Lie groups and Lie algebras (Fall 2019)

1. Terminology and notation

1.1. **Lie groups.** A Lie group is a group object in the category of manifolds:

> *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) [1]

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 $\mathrm{Mat}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ is obviously smooth – in fact, it is a polynomial.

Our second example is the *orthogonal group*

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

To see that it is a Lie group, it suffices to show

Lemma 1.3. $\mathrm{O}(n)$ *is an (embedded) submanifold of* $\mathrm{GL}(n, \mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{R})$.

Proof. This may be proved by using the regular value theorem: If we consider $A \mapsto A^\top A$ as a map to the space of symmetric $n \times n$-matrices, then I is a regular value. We'll give a somewhat longer argument, by directly constructing submanifold charts near any given $A \in \mathrm{O}(n)$: that is, local coordinate charts of $\mathrm{Mat}_n(\mathbb{R})$ around A in which $\mathrm{O}(n)$ looks like a subspace. We begin with $A = I$, using 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). Identifying $T_0 \mathrm{Mat}_n(\mathbb{R}) = \mathrm{Mat}_n(\mathbb{R})$, its differential at 0 is computed as

$$(T_0 \exp)(B) = \frac{d}{dt}\Big|_{t=0} \exp(tB) = B.$$

Hence the differential is the identity map, and in particular is invertible. The inverse function theorem tells us that there is $\epsilon > 0$ such that exp restricts to a diffeomorphism from the open

[1] There is an analogous definition of *topological group*, which is a group with a topology such that multiplication and inversion are continuous. Here, continuity of inversion does not follow from continuity of multiplication.

ϵ-ball $\{B\colon ||B|| < \epsilon\}$ around 0, onto an open neighborhood U of the identity matrix I. Here we take $||\cdot||$ to be the standard norm,

$$||B||^2 = \sum_{ij} (B_{ij})^2 \equiv \mathrm{tr}(B^\top B).$$

We claim that the inverse map

$$\log\colon U \to \{B\colon ||B|| < \epsilon\}$$

is the desired submanifold chart $(U, \log)$. In fact, for all B with $||B|| < \epsilon$,

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

where we put $\mathfrak{o}(n) = \{B|\ B + B^\top = 0\}$. So,

$$\log(\mathrm{O}(n) \cap U) = \mathfrak{o}(n) \cap \{B\colon ||B|| < \epsilon\},$$

the intersection of the range of our chart with a linear subspace.

For a more general $A \in \mathrm{O}(n)$, we use that the map

$$l_A\colon\ \mathrm{Mat}_n(\mathbb{R}) \to \mathrm{Mat}_n(\mathbb{R}),\ X \mapsto AX$$

is a diffeomorphism (since A is invertible). Hence, $l_A(U) = AU$ is an open neighborhood of A, and the map $\log \circ l_A^{-1}\colon l_A(U) \to \mathrm{Mat}_n(\mathbb{R})$ defines a submanifold chart around A. In fact, the range of this chart is the same as for $A = I$:

$$(\log \circ l_A^{-1})\big(AU \cap \mathrm{O}(n)\big) = \log\big(U \cap \mathrm{O}(n)\big).$$

$\square$

Since the group multiplication of $\mathrm{O}(n)$ is given by matrix multiplication, it is smooth. (The restriction of a smooth map to a submanifold is again smooth.) This shows that $\mathrm{O}(n)$ is a Lie group. Notice that this Lie group $\mathrm{O}(n)$ is compact: it is closed subset of $\mathrm{Mat}_{\mathbb{R}}(n)$ since it is the level set of the continuous map $A \mapsto A^\top A$, and it is also a bounded subset, since it is contained in the sphere of radius $\sqrt{n}$:

$$\mathrm{O}(n) \subseteq \{A|\ ||A||^2 = n\}$$

(using $\mathrm{tr}(A^\top A) = \mathrm{tr}(I) = n$ for $A \in \mathrm{O}(n)$).

