentation $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{R})$, the Lemma gives a Lie subgroup $G \subset \mathrm{GL}(n, \mathbb{R})$ integrating $\mathfrak{g}$. Replacing G with its universal covering, this proves:

Theorem 8.16 (Lie's third theorem)**.** *For any finite-dimensional real Lie algebra $\mathfrak{g}$, there exists a connected, simply connected Lie group G, unique up to isomorphism, having $\mathfrak{g}$ as its Lie algebra.*

The book by Duistermaat-Kolk contains a different, more conceptual proof of Cartan's theorem. This new proof has found important generalizations to the integration of *Lie algebroids*. In conjunction with the previous Theorem, Lie's third theorem gives an equivalence between the categories of finite-dimensional Lie algebras $\mathfrak{g}$ and connected, simply-connected Lie groups G.

9. Universal covering groups

Given a connected topological space X with base point x_0, one defines the covering space $\widetilde{X}$ as equivalence classes of paths $\gamma\colon [0,1] \to X$ with $\gamma(0) = x_0$. Here the equivalence is that of homotopy relative to fixed endpoints. The map taking $[\gamma]$ to $\gamma(1)$ is a covering $p\colon \widetilde{X} \to X$. The covering space carries an action of the fundamental group $\pi_1(X)$, given as equivalence classes of paths with $\gamma(1) = x_0$, i.e. $\pi_1(X) = p^{-1}(x_0)$. The group structure is given by concatenation of paths

$$(\gamma_1 * \gamma_2)(t) = \begin{cases} \gamma_1(2t) & 0 \le t \le \frac{1}{2}, \\ \gamma_2(2t-1) & \frac{1}{2} \le t \le 1 \end{cases},$$

i.e. $[\gamma_1][\gamma_2] = [\gamma_1 * \gamma_2]$ (one shows that this is well-defined). If $X = M$ is a manifold, then $\pi_1(X)$ acts on $\widetilde{X}$ by *deck transformations*, this action is again induced by concatenation of paths:

$$\mathcal{A}_{[\lambda]}([\gamma]) = [\lambda * \gamma].$$

A continuous map of connected topological spaces $\Phi\colon X \to Y$ taking x_0 to the base point y_0 lifts to a continuous map $\widetilde{\Phi}\colon \widetilde{X} \to \widetilde{Y}$ of the covering spaces, by $\widetilde{\Phi}[\gamma] = [\Phi \circ \gamma]$, and it induces a group morphism $\pi_1(X) \to \pi_1(Y)$.

If $X = M$ is a manifold, then $\widetilde{M}$ is again a manifold, and the covering map is a local diffeomorphism. For a smooth map $\Phi\colon M \to N$ of manifolds, the induced map $\widetilde{\Phi}\colon \widetilde{M} \to \widetilde{N}$ of coverings is again smooth. This construction is functorial, i.e. $\widetilde{\Psi \circ \Phi} = \widetilde{\Psi} \circ \widetilde{\Phi}$. We are interested in the case of connected Lie groups G. In this case, the natural choice of base point is the group unit $x_0 = e$. We have:

Theorem 9.1. *The universal covering $\widetilde{G}$ of a connected Lie group G is again a Lie group, and the covering map $p\colon \widetilde{G} \to G$ is a Lie group morphism. The group $\pi_1(G) = p^{-1}(\{e\})$ is a subgroup of the center of $\widetilde{G}$.*

Proof. The group multiplication and inversion lifts to smooth maps $\widetilde{Mult}\colon \widetilde{G \times G} = \widetilde{G} \times \widetilde{G} \to \widetilde{G}$ and $\widetilde{\mathrm{Inv}}\colon \widetilde{G} \to \widetilde{G}$. Using the functoriality properties of the universal covering construction, it is clear that these define a group structure on $\widetilde{G}$. A proof that $\pi_1(G)$ is central is outlined in the following exercise. □

Exercise 9.2. Recall that a subgroup $H \subset G$ is normal in G if $\mathrm{Ad}_g(H) \subset H$ for all $g \in G$.

a) Let G be a connected Lie group, and $H \subset G$ a normal subgroup that is discrete (i.e. 0-dimensional). Show that H is a subgroup of the center of G.

b) Prove that the kernel of a Lie group morphism $\phi\colon G \to G'$ is a closed normal subgroup.

The combination of these two facts shows that if a Lie group morphism is a covering, then its kernel is a central subgroup.

Example 9.3. The universal covering group of the circle group $G = \mathrm{U}(1)$ is the additive group $\mathbb{R}$.

Example 9.4. $\mathrm{SU}(2)$ is the universal covering group of $\mathrm{SO}(3)$, and $\mathrm{SU}(2)\times\mathrm{SU}(2)$ is the universal covering group of $\mathrm{SO}(4)$. In both cases, the group of deck transformations is $\mathbb{Z}_2$.

For all $n \geq 3$, the fundamental group of $\mathrm{SO}(n)$ is $\mathbb{Z}_2$. The universal cover is called the *Spin group* and is denoted $\mathrm{Spin}(n)$. We have seen that $\mathrm{Spin}(3) \cong \mathrm{SU}(2)$ and $\mathrm{Spin}(4) \cong \mathrm{SU}(2)\times\mathrm{SU}(2)$. One can also show that $\mathrm{Spin}(5) \cong \mathrm{Sp}(2)$ and $\mathrm{Spin}(6) = \mathrm{SU}(4)$. (See e.g. by lecture notes on 'Lie groups and Clifford algebras', Section III.7.6.) Starting with $n = 7$, the spin groups are 'new'.

We will soon prove that the universal covering group $\widetilde{G}$ of a Lie group G is compact if and only if G is compact with finite center.

If $\Gamma \subset \pi_1(G)$ is any subgroup, then Γ (viewed as a subgroup of $\widetilde{G}$) is central, and so $\widetilde{G}/\Gamma$ is a Lie group covering G, with $\pi_1(G)/\Gamma$ as its group of deck transformations.

10. The universal enveloping algebra

As we had seen any algebra[3] $\mathcal{A}$ can be viewed as a Lie algebra, with Lie bracket the commutator. This correspondence defines a functor from the category of algebras to the category of Lie algebras. There is also a functor in the opposite direction, associating to any Lie algebra an algebra.

Definition 10.1. The universal enveloping algebra of a Lie algebra $\mathfrak{g}$ is the algebra $U(\mathfrak{g})$, with generators $\xi \in \mathfrak{g}$ and relations, $\xi_1\xi_2 - \xi_2\xi_1 = [\xi_1, \xi_2]$.

Elements of the enveloping algebra are linear combinations words $\xi_1 \cdots \xi_r$ in the Lie algebra elements, using the relations to manipulate the words. Here we are implicitly using that the relations don't annihilate any Lie algebra elements, i.e. that the map $\mathfrak{g} \to U(\mathfrak{g}),\ \xi \mapsto \xi$ is injective. This will be justified by the Poincaré-Birkhoff-Witt theorem to be discussed below.

Example 10.2. Let $\mathfrak{g} \cong \mathfrak{sl}(2,\mathbb{R})$ be the Lie algebra with basis e, f, h and brackets

$$[e,f] = h,\ [h,e] = 2e,\ [h,f] = -2f.$$

It turns out that any element of $U(\mathfrak{sl}(2,\mathbb{R}))$ can be written as a sum of products of the form $f^k h^l e^m$ for some $k, l, m \geq 0$. Let us illustrate this for the element ef^2 (just to get used to some calculations in the enveloping algebra). We have

$$\begin{aligned} ef^2 &= [e, f^2] + f^2 e \\ &= [e,f]f + f[e,f] + f^2 e \\ &= hf + fh + f^2 e \\ &= [h,f] + 2fh + f^2 e \\ &= -2f + 2fh + f^2 e. \end{aligned}$$

More formally, the universal enveloping algebra is the quotient of the tensor algebra $T(\mathfrak{g})$ by the two-sided ideal $\mathcal{I}$ generated by all $\xi_1 \otimes \xi_2 - \xi_2 \otimes \xi_1 - [\xi_1, \xi_2]$. The inclusion map $\mathfrak{g} \hookrightarrow T(\mathfrak{g})$ descends to a map $j\colon \mathfrak{g} \to \mathrm{U}(\mathfrak{g})$. By construction, this map j is a Lie algebra morphism.

The construction of the enveloping algebra $U(\mathfrak{g})$ from a Lie algebra $\mathfrak{g}$ is functorial: Any Lie algebra morphism $\mathfrak{g}_1 \to \mathfrak{g}_2$ induces a morphism of algebras $U(\mathfrak{g}_1) \to U(\mathfrak{g}_2)$, in a way compatible with the composition of morphisms. As a special case, the zero map $\mathfrak{g} \to 0$ induces an algebra morphism $U(\mathfrak{g}) \to \mathbb{R}$, called the *augmentation map.*

Theorem 10.3 (Universal property)**.** *If $\mathcal{A}$ is an associative algebra, and $\kappa\colon \mathfrak{g} \to \mathcal{A}$ is a homomorphism of Lie algebras, then there is a unique morphism of algebras $\kappa_U\colon U(\mathfrak{g}) \to \mathcal{A}$ such that $\kappa = \kappa_U \circ j$.*

Proof. The map κ extends to an algebra homomorphism $T(\mathfrak{g}) \to \mathcal{A}$. This algebra homomorphism vanishes on the ideal $\mathcal{I}$, and hence descends to an algebra homomorphism $\kappa_U\colon U(\mathfrak{g}) \to \mathcal{A}$ with the desired property. This extension is unique, since $j(\mathfrak{g})$ generates $U(\mathfrak{g})$ as an algebra. □

By the universal property, any Lie algebra representation $\mathfrak{g} \to \mathrm{End}(V)$ extends to a representation of the algebra $U(\mathfrak{g})$. Conversely, given an algebra representation $U(\mathfrak{g}) \to \mathrm{End}(V)$

[3] Unless specified differently, we take algebra to mean associative algebra with unit.

one obtains a $\mathfrak{g}$-representation by restriction. That is, there is a 1-1 correspondence between Lie algebra representations of $\mathfrak{g}$ and algebra representations of $U(\mathfrak{g})$.

Let $\mathrm{Cent}(U(\mathfrak{g}))$ be the center of the enveloping algebra. Given a $\mathfrak{g}$-representation $\pi\colon \mathfrak{g} \to \mathrm{End}(V)$, the operators $\pi(x),\ x \in \mathrm{Cent}(U(\mathfrak{g}))$ commute with all $\pi(\xi), \xi \in \mathfrak{g}$:

$$[\pi(x), \pi(\xi)] = \pi([x, \xi]) = 0.$$

It follows that the eigenspaces of $\pi(x)$ for $x \in \mathrm{Cent}(U(\mathfrak{g}))$ are $\mathfrak{g}$-invariant.

Exercise 10.4. Let $\mathfrak{g} \cong \mathfrak{sl}(2, \mathbb{R})$ be the Lie algebra with basis e, f, h and brackets $[e, f] = h,\ [h, e] = 2e,\ [h, f] = -2f$. Show that

$$x = 2fe + \frac{1}{2}h^2 + h \in U(\mathfrak{sl}(2, \mathbb{R}))$$

lies in the center of the enveloping algebra.

The construction of the enveloping algebra works for any Lie algebra, possibly of infinite dimension. It is a non-trivial fact that the map j is always an inclusion. This is usually obtained as a corollary to the Poincaré-Birkhoff-Witt theorem. The statement of this Theorem is as follows. Note that $U(\mathfrak{g})$ has a filtration

$$\mathbb{R} = U^{(0)}(\mathfrak{g}) \subset U^{(1)}(\mathfrak{g}) \subset U^{(2)}(\mathfrak{g}) \subset \cdots,$$

where $U^{(k)}(\mathfrak{g})$ consists of linear combinations of products of at most k elements in $\mathfrak{g}$. That is, $U^{(k)}(\mathfrak{g})$ is the image of $T^{(k)}(\mathfrak{g}) = \bigoplus_{i \le k} T^i(\mathfrak{g})$.

The filtration is compatible with the product, i.e. the product of an element of filtration degree k with an element of filtration degree l has filtration degree $k + l$. Let

$$\mathrm{gr}(U(\mathfrak{g})) = \bigoplus_{k=0}^{\infty} \mathrm{gr}^k(U(\mathfrak{g}))$$

be the associated graded algebra, where $\mathrm{gr}^k(U(\mathfrak{g})) = U^{(k)}(\mathfrak{g})/U^{(k-1)}(\mathfrak{g})$.

Lemma 10.5. *The associated graded algebra* $\mathrm{gr}(U(\mathfrak{g}))$ *is commutative. Hence, the map* $j\colon \mathfrak{g} \to U(\mathfrak{g})$ *defines an algebra morphism*

$$j_S\colon S(\mathfrak{g}) \to \mathrm{gr}(U(\mathfrak{g}))$$

Proof. If $x = \xi_1 \cdots \xi_k \in U^{(k)}(\mathfrak{g})$, and $x' = \xi_{s(1)} \cdots \xi_{s(k)}$ for some permutation s, then $x' - x \in U^{(k-1)}(\mathfrak{g})$. (If s is a transposition of adjacent elements this is immediate from the definition; but general permutations are products of such transpositions.) As a consequence, the products of two elements of filtration degrees k, l is independent of their order modulo terms of filtration degree $k + l - 1$. Equivalently, the associated graded algebra is commutative. □

Explicitly, the map is the direct sum over all

$$j_S\colon S^k(\mathfrak{g}) \to U^{(k)}(\mathfrak{g})/U^{(k-1)}(\mathfrak{g}),\ \xi_1 \cdots \xi_k \mapsto \xi_1 \cdots \xi_k \mod U^{(k-1)}(\mathfrak{g}).$$

Note that the map j_S is surjective: Given $y \in U^{(k)}(\mathfrak{g})/U^{(k-1)}(\mathfrak{g})$, choose a lift $\tilde{y} \in U^{(k)}(\mathfrak{g})$ given as a linear combination of k-fold products of elements in $\mathfrak{g}$. The same linear combination, with the product now interpreted in the symmetric algebra, defines an element $x \in S^k(\mathfrak{g})$ with $j_S(x) = y$. The following important result states that j_S is also injective.

Theorem 10.6 (Poincaré-Birkhoff-Witt theorem)**.** *The map*

$$j_S\colon S\mathfrak{g} \to \mathrm{gr}(U\mathfrak{g})$$

is an isomorphism of algebras.

Corollary 10.7. *The map $j\colon \mathfrak{g} \to U(\mathfrak{g})$ is injective.*

Corollary 10.8. *Suppose $f\colon S\mathfrak{g} \to U(\mathfrak{g})$ is a filtration preserving linear map whose associated graded map $\mathrm{gr}(f)\colon S\mathfrak{g} \to \mathrm{gr}(U(\mathfrak{g}))$ coincides with j_S. Then f is an isomorphism.*

Indeed, a map of filtered vector spaces is an isomorphism if and only if the associated graded map is an isomorphism. One typical choice of f is *symmetrization*, characterized as the unique linear map $\mathrm{sym}\colon S(\mathfrak{g}) \to U(\mathfrak{g})$ such that $\mathrm{sym}(\xi^k) = \xi^k$ for all k. That is,

$$\mathrm{sym}(\xi_1, \ldots, \xi_k) = \frac{1}{k!}\sum_{s\in S_k} \xi_{s(1)} \cdots \xi_{s(k)};$$

for example,

$$\mathrm{sym}(\xi_1\xi_2) = \tfrac{1}{2}(\xi_1\xi_2 + \xi_2\xi_1) = \xi_1\xi_2 - \tfrac{1}{2}[\xi_1, \xi_2].$$

Corollary 10.9. *The symmetrization map $\mathrm{sym}\colon S(\mathfrak{g}) \to U(\mathfrak{g})$ is an isomorphism of vector spaces.*

Another choice for f is to pick a basis $e_1, \ldots, e_n$ of $\mathfrak{g}$, and define f by

$$f(e_1^{i_1} \cdots e_n^{i_n}) = e_1^{i_1} \cdots e_n^{i_n}.$$

Hence we obtain,

Corollary 10.10. *If $e_1, \ldots, e_n$ is a basis of $\mathfrak{g}$, the products $e_1^{i_1} \cdots e_n^{i_n} \in U(\mathfrak{g})$ with $i_j \geq 0$ form a basis of $U(\mathfrak{g})$.*

Corollary 10.11. *Suppose $\mathfrak{g}_1, \mathfrak{g}_2$ are two Lie subalgebras of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ as vector spaces. Then the multiplication map*

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

is an isomorphism of vector spaces.

Indeed, the associated graded map is the multiplication $S(\mathfrak{g}_1) \otimes S(\mathfrak{g}_2) \to S(\mathfrak{g})$, which is well-known to be an isomorphism. The following is left as an exercise:

Corollary 10.12. *The algebra $U(\mathfrak{g})$ has no (left or right) zero divisors.*

We will give a proof of the PBW theorem for the special case that $\mathfrak{g}$ is the Lie algebra of a Lie group G. (In particular, $\mathfrak{g}$ is finite-dimensional.) The idea is to relate the enveloping algebra to differential operators on G. For any manifold M, let

$$\mathrm{DO}^{(k)}(M) = \{D \in \mathrm{End}(C^\infty(M)) |\ \forall f_0, \ldots, f_k \in C^\infty(M),\ \mathrm{ad}_{f_0} \cdots \mathrm{ad}_{f_k} D = 0\}$$

be the differential operators of degree k on M. Here $\mathrm{ad}_f = [f, \cdot]$ is commutator with the operator of multiplication by f. [4] By polarization, $D \in \mathrm{DO}^{(k)}(M)$ if and only if $\mathrm{ad}_f^{k+1} D = 0$ for all f.

[4] In local coordinates, such operators are of the form

$$D = \sum_{i_1+\ldots+i_n\leq k} a_{i_1\cdots i_n} \frac{\partial^{i_1}}{\partial x_1^{i_1}} \cdots \frac{\partial^{i_n}}{\partial x_n^{i_n}}$$

Remark 10.13. We have $\mathrm{DO}^{(0)}(M) \cong C^\infty(M)$ by the map $D \mapsto D(1)$. Indeed, for $D \in \mathrm{DO}^{(0)}(M)$ we have $D(f) = D(f\cdot 1) = [D,f]1 + fD(1) = fD(1)$. Similarly

$$\mathrm{DO}^{(1)}(M) \cong C^\infty(M) \oplus \mathfrak{X}(M),$$

where function component of D is $D(1)$ and the vector field component is $[D,\cdot]$. Note that $[D,\cdot]$ is a vector field since $[D, f_1 f_2] = [D,f_1]f_2 + f_1[D,f_2]$ with $[D,f_i] \in C^\infty(M)$. The isomorphism follows from $D(f) = D(f\cdot 1) = [D,f]\cdot 1 + f\,D(1)$.

The algebra $\mathrm{DO}(M)$ given as the union over all $\mathrm{DO}^{(k)}(M)$ is a filtered algebra: the product of operators of degree k,l has degree $k+l$. Let $\mathrm{gr}(\mathrm{DO}(M))$ be the associated graded algebra.

Proof of the PBW theorem. (for the special case that $\mathfrak{g}$ is the Lie algebra of a Lie group G). The map $\kappa\colon \mathfrak{g} \to \mathrm{DO}(G), \xi \mapsto \xi^L$ is a Lie algebra morphism, hence by the universal property it extends to an algebra morphism

$$\kappa_U\colon U(\mathfrak{g}) \to \mathrm{DO}(G).$$

The map κ_U preserves filtrations Let $S\mathfrak{g} \cong \mathrm{Pol}(\mathfrak{g}^*),\ x \mapsto p_x$ be the identification with the algebra of polynomials on $\mathfrak{g}^*$, in such a way that $x = \xi_1 \cdots \xi_k \in S^k(\mathfrak{g})$ corresponds to the polynomial $p_x(\mu) = k!\langle \mu, \xi_1\rangle \cdots \langle \mu, \xi_k\rangle$.

Given $x \in S^k(\mathfrak{g})$, $\mu \in \mathfrak{g}^*$, choose $f \in C^\infty(G)$ and $y \in U^{(k)}(\mathfrak{g})$ such that

$$\mu = \mathrm{d}_e f\colon \mathfrak{g} \to \mathbb{R},\quad j_S(x) = y \mod U^{(k)}(\mathfrak{g}).$$

The differential operator $D = \kappa_U(y) \in \mathrm{DO}^{(k)}(G)$ satisfies

$$\mathrm{ad}_f^k(D)|_e = (-1)^k p_x(\mu),$$

by the calculation

$$\mathrm{ad}_f^k(\xi_1^L \cdots \xi_k^L)|_e = (-1)^k k! \xi_1^L(f) \cdots \xi_k^L(f)|_e = (-1)^k k!\ \langle \xi_1, \mu\rangle \cdots \langle \xi_k, \mu\rangle.$$

Hence, if $j_S(x) = 0$ so that $y \in U^{(k-1)}(\mathfrak{g})$ and hence $D \in \mathrm{DO}^{(k-1)}(G)$, i.e. $\mathrm{ad}_f^k(D) = 0$, we find $p_x = 0$ and therefore $x = 0$. □

Exercise 10.14. For any manifold M, the inclusion of vector fields is a Lie algebra morphism $\mathfrak{X}(M) = \Gamma(TM) \to \mathrm{DO}(M)$. Hence it extends to an algebra morphism $U(\mathfrak{X}(M)) \to \mathrm{DO}(M)$, which in turn gives maps

$$S^k(\mathfrak{X}(M)) \to \mathrm{gr}^k(U(\mathfrak{X}(M))) \to \mathrm{gr}^k(\mathrm{DO}(M)).$$

Show that this map descends to a map $\Gamma(S^k(TM)) \to \mathrm{gr}^k(\mathrm{DO}(M))$. Consider $\mathrm{ad}_f^k(D)$ to construct an inverse map, thus proving

$$\Gamma(S(TM)) \cong \mathrm{gr}(\mathrm{DO}(M)).$$

(This is the *principal symbol* isomorphism.)

The following is a consequence of the proof, combined with the exercise.

---

with smooth functions $a_{i_1\cdots i_n}$, but for our purposes the abstract definition will be more convenient.

Theorem 10.15. *For any Lie group G, with Lie algebra $\mathfrak{g}$, the map $\xi \mapsto \xi^L$ extends to an isomorphism*

$$U(\mathfrak{g}) \to \mathrm{DO}^L(G)$$

where $\mathrm{DO}^L(G)$ *is the algebra of left-invariant differential operators on G.*

11. Representation theory of $\mathfrak{sl}(2, \mathbb{C})$.

11.1. **Basic notions.** Until now, we have mainly considered real Lie algebras. However, the definition makes sense for any field, and in particular, we can consider *complex Lie algebras.* Given a real Lie algebra $\mathfrak{g}$, its complexification $\mathfrak{g} \otimes \mathbb{C}$ is a complex Lie algebra. Consider in particular the real Lie algebra $\mathfrak{su}(n)$. Its complexification is the Lie algebra $\mathfrak{sl}(n, \mathbb{C})$. Indeed,

$$\mathfrak{sl}(n, \mathbb{C}) = \mathfrak{su}(n) \oplus i\mathfrak{su}(n)$$

is the decomposition of a trace-free complex matrix into its skew-adjoint and self-adjoint part.

Remark 11.1. Of course, $\mathfrak{sl}(n, \mathbb{C})$ is also the complexification of $\mathfrak{sl}(n, \mathbb{R})$. We have encountered a similar phenomenon for the symplectic groups: The complexification of $\mathfrak{sp}(n)$ is $\mathfrak{sp}(n, \mathbb{C})$, which is also the complexification of $\mathfrak{sp}(n, \mathbb{R})$.

We will be interested in representations of Lie algebra $\mathfrak{g}$ on *complex* vector spaces V, i.e. Lie algebra morphisms $\mathfrak{g} \to \mathrm{End}_{\mathbb{C}}(V)$. Equivalently, this amounts to a morphism of complex Lie algebras $\mathfrak{g} \otimes \mathbb{C} \to \mathrm{End}_{\mathbb{C}}(V)$. If V is obtained by complexification of a real $\mathfrak{g}$-representation, then V carries an $\mathfrak{g}$-equivariant conjugate linear complex conjugation map $C\colon V \to V$. Conversely, we may think of real $\mathfrak{g}$-representations as complex $\mathfrak{g}$-representations with the additional structure of a $\mathfrak{g}$-equivariant conjugate linear automorphism of V.

Given a $\mathfrak{g}$-representation on V, one obtains representations on various associated spaces. For instance, one has a representation on the symmetric power $S^k(V)$ by

$$\pi(\xi)(v_1 \cdots v_k) = \sum_{j=1}^{k} v_1 \cdots \pi(\xi) v_j \cdots v_k$$

and on the exterior power $\wedge^k(V)$ by

$$\pi(\xi)(v_1 \wedge \cdots \wedge v_k) = \sum_{j=1}^{k} v_1 \wedge \cdots \pi(\xi) v_j \cdots \wedge v_k.$$

A similar formula gives a representation on the tensor powers $\bigotimes^k V$, and both $S^k(V), \wedge^k(V)$ are quotients of the representation on $\bigotimes^k(V)$. One also obtains a *dual* representation on $V^* = \mathrm{Hom}(V, \mathbb{C})$, by

$$\langle \pi(\xi)\alpha, v\rangle = -\langle \alpha, \pi(\xi)(v)\rangle, \quad \alpha \in V^*,\ v \in V.$$

A $\mathfrak{g}$-representation $\pi\colon \mathfrak{g} \to \mathrm{End}_{\mathbb{C}}(V)$ is called *irreducible* if there are no subrepresentations other than V or 0. It is *completely reducible* if it decomposes as a sum of irreducible summands. An example of a representation that is *not* completely reducible is the representation of $\mathfrak{g}$ on $U(\mathfrak{g})$ given by left multiplication.

11.2. $\mathfrak{sl}(2,\mathbb{C})$**-representations.** We are interested in the irreducible representations of $\mathfrak{sl}(2,\mathbb{C})$ (or equivalently, the irreducible complex representations of $\mathfrak{su}(2)$ or $\mathfrak{sl}(2,\mathbb{R})$, or of the corresponding simply connected Lie groups). Let e, f, h be the basis of $\mathfrak{sl}(2,\mathbb{C})$ introduced above. We have the defining representation of $\mathfrak{sl}(2,\mathbb{C})$ on $\mathbb{C}^2$, given in the standard basis $\epsilon_1, \epsilon_2 \in \mathbb{C}^2$ as follows:

$$\begin{aligned} \pi(e)\epsilon_1 &= 0, & \pi(e)\epsilon_2 &= \epsilon_1 \\ \pi(h)\epsilon_1 &= \epsilon_1, & \pi(h)\epsilon_2 &= -\epsilon_2 \\ \pi(f)\epsilon_1 &= \epsilon_2, & \pi(f)\epsilon_2 &= 0. \end{aligned}$$

It extends to a representation on the symmetric algebra $S(\mathbb{C}^2) = \bigoplus_{k\geq 0} S^k(\mathbb{C}^2)$ by derivations, preserving each of the summands $S^k(\mathbb{C}^2)$. Equivalently, this is the action on the space of homogeneous polynomials of degree k on $(\mathbb{C}^2)^* \cong \mathbb{C}^2$. Introduce the basis

$$v_j = \frac{1}{(k-j)!j!}\epsilon_1^{k-j}\epsilon_2^j, \ j = 0, \dots, k$$

of $S^k(\mathbb{C}^2)$. Then the resulting action reads,

$$\begin{aligned} \pi(f)v_j &= (j+1)v_{j+1}, \\ \pi(h)v_j &= (k-2j)v_j, \\ \pi(e)v_j &= (k-j+1)v_{j-1} \end{aligned}$$

with the convention $v_{k+1} = 0, \ v_{-1} = 0$.

Proposition 11.2. *The representations of* $\mathfrak{sl}(2,\mathbb{C})$ *on* $V(k) := S^k(\mathbb{C}^2)$ *are irreducible.*

Proof. The formulas above show that $\pi(h)$ has $k+1$ distinct eigenvalues $k, k-2, \dots, -k$, with corresponding eigenvectors $v_0, \dots, v_k$. Furthermore, if $0 \leq j < k$ the operator $\pi(f)$ restricts to an isomorphism $\pi(f)\colon \operatorname{span}(v_j) \to \operatorname{span}(v_{j+1})$, while for $0 < j \leq k$ the operator $\pi(e)$ restricts to an isomorphism $\pi(e)\colon \operatorname{span}(v_j) \to \operatorname{span}(v_{j-1})$.

Hence, if an invariant subspace contains one of the v_j's it must contain all of the v_j's and hence be equal to $V(k)$. But on any non-zero invariant subspace, $\pi(h)$ must have at least one eigenvector. This shows that any non-zero invariant subspace is equal to $V(k)$. □

Theorem 11.3. *Up to isomorphism, the representations* $V(k) = S^k(\mathbb{C}^2), \ k = 0, 1, \dots$ *are the unique* $k+1$*-dimensional irreducible representations of* $\mathfrak{sl}(2,\mathbb{C})$.

Proof. Let V be an irreducible $\mathfrak{sl}(2,\mathbb{C})$-representation. For any $s \in \mathbb{C}$, let

$$V_{[s]} = \ker(\pi(h) - s)$$

If $v \in V_{[s]}$ then

$$\pi(h)\pi(e)v = \pi([h,e])v + \pi(e)\pi(h)v = 2\pi(e)v + s\pi(e)v = (s+2)\pi(e)v.$$

Thus

$$\pi(e)\colon V_{[s]} \to V_{[s+2]},$$

and similarly

$$\pi(f)\colon V_{[s]} \to V_{[s-2]}.$$

Since $\dim V < \infty$, there exists $\lambda \in \mathbb{C}$ such that $V_{[\lambda]} \neq 0$ but $V_{[\lambda+2]} = 0$. Pick a non-zero $v_0 \in V_{[\lambda]}$, and put $v_j = \frac{1}{j!}\pi(f)^j v_0 \in V_{[\lambda-2j]}$, $j = 0, 1, \ldots$. Then

$$\pi(h)v_j = (\lambda - 2j)v_j, \quad \pi(f)v_j = (j+1)v_{j+1}.$$

We will show by induction that

$$\pi(e)v_j = (\lambda + 1 - j)v_{j-1}$$

with the convention $v_{-1} = 0$. Indeed, if the formula holds for an index $j \geq 0$ then

$$\begin{aligned}
\pi(e)v_{j+1} &= \frac{1}{j+1}\pi(e)\pi(f)v_j \\
&= \frac{1}{j+1}(\pi([e,f])v_j + \pi(f)\pi(e)v_j) \\
&= \frac{1}{j+1}(\pi(h)v_j + (\lambda+1-j)\pi(f)v_{j-1}) \\
&= \frac{1}{j+1}((\lambda-2j)v_j + (\lambda+1-j)j\, v_j) \\
&= \frac{1}{j+1}(\lambda(j+1) - j - j^2)v_j \\
&= (\lambda - j)v_j
\end{aligned}$$

which is the desired identity for $j+1$. We see that the span of the v_j is an invariant subspace, hence is all of V. Then non-zero v_j are linearly independent (since they lie in different eigenspaces for $\pi(h)$. Thus $v_0, \ldots, v_k$ is a basis of V, where $k = \dim V - 1$. In particular, $v_{k+1} = 0$. Putting $j = k+1$ in the formula for $\pi(e)v_j$, we obtain $0 = (\lambda - k)v_k$, hence $\lambda = k$. □

Remark 11.4. For any complex number $\lambda \in \mathbb{C}$, we obtain an infinite-dimensional representation $L(\lambda)$ of $\mathfrak{sl}(2,\mathbb{C})$ on $\mathrm{span}(w_0, w_1, w_2, \ldots)$, by the formulas

$$\pi(f)w_j = (j+1)w_{j+1}, \quad \pi(h)w_j = (\lambda - 2j)w_j, \quad \pi(e)w_j = (\lambda - j + 1)w_{j-1}$$

This representation is called the *Verma module* of highest weight λ. If $\lambda = k \in \mathbb{Z}_{\geq 0}$, this representation $L(k)$ has a subrepresentation $L'(k)$ spanned by $w_{k+1}, w_{k+2}, w_{k+3}, \ldots$, and

$$V(k) = L(k)/L'(k)$$

is the quotient module.

Exercise 11.5. Show that for $\lambda \notin \mathbb{Z}_{\geq 0}$, the Verma module is irreducible.

Proposition 11.6. *The* Casimir element $\mathrm{Cas} = 2fe + \frac{1}{2}h^2 + h \in U(\mathfrak{sl}(2,\mathbb{C}))$ *acts as a scalar* $\frac{1}{2}k(k+2)$ *on* $V(k)$.

Proof. We have observed earlier that elements of the center of the enveloping algebras act as scalars on irreducible representations. In this case, it is thus enough to evaluate its action on v_0. Since $\pi(h)v_0 = kv_0$ and $\pi(e)v_0 = 0$ the scalar is $\frac{1}{2}k^2 + k$. □

The representation theory of $\mathfrak{sl}(2,\mathbb{C})$ is finished off with the following result.

Theorem 11.7. *Any finite-dimensional representation* V *of* $\mathfrak{sl}(2,\mathbb{C})$ *is completely reducible.*

Equivalently, the Theorem says that any invariant subspace $V' \subset V$ has an invariant complement. We will first prove this result for the case that V' has codimension 1 in V.

Lemma 11.8. *Suppose V is a finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-representation, and $V' \subset V$ an invariant subspace with $\dim(V/V') = 1$. Then V' admits an invariant complement.*

Proof. Suppose first that V' admits a proper subrepresentation $V_1' \subset V'$. Then V'/V_1' has codimension 1 in V/V_1', hence by induction it admits an invariant complement W/V_1'. Now V_1' has codimension 1 in W, hence, using induction again it admits an invariant complement $T \subset W$. Then T is an invariant complement to V' in V.

We have thus reduced to the case that V' has no proper subrepresentation, i.e. V' is irreducible. There are now two subcases. Case 1: $V' \cong V(0)$ is the trivial $\mathfrak{sl}(2,\mathbb{C})$-representation. Since $[\mathfrak{sl}(2,\mathbb{C}), \mathfrak{sl}(2,\mathbb{C})] = \mathfrak{sl}(2,\mathbb{C})$ (each of the basis elements e, f, h may be written as a Lie bracket), and $\mathfrak{sl}(2,\mathbb{C}).V \subset V'$ (because the action on the 1-dimensional space V/V' is necessarily trivial), we have

$$\mathfrak{sl}(2,\mathbb{C}).V = [\mathfrak{sl}(2,\mathbb{C}), \mathfrak{sl}(2,\mathbb{C})].V \subset \mathfrak{sl}(2,\mathbb{C}).V' = 0.$$

Hence $\mathfrak{sl}(2,\mathbb{C})$ acts trivially, and the Lemma is obvious. Case 2: $S = V(k)$ with $k > 0$. The Casimir element Cas acts as $k(k+2)/2 > 0$ on $V(k)$. Hence the kernel of Cas on V is the desired invariant complement. □

Proof of Theorem 11.7. Suppose $V' \subset V$ is an invariant subspace. Then $\mathrm{Hom}(V, V')$ carries an $\mathfrak{sl}(2,\mathbb{C})$-representation, $\tilde{\pi}(\xi)(B) = [\pi(\xi), B]$. Let $W \subset \mathrm{Hom}(V, V')$ be the subspace of transformations $B\colon V \to V'$ that restrict to a scalar on V', and W' those that restrict to 0 on V'. Then W' is a codimension 1 subrepresentation of W, hence by the lemma it admits an invariant complement. That is, there exists $B \in \mathrm{Hom}(V, V')$ such that B restricts to 1 on V', and $[\pi(\xi), B] = 0$. Its kernel $\ker(B) \subset V$ is an invariant complement to V'. □

There is a second (much simpler) proof of Theorem 11.7 via 'Weyl's unitary trick'. However, this argument requires the passage to Lie groups, and requires the existence of a bi-invariant measure on compact Lie groups. We will review this argument in Section ... below.

Given a finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-representation $\pi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V)$, there are various methods for computing its decomposition into irreducible representations. Let n_k be the multiplicity of $V(k)$ in V.

Method 1: Determine the eigenspaces of the Casimir operator $\pi(\mathrm{Cas}) = \pi(2fe + \frac{1}{2}h^2 + h)$. The eigenspace for the eigenvalue $k(k+2)/2$ is the direct sum of all irreducible subrepresentations of type $V(k)$. Hence

$$n_k = \frac{1}{k+1} \dim\ker\Big(\pi(\mathrm{Cas}) - \frac{k(k+2)}{2}\Big).$$

Method 2: For $l \in \mathbb{Z}$, let $m_{-l} = m_l = \dim\ker(\pi(h) - l)$ be the multiplicity of the eigenvalue l of $\pi(h)$. On any irreducible component $V(k)$, the dimension of $\ker(\pi(h) - l) \cap V(k)$ is 1 if $|l| \le k$ and $k - l$ is even, and is zero otherwise. Hence $m_k = n_k + n_{k+2} + \ldots$, and consequently

$$n_k = m_k - m_{k+2}.$$

Method 3: Find $\ker(\pi(e)) =: V^{\mathfrak{n}}$, and to consider the eigenspace decomposition of $\pi(h)$ on $V^{\mathfrak{n}}$. Thus

$$n_k = \dim\ker\big((\pi(h) - k) \cap V^{\mathfrak{n}}\big),$$

i.e. the multiplicity of the eigenvalue k of the restriction of $\pi(h)$ to $V^{\mathfrak{n}}$.

Exercise 11.9. If $\pi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V)$ is a finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-representation, then we obtain a representation $\tilde{\pi}$ on $\tilde{V} = \mathrm{End}(V)$ where $\tilde{\pi}(\xi)(B) = [\pi(\xi), B]$. In particular, for every irreducible representation $\pi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}_{\mathbb{C}}(V(n))$ we obtain a representation $\tilde{\pi}$ on $\mathrm{End}_{\mathbb{C}}(V(n))$.

Determine the decomposition of $\mathrm{End}_{\mathbb{C}}(V(n))$ into irreducible representations $V(k)$, i.e determine which $V(k)$ occur and with what multiplicity. (Hint: Note that all $\pi(e^j)$ commute with $\pi(e)$.)

Later we will need the following simple consequence of the $\mathfrak{sl}(2,\mathbb{C})$-representation theory:

Corollary 11.10. *Let $\pi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V)$ be a finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-representation. The the operator $\pi(h)$ on V is diagonalizable, and its eigenvalues are integers. Moreover,*

$$r > 0 \Rightarrow \pi(f)\colon V_{[r]} \to V_{[r-2]} \text{ is injective}$$
$$r < 0 \Rightarrow \pi(e)\colon V_{[r]} \to V_{[r+2]} \text{ is injective.}$$

Proof. The statements hold true for all irreducible components, hence also for their direct sum. □

Exercise 11.11. a) Show that $\mathrm{SL}(2,\mathbb{R})$ has fundamental group $\mathbb{Z}$. (Hint: Use polar decomposition of real matrices to show that $\mathrm{SL}(2,\mathbb{R})$ retracts onto $\mathrm{SO}(2) \cong S^1$.)

b) Show that $\mathrm{SL}(2,\mathbb{C})$ is simply connected. (Hint: Use polar decomposition of complex matrices to show that $\mathrm{SL}(2,\mathbb{C})$ retracts onto $\mathrm{SU}(2) \cong S^3$.)

c) Show that the universal cover $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ is *not* a matrix Lie group. That is, there does not exist an injective Lie group morphism

$$\widetilde{\mathrm{SL}}(2,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R}),$$

for any choice of n. (Given a Lie algebra morphism $\mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{gl}(n,\mathbb{R})$, complexify to get a Lie algebra morphism $\mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(n,\mathbb{C})$.)

12. Compact Lie groups

In this section we will prove some basic facts about compact Lie groups G and their Lie algebras $\mathfrak{g}$: (i) the existence of a bi-invariant positive measure, (ii) the existence of an invariant inner product on $\mathfrak{g}$, (iii) the decomposition of $\mathfrak{g}$ into center and simple ideals, (iv) the complete reducibility of G-representations, (v) the surjectivity of the exponential map.

12.1. **Modular function.** For any Lie group G, one defines the *modular function* to be the Lie group morphism

$$\chi\colon G \to \mathbb{R}^\times,\ g \mapsto |\mathrm{det}_{\mathfrak{g}}(\mathrm{Ad}_g)|.$$

Its differential is given by

$$\mathrm{d}_e\chi\colon \mathfrak{g} \to \mathbb{R},\quad \xi \mapsto \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi),$$

by the calculation $\frac{\partial}{\partial t}|_{t=0}\mathrm{det}_{\mathfrak{g}}(\mathrm{Ad}_{\exp(t\xi)}) = \frac{\partial}{\partial t}|_{t=0}\mathrm{det}_{\mathfrak{g}}(\exp(t\,\mathrm{ad}_\xi)) = \frac{\partial}{\partial t}|_{t=0}\exp(t\,\mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi)) = \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi)$. Here we have identified the Lie algebra of $\mathbb{R}^\times$ with $\mathbb{R}$, in such a way that the exponential map is just the usual exponential of real numbers.

Lemma 12.1. *For a compact Lie group, the modular function χ is trivial.*

Proof. The range of the Lie group morphism $G \to \mathbb{R}^\times,\ g \mapsto \det_{\mathfrak{g}}(\mathrm{Ad}_g)$ (as an image of a compact set under a continuous map) is compact. But its easy to see that the only compact subgroups of $\mathbb{R}^\times$ are $\{-1,1\}$ and $\{1\}$. ∎

It also follows that, for G, compact, the infinitesimal modular function $\mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi)$ is trivial.

Remark 12.2. A Lie group whose modular function is trivial is called *unimodular*. Besides compact Lie groups, there are many other examples of unimodular lie groups. For instance, if G is a connected Lie group whose Lie algebra is *semi-simple*, i.e. $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$, then any $\xi \in \mathfrak{g}$ can be written as a sum of commutators $\xi = \sum_i [\eta_i, \zeta_i]$. But

$$\mathrm{ad}_\xi = \sum_i \mathrm{ad}_{[\eta_i,\zeta_i]} = \sum_i [\mathrm{ad}_{\eta_i},\ \mathrm{ad}_{\zeta_i}]$$

has zero trace, since the trace vanishes on commutators. Similarly, if G is a connected Lie group whose Lie algebra is *nilpotent* (i.e., the series $\mathfrak{g}_{(0)} = \mathfrak{g},\ \mathfrak{g}_{(1)} = [\mathfrak{g},\mathfrak{g}],\ \ldots, \mathfrak{g}_{(k+1)} = [\mathfrak{g},\mathfrak{g}_{(k)}], \ldots$ is eventually zero), then the operator ad_ξ is nilpotent ($\mathrm{ad}_\xi^N = 0$ for N sufficiently large). Hence its eigenvalues are all 0, and consequently $\mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi) = 0$. An example is the Lie group of upper triangular matrices with 1's on the diagonal.

An example of a Lie group that is not unimodular is the conformal group of the real line, i.e. the 2-dimensional Lie group G of matrices of the form

$$g = \begin{pmatrix} t & s \\ 0 & 1 \end{pmatrix}$$

with $t > 0$ and $s \in \mathbb{R}$. In this example, one checks $\chi(g) = t$.

12.2. **Volume forms and densities.** The modular function can also be interpreted in terms of volume forms. We begin with a review of volume forms and densities on manifolds.

Definition 12.3. Let E be a vector space of dimension n. We define a vector space $\det(E^*)$ consisting of maps $\Lambda\colon E \times \cdots \times E \to \mathbb{R}$ satisfying

$$\Lambda(Av_1, \ldots, Av_n) = \det(A)\,\Lambda(v_1, \ldots, v_n).$$

for all $A \in \mathrm{GL}(E)$ The non-zero elements of $\det(E^*)$ are called volume forms on E. We also define a space $|\det|(E^*)$ of maps $\mathsf{m}\colon E \times \cdots \times E \to \mathbb{R}$ satisfying

$$\mathsf{m}(Av_1, \ldots, Av_n) = |\det(A)|\,\mathsf{m}(v_1, \ldots, v_n)$$

for all $A \in \mathrm{GL}(E)$. The elements of $|\det|(E^*)$ are called *densities*.

Both $\det(E^*)$ and $|\det|(E^*)$ are 1-dimensional vector spaces. Of course, $\det(E^*) \equiv \wedge^n E^*$. A volume form Λ defines an orientation on E, where a basis $v_1, \ldots, v_n$ is oriented if $\Lambda(v_1, \ldots, v_n) > 0$. It also defines a non-zero density $\mathsf{m} = |\Lambda|$ by putting $|\Lambda|(v_1, \ldots, v_n) = |\Lambda(v_1, \ldots, v_n)|$. Conversely, a positive density together with an orientation define a volume form. In fact, a choice of orientation gives an isomorphism $\det(E^*) \cong |\det|(E^*)$; a change of orientation changes this isomorphism by a sign. The vector space $\mathbb{R}^n$ has a standard volume form Λ_0 (taking the oriented basis $e_1, \ldots, e_n$) to 1), hence a standard orientation and density $|\Lambda_0|$. The latter is typically denoted $\mathrm{d}^n x$, $|\mathrm{d}x|$ or similar. Given a linear map $\Phi\colon E \to E'$, one obtains pull-back maps $\Phi^*\colon \det((E')^*) \to \det(E^*)$ and $\Phi^*\colon |\det|((E')^*) \to |\det|(E^*)$; these are non-zero if and only if Φ is an isomorphism.

For manifolds M, one obtains real line bundles

$$\det(T^*M),\ |\det|(T^*M)$$

with fibers $\det(T^*_m M)$, $|\det|(T^*_m M)$. A non-vanishing section $\Lambda \in \Gamma(\det(T^*M))$ is called a volume form on M; it gives rise to an orientation on M. Hence, M is orientable if and only if $\det(T^*M)$ is trivializable. On the other hand, the line bundle $|\det|(T^*M)$ is always trivializable.

Densities on manifolds are also called *smooth measures* [5]. Being defined as sections of a vector bundle, they are a module over the algebra of functions $C^\infty(M)$: if m is a smooth measure then so is $f\mathsf{m}$. In fact, the choice of a fixed positive smooth measure m on M trivializes the density bundle, hence it identifies $C^\infty(M)$ with $\Gamma(|\det|(T^*M))$. There is an integration map, defined for measures of compact support,

$$\Gamma_{\text{comp}}(|\det|(T^*M)) \to \mathbb{R},\quad \mathsf{m} \mapsto \int_M \mathsf{m}.$$

It is characterized as the unique linear map such that for any measure m supported in a chart $\phi\colon U \to M$, with $U \subset \mathbb{R}^n$ one has

$$\int_M \mathsf{m} = \int_U \varrho(x)|\mathrm{d}x|$$

where $\varrho \in C^\infty(U)$ is the function defined by $\phi^*\mathsf{m} = \varrho|\mathrm{d}x|$.

The integral is a linear map, uniquely determined by the following property: For any coordinate chart $\phi\colon U \to M$, and any function with compact support in U,

$$\int_M f\mathsf{m} = \int_U \phi^* f\, \varrho|\mathrm{d}x|$$

(a standard n-dimensional Riemann integral) where $\varrho \in C^\infty(U)$ is defined by $\phi^*\mathsf{m} = \varrho\mathsf{m}_0$.

Given a G-action on M, a volume form is called *invariant* if $\mathcal{A}_g^*\Lambda = \Lambda$ for all $g \in G$. In particular, we can look for left-invariant volume forms on Lie groups. Any left-invariant section of $\det(T^*G)$ is uniquely determined by its value at the group unit, and any non-zero Λ_e can be uniquely extended to a left-invariant volume form. That is, $\Gamma^L(\det(T^*G)) \cong \mathbb{R}$.

Lemma 12.4. *Let G be a Lie group, and $\chi\colon G \to \mathbb{R}^\times$ its modular function. If Λ is a left-invariant volume form on G, then*

$$R_a^*\Lambda = \det(\mathrm{Ad}_{a^{-1}})\Lambda,$$

for all $a \in G$. If m is a left-invariant smooth density, we have

$$R_a^*\mathsf{m} = \chi(a)^{-1}\mathsf{m}$$

for all $a \in G$.

Proof. If Λ is left-invariant, then $R_a^*\Lambda$ is again left-invariant since left and right multiplications commute. Hence it is a multiple of Λ. To determine the multiple, note

$$R_a^*\Lambda = R_a^* L_{a^{-1}}^*\Lambda = \mathrm{Ad}_{a^{-1}}^*\Lambda.$$

[5] Note that measures in general are covariant objects: They push forward under continuous proper maps. However, the push-forward of a smooth measure is not smooth, in general. Smooth measures (densities), on the other hand, are contravariant objects.

Computing at the group unit e, we see that $\mathrm{Ad}^*_{a^{-1}}\,\Lambda_e = \det(\mathrm{Ad}_a)^{-1}\Lambda_e = \Lambda_e$. The result for densities is a consequence of that for volume forms. $\square$

Hence, if G is compact, any left-invariant density is also right-invariant. If G is compact and connected, any left-invariant volume form is also right-invariant. One can normalize the left-invariant density such that $\int_G \mathsf{m} = 1$. The left-invariant measure on a Lie group G (not necessarily normalized) is often denoted $\mathsf{m} = |\mathrm{d}g|$.

12.3. **Basic properties of compact Lie groups.** The existence of the bi-invariant measure of finite integral lies at the heart of the theory of compact Lie groups. For instance, it implies that the Lie algebra $\mathfrak{g}$ of G admits an Ad-invariant inner product B: In fact, given an arbitrary inner product B' one may take B to be its G-average:

$$B(\xi,\zeta) = \frac{1}{\mathrm{vol}(G)}\int_G B'(\mathrm{Ad}_g(\xi),\mathrm{Ad}_g(\zeta))\ |\mathrm{d}g|.$$

The Ad-invariance

$$B(\mathrm{Ad}_g\,\xi,\mathrm{Ad}_g\,\eta) = B(\xi,\eta) \tag{1}$$

follows from the bi-invariance if the measure. A symmetric bilinear form B on a Lie algebra $\mathfrak{g}$ is called ad-invariant if

$$B([\xi,\eta],\zeta) + B(\eta,[\xi,\zeta]) = 0. \tag{2}$$

for all $\xi,\eta,\zeta \in \mathfrak{g}$. If $\mathfrak{g}$ is the Lie algebra of a Lie group G, then any Ad-invariant bilinear form is also ad-invariant, by differentiating the property (1).

As an application, we obtain the following decomposition of the Lie algebra of compact Lie groups. An *ideal* in a Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h}$ with $[\mathfrak{g},\mathfrak{h}] \subset \mathfrak{h}$. (In particular, $\mathfrak{h}$ is a Lie subalgebra). For instance, $[\mathfrak{g},\mathfrak{g}]$ an ideal. An ideal is the same thing as an invariant subspace for the adjoint representation of $\mathfrak{g}$ on itself. Note that for any two ideals $\mathfrak{h}_1,\mathfrak{h}_2$, their sum $\mathfrak{h}_1+\mathfrak{h}_2$ and their intersection $\mathfrak{h}_1\cap\mathfrak{h}_2$ are again ideals.

A Lie algebra is called *simple* if it is non-abelian and does not contain non-trivial ideals, and *semi-simple* if it is a direct sum of simple ideals. For a simple Lie algebra, we must have $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$ (since $[\mathfrak{g},\mathfrak{g}]$ is a non-zero ideal), hence the same property $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$ is also true for semi-simple Lie algebras.

Proposition 12.5. *The Lie algebra $\mathfrak{g}$ of a compact Lie group G is a direct sum*

$$\mathfrak{g} = \mathfrak{z}\oplus\mathfrak{g}_1\oplus\cdots\oplus\mathfrak{g}_r,$$

where $\mathfrak{z}$ is the center of $\mathfrak{g}$, and the $\mathfrak{g}_i$ are simple ideals. One has $[\mathfrak{g},\mathfrak{g}] = \mathfrak{g}_1\oplus\cdots\oplus\mathfrak{g}_r$. The decomposition is unique up to re-ordering of the summands.

Proof. Pick an invariant Euclidean inner product B on $\mathfrak{g}$. Then the orthogonal complement (with respect to B) of any ideal $\mathfrak{h}\subset\mathfrak{g}$ is again an ideal. Indeed, $[\mathfrak{g},\mathfrak{h}]\subset\mathfrak{h}$ implies

$$B([\mathfrak{g},\mathfrak{h}^\perp],\mathfrak{h}) = B(\mathfrak{h}^\perp,[\mathfrak{g},\mathfrak{h}]) \subset B(\mathfrak{h}^\perp,\mathfrak{h}) = 0,$$

hence $[\mathfrak{g},\mathfrak{h}^\perp]\subset\mathfrak{h}^\perp$. Let $\mathfrak{z}\subset\mathfrak{g}$ be the center of $\mathfrak{g}$, and $\mathfrak{g}' = \mathfrak{z}^\perp$. Then $[\mathfrak{g},\mathfrak{g}] = [\mathfrak{g}',\mathfrak{g}']\subset\mathfrak{g}'$. The calculation

$$B([\mathfrak{g},[\mathfrak{g},\mathfrak{g}]^\perp],\mathfrak{g}) = B([\mathfrak{g},\mathfrak{g}]^\perp,[\mathfrak{g},\mathfrak{g}]) = 0$$

shows $[\mathfrak{g},[\mathfrak{g},\mathfrak{g}]^\perp]=0$, hence $[\mathfrak{g},\mathfrak{g}]^\perp\subset\mathfrak{z}=(\mathfrak{g}')^\perp$, i.e. $\mathfrak{g}'\subset[\mathfrak{g},\mathfrak{g}]$. It follows that $[\mathfrak{g},\mathfrak{g}]=\mathfrak{g}'$. This gives the desired decomposition $\mathfrak{g}=\mathfrak{z}\oplus[\mathfrak{g},\mathfrak{g}]$. By a similar argument, we may inductively decompose $[\mathfrak{g},\mathfrak{g}]$ into simple ideals. For the uniqueness part, suppose that $\mathfrak{h}\subset[\mathfrak{g},\mathfrak{g}]$ is an ideal not containing any of the $\mathfrak{g}_i$. But then $[\mathfrak{g}_i,\mathfrak{h}]\subset\mathfrak{g}_i\cap\mathfrak{h}=0$ for all i, which gives $[\mathfrak{g},\mathfrak{h}]=\bigoplus_i[\mathfrak{g}_i,\mathfrak{h}]=0$. Hence $\mathfrak{h}\subset\mathfrak{z}$. □

Exercise 12.6. Show that for any Lie group G, the Lie algebra of the center of G is the center of the Lie algebra.

12.4. **Complete reducibility of representations.** Let G be a compact Lie group, and $\pi\colon G\to\mathrm{End}(V)$ a representation on a complex vector space V. Then the vector space V admits an invariant Hermitian inner product, obtained from a given Hermitian inner product h' by averaging:

$$h(v,w)=\frac{1}{\mathrm{vol}(G)}\int_G h'(\pi(g)v,\pi(g)w)\ |\mathrm{d}g|.$$

Given a G-invariant subspace W, its orthogonal complement $W^\perp$ with respect to h is again G-invariant. As a consequence, any finite-dimensional complex G-representation is completely reducible. (A similar argument also shows that every real G-representation is completely reducible.)

Let us briefly return to the claim that every finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-representation is completely reducible. Indeed, any $\mathfrak{sl}(2,\mathbb{C})$-representation defines a representation of $\mathfrak{su}(2)$ by restriction, and may be recovered from the latter by complexification. On the other hand, since $\mathrm{SU}(2)$ is simply connected a representation of $\mathfrak{su}(2)$ is equivalent to a representation of $\mathrm{SU}(2)$. Hence, the complete reducibility of $\mathfrak{sl}(2,\mathbb{C})$-representations is a consequence of that for the Lie group $\mathrm{SU}(2)$.

12.5. **The bi-invariant Riemannian metric.** Recall some material from differential geometry. Suppose M is a manifold equipped with a pseudo-Riemannian metric B. That is, B is a family of non-degenerate symmetric bilinear forms $B_m\colon T_mM\times T_mM\to\mathbb{R}$ depending smoothly on m. A smooth curve $\gamma\colon J\to M$ (with $J\subset\mathbb{R}$ some interval) is called a *geodesic* if, for any $[t_0,t_1]\subset J$, the restriction of γ is a critical point of the energy functional

$$E(\gamma)=\int_{t_0}^{t_1}B(\dot\gamma(t),\ \dot\gamma(t))\,\mathrm{d}t.$$

That is, for any *variation* of γ, given by a smooth 1-parameter family of curves $\gamma_s\colon[t_0,t_1]\to M$ (defined for small $|s|$), with $\gamma_0=\gamma$ and with fixed end points ($\gamma_s(t_0)=\gamma(t_0),\ \gamma_s(t_1)=\gamma(t_1)$) we have

$$\frac{\partial}{\partial s}\Big|_{s=0}E(\gamma_s)=0.$$

A geodesic is uniquely determined by its values $\gamma(t_*),\ \dot\gamma(t_*)$ at any point $t_*\subset J$. It is one of the consequences of the *Hopf-Rinow theorem* that if M is a compact, connected Riemannian manifold, then any two points in M are joined by a length minimizing geodesic. The result is false in general for pseudo-Riemannian metrics, and we will encounter a counterexample at the end of this section.

Let G be a Lie group, with Lie algebra $\mathfrak{g}$. A non-degenerate symmetric bilinear form $B\colon\mathfrak{g}\times\mathfrak{g}\to\mathbb{R}$ defines, via left translation, a left-invariant pseudo-Riemannian metric (still denoted B)

on G. If the bilinear form on $\mathfrak{g}$ is Ad-invariant, then the pseudo-Riemannian metric on G is bi-invariant. In particular, any compact Lie group admits a bi-invariant Riemannian metric. As another example, the group $\mathrm{GL}(n,\mathbb{R})$ carries a bi-invariant pseudo-Riemannian metric defined by the bilinear form $B(\xi_1,\xi_2)=\mathrm{tr}(\xi_1\xi_2)$ on $\mathfrak{gl}(n,\mathbb{R})$. It restricts to a pseudo-Riemannian metric on $\mathrm{SL}(n,\mathbb{R})$.