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. Repeating the method for $\mathrm{O}(n)$, we find

$$\begin{aligned}
\exp(B) \in \mathrm{SL}(n, \mathbb{R}) &\Leftrightarrow \det(\exp(B)) = 1 \\
&\Leftrightarrow \exp(\mathrm{tr}(B)) = 1 \\
&\Leftrightarrow \mathrm{tr}(B) = 0 \\
&\Leftrightarrow B \in \mathfrak{sl}(n, \mathbb{R}).
\end{aligned}$$

with

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

where we used the identity $\det(\exp(B)) = \exp(\mathrm{tr}(B))$. The same technique works to give examples of many other *matrix Lie groups* (i.e.,submanifolds of the set of matrices which are a group under matrix multiplication). Let us now give a few more examples of Lie groups, without detailed justifications.

Examples 1.4. (a) Any finite-dimensional vector space V over $\mathbb{R}$ is a Lie group, with product Mult given by addition $V \times V \to V,\ (v,w) \mapsto v + w$.

(b) Consider a finite-dimensional associative algebra $\mathcal{A}$ over $\mathbb{R}$, with unit $1_{\mathcal{A}}$. We mostly have in mind the cases $\mathcal{A} = \mathbb{R}, \mathbb{C}, \mathbb{H}$, where $\mathbb{H}$ is the algebra of *quaternions* (due to $\mathbb{H}$amilton). 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

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

(But there are more examples, for instance the exterior algebra over a vector space, Clifford algebras and such.) For every $n \in \mathbb{N}$ we can create the algebra $\mathrm{Mat}_n(\mathcal{A})$ of matrices with entries in $\mathcal{A}$. The *general linear group*

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

is a Lie group of dimension $n^2 \dim_{\mathbb{R}}(\mathcal{A})$. Thus, we have

$$\mathrm{GL}(n,\mathbb{R}),\ \mathrm{GL}(n,\mathbb{C}), \mathrm{GL}(n,\mathbb{H})$$

as Lie groups of dimensions $n^2,\ 2n^2,\ 4n^2$.

(c) 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*

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

In particular, $\mathrm{SL}(n,\mathbb{C})$ is defined (of dimension $2n^2 - 2$). Since $\mathbb{H}$ is non-commutative (e.g. $ji = -ij$), it is not obvious how to define a determinant function on quaternionic matrices. Still, it is (unfortunately) standard to use the notation $\mathrm{SL}(n,\mathbb{H})$ for the intersection $\mathrm{GL}(n,\mathbb{H}) \cap \mathrm{SL}(2n,\mathbb{C})$ (thinking of $\mathbb{H}$ as $\mathbb{C}^2$). (But note that $\mathrm{SL}(n,\mathbb{C})$ is not $\mathrm{GL}(n,\mathbb{C}) \cap \mathrm{SL}(2n,\mathbb{R})$.)

(d) The 'absolute value' function on $\mathbb{R}, \mathbb{C}$ generalizes to $\mathbb{H}$, by setting

$$|x|^2 = a^2 + b^2 + c^2 + d^2$$

for $x = a + ib + jc + kd$, with the usual properties $|x_1 x_2| = |x_1||x_2|$, as well as $|\overline{x}| = |x|$ where $\overline{x} = a - ib - jc - kd$. The spaces $\mathbb{R}^n, \mathbb{C}^n, \mathbb{H}^n$ inherit norms, by putting

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

these are just the standard norms under the identification $\mathbb{C}^n = \mathbb{R}^{2n}$, $\mathbb{H}^n = \mathbb{R}^{4n}$. 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. Observe that

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

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

(e) Given $\mathcal{A}$ as above, we also have the Lie subgroups of $\mathrm{GL}(n,\mathcal{A})$, consisting of invertible matrices that are upper triangular, or upper triangular with positive diagonal entries, or upper triangular with 1's on the diagonal.

(f) The group $\mathrm{Aff}(n,\mathbb{R})$ of affine-linear transformations of $\mathbb{R}^n$ is a Lie group. It is the group of transformations of the form $x \mapsto Ax + b$, with $A \in \mathrm{GL}(n,\mathbb{R})$ and $b \in \mathbb{R}^n$. It is thus $\mathrm{GL}(n,\mathbb{R}) \times \mathbb{R}^n$ as a manifold, but not as a group. (As a group, it is a semidirect product $\mathbb{R}^n \rtimes \mathrm{GL}(n,\mathbb{R})$.) Note that $\mathrm{Aff}(1,\mathbb{R})$ is a 2-dimensional *non-abelian* Lie group.