Theorem 12.7. *Let G be a Lie group with a bi-invariant pseudo-Riemannian metric B. Then the geodesics on G are the left-translates (or right-translates) of the 1-parameter subgroups of G.*

Proof. Since B is bi-invariant, the left-translates or right-translates of geodesics are again geodesics. Hence it suffices to consider geodesics $\gamma(t)$ with $\gamma(0)=e$. For $\xi\in\mathfrak{g}$, let $\gamma(t)$ be the unique geodesic with $\dot\gamma(0)=\xi$ and $\gamma(0)=e$. To show that $\gamma(t)=\exp(t\xi)$, let $\gamma_s\colon[t_0,t_1]\to G$ be a 1-parameter variation of $\gamma(t)=\exp(t\xi)$, with fixed end points. If s is sufficiently small we may write $\gamma_s(t)=\exp(u_s(t))\exp(t\xi)$ where $u_s\colon[t_0,t_1]\to\mathfrak{g}$ is a 1-parameter variation of 0 with fixed end points, $u_s(t_0)=0=u_s(t_1)$. We have

$$\dot\gamma_s(t)=R_{\exp(t\xi)}L_{\exp(u_s(t))}\Big(\xi+\frac{1-e^{-\,\mathrm{ad}(u_s)}}{\mathrm{ad}(u_s)}\dot u_s(t)\Big),$$

hence, using bi-invariance of B,

$$E(\gamma_s)=\int_{t_0}^{t_1}B\Big(\xi+\frac{1-e^{-\,\mathrm{ad}(u_s)}}{\mathrm{ad}(u_s)}\dot u_s(t),\ \xi+\frac{1-e^{-\,\mathrm{ad}(u_s)}}{\mathrm{ad}(u_s)}\dot u_s(t)\Big)\mathrm{d}t$$

Notice

$$\frac{\partial}{\partial s}|_{s=0}\Big(\frac{1-e^{-\,\mathrm{ad}(u_s)}}{\mathrm{ad}(u_s)}\dot u_s(t)\Big)=\frac{\partial}{\partial s}|_{s=0}\dot u_s(t)$$

since $u_0=0,\ \dot u_0=0$. Hence, the s-derivative of $E(\gamma_s)$ at $s=0$ is

$$\begin{aligned}\frac{\partial}{\partial s}|_{s=0}E(\gamma_s)&=2\int_{t_0}^{t_1}B(\frac{\partial}{\partial s}\Big|_{s=0}\dot u_s(t),\xi)\\&=2B(\frac{\partial}{\partial s}\Big|_{s=0}u_s(t_1),\xi)-2B(\frac{\partial}{\partial s}\Big|_{s=0}u_s(t_0),\xi)=0\end{aligned}$$

□

Remark 12.8. A pseudo-Riemannian manifold is called *geodesically complete* if for any given $m\in M$ and $v\in T_mM$, the geodesic with $\gamma(0)=m$ and $\dot\gamma(0)=v$ is defined for all $t\in\mathbb{R}$. In this case one defines an *exponential map* $\mathrm{Exp}\colon TM\to M$ by taking $v\in T_mM$ to $\gamma(1)$, where $\gamma(t)$ is the geodesic defined by v. The result above shows that any Lie group G with a bi-invariant pseudo-Riemannian metric is geodesically complete, and $\mathrm{Exp}\colon TG\to G$ is the extension of the Lie group exponential map $\exp\colon\mathfrak{g}\to G$ by left translation.

Theorem 12.9. *The exponential map of a compact, connected Lie group is surjective.*

Proof. Choose a bi-invariant Riemannian metric on G. By the Hopf-Rinow theorem, any two points in G are joined by a geodesic. In particular, given $g\in G$ there exists a geodesic with $\gamma(0)=e$ and $\gamma(1)=g$. This geodesic is of the form $\exp(t\xi)$ for some ξ. Hence $\exp(\xi)=g$. □

The example of $G=\mathrm{SL}(2,\mathbb{R})$ shows that the existence of a bi-invariant *pseudo*-Riemannian metric does not suffice for this result.

12.6. **The Killing form.**

Definition 12.10. The *Killing form*[6] of a finite-dimensional Lie algebra $\mathfrak{g}$ is the symmetric bilinear form

$$\kappa(\xi,\eta) = \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi\,\mathrm{ad}_\eta).$$

Proposition 12.11. *The Killing form on a finite-dimensional Lie algebra* $\mathfrak{g}$ *is* ad*-invariant. If* $\mathfrak{g}$ *is the Lie algebra of a possibly disconnected Lie group* G*, it is furthermore* Ad*-invariant.*

Proof. The ad-invariance follows from $\mathrm{ad}_{[\xi,\zeta]} = [\mathrm{ad}_\xi, \mathrm{ad}_\zeta]$:

$$\kappa([\xi,\eta],\zeta) + \kappa(\eta,[\xi,\zeta]) = \mathrm{tr}_{\mathfrak{g}}([\mathrm{ad}_\xi,\mathrm{ad}_\eta]\,\mathrm{ad}_\zeta) + \mathrm{ad}_\eta[\mathrm{ad}_\xi,\mathrm{ad}_\zeta]) = 0.$$

The Ad-invariance is checked using $\mathrm{ad}_{\mathrm{Ad}_g(\xi)} = \mathrm{Ad}_g \circ \mathrm{ad}_\xi \circ \mathrm{Ad}_{g^{-1}}$. □

An important result of Cartan says that a Lie algebra $\mathfrak{g}$ is semi-simple (i.e. $[\mathfrak{g},\mathfrak{g}] = \mathfrak{g}$) if and only if its Killing form is non-degenerate (possibly indefinite). In this course, we will only consider this statement for Lie algebra of compact Lie groups.

Proposition 12.12. *Suppose* $\mathfrak{g}$ *is the Lie algebra of a compact Lie group* G*. Then the Killing form on* $\mathfrak{g}$ *is negative semi-definite, with kernel the center* $\mathfrak{z}$*. Thus, if in addition* $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$ *the Killing form is negative definite.*

Proof. Let B be an invariant inner product on $\mathfrak{g}$, i.e. B positive definite. The ad-invariance says that ad_ξ is skew-symmetric relative to B. Hence it is diagonalizable (over $\mathbb{C}$), and all its eigenvalues are in $i\mathbb{R}$. Consequently ad_ξ^2 is symmetric relative to B, with non-positive eigenvalues, and its kernel coincides with the kernel of ad_ξ. This shows that

$$\kappa(\xi,\xi) = \mathrm{tr}(\mathrm{ad}_\xi^2) \le 0,$$

with equality if and only if $\mathrm{ad}_\xi = 0$, i.e. $\xi \in \mathfrak{z}$. □

12.7. **Derivations.** Let $\mathfrak{g}$ be a Lie algebra. Recall that $D \in \mathrm{End}(\mathfrak{g})$ is a derivation if and only if $D([\xi,\eta]) = [D\xi,\eta] + [\xi, D\eta]$ for all $\xi,\eta \in \mathfrak{g}$, that is

$$\mathrm{ad}_{D\xi} = [D, \mathrm{ad}_\xi].$$

Let $\mathrm{Der}(\mathfrak{g})$ be the Lie algebra of derivations of a Lie algebra $\mathfrak{g}$, and $\mathrm{Inn}(\mathfrak{g})$ the Lie subalgebra of inner derivations, i.e. those of the form $D = \mathrm{ad}_\xi$.

Proposition 12.13. *Suppose the Killing form of* $\mathfrak{g}$ *is non-degenerate. Then any derivation of* $\mathfrak{g}$ *is inner. In fact,* $\mathrm{Der}(\mathfrak{g}) = \mathrm{Inn}(\mathfrak{g}) = \mathfrak{g}$.

[6]The Killing form is named after Wilhelm Killing. Killings contributions to Lie theory had long been underrated. In fact, he himself in 1880 had rediscovered Lie algebras independently of Lie (but about 10 years later). In 1888 he had obtained the full classification of Lie algebra of compact Lie groups. Killing's existence proofs contained gaps, which were later filled by E. Cartan. The Cartan matrices, Cartan subalgebras, Weyl groups, root systems Coxeter transformations etc. all appear in some form in W. Killing's work (cf. Borel 'Essays in the history of Lie groups and Lie algebras'.) According A. J. Coleman ('The greatest mathematical paper of all time'), "he exhibited the characteristic equation of the Weyl group when Weyl was 3 years old and listed the orders of the Coxeter transformation 19 years before Coxeter was born." On the other hand, the Killing form was actually first considered by E. Cartan. Borel admits that he (Borel) was probably the first to use the term 'Killing form'.

Proof. Let $D \in \mathrm{Der}(\mathfrak{g})$. Since the Killing form is non-degenerate, the exists $\xi \in \mathfrak{g}$ with

$$\kappa(\xi, \eta) = \mathrm{tr}(D \circ \mathrm{ad}_\eta)$$

for all $\eta \in \mathfrak{g}$. The derivation $D_0 = D - \mathrm{ad}_\xi$ then satisfies $\mathrm{tr}(D_0 \circ \mathrm{ad}_\xi) = 0$. For $\eta, \zeta \in \mathfrak{g}$ we obtain

$$\kappa(D_0(\eta), \zeta) = \mathrm{tr}(\mathrm{ad}_{D_0(\eta)} \mathrm{ad}_\zeta) = \mathrm{tr}([D_0, \mathrm{ad}_\eta] \mathrm{ad}_\zeta) = \mathrm{tr}(D_0[\mathrm{ad}_\eta, \mathrm{ad}_\zeta]) = \mathrm{tr}(D_0 \mathrm{ad}_{[\eta,\zeta]}) = 0.$$

This shows $D_0(\eta) = 0$ for all η, hence $D_0 = 0$. By definition, $\mathrm{Inn}(\mathfrak{g})$ is the image of the map $\mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$, $\xi \mapsto \mathrm{ad}_\xi$. The kernel of this map is the center $\mathfrak{z}$ of the Lie algebra. But if κ is non-degenerate, the center $\mathfrak{z}$ must be trivial. □

If G is a Lie group with Lie algebra $\mathfrak{g}$, we had seen that $\mathrm{Der}(\mathfrak{g})$ is the Lie algebra of the Lie group $\mathrm{Aut}(\mathfrak{g})$. The Proposition shows that if the Killing form is non-degenerate, then the differential of the map $G \to \mathrm{Aut}(\mathfrak{g})$ is an isomorphism. Hence, it defines a covering from the identity component of G to the identity component of $\mathrm{Aut}(\mathfrak{g})$.

Proposition 12.14. *Suppose the Killing form on the finite-dimensional Lie algebra $\mathfrak{g}$ is negative definite. Then $\mathfrak{g}$ is the Lie algebra of a compact Lie group.*

Proof. Since $\mathrm{Aut}(\mathfrak{g})$ preserves the Killing form, we have

$$\mathrm{Aut}(\mathfrak{g}) \subset \mathrm{O}(\mathfrak{g}, \kappa),$$

the orthogonal group relative to κ. Since κ is negative definite, $\mathrm{O}(\mathfrak{g}, \kappa)$ is compact. Hence $\mathrm{Aut}(\mathfrak{g})$ is a compact Lie group with Lie algebra $\mathrm{Inn}(\mathfrak{g}) = \mathfrak{g}$. □

Remark 12.15. The converse is true as well: That is, the Lie algebra of a Lie group G has negative definite Killing form if and only if G is compact with finite center. The most common proof of this fact is via a result that *for a compact connected Lie group with finite center, the fundamental group is finite.* This result applies to the identity component of the group $\mathrm{Aut}(\mathfrak{g})$; hence the universal cover of the identity component of $\mathrm{Aut}(\mathfrak{g})$ is compact. A different proof, not using fundamental group calculations (but instead using some facts from Riemannian geometry), may be found in Helgason's book *Differential geometry, Lie groups and symmetric spaces*, Academic Press, page 133. We may get back to this later (if time allows).

13. The maximal torus of a compact Lie group

13.1. **Abelian Lie groups.** A Lie group G is called *abelian* if $gh = hg$ for all $g, h \in G$, i.e. G is equal to its center.[7] A compact connected abelian group is called a *torus.* A Lie algebra $\mathfrak{g}$ is abelian (or commutative) if the Lie bracket is trivial, i.e. $\mathfrak{g}$ equals its center.

Proposition 13.1. *A connected Lie group G is abelian if and only if its Lie algebra $\mathfrak{g}$ is abelian. Furthermore, in this case the universal cover is*

$$\widetilde{G} = \mathfrak{g}$$

(viewed as an additive Lie group).

[7]Abelian groups are named after Nils Hendrik Abel. In the words of R. Bott, 'I could have come up with that.'

Proof. The Lie algebra $\mathfrak{g}$ is abelian if and only if $\mathfrak{X}^L(G)$ is abelian, i.e. if and only if the flows of any two left-invariant vector fields commute. Thus

$$\exp(\xi)\exp(\eta) = \exp(\xi+\eta) = \exp(\eta)\exp(\xi),$$

for all ξ, η. Hence there is a neighborhood U of e such that any two elements in U commute. Since any element of G is a product of elements in U, this is the case if and only if G is abelian. We also see that in this case, $\exp\colon \mathfrak{g} \to G$ is a Lie group morphism. Its differential at 0 is the identity, hence exp is a covering map. Since $\mathfrak{g}$ is contractible, it is the universal cover of G. □

We hence see that any abelian Lie group is of the form $G = V/\Gamma$, where $V \cong \mathfrak{g}$ is a vector space and Γ is a discrete additive subgroup of V.

Lemma 13.2. *There are linearly independent $\gamma_1, \ldots, \gamma_k \in V$ such that*

$$\Gamma = \operatorname{span}_{\mathbb{Z}}(\gamma_1, \ldots, \gamma_k).$$

Proof. Choose any $\gamma_1 \in \Gamma$ such that $\mathbb{Z}\gamma_1 = \mathbb{R}\gamma_1 \cap \Gamma$. Suppose by induction that $\gamma_1, \ldots, \gamma_l \in \Gamma$ are linearly independent vectors such that $\operatorname{span}_{\mathbb{Z}}(\gamma_1, \ldots, \gamma_l) = \operatorname{span}_{\mathbb{R}}(\gamma_1, \ldots, \gamma_l) \cap \Gamma$. If the $\mathbb{Z}$-span is all of Γ, we are done. Otherwise, let $V' = V/\operatorname{span}_{\mathbb{R}}(\gamma_1, \ldots, \gamma_l)$, $\Gamma' = \Gamma/\operatorname{span}_{\mathbb{Z}}(\gamma_1, \ldots, \gamma_l)$, and pick γ_{l+1} such that its image $\gamma'_{l+1} \in V'$ satisfies $\mathbb{Z}\gamma'_{l+1} = \mathbb{R}\gamma'_{l+1} \cap \Gamma'$. Then $\operatorname{span}_{\mathbb{Z}}(\gamma_1, \ldots, \gamma_{l+1}) = \operatorname{span}_{\mathbb{R}}(\gamma_1, \ldots, \gamma_{l+1}) \cap \Gamma$. □

Extending the γ_i to a basis of V, thus identifying $V = \mathbb{R}^n$, see that any abelian Lie group is isomorphic to $\mathbb{R}^n/\mathbb{Z}^k$ for some n, k. That is:

Proposition 13.3. *Any connected abelian Lie group is isomorphic to $(\mathbb{R}/\mathbb{Z})^k \times \mathbb{R}^l$, for some k, l. In particular, a k-dimensional torus is isomorphic to $(\mathbb{R}/\mathbb{Z})^k$.*

For a torus T, we denote by $\Lambda \subset \mathfrak{t}$ the kernel of the exponential map. We will call Λ the *integral lattice*. Thus

$$T = \mathfrak{t}/\Lambda.$$

Definition 13.4. Let H be an abelian Lie group (possibly disconnected). An element $h \in H$ is called a *topological generator* of H if the subgroup $\{h^k \mid k \in \mathbb{Z}\}$ generated by h is dense in T.

Theorem 13.5 (Kronecker lemma)**.** *Let $u = (u_1, \ldots, u_k) \in \mathbb{R}^k$, and $t = \exp(u)$ its image in $T = (\mathbb{R}/\mathbb{Z})^k$. Then t is a topological generator if and only if $1, u_1, \ldots, u_k \in \mathbb{R}$ are linearly independent over the rationals $\mathbb{Q}$. In particular, topological generators of tori exist.*

Proof. Note that $1, u_1, \ldots, u_k \in \mathbb{R}$ are linearly dependent over the rationals if and only if there exist $a_1, \ldots, a_n$, not all zero, such that $\sum_{i=1}^k a_i u_i \in \mathbb{Z}$.

Let $T = (\mathbb{R}/\mathbb{Z})^k$, and let H be the closure of the subgroup generated by t. Then T/H is a compact connected abelian group, i.e. is isomorphic to $(\mathbb{R}/\mathbb{Z})^l$ for some l. If $H \neq T$, then $l > 0$, and there exists a non-trivial group morphism $T/H \cong (\mathbb{R}/\mathbb{Z})^l \to \mathbb{R}/\mathbb{Z}$ (e.g. projection to the first factor). By composition with the quotient map, it becomes a non-trivial group morphism

$$\phi\colon T \to \mathbb{R}/\mathbb{Z}$$

that is trivial on H. Let $\mathrm{d}\phi\colon \mathbb{R}^k \to \mathbb{R}$ be its differential, and put $a_i = \mathrm{d}\phi(e_i)$. Since $\mathrm{d}\phi(\mathbb{Z}^k) \subset \mathbb{Z}$, the a_i are integers. Since ϕ vanishes on H, we have $\phi(t) = 0 \mod \mathbb{Z}$, while

$$\begin{aligned}\phi(t) &= \mathrm{d}\phi(u) \mod \mathbb{Z}\\ &= \sum_{i=1}^{k} a_i u_i \mod \mathbb{Z}.\end{aligned}$$

That is, $\sum_{i=1}^k a_i u_i \in \mathbb{Z}$. Conversely, if there are $a_i \in \mathbb{Z}$, not all zero, such that $\sum_{i=1}^k a_i u_i \in \mathbb{Z}$, define $\phi\colon T \to \mathbb{R}/\mathbb{Z}$ by $\phi(\exp(v)) = \sum_{i=1}^k a_i v_i \mod \mathbb{Z}$ for $v = (v_1, \ldots, v_l)$. Then $\phi(t) = 0 \mod \mathbb{Z}$ so that ϕ is trivial on H. Since ϕ is not trivial on T, it follows that H is a proper subgroup of T. □

Let $\mathrm{Aut}(T)$ be the group of automorphisms of the torus T. Any such automorphism ϕ induces a Lie algebra automorphism $\mathrm{d}\phi \in \mathrm{Aut}(\mathfrak{t})$ preserving Λ. Conversely, given an automorphism of the lattice Λ, we obtain an automorphism of $\mathfrak{t} = \mathrm{span}_{\mathbb{R}}(\Lambda)$ and hence of $T = \mathfrak{t}/\Lambda$. That is, $\mathrm{Aut}(T)$ can be identified with the automorphisms of the lattice Λ:

$$\mathrm{Aut}(T) = \mathrm{Aut}(\Lambda).$$

After choosing an identification $T = (\mathbb{R}/\mathbb{Z})^l$, this is the discrete group $\mathrm{GL}(n, \mathbb{Z}) = \mathrm{Mat}_n(\mathbb{Z})$ of matrices A with integer coefficients having an inverse with integer coefficients. By the formula for the inverse matrix, this is the case if and only if the determinant is ± 1:

$$\mathrm{GL}(n, \mathbb{Z}) = \{A \in \mathrm{Mat}_n(\mathbb{Z}) |\ \det(A) \pm 1\}.$$

The group $\mathrm{GL}(n, \mathbb{Z})$ contains the semi-direct product $(\mathbb{Z}_2)^n \rtimes S_n$, where S_n acts by permutation of coordinates and $(\mathbb{Z}_2)^n$ acts by sign changes. It is easy to see that the subgroup $(\mathbb{Z}_2)^n \rtimes S_n$ is $\mathrm{O}(n, \mathbb{Z}) = \mathrm{GL}(n, \mathbb{Z}) \cap \mathrm{O}(n)$, the transformations preserving also the metric. It is thus a maximal compact subgroup of $\mathrm{GL}(n, \mathbb{Z})$.

13.2. **Maximal tori.** Let G be a compact Lie group, with Lie algebra $\mathfrak{g}$. A torus $T \subset G$ is called a *maximal torus* if it is not properly contained in a larger subtorus of G.

Theorem 13.6. *(E. Cartan) Let G be a compact, connected Lie group. Then any two maximal tori of G are conjugate.*

Proof. We have to show that is $T, T' \subset G$ are two maximal tori, then there exists $g \in G$ such that $\mathrm{Ad}_a(T) = T'$. Fix an invariant inner product B on $\mathfrak{g}$. Pick topological generators t, t' of T, T', and choose $\xi, \xi' \in \mathfrak{g}$ with $\exp(\xi) = t,\ \exp(\xi') = t'$. Let $a \in G$ be a group element for which the function $g \mapsto B(\xi', \mathrm{Ad}_g(\xi))$ takes on its maximum possible value. We will show that $\mathrm{Ad}_a(T) = T'$. To see this, let $\eta \in \mathfrak{g}$. By definition of g, the function

$$t \mapsto B(\mathrm{Ad}_{\exp(t\eta)} \mathrm{Ad}_a(\xi), \xi')$$

takes on its maximum value at $t = 0$. Taking the derivative at $t = 0$, this gives

$$0 = B([\eta, \mathrm{Ad}_a(\xi)], \xi') = B(\eta, [\mathrm{Ad}_a(\xi), \xi']).$$

Since this is true for all η, we obtain $[\xi', \mathrm{Ad}_a(\xi)] = 0$. Exponentiating ξ', this shows $\mathrm{Ad}_{t'}(\mathrm{Ad}_a(\xi)) = \mathrm{Ad}_a(\xi)$. Exponentiating ξ, it follows that $\mathrm{Ad}_a(t), t'$ commute. Since these are generators, any element in $\mathrm{Ad}_a(T)$ commutes with any element in T'. The group $T' \mathrm{Ad}_a(T)$ of products of elements in $T', \mathrm{Ad}_a(T)$ is connected and abelian, hence it is a torus. Since $T', \mathrm{Ad}_a(T)$ are maximal tori, we conclude $T' = T' \mathrm{Ad}_a(T) = \mathrm{Ad}_a(T)$. □

Definition 13.7. The *rank* l of a compact, connected Lie group is the dimension of its maximal torus.

For example, $\mathrm{U}(n)$ has maximal torus given by diagonal matrices. Its rank is thus $l = n$. We will discuss the maximal tori of the classical groups further below.

Exercise 13.8. The group $\mathrm{SU}(2)$ has maximal torus T the set of diagonal matrices $\mathrm{diag}(z, z^{-1})$. Another natural choice of a maximal torus is $T' = \mathrm{SO}(2) \subset \mathrm{SU}(2)$. Find all elements $a \in G$ such that $\mathrm{Ad}_a(T) = T'$.

The Lie algebra $\mathfrak{t}$ of a maximal torus T is a maximal abelian subalgebra of $\mathfrak{g}$, where a subalgebra is called abelian if it is commutative. Conversely, for any maximal abelian subalgebra the subgroup $\exp(\mathfrak{t})$ is automatically closed, hence is a maximal torus. Cartan's theorem implies that any two maximal abelian subalgebras $\mathfrak{t}, \mathfrak{t}'$ are conjugate under the adjoint representation. That is, there exists $a \in G$ such that $\mathrm{Ad}_a(\mathfrak{t}) = \mathfrak{t}'$.

Proposition 13.9 (Properties of maximal tori)**.** (a) *Any element of a Lie group is contained in some maximal torus. That is, if T is a fixed maximal torus then*

$$\bigcup_{a \in G} \mathrm{Ad}_a(T) = G.$$

On the other hand, the intersection of all maximal tori is the center of G:

$$\bigcap_{a \in G} \mathrm{Ad}_a(T) = Z(G).$$

(b) *If $H \subset G$ is a subtorus, and $g \in Z_G(H)$ lies in the centralizer of H (the elements of G commuting with all elements in H), there exists a maximal torus T containing H and g.*

(c) *Maximal tori are maximal abelian groups: If $g \in Z_G(T)$ then $g \in T$.*

Proof. (a) Given $g \in G$ choose ξ with $\exp(\xi) = g$, and let $\mathfrak{t}$ be a maximal abelian subalgebra of $\mathfrak{g}$ containing ξ. Then $T = \exp(\mathfrak{t})$ is a maximal torus containing g. Now suppose $c \in \bigcap_{a \in G} \mathrm{Ad}_a(T)$. Since $c \in \mathrm{Ad}_a(T)$, it commutes with all elements in $\mathrm{Ad}_a(T)$. Since $G = \bigcup_{a \in G} \mathrm{Ad}_a(T)$ it commutes with all elements of G, that is, $c \in Z(G)$. This proves $\bigcap_{a \in G} \mathrm{Ad}_a(T) \subset Z(G)$; the opposite inclusion is a consequence of (b) to be proved below.

(b) Given $g \in Z_G(H)$, we obtain a closed abelian subgroup

$$B = \overline{\bigcup_{k \in \mathbb{Z}} g^k H}.$$

Let B_0 be the identity component, which is thus a torus, and let $m \geq 0$ be the smallest number with $g^m \in B_0$. Then B has connected components $B_0,\ gB_0,\ \ldots, g^{m-1}B_0$. The element $g^m \in B_0$ can be written in the form k^m with $k \in B_0$. Thus $h = gk^{-1} \in gB_0$ satisfies $h^m = e$. It follows that h generates a subgroup isomorphic to $\mathbb{Z}_m$, and the product map $B_0 \times \mathbb{Z}_m \to B,\ (t, h^i) \mapsto th^i$ is an isomorphism.

Pick a topological generator $b \in B_0$ of the torus B_0. Then b^m is again a topological generator of B_0 (by Kronecker's Lemma). Thus bh is a topological generator of B. But bh is contained in some maximal torus T. Hence $B \subset T$.

(c) By (b) there exists a maximal torus T' containing T and g. But T already is a maximal torus. Hence $g \in T' = T$.

□

Exercise 13.10. Show that the subgroup of diagonal matrices in $\mathrm{SO}(n)$, $n \geq 3$ is maximal abelian. Since this is a discrete subgroup, this illustrates that maximal abelian subgroups need not be maximal tori.

Proposition 13.11. $\dim(G/T)$ *is even.*

Proof. Fix an invariant inner product on $\mathfrak{g}$. Since G is connected, the adjoint representation takes values in $\mathrm{SO}(\mathfrak{g})$. The action of $T \subset G$ fixes $\mathfrak{t}$, hence it restricts to a representation

$$T \to \mathrm{SO}(\mathfrak{t}^{\perp})$$

8where $\mathfrak{t}^{\perp} \cong \mathfrak{g}/\mathfrak{t}$ is the orthogonal complement with respect to B. Let $t \in T$ be a topological generator. Then $\mathrm{Ad}(t)|_{\mathfrak{t}^{\perp}}$ has no eigenvalue 1. But any special orthogonal transformation on an odd-dimensional Euclidean vector space fixes at least one vector. (Exercise.) Hence $\dim(\mathfrak{g}/\mathfrak{t})$ is even. □

13.3. **The Weyl group.** For any subset $S \subset G$ of a Lie group, one defines its *normalizer* $N(S)$ (sometimes written $N_G(S)$ for clarity) to be the group of elements $g \in G$ such that $\mathrm{Ad}_g(S) \subset S$. If H is a closed subgroup of G, then $N(H)$ is a closed subgroup. Since H is a *normal* subgroup of $N(H)$, the quotient $N(H)/H$ inherits a group structure.

We are mainly interested in the normalizer of T. Thus, $N(T)$ is the stabilizer of T for the G-action on the set of maximal tori. By Cartan's theorem, this action is transitive, hence the quotient space $G/N(T)$ is identified with the set of maximal tori. The adjoint action of $T \subset N(T)$ on T is of course trivial, but there is a non-trivial action of the quotient $N(T)/T$.

Definition 13.12. Let G be a compact, connected Lie group with maximal torus T. The quotient

$$W = N_G(T)/T$$

is called the *Weyl group* of G relative to T.

Since any two maximal tori are conjugate, the Weyl groups are independent of T up to isomorphism. More precisely, if T, T' are two maximal tori, and $a \in G$ with $T' = \mathrm{Ad}_a(T)$, then $N(T') = \mathrm{Ad}_a(N(T))$, and hence Ad_a defines an isomorphism $W \to W'$. There are many natural actions of the Weyl group:

(a) W acts on T. This action is induced by the conjugation action of $N(T)$ on T (since $T \subset N(T)$ acts trivially). Note that this action on T is by Lie group automorphisms.
(b) W acts on $\mathfrak{t}$. This action is induced by the adjoint representation of $N(T) \subset G$ on T (since $T \subset N(T)$ acts trivially). Of course, the action on $\mathfrak{t}$ is just the differential of the action on T.
(c) W acts on the lattice Λ, the kernel of the exponential map $\exp\colon \to T$. Indeed, $\exp\colon \mathfrak{t} \to T$ is an $N(T)$-equivariant, hence W-equivariant, group morphism. Thus its kernel is a W-invariant subset of $\mathfrak{t}$.
(d) W acts on G/T. The action of $N(T)$ on G by multiplication from the right (i.e $n.g = gn^{-1}$) descends to a well-defined action of $W = N(T)/T$ on G/T:

$$w.(gT) = gn^{-1}T$$

where $n \in N(T)$ represents w. Note that this action is *free*, that is, all stabilizer groups are trivial. The quotient of the W-action on G/T is $G/N(T)$, the space of maximal tori of G.

Example 13.13. For $G = \mathrm{SU}(2)$, with maximal torus T consisting of diagonal matrices, we have $N(T) = T \cup nT$ where

$$n = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Thus $W = N(T)/T = \mathbb{Z}_2$, with n descending to the non-trivial generator. One easily checks that the conjugation action of n on T permutes the two diagonal entries. The action on $\mathfrak{t}$ is given by reflection, $\xi \mapsto -\xi$. The action on $S^2 = G/T$ is the antipodal map, hence $(G/T)/W = G/N(T) \cong \mathbb{R}P(2)$.

Example 13.14. Let $G = \mathrm{SO}(3)$, with maximal torus given by rotations about the 3-axis. Thus, T consists of matrices

$$g(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The normalizer $N(T)$ consist of all rotations preserving the 3-axis. The induced action on the 3-axis preserves the inner product, hence it is either trivial or the reflection. Elements in $N(T)$ fixing the axis are exactly the elements of T itself. The elements in $N(T)$ reversing the axis are the rotations by π about any axis orthogonal to the 3-axis. Thus $W = \mathbb{Z}_2$.

Theorem 13.15. *The Weyl group W of a compact, connected group G is a finite group.*

Proof. We have to show that the identity component $N(T)_0$ is T. Clearly, $T \subset N(T)_0$. For the other direction, consider the adjoint representation of $N(T)$ on $\mathfrak{t}$. As mentioned above this action preserves the lattice Λ. Since Λ is discrete, the identity component $N(T)_0$ acts trivially on Λ. It follows that $N(T)_0$ acts trivially on $\mathfrak{t} = \mathrm{span}_{\mathbb{R}}(\Lambda)$ and hence also on $T = \exp(\mathfrak{t})$. That is, $N(T)_0 \subset Z_G(T) = T$. □

Proposition 13.16. *The action of W on T (and likewise the action on $\mathfrak{t}, \Lambda$) is faithful. That is, if a Weyl group element w acts trivially then $w = 1$.*

Proof. If w acts trivially on Λ, then it also acts trivially on $\mathfrak{t} = \mathrm{span}_{\mathbb{R}}(\Lambda)$. If w acts trivially on $\mathfrak{t}$, then it also acts trivially on $T = \exp(\mathfrak{t})$. If w acts trivially on T, then any element $n \in N(T)$ representing w lies in $Z(T) = T$, thus $w = 1$. □

Exercise 13.17. a) Let $\phi\colon G \to G'$ be a surjective morphism of compact connected Lie groups. Show that if $T \subset G$ is a maximal torus in G, then $T' = \phi(T)$ is a maximal torus in G', and that the image $\phi(N(T))$ of the normalizer of T lies inside $N(T')$. Thus ϕ induces a group morphism of Weyl groups, $W \to W'$.

b) Let $\phi\colon G \to G'$ be a covering morphism of compact connected Lie groups. Let T' be a maximal torus in G'. Show that $T = \phi^{-1}(T')$ is a maximal torus in G, with normalizer $N(T) = \phi^{-1}(N(T'))$. Thus, G, G' have isomorphic Weyl groups: $W \cong W'$.

13.4. **Maximal tori and Weyl groups for the classical groups.** We will now describe the maximal tori and the Weyl groups for the classical groups. Recall that if T is a maximal torus, then the Weyl group action

$$W \to \mathrm{Aut}(T) \cong \mathrm{Aut}(\Lambda)$$

is by automorphism. Since W is finite, its image must lie in a compact subgroup of T. Recall also that for the standard torus $(\mathbb{R}/\mathbb{Z})^l$, a maximal compact subgroup of $\mathrm{Aut}((\mathbb{R}/\mathbb{Z})^l)$ is

$$\mathrm{O}(l, \mathbb{Z}) = (\mathbb{Z}_2)^l \rtimes S_l \subset \mathrm{GL}(l, \mathbb{Z}) = \mathrm{Aut}((\mathbb{R}/\mathbb{Z})^l).$$

To compute the Weyl group in the following examples of matrix Lie groups, we take into account that the Weyl group action must preserve the set of eigenvalues of matrices $t \in T$.

13.4.1. *The unitary and special unitary groups.* For $G = \mathrm{U}(n)$, the diagonal matrices

$$\mathrm{diag}(z_1, \ldots, z_n) = \begin{pmatrix} z_1 & 0 & 0 & \cdots & 0 \\ 0 & z_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & z_n \end{pmatrix}$$

with $|z_i| = 1$ define a maximal torus. Indeed, conjugation of a matrix $g \in \mathrm{U}(n)$ by $t = \mathrm{diag}(z_1, \ldots, z_n)$ gives $(tgt^{-1})_{ij} = z_i g_{ij} z_j^{-1}$. If $i \neq j$, this equals g_{ij} for all z_i, z_j if and only if $g_{ij} = 0$. Thus g is diagonal, as claimed.

The subgroup of $\mathrm{Aut}(T)$ preserving eigenvalues of matrices is the symmetric group S_n, acting by permutation of diagonal entries. Hence we obtain an injective group morphism

$$W \to S_n.$$

We claim that this map is an isomorphism. Indeed, let $n \in G$ be the matrix with $n_{i,i+1} = 1 = -n_{i+1,i}$, $n_{jj} = 1$ for $j \neq i, i+1$, and all other entries equal to zero. Conjugation by n preserves T, thus $n \in N(T)$, and the action on T exchanges the i-th and $i+1$-st diagonal entries. Hence all transpositions are in the image of W. But transpositions generate all of W.

The discussion for $G = \mathrm{SU}(n)$, is similar. The diagonal matrices $\mathrm{diag}(z_1, \ldots, z_n)$ with $|z_i| = 1$ and $\prod_{i=1}^n z_i = 1$ are a maximal torus $T \subset \mathrm{SU}(n)$, and the Weyl group is $W = S_n$, just as in the case of $\mathrm{U}(n)$.

13.4.2. *The special orthogonal groups* $\mathrm{SO}(2m)$. The group of block diagonal matrices

$$t(\theta_1, \ldots, \theta_m) = \begin{pmatrix} R(\theta_1) & 0 & 0 & \cdots & 0 \\ 0 & R(\theta_2) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & R(\theta_m) \end{pmatrix}$$

is a torus $T \subset \mathrm{SO}(2m)$. To see that it is maximal, consider conjugation of a given $g \in \mathrm{SO}(2m)$ by $t = t(\theta_1, \ldots, \theta_m)$. Writing g in block form with 2×2-blocks $g_{ij} \in \mathrm{Mat}_2(\mathbb{R})$, we have $(tgt^{-1})_{ij} = R(\theta_i) g_{ij} R(-\theta_j)$. Thus $g \in Z_G(T)$ if and only if $R(\theta_i) g_{ij} = g_{ij} R(\theta_j)$ for all i, j, and all $\theta_1, \ldots, \theta_m$. For $i \neq j$, taking $\theta_j = 0$ and $\theta_i = \pi$, this shows $g_{ij} = 0$. Thus g is block diagonal with blocks $g_{ii} \in \mathrm{O}(2)$ satisfying $R(\theta_i) g_{ii} = g_{ii} R(\theta_i)$. Since a reflection does not commute with all rotations, we must in fact have $g_{ii} \in \mathrm{SO}(2)$. This confirms that T is a maximal torus.

The eigenvalues of the element $t(\theta_1, \ldots, \theta_m)$ are $e^{i\theta_1}, e^{-i\theta_1}, \ldots, e^{i\theta_m}, e^{-i\theta_m}$. The subgroup of $\mathrm{Aut}(T)$ preserving the set of eigenvalues of matrices is thus $(\mathbb{Z}_2)^m \rtimes S_m$, where S_m acts by permutation of the θ_i, and $(\mathbb{Z}_2)^m$ acts by sign changes. That is, we have an injective group morphism

$$W \to (\mathbb{Z}_2)^m \rtimes S_m.$$

To describe its image, let $\Gamma_m \subset (\mathbb{Z}_2)^m$ be the kernel of the product map $(\mathbb{Z}_2)^m \to \mathbb{Z}_2$, corresponding to an even number of sign changes.

Theorem 13.18. *The Weyl group W of* $\mathrm{SO}(2m)$ *is the semi-direct product* $\Gamma_m \rtimes S_m$.

Proof. The matrix $g \in \mathrm{SO}(2m)$, written in block form with 2×2-blocks, with entries $g_{ij} = -g_{ji} = I$, $g_{kk} = I$ for $k \neq i, j$, and all other blocks equal to zero, lies in $N(T)$. The corresponding Weyl group element permutes the i-th and j-th blocks of any $t \in T$. Hence $S_n \subset W$. Next, observe that

$$K = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{O}(2).$$

satisfies $KR(\theta)R^{-1} = R(-\theta)$. The block diagonal matrix, with blocks K in the i-th and j-th diagonal entries, and identity matrices for the other diagonal entries, lies in $N(T)$ and its action on T changes $R(\theta_i), R(\theta_j)$ to $R(-\theta_i), R(-\theta_j)$. Hence, we obtain all even numbers of sign changes, confirming $\Gamma_n \subset W$. It remains to show that the transformation $t(\theta_1, \theta_2, \ldots, \theta_m) \mapsto t(-\theta_1, \theta_2 \ldots, \theta_m)$ does *not* lie in W. Suppose $g \in N(T)$ realizes this transformation, so that $gt(\theta_1, \theta_2, \ldots, \theta_m)g^{-1} = t(-\theta_1, \theta_2, \ldots, \theta_m)$. As above, we find $R(\theta_i)g_{ij}R(-\theta_j) = g_{ij}$ for $i \geq 2$, and $R(-\theta_1)g_{1j} = g_{1j}R(\theta_j)$ for $i = 1$, for all $\theta_1, \ldots, \theta_m$. As before, we deduce from these equations that g most be block diagonal. Thus $g \in (\mathrm{O}(2) \times \cdots \times \mathrm{O}(2)) \cap \mathrm{SO}(2m)$. From $R(\theta_i)g_{ii}R(-\theta_i) = g_{ii}$ for $i \geq 2$ we obtain $g_{ii} \in \mathrm{SO}(2)$ for $i > 1$. Since $\det(g) = 1$, this forces $g_{11} \in \mathrm{SO}(2)$, which however is incompatible with $R(-\theta_1)g_{11} = g_{11}R(\theta_1)$. $\square$

13.4.3. *The special orthogonal groups* $\mathrm{SO}(2m+1)$. Define an inclusion

$$j\colon\ \mathrm{O}(2m) \to \mathrm{SO}(2m+1),$$

placing a given orthogonal matrix A in the upper left corner and $\det(A)$ in the lower left corner. Let T' be the standard maximal torus for $\mathrm{SO}(2m)$, and $N(T')$ its normalizer. Then $T = j(T')$ is a maximal torus for $\mathrm{SO}(2m+1)$. The proof that T is maximal is essentially the same as for $\mathrm{SO}(2m)$.

Theorem 13.19. *The Weyl group of* $\mathrm{SO}(2m+1)$ *is the semi-direct product* $(\mathbb{Z}_2)^m \rtimes S_m$.

Proof. As in the case of $\mathrm{SO}(2m)$, we see that the Weyl group must be a subgroup of $(\mathbb{Z}_2)^m \rtimes S_m$. Since $j(N(T')) \subset N(T)$, we have an inclusion of Weyl groups $W' = \Gamma_m \rtimes S_m \subset W$. Hence we only need to show that the first $\mathbb{Z}_2$ is contained in W. The block diagonal matrix $g \in \mathrm{O}(2m)$ with entries $K, I, \ldots, I$ down the diagonal satisfies $gt(\theta_1, \ldots, \theta_m)g^{-1} = t(-\theta_1, \ldots, \theta_m)$. Hence $j(g) \in N(T)$ represents a generator of the $\mathbb{Z}_2$. $\square$

13.4.4. *The symplectic groups.* Recall that $\mathrm{Sp}(n)$ is the subgroup of $\mathrm{Mat}_n(\mathbb{H})^\times$ preserving the norm on $\mathbb{H}^n$. Alternatively, using the identification $\mathbb{H} = \mathbb{C}^2$, one can realize $\mathrm{Sp}(n)$ as a subgroup of $\mathrm{U}(2n)$, consisting of matrices of the form

$$\begin{pmatrix} A & -\overline{B} \\ B & \overline{A} \end{pmatrix} \tag{3}$$

with $A, B \in \mathrm{Mat}_n(\mathbb{C})$ and $A^\dagger A + B^\dagger B = I,\ A^T B = B^T A$. Let T be the torus consisting of the diagonal matrices in $\mathrm{Sp}(n)$. Letting $Z = \mathrm{diag}(z_1, \ldots, z_n)$, these are the matrices of the form

$$t(z_1, \ldots, z_n) = \begin{pmatrix} Z & 0 \\ 0 & \overline{Z} \end{pmatrix}.$$

As before, we see that a matrix in $\mathrm{Sp}(n)$ commutes with all these diagonal matrices if and only if it is itself diagonal. Hence T is a maximal torus. Note that T is the image of the maximal torus of $\mathrm{U}(n)$ under the inclusion

$$r\colon\ \mathrm{U}(n) \to \mathrm{Sp}(n),\ A \mapsto \begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix}$$

Theorem 13.20. *The Weyl group of* $\mathrm{Sp}(n)$ *is* $(\mathbb{Z}_2)^n \rtimes S_n$.

Proof. The subgroup of $\mathrm{Aut}(T)$ preserving eigenvalues is $(\mathbb{Z}_2)^n \rtimes S_n$. Hence, $W \subset (\mathbb{Z}_2)^n \rtimes S_n$. The inclusion r defines an inclusion of the Weyl group of $\mathrm{U}(n)$, hence $S_n \subset W$. On the other hand, one obtains all 'sign changes' using conjugation with appropriate matrices. E.g. the sign change $t(z_1, z_2, z_3, \ldots, z_n) \mapsto t(z_1, z_2^{-1}, z_3, \ldots, z_n)$ is obtained using conjugation by a matrix (3) with $A = \mathrm{diag}(1, 0, 1, \ldots, 1),\ B = \mathrm{diag}(0, 1, 0, \ldots, 0)$. □

13.4.5. *The spin groups.* For $n \geq 3$, the special orthogonal group $\mathrm{SO}(n)$ has fundamental group $\mathbb{Z}_2$. Its universal cover is the spin group $\mathrm{Spin}(n)$. By the general result for coverings, the pre-image of a maximal torus of $\mathrm{SO}(n)$ is a maximal torus of $\mathrm{Spin}(n)$, and the Weyl groups are isomorphic.

13.4.6. *Notation.* Let us summarize the results above, and at the same time introduce some notation. Let A_l, B_l, C_l, D_l be the Lie groups $\mathrm{SU}(l+1),\ \mathrm{Spin}(2l+1),\ \mathrm{Sp}(l),\ \mathrm{Spin}(2l)$. Here the lower index l signifies the rank. We have the following table:

	rank	name	dim	W
A_l	$l \geq 1$	$\mathrm{SU}(l+1)$	$l^2 + 2l$	S_{l+1}
B_l	$l \geq 2$	$\mathrm{Spin}(2l+1)$	$2l^2 + l$	$(\mathbb{Z}_2)^l \rtimes S_l$
C_l	$l \geq 3$	$\mathrm{Sp}(l)$	$2l^2 + l$	$(\mathbb{Z}_2)^l \rtimes S_l$
D_l	$l \geq 4$	$\mathrm{Spin}(2l)$	$2l^2 - l$	$(\mathbb{Z}_2)^{l-1} \rtimes S_l$

In the last row, $(\mathbb{Z}_2)^{l-1}$ is viewed as the subgroup of $(\mathbb{Z}_2)^l$ of tuples with product equal to 1.

Remarks 13.21. (a) Note that the groups $\mathrm{Sp}(l)$ and $\mathrm{Spin}(2l+1)$ have the same rank and dimension, and isomorphic Weyl groups.

(b) For rank $l = 1$, $\mathrm{Sp}(1) \cong \mathrm{SU}(2) \cong \mathrm{Spin}(3)$. For rank $l = 2$, it is still true that $\mathrm{Sp}(2) \cong \mathrm{Spin}(5)$. But for $l > 2$ the two groups $\mathrm{Spin}(2l+1),\ \mathrm{Sp}(l)$ are non-isomorphic. To exclude such coincidences, and to exclude the non-simple Lie groups $\mathrm{Spin}(4) = \mathrm{SU}(2) \times \mathrm{SU}(2)$, one restricts the range of l as indicated above.

(c) As we will discuss later, the table is a complete list of simple, simply connected compact Lie groups, with the exception of five aptly named *exceptional Lie groups* F_4, G_2, E_6, E_7, E_8 that are more complicated to describe.

14. Weights and roots

14.1. **Weights and co-weights.** Let T be a torus, with Lie algebra $\mathfrak{t}$.

Definition 14.1. A *weight of* T is a Lie group morphism $\mu\colon T \to \mathrm{U}(1)$. A *co-weight of* T is a Lie group morphism $\gamma\colon \mathrm{U}(1) \to T$. We denote by $X^*(T)$ the set of all weights, and by $X_*(T)$ the set of co-weights.

Let us list some properties of the weights and coweights.

- Both $X^*(T)$ and $X_*(T)$ are abelian groups: two weights μ, μ' can be added as
$$(\mu' + \mu)(t) = \mu'(t)\mu(t),$$
and two co-weights γ, γ' can be added as
$$(\gamma' + \gamma)(z) = \gamma'(z)\gamma(z).$$
- For $T = \mathrm{U}(1)$ we have a group isomorphism
$$X_*(\mathrm{U}(1)) = \mathrm{Hom}(\mathrm{U}(1), \mathrm{U}(1)) = \mathbb{Z},$$
where the last identification (the *winding number*) associates to $k \in \mathbb{Z}$ the map $z \mapsto z^k$. Likewise $X^*(\mathrm{U}(1)) = \mathbb{Z}$.
- Given tori T, T' and a Lie group morphism $T \to T'$ one obtains group morphisms
$$X_*(T) \to X_*(T'), \quad X^*(T') \to X^*(T)$$
by composition.
- For a product of two tori T_1, T_2,
$$X^*(T_1 \times T_2) = X^*(T_1) \times X^*(T_2), \quad X_*(T_1 \times T_2) = X_*(T_1) \times X_*(T_2).$$
This shows in particular $X^*(\mathrm{U}(1)^l) = \mathbb{Z}^l,\ X_*(U(1)^l) = \mathbb{Z}^l$. Since any T is isomorphic to $\mathrm{U}(1)^l$, this shows that the groups $X^*(T), X_*(T)$ are free abelian of rank $l = \dim T$. That is, they are lattices inside the vector spaces $X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$ resp. $X_*(T) \otimes_{\mathbb{Z}} \mathbb{R}$.
- The lattices $X^*(T)$ and $X_*(T)$ are dual. The pairing $\langle \mu, \gamma \rangle$ of $\mu \in X^*(T)$ and $\gamma \in X_*(T)$ is the composition $\mu \circ \gamma \in \mathrm{Hom}(\mathrm{U}(1), \mathrm{U}(1)) \cong \mathbb{Z}$.

Remark 14.2. Sometimes, it is more convenient or more natural to write the group $X^*(T)$ multiplicatively. This is done by introducing symbols e_μ corresponding to $\mu \in X^*(T)$, so that the group law becomes $\mathsf{e}_\mu \mathsf{e}_\nu = \mathsf{e}_{\mu+\nu}$.

Remark 14.3. Let $\Lambda \subset \mathfrak{t}$ be the integral lattice, and $\Lambda^* = \mathrm{Hom}(\Lambda, \mathbb{Z})$ its dual. For any weight μ, the differential of $\mu\colon T \to \mathrm{U}(1)$ is a Lie algebra morphism $\mathfrak{t} \to \mathfrak{u}(1) = i\mathbb{R}$, taking Λ to $2\pi i\mathbb{Z}$. Conversely, any group morphism $\Lambda \to 2\pi i\mathbb{Z}$ arises in this way. We may thus identify $X^*(T)$ with $2\pi i\Lambda^* \subset \mathfrak{t}^* \otimes \mathbb{C}$. Similarly, $X_*(T)$ is identified with $\frac{1}{2\pi i}\Lambda \subset \mathfrak{t} \otimes \mathbb{C}$.

Sometimes, it is more convenient or more natural to absorb the $2\pi i$ factor in the definitions, so that $X^*(T), X_*(T)$ are identified with Λ^*, Λ. For the time being, we will avoid any such identifications altogether.

Exercise 14.4. An element $t_0 \in T$ is a topological generator of T if and only if the only weight $\mu \in X^*(T)$ with $\mu(t_0) = 1$ is the zero weight.

Exercise 14.5. There is a natural identification of $X_*(T)$ with the fundamental group $\pi_1(T)$.

Exercise 14.6. Let

$$1 \to \Gamma \to T' \to T \to 1$$

be a finite cover, where T, T' are tori and $\Gamma \subset T'$ a finite subgroup. Then there is an exact sequence of groups

$$1 \to X_*(T') \to X_*(T) \to \Gamma \to 1.$$

Similarly, there is an exact sequence

$$1 \to X^*(T) \to X^*(T') \to \widehat{\Gamma} \to 1,$$

with the finite group $\widehat{\Gamma} = \mathrm{Hom}(\Gamma, \mathrm{U}(1))$.

14.2. **Schur's Lemma.** To proceed, we need the following simple but important fact.

Lemma 14.7 (Schur Lemma)**.** *Let G be any group, and $\pi\colon G \to \mathrm{GL}(V)$ a finite-dimensional irreducible complex representation.*

(a) *If $A \in \mathrm{End}(V)$ commutes with all $\pi(g)$, then A is a multiple of the identity matrix.*
(b) *If V' is another finite-dimensional irreducible G-representation, then*

$$\dim(\mathrm{Hom}_G(V, V')) = \begin{cases} 1 & \text{if } V \cong V' \\ 0 & \text{otherwise} \end{cases}$$

Proof. a) Let λ be an eigenvalue of A. Since $\ker(A - \lambda)$ is G-invariant, it must be all of V. Hence $A = \lambda I$. b) For any G-equivariant map $A\colon V \to V'$, the kernel and range of A are sub-representations. Hence $A = 0$ or A is an isomorphism. If V, V' are non-isomorphic, A cannot be an isomorphism, so $A = 0$. If V, V' are isomorphic, so that we might as well assume $V' = V$, b) follows from a). □

For any two complex G-representations V, W, one calls $\mathrm{Hom}_G(V, W)$ the space of intertwining operators. If V is irreducible, and the representation W is completely reducible (as is automatic for G a compact Lie group), the dimension $\dim \mathrm{Hom}_G(V, W)$ is the multiplicity of V in W. The range of the map

$$\mathrm{Hom}_G(V, W) \otimes V \to W,\ \ A \otimes v \mapsto A(v)$$

is the *V-isotypical subspace* of W, i.e. the sum of all irreducible components isomorphic to V.

14.3. **Weights of T-representations.** For any $\mu \in X^*(T)$, let $\mathbb{C}_\mu$ denote the T-representation on $\mathbb{C}$, with T acting via the homomorphism $\mu\colon T \to \mathrm{U}(1)$.

Proposition 14.8. *Any finite-dimensional irreducible representation of T is isomorphic to $\mathbb{C}_\mu$, for a unique weight $\mu \in X^*(T)$. Thus, $X^*(T)$ labels the isomorphism classes of finite-dimensional irreducible T-representations.*

Proof. Let $\pi\colon T \to \mathrm{GL}(V)$ be irreducible. Since T is abelian, Schur's lemma shows that all $\pi(t)$ act by scalars. Hence any $v \in V$ spans an invariant subspace. Since V is irreducible, it follows that $\dim V = 1$, and the basis vector v gives an isomorphism $V \cong \mathbb{C}$. The image $\pi(T) \subset \mathrm{GL}(V) = \mathrm{GL}(1, \mathbb{C})$ is a compact subgroup, hence it must lie in $\mathrm{U}(1)$. Thus, π becomes a morphism $\mu\colon T \to \mathrm{U}(1)$. □

Any finite-dimensional complex T-representation V has a unique direct sum decomposition $V = \bigoplus_{\mu \in X^*(T)} V_\mu$, where the V_μ are the $\mathbb{C}_\mu$-isotypical subspaces. Thus V_μ is the subspace on which elements $t \in T$ act as scalar multiplication by $\mu(t)$. Note that since $\dim C_\mu = 1$, the dimension of the space of intertwining operators coincides with the dimension of V_μ. This is called the *multiplicity* of the weight μ in V. We say that $\mu \in X^*(T)$ is a *weight* of V if $V_\mu \neq 0$, i.e. if the multiplicity is > 0, in this case V_μ is called a *weight space*.

Let $\Delta(V) \subset X^*(T)$ be the set of all weights of the representation V. Then

$$V = \bigoplus_{\mu \in \Delta(V)} V_\mu.$$

Simple properties are

$$\begin{aligned} \Delta(V_1 \oplus V_2) &= \Delta(V_1) \cup \Delta(V_2), \\ \Delta(V_1 \otimes V_2) &= \Delta(V_1) + \Delta(V_2), \\ \Delta(V^*) &= -\Delta(V). \end{aligned}$$

If V is the complexification of a *real* T-representation, or equivalently if V admits a T-equivariant conjugate linear involution $C \colon V \to V$, one has the additional property,

$$\Delta(V) = -\Delta(V).$$

Indeed, C restricts to conjugate linear isomorphisms $V_\mu \to V_{-\mu}$, hence weights appear in pairs $+\mu, -\mu$ of equal multiplicity.

14.4. **Weights of G-representations.** Let G be a compact connected Lie group, with maximal torus T. The Weyl group W acts on the coweight lattice by

$$(w.\gamma)(z) = w.\gamma(z), \quad \gamma \in X_*(T),$$

and on the weight lattice by

$$(w.\mu)(t) = \mu(w^{-1}t), \quad \mu \in X^*(T).$$

The two actions are dual, that is, the pairing is preserved: $\langle w.\mu,\ w.\gamma \rangle = \langle \mu, \gamma \rangle$.

Given a finite-dimensional complex representation $\pi \colon G \to \mathrm{GL}(V)$, we define the set $\Delta(V)$ of weights of V to be the weights of a maximal torus $T \subset G$.

Proposition 14.9. *Let G be a compact Lie group, and T its maximal torus. For any finite-dimensional G-representation $\pi \colon G \to \mathrm{End}(V)$, the set of weights is W-invariant:*

$$W.\Delta(V) = \Delta(V).$$

In fact one has $\dim V_{w.\mu} = \dim V_\mu$.

Proof. Let $g \in N(T)$ represent the Weyl group element $w \in W$. If $v \in V_\mu$ we have

$$\pi(t)\pi(g)v = \pi(g)\pi(w^{-1}(t))v = \mu(w^{-1}(t))\pi(g)v = (w.\mu)(t)\pi(g)v.$$

Thus $\pi(g)$ defines an isomorphism $V_\mu \to V_{w\mu}$. $\square$

Example 14.10. Let $G = \mathrm{SU}(2)$, with its standard maximal torus $T \cong \mathrm{U}(1)$ consisting of diagonal matrices $t = \mathrm{diag}(z, z^{-1})$, $|z| = 1$. Let ϖ be the generator of $X^*(T)$ given by $\varpi(t) = z$. The weights of the representation $V(k)$ of $\mathrm{SU}(2)$ are

$$\Delta(V(k)) = \{k\varpi,\ (k-2)\varpi, \ldots, -k\varpi\}.$$

All of these weights have multiplicity 1. The Weyl group $W = \mathbb{Z}_2$ acts sign changes of weights.

Example 14.11. Let $G = \mathrm{U}(n)$ with its standard maximal torus $T = \mathrm{U}(1)^n$ given by diagonal matrices. Let $\epsilon^i \in X^*(T)$ be the projection to the i-th factor. The defining representation of $\mathrm{U}(n)$ has set of weights,

$$\Delta(\mathbb{C}^n) = \{\epsilon^1, \ldots, \epsilon^n\},$$

all with multiplicity 1. The weights of the representation on the k-th exterior power $\wedge^k\mathbb{C}^n$ are

$$\Delta(\wedge^k\mathbb{C}^n) = \{\epsilon^{i_1} + \ldots + \epsilon^{i_k} |\ i_1 < \ldots < i_k\},$$

all with multiplicity 1. (The k-fold wedge products of basis vectors are weight vectors.) The weights for the action on $S^k\mathbb{C}^n$ are

$$\Delta(S^k\mathbb{C}^n) = \{\epsilon^{i_1} + \ldots + \epsilon^{i_k} |\ i_1 \le \ldots \le i_k\}.$$

(The k-fold products of basis vectors, possibly with repetitions, are weight vectors.) The multiplicity of the weight μ is the number of ways of writing it as a sum $\mu = \epsilon^{i_1} + \ldots + \epsilon^{i_k}$.

14.5. **Roots.** The adjoint representation is of special significance, as it is intrinsically associated to any Lie group. Let G be compact, connected, with maximal torus T.

Definition 14.12. A *root* of G is a non-zero weight for the adjoint representation on $\mathfrak{g}^{\mathbb{C}}$. The set of roots is denoted $\mathfrak{R} \subset X^*(T)$.