We'll see that for all matrix Lie groups, the 'exponential charts' will always work as submanifold charts. But even without any explicit construction, we can see that these are all Lie groups, by using 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 later, once we have developed some more basics of Lie group theory. Let us finally remark that not every Lie group is realized as a matrix Lie group. For example, we will see that the universal covering space of any Lie group G is a Lie group $\widetilde{G}$; but it may be shown that

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

(or already the connected double cover of $\mathrm{SL}(2,\mathbb{R})$) is *not* isomorphic to a matrix Lie group.

1.2. **Lie algebras.** We start out with the definition:

Definition 1.5. 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.

A first example of a Lie algebra is the space

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

of square matrices, with bracket the commutator of matrices. (The notation $\mathfrak{gl}(n,\mathbb{R})$ indicates that we think of it 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,

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

and

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

are matrix Lie algebras (as one easily verifies). In contrast to the situation for Lie groups, it turns out that every finite-dimensional real Lie algebra is isomorphic to a matrix Lie algebra (*Ado's theorem*). 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 1.6. (a) Any vector space V is a Lie algebra for the zero bracket.

(b) For any associative unital algebra $\mathcal{A}$ over $\mathbb{R}$, the space of matrices with entries in $\mathcal{A}$, $\mathfrak{gl}(n,\mathcal{A}) = \mathrm{Mat}_n(\mathcal{A})$, is a Lie algebra, with bracket the commutator. In particular, we have Lie algebras

$$\mathfrak{gl}(n,\mathbb{R}),\ \mathfrak{gl}(n,\mathbb{C}),\ \mathfrak{gl}(n,\mathbb{H}).$$

(c) If $\mathcal{A}$ is commutative, then the subspace $\mathfrak{sl}(n,\mathcal{A}) \subseteq \mathfrak{gl}(n,\mathcal{A})$ of matrices of trace 0 is a Lie subalgebra. In particular,

$$\mathfrak{sl}(n,\mathbb{R}),\ \mathfrak{sl}(n,\mathbb{C})$$

are defined. The space of trace-free matrices in $\mathfrak{gl}(n,\mathbb{H})$ is *not* a Lie subalgebra; however, one may define $\mathfrak{sl}(n,\mathbb{H})$ to be the subalgebra *generated by* trace-free matrices; equivalently, this is the space of quaternionic matrices whose trace takes values in $i\mathbb{R} + j\mathbb{R} + k\mathbb{R} \subseteq \mathbb{H}$.

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

(e) Given $\mathcal{A}$, we can also consider the Lie subalgebras of $\mathfrak{gl}(n, \mathcal{A})$ that are upper triangular, or upper triangular with real diagonal entries, or strictly upper triangular, and many more.

Exercise 1.7. 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 1.8. 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 1.9. 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 1.10. Let $\phi\colon G \to H$ be a morphism of connected Lie groups, and assume that the differential $T_e\phi\colon T_eG \to T_eH$ is bijective (resp. surjective). Show that ϕ is a covering (resp. surjective). Hint: Use Exercise 1.9.

2. The groups SU(2) and SO(3)

A great deal of Lie theory depends on a good understanding of the low-dimensional Lie groups. Let us focus, in particular, on the groups SO(3) and SU(2), and their topology.

The Lie group SO(3) consists of rotations in 3-dimensional space. Let $D \subseteq \mathbb{R}^3$ 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 obtain a real projective space. Thus

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

(at least, topologically). With a little extra effort (which we'll make below) one can make this into a diffeomorphism of manifolds. There are many nice illustrations of the fact that the rotation group has fundamental group $\mathbb{Z}_2$, known as the 'Dirac belt trick'. See for example the left two columns of `https://commons.wikimedia.org/wiki/User:JasonHise`

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, $\mathrm{SU}(2) \cong S^3$ 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,

$$\det(B) = t^2 + |u|^2 = 2||B||^2.$$

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, the corresponding action on $\mathbb{R}^3 \cong \mathfrak{su}(2)$ preserves 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. This defines

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

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. Now, any morphism of Lie groups $\phi \colon G \to G'$ has constant rank, due to the symmetry: In fact, the kernel of the differential $T\phi$ is left-invariant, as a consequence of $\phi \circ l_a = l_{\phi(a)} \circ \phi$. Hence, in our case we may conclude that ϕ must be a double covering. This exhibits $\mathrm{SU}(2) = S^3$ as the double cover of SO(3). In particular, $\mathrm{SO}(3) = S^3/\pm = \mathbb{R}P^3$.

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

$$x = a + ib + jc + kd \mapsto x = \left(\begin{array}{cc} a+ib & c+id \\ -c+id & a-ib \end{array} \right).$$

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

We have encountered another important 3-dimensional Lie group: $\mathrm{SL}(2, \mathbb{R})$. This acts naturally on $\mathbb{R}^2$, and has a subgroup $\mathrm{SO}(2)$ of rotations. It turns out that as a manifold (not as a group),

$$\mathrm{SL}(2, \mathbb{R}) \cong \mathrm{SO}(2) \times \mathbb{R}^2 = S^1 \times \mathbb{R}^2.$$

One may think of $\mathrm{SL}(2, \mathbb{R})$ as the interior of a solid 2-torus. Here is how this goes: Consider the set of non-zero 2×2-matrices of zero determinant. Under the map

$$q\colon \mathrm{Mat}_2(\mathbb{R}) - \{0\} \cong \mathbb{R}^4 - \{0\} \to S^3,$$

these map onto a 2-torus inside S^3, splitting S^3 into two solid 2-tori $M^{\pm}$, given as the images of matrices of non-negative and non-positive determinant, respectively:

$$S^3 = M^+ \cup_{T^2} M^-.$$

(This is an example of a Heegard splitting.) The group $\mathrm{SL}(2, \mathbb{R})$ maps diffeomorphically onto the interior of the first solid torus M^+. Indeed, for any matrix with $\det(A) > 0$ there is a unique $\lambda > 0$ such that $\lambda A \in \mathrm{SL}(2, \mathbb{R})$; this defines a section $\mathrm{Mat}_2(\mathbb{R}) - \{0\} \to S^2$ over the interior of M^+.

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.

(a) For $m \in M$, we define the *tangent space* T_mM to be the space of directional derivatives:

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

It is automatic that v is *local*, in the sense that $T_mM = T_mU$ for any open neighborhood U of m. A smooth map of manifolds $\phi\colon M \to M'$ defines a *tangent map*:

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

(b) For $x \in U \subseteq \mathbb{R}^n$, the space $T_xU = T_x\mathbb{R}^n$ has basis the partial derivatives $\frac{\partial}{\partial x_1}|_x, \ldots, \frac{\partial}{\partial x_n}|_x$. Hence, any coordinate chart $\phi\colon U \to \phi(U) \subseteq \mathbb{R}^n$ gives an isomorphism

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

(c) 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 collection of all maps $T_m\phi$ defines a smooth vector bundle map

$$T\phi\colon TM \to TM'.$$

(d) A *vector field* on M is a collection of tangent vectors $X_m \in T_mM$ depending smoothly on m, in the sense that $\forall f \in C^\infty(M)$ the map $m \mapsto X_m(f)$ is smooth. The collection of all these tangent vectors defines 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 $X|_U(f|_U) = (X(f))|_U$. 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.

(e) If $\gamma\colon J \to M$, $J \subseteq \mathbb{R}$ is a smooth curve we obtain tangent vectors to the curve,

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

(For example, if $x \in U \subseteq \mathbb{R}^n$, the tangent vector corresponding to $a \in \mathbb{R}^n \cong T_xU$ is represented by the curve $x + ta$.) A curve $\gamma(t),\ t \in J \subseteq \mathbb{R}$ is called an *integral curve* of $X \in \mathfrak{X}(M)$ if for all $t \in J$,

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

In local coordinates, this is an ODE $\frac{dx_i}{dt} = a_i(x(t))$. 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 $\gamma(0) = m$.

(f) 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 smooth map, called the *flow of X*

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

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 [2][3]

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

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

(g) 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

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

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

$$[X, Y] = 0 \quad \Leftrightarrow \quad \Phi_t^X \circ \Phi_s^Y = \Phi_s^Y \circ \Phi_t^X.$$

(h) In general, smooth maps $\phi\colon M \to N$ of manifolds do not induce maps between their spaces of vector fields (unless ϕ is a diffeomorphism). Instead, one has the notion of *related vector fields* $X \in \mathfrak{X}(M),\ Y \in \mathfrak{X}(N)$ where

$$X \sim_\phi Y \quad \Leftrightarrow \quad \forall m\colon Y_{\phi(m)} = T_m\phi(X_m) \quad \Leftrightarrow \quad X \circ \phi^* = \phi^* \circ Y$$

From the definitions, one checks

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

---

[2] For $\phi\colon M \to N$ we denote by $\phi^*\colon C^\infty(N) \to C^\infty(M)$ the pullback.