The weight spaces $\mathfrak{g}_\alpha \subset \mathfrak{g}^{\mathbb{C}}$ for roots $\alpha \in \mathfrak{R}$ are called the *root spaces*. As remarked above, $\mathfrak{g}_{-\alpha}$ is obtained from $\mathfrak{g}_\alpha$ by complex conjugation. The weight space $\mathfrak{g}_0$ for the weight 0 is the subspace fixed under the adjoint action of T, that is, $\mathfrak{t}^{\mathbb{C}}$. Hence

$$\mathfrak{g}^{\mathbb{C}} = \mathfrak{t}^{\mathbb{C}} \oplus \bigoplus_{\alpha\in\mathfrak{R}} \mathfrak{g}_\alpha.$$

The set $\mathfrak{R} = \Delta(\mathfrak{g}^{\mathbb{C}})\backslash\{0\}$ is a finite W-invariant subset of $\mathfrak{t}^*$.

Example 14.13. Let $G = \mathrm{U}(n)$, and $T = \mathrm{U}(1) \times \cdots \times \mathrm{U}(1)$ its standard maximal torus. Denote by $\epsilon^1, \ldots, \epsilon^n \in X^*(T)$ the standard basis. That is, writing $t = \mathrm{diag}(z_1, \ldots, z_n) \in T$ we have

$$\epsilon^i(t) = z_i.$$

We have $\mathfrak{g} = \mathfrak{u}(n)$, the skew-adjoint matrices, with complexification $\mathfrak{g}^{\mathbb{C}} = \mathfrak{gl}(n, \mathbb{C}) = \mathrm{Mat}_n(\mathbb{C})$ all $n \times n$-matrices. Conjugation of a matrix ξ by t multiplies the i-th row by z_i and the j-th column by z_j^{-1}. Hence, if ξ is a matrix having entry 1 in one (i, j) slot and 0 everywhere else, then $\mathrm{Ad}_t(\xi) = z_i z_j^{-1}\xi$. That is, if $i \neq j$, ξ is a root vector for the root $\epsilon^i - \epsilon^j$. We conclude that the set of roots of $\mathrm{U}(n)$ is

$$\mathfrak{R} = \{\epsilon^i - \epsilon^j |\ i \neq j\} \subset X^*(T).$$

Example 14.14. For $G = \mathrm{SU}(n)$, let T be the maximal torus given by diagonal matrices. Let T' be the maximal torus of $\mathrm{U}(n)$, again consisting of the diagonal matrices. Then $X_*(T) \subset X_*(T')$. In terms of the standard basis $\epsilon_1, \ldots, \epsilon_n$ of $X_*(T')$, the lattice $X_*(T)$ has basis $\epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3, \ldots, \epsilon_{n-1} - \epsilon_n$. The kernel of the projection map $X^*(T') \to X^*(T)$ is a rank 1 lattice generated by $\epsilon^1 + \ldots + \epsilon^n$. Thus, we can think of $X^*(T)$ as a quotient lattice

$$X^*(T) = \mathrm{span}_{\mathbb{Z}}(\epsilon^1, \ldots, \epsilon^n)/\,\mathrm{span}_{\mathbb{Z}}(\epsilon^1 + \ldots + \epsilon^n).$$

The images of $\epsilon^i - \epsilon^j$ under the quotient map are then the roots of $\mathrm{SU}(n)$. The root vectors are the same as for $\mathrm{U}(n)$ (since they all lie in $\mathfrak{sl}(n, \mathbb{C})$).

One can get a picture of the rot system by identifying $X^*(T)$ with the orthogonal projection of $X^*(T')$ to the subspace orthogonal to $\epsilon^1 + \ldots + \epsilon^n$, using the standard inner product on $X^*(T') \otimes_{\mathbb{Z}} \mathbb{R} = \mathbb{R}^n$. Note that the standard inner product is $W = S_n$-invariant, hence this identification respects the Weyl group action. The projections of the ϵ^i are

$$\sigma^i = \epsilon^i - \frac{1}{n}(\epsilon^1 + \ldots + \epsilon^n), \quad i = 1, \ldots, n.$$

Then $\sigma^1, \ldots, \sigma^{n-1}$ are a lattice basis of $X^*(T)$, and $\sigma^n = -(\sigma^1 + \ldots + \sigma^{n-1})$. The roots are $\sigma^i - \sigma^j$ for $i \neq j$. Here is a picture of the weights and roots of SU(3):

PICTURE

Example 14.15. Let $G = \mathrm{SO}(2m)$ with its standard maximal torus $T \cong \mathrm{U}(1)^m$ given by the block diagonal matrices $t(\theta_1, \ldots, \theta_m)$. Let $\epsilon^i \in X^*(T)$ be the standard basis of the weight lattice. Thus $\epsilon^j(t(\theta_1, \ldots, \theta_m)) = e^{i\theta_j}$. The complexified Lie algebra $\mathfrak{g}^{\mathbb{C}} = \mathfrak{so}(2m) \otimes \mathbb{C} =: \mathfrak{so}(2m, \mathbb{C})$ consists of skew-symmetric complex matrices. To find the root vectors, write the elements $X \in \mathfrak{so}(2m, \mathbb{C})$ in block form, with 2×2-blocks $X_{ij} = -X_{ji}^T$. Conjugation by $t(\theta_1, \ldots, \theta_m)$ changes the (i, j)-block to $R(\theta_i) X_{ij} R(-\theta_j)$. Let $v_+, v_- \in \mathbb{C}^2$ be non-zero column vectors with

$$R(\theta)v_+ = e^{i\theta}v_+, \quad R(\theta)v_- = e^{-i\theta}v_-.$$

For instance, we may take

$$v_+ = \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad v_- = \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

The matrices $v_\pm v_\pm^T \in \mathrm{Mat}_2(\mathbb{C})$ satisfy

$$\begin{aligned}
R(\theta)(v_+v_+^T)R(-\psi) &= e^{i(\phi+\psi)}, \\
R(\theta)(v_-v_+^T)R(-\psi) &= e^{i(-\phi+\psi)}, \\
R(\theta)(v_-v_+^T)R(-\psi) &= e^{i(-\phi+\psi)}, \\
R(\theta)(v_-v_-^T)R(-\psi) &= e^{i(-\phi-\psi)}.
\end{aligned}$$

Let $i < j$ be given. Putting $v_\pm v_\pm^T$ in the (i, j) position, and its negative transpose in the (j, i) position, and letting all other entries of X be zero, we obtain root vectors for the roots $\pm\epsilon^i \pm \epsilon^j$.

To summarize, $\mathrm{SO}(2m)$ has $2m(m-1)$ roots

$$\mathfrak{R} = \{\pm\epsilon^i \pm \epsilon^j, \ i < j\}.$$

This checks with dimensions, since $\dim T = m$, $\dim \mathrm{SO}(2m) = 2m^2 - m$, so $\dim \mathrm{SO}(2m)/T = 2(m^2 - m)$.

PICTURE

Example 14.16. Let $G = SO(2m+1)$. We write matrices in block form, corresponding to the decomposition $\mathbb{R}^{2m+1} = \mathbb{R}^2 \oplus \cdots \oplus \mathbb{R}^2 \oplus \mathbb{R}$. Thus, $X \in \mathrm{Mat}_{2m+1}(\mathbb{C})$ has 2×2-blocks X_{ij} for $i, j \le m$, a 1×1-block $X_{m+1,m+1}$, 2×1-blocks $X_{i,m+1}$ for $i \le m$, and 1×2-blocks $X_{m+1,i}$ for $i \le m$. As we saw earlier, the inclusion $\mathrm{SO}(2m) \hookrightarrow \mathrm{SO}(2m+1)$ defines an isomorphism from the maximal torus T' of $\mathrm{SO}(2m)$ to a maximal torus T of $\mathrm{SO}(2m+1)$. The latter consists of block diagonal matrices, with 2×2-blocks $g_{ii} = R(\theta_i)$ for $i = 1, \ldots, m$ and 1×1-block $g_{m+1,m+1} = 1$. Under the inclusion $\mathfrak{so}(2m, \mathbb{C}) \hookrightarrow \mathfrak{so}(2m+1, \mathbb{C})$, root vectors for the former become root vectors for the latter. Hence, all $\pm\epsilon^i \pm \epsilon^j$ are roots, as before. Additional root vectors X are obtained by putting $v_\pm$ as the $X_{i,m+1}$ block and its negative transpose in the $X_{m+1,i}$ block, and letting all other entries be zero. The corresponding roots are $\pm\epsilon^i$. In summary, $\mathrm{SO}(2m+1)$ has roots

$$\mathfrak{R} = \{\pm\epsilon^i \pm \epsilon^j,\ 1 \le i < j \le m\} \cup \{\pm\epsilon^i,\ i = 1, \ldots, m\}.$$

PICTURE

This checks with dimensions: We have found $2m(m-1)+2m = 2m^2$ roots, while $\dim \mathrm{SO}(2m+1)/T = (2m^2+m) - m = 2m^2$. Note that in this picture, the root system for $\mathrm{SO}(2m+1)$ naturally contains that for $\mathrm{SO}(2m)$. Note also the invariance under the Weyl group action in both cases.

Example 14.17. Let $G = \mathrm{Sp}(n)$, viewed as the matrices in $\mathrm{SU}(2n)$ of the form

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

and let T be its standard maximal torus consisting of the diagonal matrices

$$t = \begin{pmatrix} Z & 0 \\ 0 & \overline{Z} \end{pmatrix}$$

with $Z = \mathrm{diag}(z_1, \ldots, z_n)$. Recall that we may view T as the image of the maximal torus $T' \subset \mathrm{U}(n)$ under the inclusion $\mathrm{U}(n) \to \mathrm{Sp}(n)$ taking A to $\begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix}$. As before, we have

$$X^*(T) = \mathrm{span}_{\mathbb{Z}}(\epsilon^1, \ldots, \epsilon^n).$$

To find the roots, recall that the Lie algebra $\mathfrak{sp}(n)$ consists of complex matrices of the form

$$\xi = \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix},$$

with $a^T = \overline{a},\ b^T = b$. Hence its complexification $\mathfrak{sp}(n) \otimes \mathbb{C}$ consists of complex matrices of the form

$$\xi = \begin{pmatrix} a & b \\ c & -a^T \end{pmatrix},$$

with $b^T = b,\ c^T = c$. Thus

$$t\xi t^{-1} = \begin{pmatrix} ZaZ^{-1} & ZbZ \\ Z^{-1}cZ^{-1} & -Z^{-1}a^T Z \end{pmatrix}$$

We can see the following root vectors:

- Taking $a = 0, c = 0$ and letting b be a matrix having 1 in the (i, j) slot and zeroes elsewhere, we obtain a root vector ξ for the root $\epsilon^i + \epsilon^j$.
- Letting $a = 0,\ b = 0$, and letting c be a matrix having 1 in the (i, j) slot and zeroes elsewhere, we obtain a root vector ξ for the root $-\epsilon^i - \epsilon^j$.
- Letting $b = 0,\ c = 0$ and taking for a the matrix having $a_{ij} = 1$ has its only non-zero entry, we obtain a root vector for $\epsilon^i - \epsilon^j$ (provided $i \neq j$).

Hence we have found

$$\frac{n(n+1)}{2} + \frac{n(n+1)}{2} + (n^2 - n) = 2n^2$$

roots:

$$\mathfrak{R} = \{\pm\epsilon^i \pm \epsilon^j | 1 \leq i < j \leq m\} \cup \{\pm 2\epsilon^i |\ i = 1, \ldots, m\}$$

PICTURE

This checks with dimensions: $\dim(\mathrm{Sp}(n)/T) = (2n^2 + n) - n = 2n^2$. Observe that the inclusion $\mathfrak{u}(n) \to \mathfrak{sp}(n)$ takes the root spaces of $\mathrm{U}(n)$ to root spaces of $\mathrm{Sp}(n)$. Hence, the set of roots of $\mathrm{U}(n)$ is naturally a subset of the set of roots of $\mathrm{Sp}(n)$.

Suppose G, G' are compact, connected Lie groups, and $\phi\colon G \to G'$ is a covering map, with kernel Γ. Then ϕ restricts to a covering of the maximal tori, $1 \to \Gamma \to T \to T' \to 1$. hence $X_*(T)$ is a sublattice of $X_*(T')$, with quotient Γ, while $X^*(T')$ is a sublattice of $X^*(T)$, with quotient $\widehat{\Gamma}$. The roots of G are identified with the roots of G' under the inclusion $X^*(T') \to X^*(T)$.

Example 14.18. Let $G' = \mathrm{SO}(2m)$, and $G = \mathrm{Spin}(2m)$ its double cover. Let $\epsilon_1, \ldots, \epsilon_m$ be the standard basis of the maximal torus $T' \cong \mathrm{U}(1)^m$. Each $\epsilon_i\colon \mathrm{U}(1) \to T'$ may be regarded as a loop in $\mathrm{SO}(2m)$, and in fact any of these represents a generator $\pi_1(\mathrm{SO}(2m)) = \mathbb{Z}_2$. With a little work, one may thus show that $X_*(T)$ is the sublattice of $X_*(T')$ consisting of linear combinations $\sum_{i=1}^m a_i \epsilon^i$ with integer coefficients, such that $\sum_{i=1}^m a_i$ is even. Generators for this lattice are, for example, $\epsilon_1 - \epsilon_2,\ \epsilon_2 - \epsilon_3, \ldots, \epsilon_{n-1} - \epsilon_n, \epsilon_{n-1} + \epsilon_n$. Dually, $X^*(T)$ is a lattice containing $X^*(T') = \mathrm{span}_{\mathbb{Z}}(\epsilon^1, \ldots, \epsilon^m)$ as a sublattice. It is generated by $X^*(T')$ together with the vector $\frac{1}{2}(\epsilon^1 + \ldots + \epsilon^n)$. The discussion for $\mathrm{Spin}(2m+1)$ is similar. Here is, for example, a picture of the root system for $\mathrm{Spin}(5)$:

PICTURE

15. Properties of root systems

Let G be a compact, connected Lie group with maximal torus T. We will derive general properties of the set of roots $\mathfrak{R} \subset X^*(T)$ of G, and of the decomposition

$$\mathfrak{g}^{\mathbb{C}} = \mathfrak{t}^{\mathbb{C}} \oplus \bigoplus_{\alpha \in \mathfrak{R}} \mathfrak{g}_\alpha.$$

15.1. **First properties.** We have already seen that the set of roots is W-invariant, and that the roots come in pairs $\pm\alpha$, with complex conjugate root spaces $\mathfrak{g}_{-\alpha} = \overline{\mathfrak{g}_\alpha}$. Another simple property is

Proposition 15.1. *For all* $\alpha, \beta \in \Delta(\mathfrak{g}^{\mathbb{C}}) = \mathfrak{R} \cup \{0\}$,

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$$

In particular, if $\alpha+\beta\notin\Delta(\mathfrak{g}^{\mathbb{C}})$ *then* $[\mathfrak{g}_\alpha,\mathfrak{g}_\beta]=0$. *Furthermore,*

$$[\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}]\subset\mathfrak{t}^{\mathbb{C}}$$

for all roots α.

Proof. The last claim follows from the first, since $\mathfrak{g}_0=\mathfrak{t}^{\mathbb{C}}$. For $X_\alpha\in\mathfrak{g}_\alpha,\ X_\beta\in\mathfrak{g}_\beta$ we have

$$\mathrm{Ad}(t)[X_\alpha,X_\beta]=[\mathrm{Ad}(t)X_\alpha,\mathrm{Ad}(t)X_\beta]=\alpha(t)\beta(t)[X_\alpha,X_\beta]=(\alpha+\beta)(t)[X_\alpha,X_\beta].$$

This shows $[X_\alpha,X_\beta]\in\mathfrak{g}_{\alpha+\beta}$. □

Let us fix a non-degenerate Ad-invariant inner product B on $\mathfrak{g}$. Its restriction to $\mathfrak{t}$ is a W-invariant inner product on $\mathfrak{t}$. We use the same notation B for its extension to a non-degenerate symmetric complex-bilinear form on $\mathfrak{g}^{\mathbb{C}}$, respectively $\mathfrak{t}^{\mathbb{C}}$.

Proposition 15.2. *The spaces* $\mathfrak{g}_\alpha,\mathfrak{g}_\beta$ *for* $\alpha+\beta\neq0$ *are* B*-orthogonal, while* $\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}$ *are non-singularly paired.*

Proof. If $X_\alpha\in\mathfrak{g}_\alpha,\ X_\beta\in\mathfrak{g}_\beta$, then

$$B(X_\alpha,\ X_\beta)=B(\mathrm{Ad}(t)X_\alpha,\ \mathrm{Ad}(t)X_\beta)=(\alpha+\beta)(t)\ B(X_\alpha,X_\beta),$$

hence $\alpha+\beta=0$ if $B(X_\alpha,X_\beta)\neq0$. □

15.2. **The Lie subalgebras $\mathfrak{sl}(2,\mathbb{C})_\alpha\subset\mathfrak{g}^{\mathbb{C}},\ \mathfrak{su}(2)_\alpha\subset\mathfrak{g}$.** To proceed, we need to use some representation theory of $\mathfrak{sl}(2,\mathbb{C})$. Recall that $\mathfrak{sl}(2,\mathbb{C})$ has the standard basis e,f,h with bracket relations $[e,f]=h,\ [h,e]=2e,\ [h,f]=-2f$. Any finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-representation is a direct sum of irreducible representations. Furthermore, we had an explicit description of the irreducible $\mathfrak{sl}(2,\mathbb{C})$ representations. From this we read off:

Lemma 15.3. *Let* $\pi\colon\mathfrak{sl}(2,\mathbb{C})\to\mathrm{End}(V)$ *be a finite-dimensional complex* $\mathfrak{sl}(2,\mathbb{C})$*-representation. Then* $\pi(h)$ *has integer eigenvalues, and* V *is a direct sum of the eigenspaces* $V_m=\ker(\pi(h)-m)$. *For* $m>0$, *the operator* $\pi(f)$ *gives an injective map*

$$\pi(f)\colon V_m\to V_{m-2}.$$

For $m<0$, *the operator* $\pi(e)$ *gives an injective map*

$$\pi(e)\colon V_m\to V_{m+2}.$$

One has direct sum decompositions

$$V=\ker(e)\oplus\mathrm{ran}(f)=\ker(f)\oplus\mathrm{ran}(e).$$

Proof. All these claims are evident for irreducible representations $V(k)$, hence they also hold for direct sums of irreducibles. □

The Lie algebra $\mathfrak{su}(2)$ can be regarded as fixed point set of the *conjugate* linear involution of $\mathfrak{sl}(2,\mathbb{C})$, given by

$$h\mapsto-h,\ e\mapsto-f,\ f\mapsto-e.$$

Indeed, the fixed point set of this involution is spanned by

$$X=i(e+f),\ Y=f-e,\ Z=i\,h,$$

and these vectors satisfy $[X,Y]=2Z,\ [Y,Z]=2X,\ [Z,X]=2Y$. In the matrix representation,

$$X=\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},\quad Y=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\quad Z=\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

For any weight $\mu \in X^*(T)$, let $\mathrm{d}\mu\colon \mathfrak{t} \to \mathfrak{u}(1) = i\mathbb{R}$ be the infinitesimal weight. In particular, we have the infinitesimal roots $\mathrm{d}\alpha$ satisfying

$$[h, X_\alpha] = \mathrm{d}\alpha(h) X_\alpha$$

for all $X_\alpha \in \mathfrak{g}_\alpha$.

Theorem 15.4. (a) *For every root $\alpha \in \mathfrak{R}$, the root space $\mathfrak{g}_\alpha$ is 1-dimensional.*

(b) *The subspace $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{t}^{\mathbb{C}}$ is 1-dimensional, and contains a unique element h_α such that*

$$d\alpha(h_\alpha) = 2.$$

(c) *Let $e_\alpha \in \mathfrak{g}_\alpha$ be non-zero, and normalize $f_\alpha \in \mathfrak{g}_{-\alpha}$ by the condition $[e_\alpha, f_\alpha] = h_\alpha$. Then $e_\alpha, f_\alpha, h_\alpha$ satisfy the bracket relations of $\mathfrak{sl}(2,\mathbb{C})$:*

$$[h_\alpha, e_\alpha] = 2e_\alpha, \quad [h_\alpha, f_\alpha] = -2f_\alpha, \quad [e_\alpha, f_\alpha] = h_\alpha.$$

Thus $\mathfrak{sl}(2,\mathbb{C})_\alpha = \mathrm{span}_{\mathbb{C}}(e_\alpha, f_\alpha, h_\alpha) \subset \mathfrak{g}_{\mathbb{C}}$ is a complex Lie subalgebra isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.

(d) *The real Lie subalgebra $\mathfrak{su}(2)_\alpha := \mathfrak{sl}(2,\mathbb{C})_\alpha \cap \overline{\mathfrak{sl}(2,\mathbb{C})_\alpha} \subset \mathfrak{g}$ is isomorphic to $\mathfrak{su}(2)$.*

Proof. We begin by showing that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is 1-dimensional. Pick an invariant inner product B on $\mathfrak{g}$. Let $H_\alpha \in \mathfrak{t}^{\mathbb{C}},\ \alpha \in \mathfrak{R}$ be defined by

$$(\mathrm{d}\alpha)(h) = B(H_\alpha, h)$$

for all $h \in \mathfrak{t}^{\mathbb{C}}$. Since $\mathrm{d}\alpha(h) \in i\mathbb{R}$ for $h \in \mathfrak{t}$, we have $H_\alpha \in i\mathfrak{t}$, hence $B(H_\alpha, H_\alpha) < 0$, hence $\mathrm{d}\alpha(H_\alpha) < 0$. Let $e_\alpha \in \mathfrak{g}_\alpha,\ e_{-\alpha} \in \mathfrak{g}_{-\alpha}$ be non-zero. For all $h \in \mathfrak{t}^{\mathbb{C}}$ we have

$$B([e_\alpha, e_{-\alpha}], h) = B(e_{-\alpha}, [h, e_\alpha]) = \mathrm{d}\alpha(h) B(e_{-\alpha}, e_\alpha) = B(e_{-\alpha}, e_\alpha) B(H_\alpha, h).$$

This shows

$$[e_\alpha, e_{-\alpha}] = B(e_{-\alpha}, e_\alpha) H_\alpha, \tag{4}$$

proving that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathrm{span}_{\mathbb{C}}(H_\alpha)$. Taking $e_{-\alpha} = \overline{e_\alpha}$, we have $B(e_{-\alpha}, e_\alpha) > 0$, hence the equality $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathrm{span}_{\mathbb{C}}(H_\alpha)$. This proves $\dim_{\mathbb{C}}[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = 1$.

Since $\mathrm{d}\alpha(H_\alpha) < 0$, we may take a multiple h_α of H_α such that $\mathrm{d}\alpha(h_\alpha) = 2$. Clearly, h_α is uniquely element of $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ with this normalization.

Let f_α be the unique multiple of $\overline{e_\alpha}$ so that $[e_\alpha, f_\alpha] = h_\alpha$. Since

$$[h_\alpha, e_\alpha] = \mathrm{d}\alpha(h_\alpha) e_\alpha = 2e_\alpha,$$

and similarly $[h_\alpha, f_\alpha] = -2f_\alpha$, we see that $e_\alpha, f_\alpha, h_\alpha$ span an $\mathfrak{sl}(2,\mathbb{C})$ subalgebra.

Let us now view $\mathfrak{g}^{\mathbb{C}}$ as a complex representation of this $\mathfrak{sl}(2,\mathbb{C})$ subalgebra, by restriction of the adjoint representation. The operator $\mathrm{ad}(h_\alpha)$ acts on $\mathfrak{g}_\alpha$ as the scalar $\mathrm{d}\alpha(h_\alpha) = 2$. Hence $\mathrm{ad}(f_\alpha)\colon \mathfrak{g}_\alpha \to \mathfrak{g}_0$ is injective. Since its image $\mathrm{span}_{\mathbb{C}}(h_\alpha)$ is 1-dimensional, this proves that $\mathfrak{g}_\alpha$ is 1-dimensional, hence spanned by e_α.

Since $[e_\alpha, \overline{e_\alpha}]$ is a positive multiple of H_α, hence a negative multiple of h_α, we may normalize e_α (up to a scalar in U(1)) by requiring that $[e_\alpha, \overline{e_\alpha}] = -h_\alpha$. Then

$$\overline{e_\alpha} = -f_\alpha,\ \overline{f_\alpha} = -e_\alpha,\ \overline{h_\alpha} = -h_\alpha$$

confirming that $\mathfrak{sl}(2,\mathbb{C})_\alpha$ is invariant under complex conjugation, and its real part is isomorphic to $\mathfrak{su}(2)$. ☐

Theorem 15.5. *If $\alpha \in \mathfrak{R}$, then $\mathbb{R}\alpha \cap \mathfrak{R} = \{\alpha, -\alpha\}$.*

Proof. We may assume that α is a shortest root in the line $\mathbb{R}\alpha$. We will show that $t\alpha$ is not a root for any $t > 1$. Suppose on the contrary that $t\alpha$ is a root for some $t > 1$, and take the smallest such t. The operator $\mathrm{ad}(h_\alpha)$ acts on $\mathfrak{g}_{t\alpha}$ as a positive scalar $t\mathrm{d}\alpha(h_\alpha) = 2t > 0$. By the $\mathfrak{sl}(2,\mathbb{C})$-representation theory, it follows that $\mathrm{ad}(f_\alpha)\colon \mathfrak{g}_{t\alpha} \to \mathfrak{g}_{(t-1)\alpha}$ is injective. Since $t > 1$, and since there are no smaller multiples of α that are roots, other than α itself, this implies that $t = 2$, and $\mathrm{ad}(f_\alpha)\colon \mathfrak{g}_{2\alpha} \to \mathfrak{g}_\alpha$ is injective, hence an isomorphism. But this contradicts $\mathrm{ran}(f_\alpha) \cap \ker(e_\alpha) = 0$. ☐

15.3. **Co-roots.** The Lie subalgebra $\mathfrak{su}(2)_\alpha \subset \mathfrak{g}$ is spanned by

$$i(e_\alpha + f_\alpha),\ f_\alpha - e_\alpha,\ ,\ i\,h_\alpha.$$

Let $\mathrm{SU}(2)_\alpha \to G$ be the Lie group morphism exponentiating the inclusion $\mathfrak{su}(2)_\alpha \subset \mathfrak{g}$. Since SU(2) is simply connected, with center $\mathbb{Z}_2$, the image is isomorphic either to SU(2) or to SO(3). Let $T_\alpha \subset \mathrm{SU}(2)_\alpha$ be the maximal torus obtained by exponentiating $\mathrm{span}_\mathbb{R}(ih_\alpha) \subset \mathfrak{t}$. The morphism $T_\alpha \to T$ defines an injective map of the coweight lattices,

$$X_*(T_\alpha) \to X_*(T). \tag{5}$$

But $T_\alpha \cong \mathrm{U}(1)$, by exponentiating the isomorphism $\mathfrak{t}_\alpha \to \mathfrak{u}(1) = i\mathbb{R},\ ish_\alpha \mapsto is$. Hence $X_*(T_\alpha) = X_*(\mathrm{U}(1)) = \mathbb{Z}$.

Definition 15.6. The co-root $\alpha^\vee \in X_*(T)$ corresponding to a root α is the image of $1 \in \mathbb{Z} \cong X_*(T_\alpha)$ under the inclusion (5). The set of co-roots is denoted $\mathfrak{R}^\vee \subset X_*(T)$.

Note that $2\pi i h_\alpha \in \mathfrak{t}_\alpha$ generates the integral lattice Λ_α of T_α. Thus, $\alpha^\vee$ corresponds to h_α under the identification $X_*(T) \otimes_\mathbb{Z} \mathbb{R} = i\mathfrak{t}$. That is,

$$\mathrm{d}\mu(h_\alpha) = \langle \alpha^\vee, \mu \rangle$$

for all $\mu \in X^*(T)$.

Remark 15.7.
(a) Note $\langle \beta^\vee, \alpha \rangle \in \mathbb{Z}$ for all $\alpha, \beta \in \mathfrak{R}$. In particular,

$$\langle \alpha^\vee, \alpha \rangle = 2,$$

as a consequence of the equation $\mathrm{d}\alpha(h_\alpha) = 2$.

(b) Let $(\cdot,\cdot)$ be a W-invariant inner product $(\cdot,\cdot)$ on $X_*(T) \otimes_\mathbb{Z} \mathbb{R}$. Using this inner product to identify $X_*(T) \otimes_\mathbb{Z} \mathbb{R}$ with $X^*(T) \otimes_\mathbb{Z} \mathbb{R}$, the co-roots are expressed in terms of the roots as

$$\alpha^\vee = \frac{2\alpha}{(\alpha,\alpha)}.$$

This is often used as the definition of $\alpha^\vee$, and in any case allows us to find the co-roots in all our examples $\mathrm{U}(n),\ \mathrm{SU}(n),\ \mathrm{SO}(n),\ \mathrm{Sp}(n)$.

Example 15.8. Recall that SO$(2m)$ has roots $\alpha = \pm\epsilon^i \pm \epsilon^j$ for $i \neq j$, together with roots $\beta = \epsilon^i$. In terms of the standard inner product $(\cdot,\cdot)$ on $X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$, the co-roots for roots of the first type are $\alpha^\vee = \pm\epsilon^i \pm \epsilon^j$, while for the second type we get $\beta^\vee = 2\epsilon^i$. Note that these co-roots for SO$(2m)$ are precisely the roots for Sp(m). This is an example of *Langlands duality.*

15.4. **Root lengths and angles.** Choose a W-invariant inner product $(\cdot,\cdot)$ on $E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$.