[3] The minus sign is convention. 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 $\Phi_t^{-1} = \Phi_{-t}$.

3.2. **The Lie algebra of a Lie group.** Let G be a Lie group, and TG its tangent bundle. Denote by

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

the tangent space to the group unit. For all $a \in G$, the left translation

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

and the right translation

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

are smooth maps. Their differentials at g define isomorphisms of vector spaces $T_g L_a \colon T_g G \to T_{ag} G$; in particular

$$T_e L_a \colon \mathfrak{g} \to T_a G.$$

Taken together, they define a vector bundle isomorphism

$$G \times \mathfrak{g} \to TG,\ (g, \xi) \mapsto (T_e L_g)(\xi)$$

called *left trivialization*. The fact that this is smooth follows because it is the restriction of $T\,\mathrm{Mult} \colon TG \times TG \to TG$ to $G \times \mathfrak{g} \subseteq TG \times TG$, and hence is smooth. Using right translations instead, we get another vector bundle isomorphism

$$G \times \mathfrak{g} \to TG,\ (g, \xi) \mapsto (T_e R_g)(\xi)$$

called *right trivialization*.

Definition 3.1. A vector field $X \in \mathfrak{X}(G)$ is called *left-invariant* if it has the property

$$X \sim_{L_a} X$$

for all $a \in G$, i.e. if it commutes with the pullbacks $(L_a)^*$. Right-invariant vector fields are defined similarly.

The space $\mathfrak{X}^L(G)$ of left-invariant vector fields is thus a Lie subalgebra of $\mathfrak{X}(G)$. Similarly the space $\mathfrak{X}^R(G)$ of right-invariant vector fields is a Lie subalgebra. In terms of left trivialization of TG, the left-invariant vector fields are the constant sections of $G \times \mathfrak{g}$. In particular, we see that both maps

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

are isomorphisms of vector spaces. For $\xi \in \mathfrak{g}$, we denote by $\xi^L \in \mathfrak{X}^L(G)$ the unique left-invariant vector field such that $\xi^L|_e = \xi$. Similarly, ξ^R denotes the unique right-invariant vector field such that $\xi^R|_e = \xi$.

Definition 3.2. The Lie algebra of a Lie group G is the vector space $\mathfrak{g} = T_e G$, equipped with the unique Lie bracket such that the map $\mathfrak{X}(G)^L \to \mathfrak{g},\ X \mapsto X_e$ is an isomorphism of Lie algebras.

So, by definition, $[\xi, \eta]^L = [\xi^L, \eta^L]$. Of course, we could also use right-invariant vector fields to define a Lie algebra structure; it turns out (we will show this below) that the resulting bracket is obtained simply by a sign change.

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

Theorem 3.3. *For any morphism of Lie groups $\phi\colon G \to G'$, the tangent map $T_e\phi\colon \mathfrak{g} \to \mathfrak{g}'$ is a morphism of Lie algebras.*

Proof. Given $\xi \in \mathfrak{g}$, let $\xi' = T_e\phi(\xi) \in \mathfrak{g}'$. 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 $(T_eL_{\phi(a)})\xi' = (T_a\phi)(T_eL_a(\xi))$ hence $(\xi')^L_{\phi(a)} = (T_a\phi)(\xi^L_a)$. That is,

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

Hence, given $\xi_1, \xi_2 \in \mathfrak{g}$ we have

$$[\xi_1, \xi_2]^L = [\xi_1^L, \xi_2^L] \sim_\phi [\xi_1'^L, \xi_2'^L] = [\xi_1', \xi_2']^L.$$

In particiular, $T_e\phi[\xi_1, \xi_2] = [\xi_1', \xi_2']$. It follows that $T_e\phi$ is a Lie algebra morphism. $\square$

Remark 3.4. Two special cases are worth pointing out.

(a) A *representation of a Lie group* G on a finite-dimensional (real) vector space V is a Lie group morphism

$$\pi\colon G \to \mathrm{GL}(V).$$

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 $T_e\pi$ of any Lie group representation π 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

$$T_e\phi\colon \mathfrak{g} \to \mathfrak{g}$$

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.

3.3. **Properties of left-invariant and right-invariant vector fields.** A *1-parameter subgroup* of a Lie group G is a smooth curve $\gamma\colon \mathbb{R} \to G$ which is a group homomorphism from $\mathbb{R}$ (as an additive Lie group) to G.

Theorem 3.5. *The left-invariant vector fields ξ^L are complete, hence it defines a flow on G given by a 1-parameter group of diffeomorphisms. The unique integral curve $\gamma^\xi(t)$ of ξ^L with initial condition $\gamma^\xi(0) = e$ is a 1-parameter subgroup, and the flow of ξ^L is given by right translations:*

$$(t, g) \mapsto g\ \gamma^\xi(-t).$$

Proof. If $\gamma(t),\ t \in J \subseteq \mathbb{R}$ is any integral curve of 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. By uniqueness of integral curves of vector fields with given initial conditions, 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 $t \in \mathbb{R}$, i.e. ξ^L is complete. Let Φ^ξ_t be its flow. Thus $\Phi^\xi_t(e) = \gamma^\xi(-t)$. Since ξ^L is left-invariant, its flow commutes with left translations. Hence

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

The property $\Phi^\xi_{t_1+t_2} = \Phi^\xi_{t_1}\Phi^\xi_{t_2}$ shows that $\gamma^\xi(t_1+t_2) = \gamma^\xi(t_1)\gamma^\xi(t_2)$. □

Of course, a similar result will apply to right-invariant vector fields. Essentially the same 1-parameter subgroups will appear. To see this, note:

Lemma 3.6. *Under group inversion,* $\xi^R \sim_{\mathrm{Inv}} -\xi^L,\ \ \xi^L \sim_{\mathrm{Inv}} -\xi^R$.

Proof. The inversion map $\mathrm{Inv}\colon G \to G$ interchanges left translations and right translations:

$$\mathrm{Inv} \circ L_a = R_{a^{-1}} \circ \mathrm{Inv}\,.$$

Hence, $\xi^R = \mathrm{Inv}^* \zeta^L$ for some ζ. Since $T_e \mathrm{Inv} = -\mathrm{Id}$, we see $\zeta = -\xi$. □

As a consequence, we see that $t \mapsto \gamma(t)$ is an integral curve for ξ^R, if and only if $t \mapsto \gamma(t)^{-1}$ is an integral curve of $-\xi^L$, if and only if $t \mapsto \gamma(-t)^{-1}$ is an integral curve of ξ^L. In particular, the 1-parameter subgroup $\gamma^\xi(t)$ is an integral curve for ξ^R as well, and the flow of ξ^R is given by left translations, $(t,g) \mapsto \gamma^\xi(t)g$.

Proposition 3.7. *The left-invariant and right-invariant vector fields satisfy the bracket relations,*

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

Proof. The first relation holds by definition of the bracket on $\mathfrak{g}$. The second relation holds because the flows of ξ^L is given by right translations, the flow of ξ^R is given by left translations. Since these flows commute, teh vector fields commute. The third relation follows by applying Inv^* to the first relation, using that $\mathrm{Inv}^* \xi^L = -\xi^R$ for all ξ. □

4. The exponential map

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

Proposition 4.1. *If* $\gamma\colon \mathbb{R} \to G$ *is a 1-parameter subgroup of* G*, then* $\gamma = \gamma^\xi$ *where* $\xi = \dot\gamma(0) \in T_eG = \mathfrak{g}$*. One has*

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

The map $\mathbb{R} \times \mathfrak{g} \to G,\ (t,\xi) \mapsto \gamma^\xi(t)$ *is smooth.*

Proof. Let $\gamma(t)$ be a 1-parameter group. Then $\Phi_t(g) := g\gamma(-t)$ 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 $\Phi_{-t}(e) = \gamma(t)$ at $t = 0$: Thus $\xi = \dot{\gamma}(0)$. This shows $\gamma = \gamma^\xi$.