Proposition 15.9. *Let $\alpha, \beta \in \mathfrak{R}$ be two roots, with $||\beta|| \geq ||\alpha||$. Suppose the angle θ between α, β is not a multiple of $\frac{\pi}{2}$. Then one of the following three cases holds true:*

$$\frac{||\beta||^2}{||\alpha||^2} = 1, \qquad \theta = \pm\frac{\pi}{3} \mod \pi,$$
$$\frac{||\beta||^2}{||\alpha||^2} = 2, \qquad \theta = \pm\frac{\pi}{4} \mod \pi,$$
$$\frac{||\beta||^2}{||\alpha||^2} = 3, \qquad \theta = \pm\frac{\pi}{6} \mod \pi.$$

Proof. Since $(\alpha,\beta) = ||\alpha||\ ||\beta||\ \cos(\theta)$, we have

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

Multiplying, this shows

$$\langle \alpha^\vee, \beta\rangle\ \langle \beta^\vee, \alpha\rangle = 4\cos^2\theta.$$

The right hand side takes values in the open interval $(0,4)$. The left hand side is a product of two integers, with $|\langle \alpha^\vee, \beta\rangle| \geq |\langle \beta^\vee, \alpha\rangle|$. If $\cos\theta > 0$ the possible scenarios are:

$$1 \cdot 1 = 1, \quad 2 \cdot 1 = 2, \quad 3 \cdot 1 = 3,$$

while for $\cos\theta < 0$ the possibilities are

$$(-1)\cdot(-1) = 1, \quad (-2)\cdot(-1) = 2, \quad (-3)\cdot(-1) = 3.$$

Since

$$\frac{||\beta||^2}{||\alpha||^2} = \frac{\langle \alpha^\vee, \beta\rangle}{\langle \beta^\vee, \alpha\rangle},$$

we read off the three cases listed in the Proposition. □

These properties of the root systems are nicely illustrated for the classical groups. Let us also note the following consequence of this discussion:

Lemma 15.10. *For all roots $\alpha, \beta \in \mathfrak{R}$, the integer $\langle \alpha^\vee, \beta\rangle$ lies in the interval* $[-3, 3]$.

15.5. **Root strings.**

Theorem 15.11. *Let $\alpha \in \mathfrak{R}$ be a root. Then:*

(a) *Let $\beta \in \mathfrak{R}$, with $\beta \neq \pm\alpha$. Then*

$$\langle \alpha^\vee, \beta \rangle < 0 \Rightarrow \alpha + \beta \in \mathfrak{R}.$$

(b) *(Root strings.) Given $\beta \in \mathfrak{R}$, with $\beta \neq \pm\alpha$, there exist integers $q, p \geq 0$ such that for any integer r, $\beta + r\alpha \in \mathfrak{R}$ if and only if $-q \leq r \leq p$. These integers satisfy*

$$q - p = \langle \alpha^\vee, \beta \rangle.$$

The direct sum $\bigoplus_{j=-q}^{p} \mathfrak{g}_{\beta+j\alpha}$ is an irreducible $\mathfrak{sl}(2,\mathbb{C})_\alpha$-representation of dimension $p+q+1$.

(c) *If $\alpha, \beta, \alpha+\beta$ are all roots, then*

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}.$$

Proof. We will regard $\mathfrak{g}$ as an $\mathfrak{sl}(2,\mathbb{C})_\alpha$-representation. By definition of the co-roots, we have

$$\mathrm{ad}(h_\alpha) e_\beta = \langle \alpha^\vee, \beta \rangle e_\beta$$

for $e_\beta \in \mathfrak{g}_\beta$.

(a) Suppose $\beta \neq -\alpha$ is a root with $\langle \alpha^\vee, \beta \rangle < 0$. Since $\mathrm{ad}(h_\alpha)$ acts on $\mathfrak{g}_\beta$ as a negative scalar $\langle \alpha^\vee, \beta \rangle < 0$, the $\mathfrak{sl}(2,\mathbb{C})$-representation theory shows that $\mathrm{ad}(e_\alpha)\colon \mathfrak{g}_\beta \to \mathfrak{g}_{\alpha+\beta}$ is injective. In particular, $\mathfrak{g}_{\alpha+\beta}$ is non-zero.

(b) Suppose $\pi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V)$ is a finite-dimensional complex $\mathfrak{sl}(2,\mathbb{C})$-representation. By the representation theory of $\mathfrak{sl}(2,\mathbb{C})$, the number of odd-dimensional irreducible components occurring in V is equal to $\dim\ker(\pi(h))$, while the number of odd-dimensional irreducible components is equal to $\dim\ker(\pi(h)-1)$. Apply this to $V = \bigoplus_{j\in\mathbb{Z}} \mathfrak{g}_{\beta+j\alpha}$ as an $\mathfrak{sl}(2,\mathbb{C})_\alpha$-representation. $\mathrm{ad}(h_\alpha)$ acts on $\mathfrak{g}_{\beta+j\alpha}$ as $\langle \alpha^\vee, \beta \rangle + 2j$. The eigenvalues on V are hence either all even or all odd, and their multiplicities are all one. This shows that $V = \bigoplus_{j\in\mathbb{Z}} \mathfrak{g}_{\beta+j\alpha}$ is irreducible. Let q, p be the largest integers such that $\mathfrak{g}_{\beta+p\alpha} \neq 0$, respectively $\mathfrak{g}_{\beta-q\alpha} \neq 0$. Thus

$$V = \bigoplus_{j=-q}^{p} \mathfrak{g}_{\beta+j\alpha}.$$

Let $k+1 = \dim V$. Then k is the eigenvalue of $\mathrm{ad}(h_\alpha)$ on $\mathfrak{g}_{\beta+p\alpha}$, while $-k$ is its eigenvalue on $\mathfrak{g}_{\beta-q\alpha}$. This gives,

$$k = \langle \alpha^\vee, \beta \rangle + 2p, \quad -k = \langle \alpha^\vee, \beta \rangle - 2q.$$

Hence $k = q + p$ and $q - p = \langle \alpha^\vee, \beta \rangle$.

(c) follows from (b), since $\mathrm{ad}(e_\alpha)\colon \mathfrak{g}_\beta \to \mathfrak{g}_{\beta+\alpha}$ for non-zero $e_\alpha \in \mathfrak{g}_\alpha$ is an isomorphism if $\mathfrak{g}_\beta$, $\mathfrak{g}_{\beta+\alpha}$ are non-zero.

□

The set of roots $\beta + j\alpha$ with $-q \leq j \leq p$ is called the *α-root string through β*. If β is such that $\beta - \alpha$ is *not* a root, we have $q = 0$, $p = -\langle \alpha^\vee, \beta \rangle$. As we had seen, this integer is ≤ 3. Hence, the length of any root string is at most 4.

15.6. **Weyl chambers.** Let

$$E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$$

be the real vector space spanned by the weight lattice. Its dual is identified with the vector space spanned by the coweight lattice, $E^* = X_*(T) \otimes_{\mathbb{Z}} \mathbb{R}$. The Weyl group $W = N(T)/T$ acts faithfully on E (and dually on E^*), hence it can be regarded as a subgroup of $\mathrm{GL}(E)$. We will now now realize this subgroup as a reflection group.

Let $\alpha \in \mathfrak{R}$ be a root, and $j_\alpha \colon \mathrm{SU}(2)_\alpha \to G$ the corresponding rank 1 subgroup. Let $T_\alpha \subset \mathrm{SU}(2)_\alpha$ be the maximal torus as before, $N(T_\alpha)$ its normalizer in $\mathrm{SU}(2)_\alpha$, and $W_\alpha = N(T_\alpha)/T \cong \mathbb{Z}_2$ the Weyl group.

Proposition 15.12. *The morphism j_α takes $N(T_\alpha)$ to $N(T)$. Hence it descends to a morphism of the Weyl groups, $W_\alpha \to W$. Letting $w_\alpha \in W$ be the image of the non-trivial element in W_α, its action on E is given by*

$$w_\alpha \, \mu = \mu - \langle \alpha^\vee, \mu \rangle \alpha, \;\; \mu \in E$$

and the dual action on E^* reads $w_\alpha \, \gamma = \gamma - \langle \gamma, \alpha \rangle \alpha^\vee$.

Proof. Consider the direct sum decomposition

$$\mathfrak{t}^{\mathbb{C}} = \mathrm{span}_{\mathbb{C}}(h_\alpha) + \ker(\mathrm{d}\alpha).$$

Elements $h \in \ker(\mathrm{d}\alpha)$ commute with $e_\alpha, f_\alpha, h_\alpha$, hence $[\ker(\mathrm{d}\alpha), \mathfrak{sl}(2,\mathbb{C})_\alpha] = 0$. It follows that the adjoint representation of $j_\alpha(\mathrm{SU}(2)_\alpha)$ on $\ker(\mathrm{d}\alpha)$ is trivial. On the other hand, $\mathrm{span}_{\mathbb{C}}(h_\alpha)$ is preserved under $j_\alpha(N(T_\alpha))$. Hence, all $\mathfrak{t}$ is preserved under $j_\alpha(N(T_\alpha))$, proving that $j_\alpha(N(T_\alpha)) \subset T$. We also see that w_α acts trivially on $\ker(\mathrm{d}\alpha)$, and as -1 on $\mathrm{span}(h_\alpha)$.

Hence, the action on $E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$ is as -1 on the line $\mathbb{R}\alpha$, and is trivial on $\ker(\alpha^\vee)$. This is precisely the action $\mu \mapsto \mu - \langle \mu, \alpha^\vee \rangle \alpha$. The statement for the co-weight lattice follows by duality. ☐

Remark 15.13. Explicitly, using the basis $e_\alpha, f_\alpha, h_\alpha$ to identify $\mathrm{SU}(2)_\alpha \cong \mathrm{SU}(2)$, the element w_α is represented by

$$j_\alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in N(T).$$

Let us now use a W-invariant inner product on E to identify $E^* = E$. Recall that under this identification, $\alpha^\vee = \frac{2\alpha}{(\alpha,\alpha)}$. The transformation

$$w_\alpha(\mu) = \mu - 2\frac{(\alpha, \mu)}{(\alpha, \alpha)} \alpha$$

is reflection relative to the *root hyperplane*

$$H_\alpha = \mathrm{span}_{\mathbb{R}}(\alpha)^\perp \subset E.$$

It is natural to ask if the full Weyl group W is generated by the reflections $w_\alpha, \alpha \in \mathfrak{R}$. This is indeed the case, as we will now demonstrate with a series of Lemmas.

An element $x \in E$ is called *regular* if it does not lie on any of these hyperplanes, and *singular* if it does. Let

$$E^{\mathrm{reg}} = E \backslash \bigcup_{\alpha \in \mathfrak{R}} H_\alpha, \quad E^{\mathrm{sing}} = \bigcup_{\alpha \in \mathfrak{R}} H_\alpha$$

be the set of regular elements, respectively singular elements.

Lemma 15.14. *An element $x \in E$ is regular if and only if its stabilizer under the action of W is trivial.*

Proof. If x is not regular, there exists a root α with $(\alpha, x) = 0$. It then follows that $w_\alpha(x) = x$.

If x is regular, and $w(x) = x$, we will show that $w = 1$. Denote by $h \in i\mathfrak{t}$ be the element corresponding to x under the identification $E = \mathrm{Hom}(\mathfrak{u}(1), \mathfrak{t}) \cong i\mathfrak{t}$; thus $\mathrm{d}\mu(h) = (\mu, x)$ for all $\mu \in X^*(T)$. Since $\ker(\mathrm{ad}(h)) \subset \mathfrak{g}^{\mathbb{C}}$ is invariant under the adjoint representation of $T \subset G$, it is a sum of weight spaces. But $\mathrm{ad}(h)$ acts on the root space $\mathfrak{g}_\alpha$ as a non-zero scalar $\mathrm{d}\alpha(h) = (\alpha, x)$. Thus $\ker(\mathrm{ad}(h))$ does not contain any of the root spaces, proving that $\ker(\mathrm{ad}(h)) = \mathfrak{g}_0 = \mathfrak{t}^{\mathbb{C}}$. It follows that $\mathfrak{t}$ is the unique maximal abelian subalgebra containing ih. Equivalently, the 1-parameter subgroup S generated by the element $ih \in \mathfrak{t}$ is contained in a unique maximal torus, given by T itself. Suppose now that $w \in W$ with $wx = x$, and let $g \in N(T)$ be a lift of w. Then $\mathrm{Ad}_g(h) = h$, so that $g \in Z_G(S)$. By our discussion of maximal tori, there exists a maximal torus T' containing $S \cup \{g\}$. But we have seen that T is the unique maximal torus containing S. Hence $g \in T' = T$, proving that $w = 1$. □

Remark 15.15. This result (or rather its proof) also has the following consequence. Let $\mathfrak{g}^{\mathrm{reg}} \subset \mathfrak{g}$ be the set of Lie algebra elements whose stabilizer group

$$G_\xi = \{g \in G |\, \mathrm{Ad}_g(\xi) = \xi\}$$

under the adjoint action is a maximal torus, and $\mathfrak{g}^{\mathrm{sing}} = \mathfrak{g} \backslash \mathfrak{g}^{\mathrm{reg}}$ those elements whose stabilizer is strictly larger than a maximal torus. Then

$$\mathfrak{g}^{\mathrm{reg}} \cap \mathfrak{t}$$

is the set of all $\xi \in \mathfrak{t}$ such that $\mathrm{d}\alpha(\xi) \neq 0$ for all roots α.

Exercise 15.16. For arbitrary $\xi \in \mathfrak{t}$, the stabilizer $G_\xi = \{g \in G |\, \mathrm{Ad}_g(\xi) = \xi\}$ contains T, hence $\mathfrak{g}_\xi^{\mathbb{C}}$ is a sum of weight spaces. Which roots of G are also roots of G_ξ? What can you say about the dimension of G_ξ?

The connected components of the set E^{reg} are called the *open Weyl chambers*, their closures are called the *closed Weyl chambers*. Unless specified differently, we will take Weyl chamber to mean closed Weyl chamber. Note that the Weyl chambers C are closed convex cones. (That is, if $x, y \in C$ then $rx + sy \in C$ for all $r, s \geq 0$.) The Weyl group permutes the set of roots, hence it acts by permutation on the set of root hyperplanes H_α and on the set of Weyl chambers.

Lemma 15.17. *The Weyl group acts freely on the set of Weyl chambers. That is, if C is a chamber and $w \in W$ with $wC \subset C$ then $w = 1$.*

Proof. If $wC = C$, then w preserves the interior of C. Let $x \in \mathrm{int}(C)$. Then $w^i x \in \mathrm{int}(C)$ for all $i \geq 0$. Letting k be the order of w, the element $x' := x + wx + \ldots w^{k-1}x \in \mathrm{int}(C)$ satisfies $wx' = x'$. By the previous Lemma this means $w = 1$. □

Exercise 15.18. Let C be a fixed (closed) Weyl chamber. a) Let $D \subset C$ one of its 'faces'. (Thus D is the intersection of C with some of the root hyperplanes.). Show that if $w(D) \subset D$, then $wx = x$ for all $x \in D$. (Hint: D can be interpreted as the Weyl chamber of a subgroup of G.) b) Show that if $w \in W$ takes $x \in C$ to $x' \in C$ then $x' = x$.

We say that a root hyperplane H_α *separates the chambers* $C, C' \subset E$ if for points x, x' in the interior of the chambers, (x, α) and (x', α) have opposite signs, but (x, β) and (x', β) have equal sign for all roots $\beta \neq \pm\alpha$. Equivalently, the line segment from x to x' meets H_α, but does not meet any of the hyperplanes H_β for $\beta \neq \pm\alpha$.

Lemma 15.19. *Suppose the root hyperplane H_α separates the Weyl chambers C, C'. Then w_α interchanges C, C'.*

Proof. This is clear from the description of w_α as reflection across H_α, and since w_α must act as a permutation on the set of Weyl chambers. □

Since any two Weyl chambers are separated by finitely many root hyperplanes, it follows that any two Weyl chambers are related by some $w \in W$. To summarize, we have shown:

Theorem 15.20. *The Weyl group W acts simply transitively on the set of Weyl chambers. That is, for any two Weyl chambers C, C' there is a* unique *Weyl group element $w \in W$ with $w(C) = C'$. In particular, the cardinality $|W|$ equals the number of Weyl chambers.*

Corollary 15.21. *Viewed as a subgroup of $\mathrm{GL}(E)$, the Weyl group W coincides with the group generated by the reflections across root hyperplanes H_α. In fact, W is already generated by reflections across the hyperplanes H_α supporting any fixed Weyl chamber C.*

The proof of the last part of this corollary is left as an exercise.

16. Simple roots, Dynkin diagrams

Let us fix a Weyl chamber C_+, called the *positive* or *fundamental* Weyl chamber. Then any Weyl chamber is of the form $C = wC_+$ for $w \in W$. The choice of C_+ determines a decomposition

$$\mathfrak{R} = \mathfrak{R}_+ \cup \mathfrak{R}_-$$

into *positive roots* and *negative roots*, where $\mathfrak{R}_\pm$ are the roots α with $(\alpha, x) > 0$ (resp. < 0) for $x \in \mathrm{int}(C)$. For what follows, it is convenient to fix some choice $x_* \in \mathrm{int}(C_+)$.

Definition 16.1. A *simple root* is a positive root that cannot be written as a sum of two positive roots. We will denote the set of simple roots by Π.

Proposition 16.2 (Simple roots)**.** *The set $\Pi = \{\alpha_1, \dots \alpha_l\}$ of simple roots has the following properties.*

(a) Π *is a basis of the root lattice* $\mathrm{span}_{\mathbb{Z}} \mathfrak{R}$.
(b) *Let $\alpha = \sum_{i=1}^{l} k_i \alpha_i \in \mathfrak{R}$. Then $\alpha \in \mathfrak{R}_+$ if and only if all $k_i \geq 0$, and $\alpha \in \mathfrak{R}_-$ if and only if all $k_i \leq 0$.*
(c) *One has $\langle \alpha_i^\vee, \alpha_j \rangle \leq 0$ for $i \neq j$.*

Proof. If α_i, α_j are distinct simple roots, then their difference $\alpha_i - \alpha_j$ is *not* a root. (Otherwise, either $\alpha_i = \alpha_j + (\alpha_i - \alpha_j)$ or $\alpha_j = \alpha_i + (\alpha_j - \alpha_i)$ would be a sum of two positive roots.) On the other hand, we had shown that if two roots α, β form an obtuse angle (i.e. $(\alpha, \beta) < 0$), then their sum is a root. Applying this to $\alpha = \alpha_i,\ \beta = -\alpha_j$ it follows that $(\alpha_i, -\alpha_j) \geq 0$, hence $\langle \alpha_i^\vee, \alpha_j \rangle \leq 0$, proving (c).

We next show that the α_i are linearly independent. Indeed suppose $\sum_i k_i\alpha_i = 0$ for some $k_i \in \mathbb{R}$. Let

$$\mu := \sum_{k_i>0} k_i\alpha_i = -\sum_{k_j<0} k_j\alpha_j.$$

Taking the scalar product with itself, and using $(\alpha_i, \alpha_j) \le 0$ for $i \neq j$ we obtain

$$0 \le (\mu, \mu) = -\sum_{k_i>0,\ k_j<0} k_i k_j(\alpha_i, \alpha_j) \le 0.$$

Hence $\mu = 0$. Taking the inner product with any $x_* \in \mathrm{int}(C_+)$, we get $0 = \sum_{k_i>0} k_i(\alpha_i, x_*) = -\sum_{k_j<0} k_j(\alpha_j, x_*)$. Hence the set of i with $k_i > 0$ is empty, and so is the set of i with $k_i < 0$. Thus all $k_i = 0$, proving that the α_i are linearly independent.

We claim that any $\alpha \in \mathfrak{R}_+$ can be written in the form $\alpha = \sum k_i\alpha_i$ for some $k_i \in \mathbb{Z}_{\ge 0}$. This will prove (b), and also finish the proof of (a). Suppose the claim is false, and let α be a counterexample with (α, x_*) as small as possible. Since α is not a simple root, it can be written as a sum $\alpha = \alpha' + \alpha''$ of two positive roots. Then (α', x_*), (α'', x_*) are both strictly positive, hence they are strictly smaller than their sum (α, x_*). Hence, neither α' nor α'' is a counterexample, and each can be written as a linear combination of α_i's with coefficients in $\mathbb{Z}_{\ge 0}$. Hence the same is true of α, hence α is not a counterexample. Contradiction. □

Corollary 16.3. *The simple co-roots* $\mathfrak{R}^\vee = \{\alpha_1^\vee, \ldots, \alpha_l^\vee\}$ *are a basis of the co-root lattice* $\mathrm{span}_\mathbb{Z}\, \mathfrak{R}^\vee \subset X_*(T)$.

Definition 16.4. The $l \times l$-matrix with entries $A_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle$ is called the *Cartan matrix* of G (or of the root system $\mathfrak{R} \subset E$).

Note that the diagonal entries of the Cartan matrix are equal to 2, the and that the off-diagonal entries are ≤ 0.

Example 16.5. Let $G = \mathrm{U}(n)$, and use the standard inner product on $E = X^*(T) \otimes_\mathbb{Z} \mathbb{R} = \mathrm{span}_\mathbb{R}(\epsilon^1, \ldots, \epsilon^n)$ to identify $E \cong E^*$. Recall that $\mathrm{U}(n)$ has roots $\alpha = \epsilon^i - \epsilon^j$ for $i \neq j$. The roots coincide with the coroots, under the identification $E = E^*$.

Let $x_* = n\epsilon_1 + (n-1)\epsilon_2 + \ldots + \epsilon_n$. Then $\langle \alpha, u \rangle \neq 0$ for all roots. The positive roots are $\epsilon^i - \epsilon^j$ with $i < j$, the negative roots are those with $i > j$. The simple roots are

$$\Pi = \{\epsilon^1 - \epsilon^2,\ \epsilon^2 - \epsilon^3, \ldots,\ \epsilon^{n-1} - \epsilon^n\},$$

and are equal to the simple co-roots $\Pi^\vee$. For the Cartan matrix we obtain the $(n-1) \times (n-1)$-matrix,

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & 2 & -1 \\ \cdots & \cdots & \cdots & \cdots & -1 & 2 \end{pmatrix}$$

This is also the Cartan matrix for $\mathrm{SU}(n)$ (which has the same roots as $\mathrm{U}(n)$).

Example 16.6. Let $G = \mathrm{SO}(2l+1)$. Using the standard maximal torus and the basis $X^*(T) = \mathrm{span}_{\mathbb{Z}}(\epsilon_1, \ldots, \epsilon^l)$, we had found that the roots are $\pm\epsilon^i \pm \epsilon^j$ for $i \neq j$, together with the set of all $\pm\epsilon^i$. Let $x_* = n\epsilon_1 + (n-1)\epsilon_2 + \ldots + \epsilon_n$. Then $(x_*, \alpha) \neq 0$ for all roots α. The positive roots are the set of all $\epsilon^i - \epsilon^j$ with $i < j$, together with all $\epsilon^i + \epsilon^j$ for $i \neq j$, together with all ϵ^i. The simple roots are

$$\Pi = \{\epsilon^1 - \epsilon^2, \epsilon^2 - \epsilon^3, \ldots, \epsilon^{l-1} - \epsilon^l, \epsilon^l\}.$$

Here is the Cartan matrix for $l = 4$:

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -2 & 2 \end{pmatrix}$$

The calculation for $\mathrm{SO}(2l+1)$ is similar: One has

$$\Pi = \{\epsilon^1 - \epsilon^2, \epsilon^2 - \epsilon^3, \ldots, \epsilon^{l-1} - \epsilon^l, 2\epsilon^l\},$$

with Cartan matrix the transpose of that of $\mathrm{SO}(2l+1)$.

A more efficient way of recording the information of a Cartan matrix is the *Dynkin diagram*[8].
The Dynkin diagram is a graph, with vertices (nodes) the simple roots, connected by edges if $i \neq j$ and $(\alpha_i, \alpha_j) \neq 0$. One gives each edge a multiplicity of one, two, or three according to whether $\frac{||\alpha_i||^2}{||\alpha_j||^2}$ equals $1, 2$ or 3. For edges with multiplicity 2 or 3, one also puts an arrow from longer roots down to shorter roots. Note that the Dynkin diagram contains the full information of the Cartan matrix.

Example 16.7. There are only four possible Dynkin diagrams with 2 nodes: a disconnected Dynkin diagram (corresponding to $\mathrm{SU}(2) \times \mathrm{SU}(2)$ or $\mathrm{SO}(4)$), and three connected ones, for the three possible multiplicities of the edge connecting the two nodes. The Dynkin diagram for multiplicity 1 is that of $\mathrm{SU}(3)$, the one for multiplicity 2 is that of $\mathrm{SO}(5)$. It turns out that there is a unique compact, simple Lie group corresponding to the Dynkin diagram with two nodes and an edge with multiplicity 3: This is the exceptional Lie group G_2.

Exercise 16.8. Using only the information from the Dynkin diagram for G_2, give a picture of the root system for G_2. Use the root system to read off the dimension of G_2 and the order of its Weyl group. Show that the dual root system $\mathfrak{R}^\vee$ for G_2 is isomorphic to $\mathfrak{R}$.

Proposition 16.9. *The positive Weyl chamber can be described in terms of the simple roots as*

$$C_+ = \{x \in E |\ (\alpha_i, x) \geq 0,\ i = 1, \ldots, l\}.$$

Proof. By definition, C_+ is the set of all x with $(\alpha, x) \geq 0$ for $\alpha \in \mathfrak{R}_+$. But since every positive root is a linear combination of simple roots with non-negative coefficients, it suffices to require the inequalities for the simple roots. □

[8]Dynkin diagrams were used by E. Dynkin in his 1946 papers. Similar diagrams had previously been used by Coxeter in 1934 and Witt 1941.

Thus, in particular C_+ is a simple polyhedral cone, cut out by l inequalities.

We had remarked that W is generated by reflections across boundary hyperplanes H_α for C_+. Hence it is generated by the simple reflections $s_i = w_{\alpha_i},\ i = 1, \ldots, l$. Since every H_α bounds some C, it follows that every α is W-conjugate to some α_i. This essentially proves:

Proposition 16.10. *The Dynkin diagram determines the root system $\mathfrak{R}$, up to isomorphism.*

Proof. The Dynkin diagram determines the set Π of simple roots, as well as their angles and relative lengths. Indeed, by Proposition [...], and using the fact that angles between non-orthogonal roots are obtuse, we have, for the case $\langle \alpha_i^\vee, \alpha_j \rangle \neq 0$,

$$\begin{aligned} ||\alpha_i||^2 = ||\alpha_j||^2 \quad &\Rightarrow \quad \theta = \frac{2\pi}{3}, \\ ||\alpha_i||^2 = 2||\alpha_j||^2 \quad &\Rightarrow \quad \theta = \frac{3\pi}{4}, \\ ||\alpha_i||^2 = 3||\alpha_j||^2 \quad &\Rightarrow \quad \theta = \frac{5\pi}{6}, \end{aligned}$$

The Weyl group W is recovered as the group generated by the simple reflections $s_i = w_{\alpha_i}$, and

$$\mathfrak{R} = W\Pi.$$

□

Hence, given the Dynkin diagram one may recover the root system, the Weyl group, the Weyl chamber etc.

Example 16.11. The Dynkin diagram of SO(5) has two vertices α_1, α_2, connected by an edge of multiplicity 2 directed from α_1 to α_2. Thus $||\alpha_1||^2 = 2||\alpha_2||^2$, and the angle between α_1, α_2 is $\frac{3\pi}{4}$. It is standard to work with a normalization where the long roots satisfy $||\alpha||^2 = 2$. A concrete realization as a root system in $\mathbb{R}^2$ is given by $\alpha_1 = \epsilon^1 - \epsilon^2$ and $\alpha_2 = \epsilon^2$; other realizations are related by an orthogonal transformation of $\mathbb{R}^2$.

The corresponding co-roots are $\alpha_1^\vee = \epsilon^1 - \epsilon^2$ and $\alpha_2^\vee = 2\epsilon^2$. Let s_1, s_2 be the simple reflections corresponding to α_1, α_2. One finds

$$s_1(k_1\epsilon^1 + k_2\epsilon^2) = k_1\epsilon^2 + k_2\epsilon^1, \qquad s_2(l_1\epsilon^1 + l_2\epsilon^2) = l_1\epsilon^1 - l_2\epsilon^2,$$

Hence

$$\begin{aligned} s_1(\alpha_1) &= -\alpha_1 = -\epsilon^1 + \epsilon^2, \\ s_1(\alpha_2) &= \epsilon^1, \\ s_2(\alpha_1) &= \epsilon^1 + \epsilon^2 \\ s_2(\alpha_2) &= -\epsilon^2, \\ s_2 s_1(\alpha_1) &= -\epsilon^1 - \epsilon^2, \\ s_1 s_2(\alpha_2) &= -\epsilon^1, \end{aligned}$$

which recovers all the roots. The Weyl group is the reflection group generated by s_1, s_2. As an abstract group, it is the group generated by s_1, s_2 with the single relation $(s_1 s_2)^3 = 1$.

For any root $\alpha = \text{sum}_{i=1}^{l} k_i\alpha_i \in \mathfrak{R}$ (or more generally for any element of the root lattice), one defines its *height* by

$$\text{ht}(\alpha) = \sum_{i=1}^{l} k_i.$$

In terms of the fundamental coweights (cf. below),

$$\text{ht}(\alpha) = \sum_{i=1}^{l} \langle \varpi_i^\vee, \alpha \rangle.$$

Proposition 16.12. *For any $\alpha \in \mathfrak{R}_+ \backslash \Pi$ there exists $\beta \in \mathfrak{R}_+$ with $\text{ht}(\beta) = \text{ht}(\alpha) - 1$.*

Proof. Choose a W-invariant inner product on E. Write $\alpha = \sum_i k_i\alpha_i$. Then

$$0 < ||\alpha||^2 = \sum_i k_i(\alpha, \alpha_i).$$

Since all $k_i \geq 0$, there must be at least one index r with $(\alpha, \alpha_r) > 0$. This then implies that $\alpha - \alpha_r \in \mathfrak{R}$. Since $\alpha \notin \Pi$, there must be at least one index $i \neq r$ with $k_i > 0$. Since the coefficient of this α_i in $\alpha - \alpha_r$ is again $k_i > 0$, it follows that $\alpha - \alpha_r \in \mathfrak{R}_+$. □

17. Serre relations

Let G be a compact connected semi-simple Lie group, with given choice of maximal torus T and positive Weyl chamber C_+. Let $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ be the set of simple roots, and $a_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle$ be the entries of the Cartan matrix. Let $h_i \in [\mathfrak{g}_{\alpha_i}, \mathfrak{g}_{-\alpha_i}]$ with normalization $\mathrm{d}\alpha_i(h_i) = 2$. Pick $e_i \in \mathfrak{g}_{\alpha_i}$, normalized up to $\mathrm{U}(1)$ by the condition $[e_i, \overline{e_i}] = -h_i$, and put $f_i = -\overline{e_i}$.

Proposition 17.1. *The elements e_i, f_i, h_i generate $\mathfrak{g}^{\mathbb{C}}$. They satisfy the* Serre relations*,*

$$\begin{aligned}
&(S1) \quad [h_i, h_j] = 0,\\
&(S2) \quad [e_i, f_j] = \delta_{ij} h_i,\\
&(S3) \quad [h_i, e_j] = a_{ij} e_j,\\
&(S4) \quad [h_i, f_j] = -a_{ij} f_j,\\
&(S5) \quad \text{ad}(e_i)^{1-a_{ij}}(e_j) = 0,\\
&(S6) \quad \text{ad}(f_i)^{1-a_{ij}}(f_j) = 0
\end{aligned}$$

Proof. Induction on height shows that all root spaces $\mathfrak{g}_\alpha$ for positive roots are in the subalgebras generated by the e_i, f_i, h_i. Indeed, if $\alpha \in \mathfrak{R}_+$ we saw that $\alpha = \beta + \alpha_r$ for some $\beta \in \mathfrak{R}_+$ with $\text{ht}(\beta) = \text{ht}(\alpha) = 1$, and $[e_r, \mathfrak{g}_\beta] = \mathfrak{g}_{\beta+\alpha_r}$ since α_r, α, β are all roots). Similarly the root spaces for the negative roots are contained in this subalgebra, and since the h_i span $\mathfrak{t}^{\mathbb{C}}$, it follows that the subalgebra generated by the e_i, f_i, h_i is indeed all of $\mathfrak{g}^{\mathbb{C}}$. Consider next the relations. (S1) is obvious. (S2) holds true for $i = j$ by our normalizations of e_i, f_i, h_i, and for $i \neq j$ because $[\mathfrak{g}_{\alpha_i}, \mathfrak{g}_{-\alpha_j}] \subset \mathfrak{g}_{\alpha_i - \alpha_j} = 0$ since $\alpha_i - \alpha_j$ is not a root. (S3) and (S4) follow since e_j, f_j are in the root spaces $\mathfrak{g}_{\pm\alpha_j}$:

$$[h_i, e_j] = \mathrm{d}\alpha_j(h_i)e_j = \langle \alpha_i^\vee, \alpha_j \rangle e_j = a_{ij} e_j$$

and similarly for $[h_i, f_j]$. For (S5), consider the α_i-root string through α_j. Since $\alpha_j - \alpha_i$ is not a root, the length of the root string is equal to $k+1$ where $-k$ is the eigenvalue of $\mathrm{ad}(h_i)$ on $\mathfrak{g}_{\alpha_j}$. But this eigenvalue is $\mathrm{d}\alpha_j(h_i) = a_{ij}$. Hence root string has length $1 - a_{ij}$, and consists of the roots

$$\alpha_j,\ \alpha_j + \alpha_i, \ldots,\ \alpha_j - a_{ij}\alpha_i.$$

In particular, $\alpha_j + (1 - a_{ij})\alpha_i$ is not a root. This proves (S5), and (S6) is verified similarly. □

The elements e_i, f_i, h_i are called the *Chevalley generators* of the complex Lie algebra $\mathfrak{g}^{\mathbb{C}}$. It turns out that the relations (S1)-(S6) are in fact a complete system of relations. This is a consequence of Serre's theorem, stated below. Hence, one may reconstruct $\mathfrak{g}^{\mathbb{C}}$ from the information given by the Dynkin diagram, or equivalently the Cartan matrix $a_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle$. In fact, we may start out with any 'abstract' Root system.

We begin with the abstract notion of a root system.

Definition 17.2. Let E be a Euclidean vector space, and $\mathfrak{R} \subset E \backslash \{0\}$. For $\alpha \in \mathfrak{R}$ define $\alpha^\vee = 2\alpha/(\alpha, \alpha)$. Then $\mathfrak{R}$ is called a *(reduced) root system* if

(a) $\mathrm{span}_{\mathbb{R}}(\mathfrak{R}) = E$.
(b) The reflection $s_\alpha \colon \mu \mapsto \mu - \langle \alpha^\vee, \mu \rangle \alpha$ preserves $\mathfrak{R}$.
(c) For all $\alpha, \beta \in \mathfrak{R}$, the number $(\alpha^\vee, \beta) \in \mathbb{Z}$,
(d) For all $\alpha \in \mathfrak{R}$, we have $\mathbb{R}\alpha \cap \mathfrak{R} = \{\alpha, -\alpha\}$.

The *Weyl group* of a reduced root system is defined as the group generated by the reflections s_α.

As in the case of root systems coming from compact Lie groups, one can define Weyl chambers, positive roots, simple roots, and a Cartan matrix and Dynkin diagram.

Theorem 17.3 (Serre). *Let $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ be the set of simple roots of a reduced root system of rank l, and let $a_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle$ be the Cartan matrix. The complex Lie algebra with generators $e_i, f_i, h_i,\ i = 1, \ldots, l$ and relations (S1)-(S6) is finite-dimensional and semi-simple. It carries a conjugate linear involution ω_0, given on generators by*

$$\omega_0(e_i) = -f_i,\ \omega_0(f_i) = -e_i,\ \omega_0(h_i) = -h_i,$$

hence may be regarded as the complexification of a real semi-simple Lie algebra $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ integrates to a compact semi-simple Lie group G, with the prescribed root system.

For a proof of this result, see e.g. V. Kac '*Infinite-dimensional Lie algebras*' or A. Knapp, '*Lie groups beyond an introduction*'.

18. Classification of Dynkin diagrams

There is an obvious notion of *sum* of root systems $\mathfrak{R}_1 \subset E_1,\ \mathfrak{R}_2 \subset E_2$, as the root system $\mathfrak{R}_1 \cup \mathfrak{R}_2$ in $E_1 \oplus E_2$. A root system is *irreducible* if it is not a sum of two root systems.

Given an abstract root system, we may as before define Weyl chambers, and the same proof as before shows that for non-orthogonal roots α, β with $||\alpha|| \geq ||\beta||$, the ratio of the root lengths is given by $||\alpha||^2/||\beta||^2 \in \{1, 2, 3\}$, and the angles in the three cases are $\pm\frac{\pi}{3}, \pm\frac{\pi}{4}, \pm\frac{\pi}{6} \mod \pi$. Hence, we may define simple roots and a Dynkin diagram as before.

Proposition 18.1. *A root system is irreducible if and only if its Dynkin diagram is connected.*

Proof. Let Π be a set of simple roots for $\mathfrak{R}$. If $\mathfrak{R}$ is a sum of root systems $\mathfrak{R}_1$ and $\mathfrak{R}_2$, then $\Pi_1 = \mathfrak{R}_1 \cap \Pi$ and $\Pi_2 = \mathfrak{R}_2 \cap \Pi$ are simple roots for $\mathfrak{R}_i$. Since all roots in Π_1 and orthogonal to all roots in Π_2, the Dynkin diagram is disconnected. Conversely, given a root system $\mathfrak{R} \subset E$ with disconnected Dynkin diagram, then $\Pi = \Pi_1 \cup \Pi_2$ where all roots in Π_1 are orthogonal to all roots in Π_2. This gives an orthogonal decomposition $E = E_1 \oplus E_2$ where E_1, E_2 is the space spanned by roots in Π_1, Π_2. The simple reflections s_i for roots $\alpha_i \in \Pi_1$ commute with those of roots $\alpha_j \in \Pi_2$, hence the Weyl group is a direct product $W = W_1 \times W_2$, and $\mathfrak{R}$ is the sum of $\mathfrak{R}_1 = W_1\Pi_1$ and $\mathfrak{R}_2 = W_2\Pi_2$. □

Hence, we will only consider connected Dynkin diagrams. The main theorem is as follows:

Theorem 18.2. *Let $\mathfrak{R}$ be an irreducible root system. Then the Dynkin diagram is given by exactly one of the following diagrams*

$$A_l \ (l \geq 1), B_l \ (l \geq 2), C_l \ (l \geq 3), D_l \ (l \geq 4), E_6, E_7, E_8, F_4, G_2.$$

PICTURE

Here the subscript signifies the rank, i.e. the number of vertices of the Dynkin diagram.

We will sketch the proof in the case that the root system is *simply laced*, i.e. all roots have the same length and hence the Dynkin diagram has no multiple edges. We will thus show that all simply laced connected Dynkin diagrams are of one of the types A_l, D_l, E_6, E_7, E_8.

We will use the following elementary Lemma:

Lemma 18.3. *Let $u_1, \ldots, u_k$ be pairwise orthogonal vectors in a Euclidean vector space E. For all $v \in E$ we have*

$$||v||^2 > \sum_{i=1}^{k} \frac{(v, u_i)^2}{||u_i||^2},$$

with equality if and only if v lies in $\mathrm{span}(u_1, \ldots, u_k)$.

Proof in the simply laced case. We normalize the inner product on E so that all roots satisfy $||\alpha||^2 = 2$. Since all roots have equal length, the angle between non-orthogonal simple roots is $\frac{2\pi}{3}$. Since $\cos(\frac{2\pi}{3}) = -\frac{1}{2}$, it follows that

$$(\alpha_i, \alpha_j) = -1$$

if α_i, α_j are connected by an edge of the Dynkin diagram.

A *subdiagram* of a Dynkin diagram is obtained by taking a subset $\Pi' \subset \Pi$ of vertices, together with the edges connecting any two vertices in Π'. It is clear that such a subdiagram is again a Dynkin diagram. (If Π corresponds to the root system $\mathfrak{R}$, then Π' corresponds to a root system $\mathfrak{R} \cap \mathrm{span}_{\mathbb{R}} \Pi'$.)

The first observation is that the number of edges in the Dynkin diagram is $< l$. Indeed,

$$0 < ||\sum_{i=1}^{l} \alpha_i||^2 = 2l + 2\sum_{i<j} (\alpha_i, \alpha_j) = 2l - 2\#\{edges\}.$$

Hence $\#\{edges\} < l$. Since this also applies to subdiagrams of the Dynkin diagram, it follows in particular that the diagram cannot contain any loops.

One next observes that the number of edges originating at a vertex is at most 3. Otherwise, there would be a star-shaped subdiagram with 5 vertices, with $\alpha_1, \ldots, \alpha_4$ connected to the central vertex ψ. In particular, $\alpha_1, \ldots, \alpha_4$ are pairwise orthogonal. Since ψ is linearly independent of $\alpha_1, \ldots, \alpha_4$, we have

$$2 = ||\psi||^2 > \sum_{i=1}^{4} \frac{(\psi, \alpha_i)^2}{||\alpha_i||^2} = \sum_{i=1}^{4} (\frac{-1}{2})^2 = 2,$$

a contradiction. (To get the inequality $<$, note that $||\psi||^2$ is the sum of squares of its coefficients in an orthonormal basis. The $\alpha_i/||\alpha_i||, \ i \leq 4$ is part of such a basis, but since ψ is not in their span we have the strict inequality.)

Next, one shows that the Dynkin diagram cannot contain more than one 3-valent vertex. Otherwise it contains a subdiagram with a chain $\alpha_1, \ldots, \alpha_n$, and two extra vertices β_1, β_2 connected to α_1 and two extra vertices β_3, β_4 connected to α_n. Let $\alpha = \alpha_1 + \ldots + \alpha_n$. Then $||\alpha||^2 = 2n - 2\sum_{i=1}^{n-1}(\alpha_i, \alpha_{i+1}) = 2$, and $(\alpha, \beta_i) = -1$. Hence, the same argument as in the previous step (with α here playing the role of α_5 there) gives a contradiction:

$$2 = ||\alpha||^2 > \sum_{i=1}^{4} \frac{(\alpha, \beta_i)^2}{||\beta_i||^2} = \sum_{i=1}^{4} (\frac{-1}{2})^2 = 2.$$

Thus, the only type of diagrams that remain are chains, i.e. diagrams of type A_l, or star-shaped diagrams with a central vertex ψ and three 'branches' of length r, s, t emanating from ψ. Label the vertices in these branches by $\alpha_1, \ldots, \alpha_{r-1}$, $\beta_1, \ldots, \beta_{s-1}$ and $\gamma_1, \ldots, \gamma_{t-1}$ in such a way that $(\alpha_1, \alpha_2) \neq 0, \ldots, (\alpha_{r-1}, \psi) \neq 0$ and similarly for the other branches. Let

$$\alpha = \sum_{j=1}^{r-1} j\alpha_j, \quad \beta = \sum_{j=1}^{s-1} j\beta_j, \quad \gamma = \sum_{j=1}^{t-1} j\gamma_j.$$

Then α, β, γ are pairwise orthogonal, and $\alpha, \beta, \gamma, \psi$ are linearly independent. We have $||\alpha||^2 = r(r-1)$ and $(\alpha, \psi) = -(r-1)$, and similarly for β, γ. Hence

$$2 = ||\psi||^2 > \frac{(\alpha, \psi)^2}{||\alpha||^2} + \frac{(\beta, \psi)^2}{||\beta||^2} + \frac{(\gamma, \psi)^2}{||\gamma||^2} = \frac{r-1}{r} + \frac{s-1}{s} + \frac{t-1}{t}.$$

Equivalently,

$$\frac{1}{r} + \frac{1}{s} + \frac{1}{t} > 1.$$

One easily checks that the only solutions with $r, s, t \geq 2$ and (with no loss of generality) $r \leq s \leq t$ are:

$$(2, 2, l-2), \ l \geq 4, \ (2, 3, 3), \ (2, 3, 4), (2, 3, 5).$$

These are the Dynkin diagrams of type $D_l, \ E_6, E_7, E_8$. It remains to show that these Dynkin diagrams correspond to root systems, but this can be done by explicit construction of the root systems. □

Consider the Dynkin diagram of E_8, with vertices of the long chain labeled as $\alpha_1, \ldots, \alpha_7$, and with the vertex α_5 connected to α_8. It may be realized as the following set of vectors in $\mathbb{R}^8$:

$$\alpha_i = \epsilon^i - \epsilon^{i+1}, \ i = 1, \ldots, 7$$

together with

$$\alpha_8 = \tfrac{1}{2}(\epsilon^1 + \ldots + \epsilon^5) - \tfrac{1}{2}(\epsilon^6 + \epsilon^7 + \epsilon^8).$$

(Indeed, this vectors have length squared equal to 2, and the correct angles.) The reflection s_i for $i \leq 7$ acts as transposition of indices $i, i+1$. Hence S_8 is embedded as a subgroup of the Weyl group. Hence,

$$\beta = -\tfrac{1}{2}(\epsilon_1 + \epsilon_2 + \epsilon_3) + \tfrac{1}{2}(\epsilon_4 + \ldots + \epsilon_8)$$

is also a root, obtained from α_8 by permutation of $1,2,3$ with $4,5,6$. Applying s_8, we see that

$$s_8(\beta) = \beta + \alpha_8 = \epsilon^4 + \epsilon^5$$

is a root. Hence, the set of roots contains all $\pm\epsilon^i \pm \epsilon^j$ with $i < j$, and the Weyl group contains all even numbers of sign changes. (In fact, we have just seen that the root system of E_8 contains that of D_8.) We conclude that

$$\mathfrak{R} = \{\pm\epsilon^i \pm \epsilon^j\} \cup \{\tfrac{1}{2}(\pm\epsilon^1 \pm \epsilon^2 \cdots \pm \epsilon^8)\}$$

where the second set has all sign combinations with an odd number of minus signs. Note that there are $2l(l-1) = 112$ roots of the first type, and $2^7 = 128$ roots of the second type. Hence the dimension of the Lie group with this root system is $112 + 128 + 8 = 248$. With a little extra effort, one finds that the order of the Weyl group is $|W| = 696{,}729{,}600$.

19. Fundamental weights

Let G be a semi-simple compact Lie group. Inside $E^* = X_*(T) \otimes_{\mathbb{Z}} \mathbb{R}$, we have three lattices

$$\mathrm{span}_{\mathbb{Z}}(\mathfrak{R}^\vee) \subset X_*(T) \subset \mathrm{span}_{\mathbb{Z}}(\mathfrak{R})^*,$$

where $\mathrm{span}_{\mathbb{Z}}(\mathfrak{R})^* = \mathrm{Hom}(\mathrm{span}_{\mathbb{Z}}(\mathfrak{R}), \mathbb{Z})$ is the dual of the root lattice.

Theorem 19.1. *There are canonical isomorphism,*

$$\mathrm{span}_{\mathbb{Z}}(\mathfrak{R})^*/X_*(T) \cong Z(G)$$

and

$$X_*(T)/\mathrm{span}_{\mathbb{Z}}(\mathfrak{R}^\vee) \cong \pi_1(G).$$

In particular, if G is simply connected the co-weight lattice agrees with the co-root lattice, and hence has basis $\alpha_i^\vee$ the simple co-roots. Furthermore, the center is then given in terms of root data by

$$Z(G) = \mathrm{span}_{\mathbb{Z}}(\mathfrak{R})^*/\mathrm{span}_{\mathbb{Z}}(\mathfrak{R}^\vee).$$

We will give a full proof of the first part, but only an incomplete proof of the second part.

Partial proof. We have

$$Z(G) = \bigcap_{\alpha \in \mathfrak{R}} \ker(\alpha),$$

since $t \in T$ lies in $Z(G)$ if and only if its action on all root spaces $\mathfrak{g}_\alpha$ is trivial. Under the identification $X_*(T) \otimes_{\mathbb{Z}} \mathbb{R} \cong i\mathfrak{t}$, the elements of $\mathrm{span}_{\mathbb{Z}}(\mathfrak{R})^*$ consist of all $\xi \in i\mathfrak{t}$ such that $\mathrm{d}\alpha(\xi) \in \mathbb{Z}$ for all $\alpha \in \mathfrak{R}$, hence $\exp(2\pi i \xi) \in Z(G)$. On the other hand, $X_*(T)$ corresponds to elements such that $\exp(2\pi i\xi) = 1$. This shows $\mathrm{span}_{\mathbb{Z}}(\mathfrak{R})^*/X_*(T) = Z(G)$.

The inclusion $T \to G$ gives a group morphism

$$X_*(T) = \pi_1(T) \to \pi_1(G).$$

Choose an invariant inner product B on $\mathfrak{g}$, and let G carry the corresponding Riemannian metric. Think of $\pi_1(G)$ as a subgroup of $\widetilde{G}$. Given $x \in \pi_1(G)$, let $\tilde{\gamma}(t)$ be a geodesic in $\widetilde{G}$ from 0 to x. Its image $\gamma(t) = \exp_G(t\xi)$ is then a closed geodesic in G representing the given element of the fundamental class. But any geodesic of G is of the form $\exp(t\xi)$ for some $\xi \in \mathfrak{g}$. Choose $g \in G$ such that $\mathrm{Ad}_g(\xi) \in \mathfrak{t}$. Then $\mathrm{Ad}_g \gamma(t)$ is homotopic to $\gamma(t)$ (since G is connected), and lies in T. This shows that the map $\pi_1(T) \to \pi_1(G)$ is *surjective*.

Recall next that we defined the co-roots $\alpha^\vee$ as the composition of $\mathrm{U}(1) \cong T_\alpha \subset \mathrm{SU}(2)_\alpha$ with the map $j_\alpha\colon \mathrm{SU}(2)_\alpha \to G$. But $\mathrm{SU}(2)_\alpha$ is simply connected. Hence the loop in $\mathrm{SU}(2)_\alpha$ is contractible, and so is its image under j_α. This shows that $\mathrm{span}_\mathbb{Z}(\mathfrak{R}^\vee)$ lies in the kernel of the map $X_*(T) \to \pi_1(G)$. The harder part is to show that it is actually equal to the kernel. For a proof of this part, see e.g. Broecker-tom Dieck, chapter V.7. □

Exercise 19.2. Find a generalization of this Theorem to arbitrary compact connected Lie groups.

If G is semi-simple, define elements $\varpi_i \in E = X^*(T) \otimes_\mathbb{Z} \mathbb{R}$ by $\langle \alpha_i^\vee, \varpi_j \rangle = \delta_{ij}$. By definition, these elements span the positive Weyl chamber $C_+ \subset E$. (If G simply connected, the Theorem above says that the co-roots $\alpha_i^\vee$ are a $\mathbb{Z}$-basis of $X_*(T)$, hence the ϖ_i are a $\mathbb{Z}$-basis of $X^*(T)$ in that case.) One calls $\varpi_1, \ldots, \varpi_l$ the *fundamental weights*. One also defines fundamental co-weights $\varpi_1^\vee, \ldots, \varpi_l^\vee \in E^*$, by duality to the simple roots. The $\varpi_i^\vee$ are then a basis of $\mathrm{span}_\mathbb{Z}(\mathfrak{R})^*$.

Example 19.3. The root system of type D_l is realized as the set of all $\pm\epsilon^i \pm \epsilon^j$ with $i < j$. Since all roots have $||\alpha||^2 = 2$, we may identify roots and co-roots, fundamental weights and fundamental co-weights. Recall that we may take $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ to be

$$\alpha_i = \alpha_i^\vee = \begin{cases} \epsilon^i - \epsilon^{i+1} & i \le l-1 \\ \epsilon^{l-1} + \epsilon^l & i = l. \end{cases}$$

The corresponding set of fundamental weights is

$$\varpi_i = \varpi_i^\vee = \begin{cases} \epsilon^1 + \ldots + \epsilon^i & i < l-1, \\ \frac{1}{2}(\epsilon^1 + \ldots + \epsilon^{l-1} - \epsilon^l) & i = l-1, \\ \frac{1}{2}(\epsilon^1 + \ldots + \epsilon^{l-1} + \epsilon^l) & i = l. \end{cases}$$

The root lattice $\mathrm{span}_\mathbb{Z}(\mathfrak{R})$ consists of all $\sum k_i \epsilon^i$ with integer coefficients such that $\sum k_i \in 2\mathbb{Z}$. If l is odd, the quotient $\mathrm{span}_\mathbb{Z}(\mathfrak{R})/\mathrm{span}_\mathbb{Z}(\mathfrak{R})^*$ is isomorphic to $\mathbb{Z}_4$; it is generated by the image of ϖ_l. If l is even, we have $\mathrm{span}_\mathbb{Z}(\mathfrak{R})/\mathrm{span}_\mathbb{Z}(\mathfrak{R})^* = \mathbb{Z}_2 \times \mathbb{Z}_2$, generated by the images of ϖ_{l-1} and ϖ_l. Hence, the simply connected Lie group corresponding to D_l has center $\mathbb{Z}_4$ if l is odd, and $Z_2 \times \mathbb{Z}_2$ if l is even.

20. More on the Weyl group

We will need a few more facts about the relation of the Weyl group with the root system. Recall that $s_i = w_{\alpha_i}$ are the simple reflections corresponding to the simple roots. We have:

Lemma 20.1. *The simple reflection s_i preserves the set $\mathfrak{R}_+ \backslash \{\alpha_i\}$.*

Proof. Suppose $\alpha \in \mathfrak{R}_+\backslash\{\alpha_i\}$. Write $\alpha = \sum_j k_j\alpha_j \in \mathfrak{R}_+$, so that all $k_j \geq 0$. The root

$$s_i\alpha = \alpha - \langle \alpha, \alpha_i^\vee \rangle \alpha_i = \sum_j k'_j \alpha_j. \tag{6}$$

has coefficients $k'_j = k_j$ for $j \neq i$. Since α is not a multiple of α_i it follows that $k'_j = k_j > 0$ for some $j \neq i$. This shows that $s_i\alpha$ is positive. □

Remark 20.2. This Lemma may also be understood geometrically. Fix $\xi \in \operatorname{int}(C_+)$, so that a root is positive if and only if $(\alpha, \xi) > 0$. The Weyl group element s_i acts by reflection across the hyperplane corresponding to α_i. Then $s_i(\xi)$ lies in the adjacent chamber, and the line segment from ξ to $s_i(\xi)$ does not meet any root hyperplane other than $H_{\alpha_i} = H_{-\alpha_i}$. Hence, for all roots $\alpha \neq \pm\alpha_i$, the inner products (α, ξ) and $(\alpha, s_i\xi) = (s_i\alpha, \xi)$ have the same sign.

We have seen that the Weyl group is generated by simple reflections. We use this to define:

Definition 20.3. The *length* $l(w)$ of a Weyl group element $w \in W$ is the smallest number r such that w can be written in the form

$$w = s_{i_1} \ldots s_{i_r}. \tag{7}$$

If $r = l(w)$ the expression (7) is called *reduced*.

Simple properties of the length function are

$$l(w^{-1}) = l(w), \quad l(ww') \leq l(w) + l(w'), \ (-1)^{l(ww')} = (-1)^{l(w)}(-1)^{l(w')}.$$

(The last identity follows e.g. since $(-1)^{l(w)}$ is the determinant of w as an element of $\mathrm{GL}(E)$.) It is also clear that for any i, $l(ws_i) - l(w) = \pm 1$. The sign can be determined as follows.

Proposition 20.4. *For any Weyl group element w, and any simple root α_i, we have*

$$l(ws_i) = \begin{cases} l(w) + 1 & w\alpha_i \in \mathfrak{R}_+ \\ l(w) - 1 & w\alpha_i \in \mathfrak{R}_- \end{cases}$$

Proof. Consider a reduced expression (7) for w. Suppose $w\alpha_i \in \mathfrak{R}_-$. Then there is an index $m \leq r$ such that

$$s_{i_{m+1}} \cdots s_{i_r}\alpha_i \in \mathfrak{R}_+, \quad s_{i_m} \cdots s_{i_r}\alpha_i \in \mathfrak{R}_-.$$

In terms of $u = s_{i_{m+1}} \cdots s_{i_r}$, these equations read

$$u\alpha_i \in \mathfrak{R}_+, \quad s_{i_m}u\alpha_i \in \mathfrak{R}_-.$$

Since α_{i_m} is the unique positive root that becomes negative under s_{i_m}, it follows that $\alpha_{i_m} = u\alpha_i$. Consequently,

$$s_{i_m} = us_iu^{-1}.$$

Multiplying from the right by u, and from the left by $s_{i_1} \cdots s_{i_m}$, we obtain

$$s_{i_1} \cdots s_{i_{m-1}} s_{i_{m+1}} \cdots s_{i_r} = ws_i.$$

This shows $l(ws_i) = l(w) - 1$. The case that $w\alpha_i \in \mathfrak{R}_+$ is reduced to the previous case, since $w'\alpha_i$ with $w' = ws_i$ is negative. □

For any $w \in W$, let $\mathfrak{R}_{+,w}$ be the set of positive roots that are made negative under w^{-1}. That is,

$$\mathfrak{R}_{+,w} = \mathfrak{R}_+ \cap w\mathfrak{R}_-. \tag{8}$$

Proposition 20.5. *For any $w \in W$,*

$$l(w) = |\mathfrak{R}_{+,w}|. \tag{9}$$

In fact, given a reduced expression (7) *one has*

$$\mathfrak{R}_{+,w} = \{\alpha_{i_1},\ s_{i_1}\alpha_{i_2},\ \cdots,\ s_{i_1}\cdots s_{i_{r-1}}\alpha_{i_r}\} \tag{10}$$

Proof. Consider $w' = ws_i$. Suppose a positive root α is made negative under $(w')^{-1} = s_i w^{-1}$. Since s_i changes the sign of $\pm\alpha_i$, and preserves both $\mathfrak{R}_+\backslash\{\alpha_i\}$ and $\mathfrak{R}_-\backslash\{-\alpha_i\}$, we see that $w^{-1}\alpha$ is negative if and only if $w^{-1}\alpha \neq -\alpha_i$. That is,

$$\mathfrak{R}_{+,ws_i} = \begin{cases} \mathfrak{R}_{+,w} \cup \{w\alpha_i\}, & \text{if } w\alpha_i \in \mathfrak{R}_+ \\ \mathfrak{R}_{+,w}\backslash\{-w\alpha_i\}, & \text{if } w\alpha_i \in \mathfrak{R}_-. \end{cases}$$

(10) now follows by induction on $l(w)$, using that for a reduced expression $w = s_{i_1}\cdots s_{i_r}$, all $s_{i_1}\cdots s_{i_{k-1}}\alpha_{i_k}$ are positive. □

This result has the following geometric interpretation in terms of Weyl chambers.