For fixed s, the map $t \mapsto \lambda(t) = \gamma^\xi(st)$ is a 1-parameter group with $\dot{\lambda}(0) = s\dot{\gamma}^\xi(0) = s\xi$, so $\lambda(t) = \gamma^{s\xi}(t)$. This proves $\gamma^{s\xi}(t) = \gamma^\xi(st)$. Smoothness of the map $(t,\xi) \mapsto \gamma^\xi(t)$ follows from the smooth dependence of solutions of ODE's on parameters. $\square$

Definition 4.2. The *exponential map* for the Lie group G is the smooth map defined by

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

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

Note

$$\gamma^\xi(t) = \exp(t\xi)$$

by setting $s = 1$ in $\gamma^{t\xi}(1) = \gamma^\xi(st)$. One reason for the terminology is the following

Proposition 4.3. *If* $[\xi, \eta] = 0$ *then* $\exp(\xi + \eta) = \exp(\xi)\exp(\eta)$.

Proof. The condition $[\xi, \eta] = 0$ means that ξ^L, η^L commute. Hence their flows $\Phi_t^\xi,\ \Phi_t^\eta$ commute. The map $t \mapsto \Phi_t^\xi \circ \Phi_t^\eta$ is the flow of $\xi^L + \eta^L$. Hence it coincides with $\Phi_t^{\xi+\eta}$. Applying to e (and replacing t with $-t$), this shows $\gamma^\xi(t)\gamma^\eta(t) = \gamma^{\xi+\eta}(t)$. Now put $t = 1$. $\square$

In terms of the exponential map, we may now write the flow of ξ^L as

$$(t, g) \mapsto= g\exp(-t\xi),$$

and similarly for the flow of ξ^R as

$$(t, g) \mapsto \exp(-t\xi)g.$$

That is, as operators on functions,

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

Proposition 4.4. *The exponential map is functorial with respect to Lie group homomorphisms* $\phi\colon G \to H$. *That is, we have a commutative diagram*

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

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

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

Hence $\phi(\exp(t\xi)) = \exp(tT_e\phi(\xi))$. Now put $t = 1$. □

Some of our main examples of Lie groups are matrix Lie groups. In this case, the exponential map is just the usual exponential map for matrices:

Proposition 4.5. *Let $G \subseteq \mathrm{GL}(n, \mathbb{R})$ be a matrix Lie group, and $\mathfrak{g} \subseteq \mathfrak{gl}(n, \mathbb{R})$ 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

$$\gamma(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!}\xi^n$$

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 $t = 1$. □

Remark 4.6. This result shows, in particular, that the exponentiation of matrices takes $\mathfrak{g} \subseteq \mathfrak{gl}(n, \mathbb{R}) = \mathrm{Mat}_n(\mathbb{R})$ to $G \subseteq \mathrm{GL}(n.\mathbb{R})$.

Using this result, we can also prove:

Proposition 4.7. *For a matrix Lie group $G \subseteq \mathrm{GL}(n, \mathbb{R})$, the Lie bracket on $\mathfrak{g} = T_I G$ is just the commutator of matrices.*

Proof. Since the exponential map for $G \subseteq \mathrm{GL}(n, \mathbb{R})$ is just the usual exponential map for matrices, we have, by Taylor expansions,

$$\begin{aligned}\exp(t\xi)\exp(s\eta)\exp(-t\xi)\exp(-s\eta) &= I + st(\xi\eta - \eta\xi) + \text{terms cubic or higher in } s, t \\ &= \exp\big(st(\xi\eta - \eta\xi)\big) + \text{terms cubic or higher in } s, t\end{aligned}$$

(note the coefficients of t, s, t^2, s^2 are zero). This formula relates the exponential map with the matrix commutator $\xi\eta - \eta\xi$. [4] A version of this formula holds for arbitrary Lie groups, as follows. For $g \in G$, let $\rho(g)$ be the operator on $C^\infty(G)$ given as R_g^*, thus

$$\rho(g)(f)(a) = f(ag).$$

[4]The expression on the left is an example of a 'group commutator' $ghg^{-1}h^{-1}$. Note the group commutator of g, h is trivial exactly of g, h commute; in this sense it is the group analogue to the Lie bracket.