Corollary 20.6. *Let $w \in W$. A line segment from a point in $\mathrm{int}(C_+)$ to a point in $\mathrm{int}(w(C_+))$ meets the hyperplane $H_\alpha, \alpha \in \mathfrak{R}_+$ if and only if $\alpha \in \mathfrak{R}_{+,w}$. Hence, $l(w)$ is the number of hyperplanes crossed by such a line segment.*

The proof is left as an exercise. As a special case, there is a unique Weyl group element of length $l(w) = |\mathfrak{R}|$. It satisfies $\mathfrak{R}_{w,+} = \mathfrak{R}_+$, hence $w(C_+) = -C_+$.

We have the following explicit formula for the action of $w \in W$ on E.

Lemma 20.7. *For all $\mu \in E$, and all $w \in W$ with reduced expression* (7)*,*

$$w\mu = \mu - \sum \langle \alpha_{i_j}^\vee, \mu\rangle s_{i_1}\cdots s_{i_{j-1}}\alpha_{i_j}.$$

Proof. We calculate,

$$\begin{aligned} \mu - w\mu &= \mu - s_{i_1}\mu + s_{i_1}(\mu - s_{i_2}\mu) + \cdots + s_{i_1}\cdots s_{i_{r-1}}(\mu - s_{i_r}\mu) \\ &= \sum \langle \alpha_{i_j}^\vee, \mu\rangle s_{i_1}\cdots s_{i_{j-1}}\alpha_{i_j}. \end{aligned}$$ □

Observe this formula expresses $\mu - w\mu$ as a linear combination of elements in $\mathfrak{R}_{+,w}$. If $\mu \in C_+$ (resp. $\mu \in X^*(T)$), then all the coefficients are ≥ 0 (resp. $\in \mathbb{Z}$). An interesting special cases arises for the element

$$\rho = \sum_{i=1}^{l} \varpi_i.$$

Note that $\rho \in \mathrm{int}(C_+)$, and that $\langle \alpha_i^\vee, \rho\rangle = 1$ for all i. Hence, the formula shows

$$\rho - w\rho = \sum_{\alpha \in \mathfrak{R}_{+,w}} \alpha.$$

Taking w to be the longest Weyl group element, we have $w\Pi = -\Pi$, hence $w\rho = -\rho$. Consequently,

$$\rho = \tfrac{1}{2} \sum_{\alpha \in \mathfrak{R}_+} \alpha.$$

In particular, the expression on the right hand side lies in the lattice $\mathrm{span}_{\mathbb{Z}}(\mathfrak{R}^\vee)^*$ spanned by the fundamental weights (equal to $X^*(T)$ is G is simply connected). The element ρ plays an important role in representation theory.

21. Representation theory

Let G be a compact, connected Lie group. Fix a choice of maximal torus T and a positive Weyl chamber $C_+ \subset E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$, and let $\Pi \subset \mathfrak{R}_+$ be the set of simple roots and positive roots, respectively. Recall that the Weyl chamber is given in terms of simple co-roots by

$$C_+ = \{\mu \in E |\ \langle \alpha^\vee, \mu \rangle \geq 0, \quad \alpha \in \Pi\},$$

and that any $\mu \in E$ is W-conjugate to a unique element of C_+. In particular, every $\mu \in X^*(T)$ is W-conjugate to a unique element in

$$X^*(T)_+ := C_+ \cap X^*(T).$$

We call these the *dominant weights* of G. If G is simply connected, so that $X^*(T)$ has basis the fundamental weights, the dominant weights are those of the form $\sum_{i=1}^{l} m_i \varpi_i$ with $m_i \in \mathbb{Z}_{\geq 0}$. That is, $X^*(T)_+$ is a free monoid with basis the fundamental weights.

Similar to the representation theory of $\mathfrak{sl}(2, \mathbb{C})$ (or equivalently $\mathrm{SU}(2)$), the representation theory of compact Lie groups relies on the concept of a highest weight. It is convenient to introduce the nilpotent subalgebras

$$\mathfrak{n}^+ = \bigoplus_{\alpha \in \mathfrak{R}_+} \mathfrak{g}_\alpha, \quad \mathfrak{n}^- = \bigoplus_{\alpha \in \mathfrak{R}_-} \mathfrak{g}_\alpha.$$

Then

$$\mathfrak{g}^{\mathbb{C}} = \mathfrak{n}^- \oplus \mathfrak{t}^{\mathbb{C}} \oplus \mathfrak{n}^+$$

as vector spaces. In terms of the Chevalley generators e_i, f_i, h_i, $\mathfrak{n}_+$ is generated by the e_i, $\mathfrak{n}_-$ is generated by the f_i.

Recall that for any G-representation, we denote by $\Delta(V)$ the W-invariant set of its weights.

Theorem 21.1. *Let $\pi \colon G \to \mathrm{End}(V)$ be an irreducible complex G-representation. Then there is a unique weight $\mu \in \Delta(V) \subset X^*(T)$, with the property that $\mu + \alpha$ is not a weight for all positive roots $\alpha \in \mathfrak{R}_+$. The weight μ has multiplicity* 1. *All other weights $\nu \in \Delta(V)$ are of the form*

$$\nu = \mu - \sum_i k_i \alpha_i$$

where all $k_i \in \mathbb{Z}_{\geq 0}$.

Proof. Choose $\mu \in \Delta(V) \cap C_+$ with $||\mu||$ as large as possible. Then $\mu + \alpha \notin \Delta(V)$, for all $\alpha \in \mathfrak{R}_+$, because

$$||\mu + \alpha||^2 = ||\mu||^2 + ||\alpha||^2 + 2(\mu, \alpha) > ||\mu||^2.$$

Let $v \in V_\mu$ be a non-zero vector.

By one of the corollaries to the Poincare-Birkhoff-Witt theorem, the multiplication map

$$U(\mathfrak{n}^-) \otimes U(\mathfrak{t}^{\mathbb{C}}) \otimes U(\mathfrak{n}^+) \to U(\mathfrak{g}^{\mathbb{C}})$$

is an isomorphism of vector spaces. The 1-dimensional subspace $\mathrm{span}(v)$ is invariant under the action of $U(\mathfrak{n}^+)$ (since v is annihilated by all $\mathrm{d}\pi(e_\alpha)$ with $\alpha \in \mathfrak{R}_+$), and also under $U(\mathfrak{t}^{\mathbb{C}})$, since v is a weight vector. Omitting $\mathrm{d}\pi$ from the notation, this shows

$$U(\mathfrak{g}^{\mathbb{C}})v = U(\mathfrak{n}^-)U(\mathfrak{t}^{\mathbb{C}})U(\mathfrak{n}^+)v = U(\mathfrak{n}^-)v.$$

Hence $U(\mathfrak{n}^-)v$ is a non-zero sub-representation, and is therefore equal to V. Since $\mathfrak{n}^-$ is generated by the Chevalley generators f_i, we see that V is spanned by vectors of the form

$$f_{i_1} \cdots f_{i_r} v$$

Since f_i takes V_ν to $V_{\nu-\alpha_i}$, each $f_{i_1} \cdots f_{i_r} v$ is a weight vector, of weight $\mu - \sum_{i=1}^l k_i\alpha_i$, where $k_i = \#\{j : i_j = i\}$. If $r > 0$, we have $\sum_{i=1}^l k_i > 0$. This shows that μ has multiplicity 1, and that it is the unique highest weight of V. □

We introduce a partial order on E, by declaring that $\mu \succeq \mu'$ if $\mu - \mu'$ lies in the cone spanned by the positive roots:

$$\mu \succeq \mu' \Leftrightarrow \mu = \mu' + \sum_{i=1}^{l} k_i\alpha_i, \ k_i \in \mathbb{R}_{\geq 0}.$$

Equivalently, in terms of the fundamental coweights, $\langle \mu - \mu', \varpi_i^\vee \rangle \geq 0$ for $i = 1, \ldots, l$. The formula for the action of W (cf. Lemma 20.7) shows that for all $\mu \in C_+$, and all $w \in W$,

$$\mu \succeq w\mu.$$

In the notation of the Theorem, any irreducible representation V has a unique weight $\mu \in \Delta(V)$ with the property that $\mu \succ \nu$ for all other weights $\nu \in \Delta(V) - \{\mu\}$. It is called the *highest weight* of the irreducible representation V.

Exercise 21.2. Show that the highest weight may also be characterized as the unique weight $\mu \in \Delta(V)$ where the function $\nu \mapsto ||\nu + \rho||$ takes on its maximum.

Theorem 21.3. *Any irreducible G-representation is uniquely determined, up to isomorphism, by its highest weight.*

Proof. Let V, V' be two irreducible representations, both of highest weight μ. Let $v \in V_\mu$ and $v' \in V'_\mu$ be (non-zero) highest weight vectors. The element $(v, v') \in (V \oplus V')_\mu$ generates a sub-representation

$$W = U(\mathfrak{g}^{\mathbb{C}})(v, v') \subset V \oplus V'.$$

The projection $W \to V$ is a G-equivariant surjective map, since its image is $U(\mathfrak{g}^{\mathbb{C}})v = V$. Since $W \cap 0 \oplus V'$ is a proper subrepresention of V', it must be zero. Hence, the projection map $W \to V$ is a G-equivariant isomorphism, and so is the projection $W \to V'$. □

The discussion above has the following consequence.

Corollary 21.4. *Let V be a finite-dimensional G-representation, and let*

$$V^{\mathfrak{n}} = \{v \in V |\ \pi(e_i)v = 0, \ i = 1, \ldots, l\}.$$

Then $\dim V^{\mathfrak{n}}$ is the number of irreducible components of V. If $v \in V^{\mathfrak{n}}$ is a weight vector of weight μ, then $U(\mathfrak{g}^{\mathbb{C}})v \subset V$ is an irreducible representation of highest weight μ.

The much deeper result is the converse to Theorem 21.3:

Theorem 21.5 (H. Weyl)**.** *For any dominant weight $\mu \in X^*(T)_+$, there is an irreducible G-representation $V(\mu)$, unique up to isomorphism, having μ as its highest weight.*

Remark 21.6. There is a beautiful geometric construction of these representations, called the *Borel-Weil-Bott* construction. In a nutshell, let C_μ be the 1-dimensional T-representation of weight $\mu \in X^*(T)_+$. The quotient $L = (G \times \mathbb{C}_\mu)/T$, where T acts as $t.(g,z) = (gt^{-1}, \mu(t)z)$, is a complex line bundle over G/T. Using the identification of $T_e(G/T) = \mathfrak{g}/\mathfrak{t} = \bigoplus_{\alpha \in \mathfrak{R}_+} \mathfrak{g}_\alpha$, one gets a T-invariant complex structure on $T_e(G/T)$, which extends to a G-invariant complex structure on the fibers of the tangent bundle $T(G/T)$. This turns out to be integrable, thus G/T becomes a complex manifold (with holomorphic transition functions). Likewise the line bundle acquires a holomorphic structure. The space of holomorphic sections $V = \Gamma_{\mathrm{hol}}(G/T, L)$ is a finite-dimensional G-representation. One may show that it is irreducible of highest weight μ.

Suppose G is simply connected. The representations $V(\varpi_j)$ corresponding to the fundamental weights are called the fundamental representations. One these are known, one obtains a model of the irreducible representation of highest weight $\mu = \sum_{i=1}^{l} k_i \varpi_i$, for any $k_i \geq 0$: The tensor product $V(\varpi_1)^{\otimes k_1} \otimes \cdots \otimes V(\varpi_l)^{\otimes k_l}$ has a highest weight vector of weight $\sum k_i \varpi_i$, generating an irreducible subrepresentation $V(\mu)$.

Let us consider some representations of $\mathrm{SU}(l+1)$. We work with the standard choice of maximal torus $T \subset \mathrm{SU}(l+1)$. Recall that $E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$ may be regarded as the subspace of elements of coordinate sum 0,

$$E = \{\sum k_i \epsilon^i | \ \sum k_i = 0\}.$$

The weight lattice $X^*(T)$ is obtained from the weight lattice $X^*(T') = \mathrm{span}_{\mathbb{Z}}(\epsilon^1, \ldots, \epsilon^{l+1})$ by projection along the direction $\sum_{j=1}^{l+1} \epsilon^j$.

The simple roots are $\alpha_i = \epsilon^i - \epsilon^{i+1}, \ i = 1, \ldots, l$. For the standard inner product on E these have length squared equal to 2, hence they coincide with the co-roots. The fundamental weights dual to the co-roots are

$$\varpi^i = (\epsilon^1 + \cdots + \epsilon^i) - \frac{i}{l+1} \sum_{j=1}^{l+1} \epsilon^j.$$

The defining action of $\mathrm{U}(l+1)$ on $\mathbb{C}^{l+1}$ has the set of weights $\Delta'(V) = \{\epsilon^1, \ldots, \epsilon^{l+1}\} \subset X^*(T')$. To get the corresponding set of weights $\Delta(V)$ of $\mathrm{SU}(l+1)$, we have to project along $\epsilon^1 + \ldots + \epsilon^{l+1}$. That is,

$$\Delta(\mathbb{C}^{l+1}) = \{\epsilon^i - \frac{1}{l+1}(\epsilon^1 + \ldots + \epsilon^{l+1}), \ i = 1, \ldots, l+1\}.$$

Taking inner products with the simple roots, we see that only one of these weights lies in the Weyl chamber C_+: This is the weight $\epsilon^1 - \frac{1}{l+1}(\epsilon^1 + \ldots + \epsilon^{l+1}) = \varpi_1$. We hence see that $\mathbb{C}^{l+1}$ is the irreducible representation of highest weight ϖ_1.

Consider more generally the k-th exterior power of the defining representation, $\wedge^k(\mathbb{C}^{l+1})$. The weights for the $\mathrm{U}(l+1)$-action are all $\epsilon_{i_1} + \ldots + \epsilon_{i_k}$ such that $i_1 < \cdots < i_k$. Each of these weights has multiplicity 1. The corresponding weights of $\mathrm{SU}(l+1)$ are their projections:

$$\epsilon^{i_1} + \ldots + \epsilon^{i_k} - \frac{k}{l+1}(\epsilon^1 + \ldots + \epsilon^{l+1}).$$

Again, we find find that only one of the weights lies in C_+, this is the weight

$$\epsilon^1 + \cdots + \epsilon^k - \frac{k}{l+1}(\epsilon^1 + \ldots + \epsilon^{l+1}) = \varpi_k.$$

That is, the irreducible representation of highest weight ϖ_k is realized as $\wedge^k(\mathbb{C}^{l+1})$.

Proposition 21.7. *The irreducible representation of* $\mathrm{SU}(l+1)$ *of highest weight* ϖ_k *is realized as the k-th exterior power* $\wedge^k(\mathbb{C}^{l+1})$ *of the defining representation on* $\mathbb{C}^{l+1}$*.*

It is also interesting to consider the symmetric powers of the defining representation.

Proposition 21.8. *The k-th symmetric power* $S^k(\mathbb{C}^{l+1})$ *of the defining representation of* $\mathrm{SU}(l+1)$ *is irreducible, of highest weight* $k\varpi_1$*. Similarly the representation* $S^k((\mathbb{C}^{l+1})^*)$ *is irreducible, of highest weight* $k\varpi_l$*.*

The proof is left as an exercise.

The complexified adjoint representation on the complexified Lie algebra $\mathfrak{su}(l+1)^{\mathbb{C}}$ is irreducible, since $\mathrm{SU}(l+1)$ is simple. Its weights are, by definition, the roots together with 0. Hence, there must be a *unique root* $\alpha_{\max}$ such that $\alpha + \alpha_i$ is not a root, for any i. Indeed, one observes that

$$\alpha_{\max} = \epsilon^1 - \epsilon^{l+1}$$

is such a root. It is the *highest root*, i.e. the root for which $\mathrm{ht}(\alpha)$ takes on its maximum. In fact, $\mathrm{ht}(\alpha_{\max}) = l$ since it is the sum of the simple roots.

22. Highest weight representations

In this section we will present a proof of Weyl's theorem for the case of a *simply connected* compact Lie group G. The proof is on the level of Lie algebras. Hence, we will consider complex representations $\pi\colon \mathfrak{g} \to \mathrm{End}(V)$ of the Lie algebra - if V is finite-dimensional, any such representation integrates to the group level.

Let $E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$. It is convenient to identify $E^{\mathbb{C}} \cong (\mathfrak{t}^{\mathbb{C}})^*$, by the linear map taking $\mu \in X^*(T)$ to the linear functional $\mathrm{d}\mu\colon \mathfrak{t}^{\mathbb{C}} \to \mathbb{C}$. Using this identification, we will simply write μ in place of $\mathrm{d}\mu$, and likewise for the roots.

As before, we let $\mathfrak{n}^{\pm}$ denote the nilpotent Lie subalgebra of $\mathfrak{g}^{\mathbb{C}}$ given as the direct sum of root spaces $\mathfrak{g}_\alpha$ for $\alpha \in \mathfrak{R}_{\pm}$. Their direct sum with $\mathfrak{t}^{\mathbb{C}}$ is still a Lie subalgebra of $\mathfrak{g}^{\mathbb{C}}$, called the *Borel subalgebras* $\mathfrak{b}^{\pm} = \mathfrak{t}^{\mathbb{C}} \oplus \mathfrak{n}^{\pm}$.

Definition 22.1. Let $\pi\colon \mathfrak{g}^{\mathbb{C}} \to \mathrm{End}(V)$ be a complex representation, possibly infinite-dimensional. A non-zero vector $v \in V$ is called a *weight vector* if $\mathrm{span}(v)$ is invariant under the action of $\mathfrak{t}^{\mathbb{C}}$.

For $\mu \in (\mathfrak{t}^{\mathbb{C}})^*$ we denote $V_\mu = \{v \in V \mid \pi(\xi)v = \mu(\xi)v,\ \xi \in \mathfrak{t}^{\mathbb{C}}\}$. Then the weight vectors are the non-zero elements of the V_μ's. Note that μ need not lie in $X^*(T)$ in general.

Definition 22.2. A $\mathfrak{g}^{\mathbb{C}}$-representation V is called a *highest weight representation* if there is a weight vector v with $\pi(\mathfrak{n}^+)v = 0$ and such that

$$V = \pi(U\mathfrak{g}^{\mathbb{C}})v.$$

Thus, span(v) is invariant under the action of $U(\mathfrak{b}^+)$. Using the PBW isomorphism $U(\mathfrak{n}^-)\otimes U(\mathfrak{b}^+)\to U(\mathfrak{g}^\mathbb{C})$ given by multiplication, it hence follows that

$$V = U(\mathfrak{n}^-).v.$$

Thus, any highest weight representation is spanned by weight vectors vectors

$$f_{i_1}\cdots f_{i_r}.v \tag{11}$$

The weight of such a weight vector (11) is $\mu - \sum k_i\alpha_i$ where $k_i = \#\{j|\ i_j = i\}$, and the multiplicity of any weight ν is bounded above by the of sequences $i_1,\ldots,i_r$ with $\nu = \mu - \sum_{j=1}^r \alpha_{i_j}$. As a very rough estimate the multiplicity is $\le r^l$ with $r = \mathrm{ht}(\mu-\nu)$. In particular, the highest weight μ.

An important example of a highest weight representation is given as follows. Given $\mu \in (\mathfrak{t}^\mathbb{C})^*$, define a 1-dimensional representation $\mathbb{C}_\mu$ of $\mathfrak{b}^+$ by letting $\xi\in\mathfrak{t}^\mathbb{C}$ act as a scalar $\langle\mu,\xi\rangle$ and letting $\mathfrak{n}^+$ act as zero.

Definition 22.3. The induced $\mathfrak{g}^\mathbb{C}$-representation

$$L(\mu) = U(\mathfrak{g}^\mathbb{C})\otimes_{U(\mathfrak{b}^+)}\mathbb{C}_\mu$$

(where $\mathfrak{g}^\mathbb{C}$-acts by the left regular representation on $U(\mathfrak{g}^\mathbb{C})$) is called the *Verma module* of highest weight $\mu\in(\mathfrak{t}^\mathbb{C})^*$.

Using the PBW theorem to write $U(\mathfrak{g}^\mathbb{C}) = U(\mathfrak{n}^-)\otimes U(\mathfrak{b}^+)$, we see that

$$L(\mu)\cong U(\mathfrak{n}^-)$$

as a vector space. Here $1\in U(\mathfrak{n}_-)$ corresponds to the highest weight vector. The PBW basis of $U(\mathfrak{n}^-)$ (defined by some ordering of the set $\mathfrak{R}_+$ of roots, and the corresponding basis of $\mathfrak{n}^-$) defines a basis of $L(\mu)$, consisting of weight vectors. We thus see that the set of weights of the Verma module is

$$\Delta(L(\mu)) = \{\mu - \sum_i k_i\alpha_i|\ k_i\ge 0\},$$

and the multiplicity of a weight ν is the number of distinct ways of writing $\nu = \mu - \sum k_\alpha\alpha$ as a sum over positive roots.

The *Verma module* is the universal highest weight module, in the following sense.

Proposition 22.4. *Let V be a highest weight representation, of highest weight $\mu\in(\mathfrak{t}^\mathbb{C})^*$. Then there exists a surjective $\mathfrak{g}$-module morphism $L(\mu)\to V$.*

Proof. Let $v\in V$ be a highest weight vector. The map $\mathbb{C}_\mu\to V,\ \lambda\mapsto\lambda v$ is $\mathfrak{b}^+$-equivariant. Hence, the surjective $\mathfrak{g}$-map $U(\mathfrak{g}^\mathbb{C})\to V\ x\mapsto\pi(x)v$ descends to a surjective $\mathfrak{g}$-map $L(\mu)\to V$. ∎

Lemma 22.5. *The Verma module $L(\mu)$ for $\mu\in(\mathfrak{t}^\mathbb{C})^*$ has a unique maximal proper submodule $L'(\mu)$. The quotient module*

$$V(\mu) = L(\mu)/L'(\mu)$$

is an irreducible highest weight module.

Proof. Any submodule is a sum of weight spaces; the submodule is proper if and only if μ does not appear as a weight. Hence, the sum of two proper submodules of $L(\mu)$ is again a proper submodule. Taking the sum of all proper submodules, we obtain a maximal proper submodule $L'(\mu)$.

Now let $V(mu) = L(\mu)/L'(\mu)$. The preimage of a proper submodule $W \subset V(\mu)$ is a proper submodule in $L(\mu)$, hence contained in $L'(\mu)$. Thus $W = 0$. This shows that $V(\mu)$ is irreducible. □

Proposition 22.6. *Let V be an irreducible $\mathfrak{g}$-representation of highest weight $\mu \in (\mathfrak{t}^{\mathbb{C}})^*$. Then V is isomorphic to $V(\mu)$; the isomorphism is unique up to a non-zero scalar.*

Proof. This is proved by a similar argument as in the Theorem for G-representations; see Theorem 21.3. □

It remains to investigate which of the irreducible $\mathfrak{g}^{\mathbb{C}}$-modules $V(\mu)$ are finite-dimensional. To this end we need:

Proposition 22.7. *Let $\pi\colon \mathfrak{sl}(2,\mathbb{C}) = \mathrm{span}_{\mathbb{C}}(e,f,h) \to \mathrm{End}(V)$ be a finite-dimensional representation. Then the transformation*

$$\Theta = \exp(\pi(e))\exp(-\pi(f))\exp(\pi(e)) \in \mathrm{GL}(V)$$

is well-defined. It implements the non-trivial Weyl group element, in the sense that

$$\Theta \circ \pi(h) \circ \Theta^{-1} = -\pi(h), \tag{12}$$

and takes the weight space V_l to V_{-l}.

Proof. Θ is well-defined, since $\pi(e), \pi(f)$ are nilpotent on V. For (12), consider first the defining representation on $\mathbb{C}^2$, given by matrices

$$\pi(e) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \pi(h) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \pi(f) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Here

$$\exp(\pi(e))\exp(-\pi(f))\exp(\pi(e)) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix};$$

conjugation of $\pi(h) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ by this matrix gives $-\pi(h)$.

Equation (12) for $V(1) = \mathbb{C}^2$ implies the result for each $V(k)$, since we may think of $V(k)$ as a sub-representation of $V(1)^{\otimes k}$. Taking direct sums, one gets the result for any finite-dimensional V. The last part is a consequence of (12): If $v \in V_l$, then $\pi(h)\Theta(v) = -\Theta(\pi(h)v) = -l\Theta(v)$, proving the claim. □

Proposition 22.8. *Let V be an irreducible highest weight representation, of highest weight $\mu \in (\mathfrak{t}^{\mathbb{C}})^*$. Then*

$$\dim V < \infty \Leftrightarrow \mu \in \mathrm{span}_{\mathbb{Z}}(\varpi_1, \ldots, \varpi_l) \cap C_+.$$

Proof. A finite-dimensional irreducible $\mathfrak{g}$-representation integrates to a G-representation, and we have already seen that $\mu \in X^*(T)_+ = \mathrm{span}_{\mathbb{Z}}(\varpi_1, \ldots, \varpi_l) \cap C_+$ in that case. Conversely, suppose $\mu \in \mathrm{span}_{\mathbb{Z}}(\varpi_1, \ldots, \varpi_l) \cap C_+$, write $V = V(\mu)$, and let $v \in V_\mu$ be the highest weight vector. Given $\alpha \in \mathfrak{R}_+$, consider the corresponding subalgebra $\mathfrak{sl}(2,\mathbb{C})_\alpha$.

Claim: Any element of $V(\mu)$ is contained in a finite-dimensional $\mathfrak{sl}(2,\mathbb{C})_\alpha$-subrepresentation of V. Proof of claim: The $\mathfrak{g}^{\mathbb{C}}$-module V is filtered by finite-dimensional subspaces

$$V^{(r)} = \pi(U^{(r)}(\mathfrak{g}^{\mathbb{C}}))v.$$

Each such subspace generates an $\mathfrak{sl}(2,\mathbb{C})_\alpha$-subrepresentation $\pi(U(\mathfrak{sl}(2,\mathbb{C})_\alpha))V^{(r)}$. Since

$$\mathfrak{sl}(2,\mathbb{C})_\alpha\, U^{(r)}(\mathfrak{g}^{\mathbb{C}}) = U^{(r)}(\mathfrak{g}^{\mathbb{C}})\mathfrak{sl}(2,\mathbb{C})_\alpha \mod U^{(r)}(\mathfrak{g}^{\mathbb{C}}),$$

we see by induction that

$$V^{(r)} \subset \pi(U^{(r)}(\mathfrak{g}^{\mathbb{C}}))(\pi(U(\mathfrak{sl}(2,\mathbb{C})_\alpha))v).$$

Hence it is enough to show that the subspace $\pi(U(\mathfrak{sl}(2,\mathbb{C})_\alpha))v$ is finite-dimensional. But this follows since

$$\pi(e_\alpha)v = 0,\ \pi(h_\alpha)v = \langle \mu, \alpha^\vee \rangle v,$$

implies that $\pi(U(\mathfrak{sl}(2,\mathbb{C})_\alpha))v$ is an irreducible $\mathfrak{sl}(2,\mathbb{C})_\alpha$-representation of dimension $\langle \mu, \alpha^\vee \rangle + 1$. (Here we are using that $\mu \in \mathrm{span}_{\mathbb{Z}}(\varpi_1, \ldots, \varpi_l) \cap C_+$.)

In particular, the operators $\pi(e_\alpha), \pi(f_\alpha)$ are *locally nilpotent*. (That is, for all $w \in V$ there exists $N > 0$ such that $\pi(e_\alpha)^N w = 0$ and $\pi(f_\alpha)^N w = 0$.) As a consequence, the transformation

$$\Theta_\alpha = \exp(\pi(e_\alpha))\exp(-\pi(f_\alpha))\exp(\pi(e_\alpha)) \in \mathrm{GL}(V)$$

is a well-defined automorphism of V. It satisfies

$$\Theta_\alpha \circ \pi(h) \circ \Theta_\alpha^{-1} = \pi(w_\alpha h)$$

for all $h \in \mathfrak{t}$. It follows that the set of weights $\Delta(V)$ is w_α-invariant. Since α was arbitrary, this proves that $\Delta(V)$ is W-invariant. But $\Delta(V) \subset \mu - \mathrm{cone}\,\mathfrak{R}_+$ has compact intersection with P_+. We conclude that $\Delta(V)$ is finite. Since the weights have finite multiplicity, it then follows that $\dim V < \infty$. ☐

In summary, we have constructed an irreducible finite-dimensional $\mathfrak{g}^{\mathbb{C}}$-representation $V(\mu)$, for any dominant weight μ. The action of $\mathfrak{g} \subset \mathfrak{g}^{\mathbb{C}}$ exponentiates to G, making $V(\mu)$ an irreducible G-representation.