Since the flow of ξ^L is by right translations, $t \mapsto R_{\exp(-t\xi)}$, we have [5]

$$\frac{d}{dt}\rho(\exp t\xi) = \rho(\exp t\xi) \circ \xi^L.$$

Hence,

$$\rho(\exp t\xi) = 1 + t\xi^L + \frac{t^2}{2}(\xi^L)^2 + \dots$$

as operators on $C^\infty(G)$, and consequently

$$\begin{aligned}
&\rho\big(\exp(t\xi)\exp(s\eta)\exp(-t\xi)\exp(-s\eta)\big)\\
&\qquad = \rho\big(\exp(t\xi)\big)\rho\big(\exp(s\eta)\big)\rho\big(\exp(-t\xi)\big)\rho\big(\exp(-s\eta)\big)\\
&\qquad = 1 + st(\xi^L\eta^L - \eta^L\xi^L) + \text{terms cubic or higher in } s,t\\
&\qquad = 1 + st[\xi,\eta]^L + \text{terms cubic or higer in } s,t.\\
&\qquad = \rho(\exp(st[\xi,\eta])) + \text{terms cubic or higher in } s,t.
\end{aligned}$$

We'll prove the proposition by comparing the two results. Here we may assume that $G = \mathrm{GL}(n,\mathbb{R})$ (by functoriality with respect to inclusions of subgroups), so that G is an open subset of the vector space $\mathrm{Mat}_n(\mathbb{R})$. By the matrix calculation above, the operator

$$\rho\big(\exp(t\xi)\exp(s\eta)\exp(-t\xi)\exp(-s\eta)\big) = \rho(\exp(st(\xi\eta - \eta\xi))) + \text{terms cubic or higher in } s,t.$$

But

$$\rho(\exp(st(\xi\eta - \eta\xi))) = 1 + (st)(\xi\eta - \eta\xi)^L + \dots$$

by the matrix calculation above. Comparing the coefficients of st, we see $[\xi,\eta]^L = (\xi\eta - \eta\xi)^L$, hence $[\xi,\eta] = \xi\eta - \eta\xi$. $\square$

Remark 4.8. This proves in particular, that for any matrix Lie group G, the space $\mathfrak{g} = T_I G$ is closed under commutation of matrices.

Remark 4.9. Had we defined the Lie algebra using right-invariant vector fields, we would have obtained *minus* the commutator of matrices. [6]

In the case of matrix Lie groups, we used the exponential map (of matrices) to construct local charts. This works in general, using the following fact:

Proposition 4.10. *The differential of the exponential map at the origin is*

$$T_0 \exp = \mathrm{id}\,.$$

As a consequence, there is an open neighborhood U of $0 \in \mathfrak{g}$ such that the exponential map restricts to a diffeomorphism $U \to \exp(U)$.

---

[5] If Φ_t is the flow of a vector field X, then

$$\frac{d}{dt}\Phi_{-t}^* = \frac{d}{ds}|_{s=0}\Phi_{-(t+s)}^* = \Phi_{-t}^* \circ X.$$

[6] Nonetheless, some authors use that convention.

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

$$(T_0 \exp)(\xi) = \frac{\partial}{\partial t}\Big|_{t=0} \exp(t\xi) = \xi.$$

□

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 $G \subseteq \mathrm{GL}(n, \mathbb{R})$ 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 E. Cartan's theorem on (topologically) 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 real vector space, and let

$$S(V) = (V\backslash\{0\})/\sim$$

(where $v' \sim v \Leftrightarrow v' = \lambda v,\ \lambda > 0$). For $v \in V\backslash\{0\}$, let $[v] \in S(V)$ be its equivalence class.

Lemma 5.2. *Let $v_1, v_2, \ldots \in V\backslash\{0\}$ be a sequence with $\lim_{n\to\infty} v_n = 0$, and let $v \in V\backslash\{0\}$. 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 $\lim_{n\to\infty}[v_n] = [v]$. It is convenient to introduce an inner product on V, so that we can think of $S(V)$ as the unit sphere and $[v] = v/||v||$. Let $\epsilon_n = ||v_n||/||v||$, and let $a_n \in \mathbb{N}$ be the smallest natural number satisfying $a_n\epsilon_n \geq 1$. Then $1 \leq a_n\epsilon_n < 1 + \epsilon_n$, which shows $\lim_{n\to\infty} a_n\epsilon_n = 1$. It follows that

$$\lim_{n\to\infty} a_n v_n = \lim_{n\to\infty} a_n||v_n||\ [v_n] = \lim_{n\to\infty} a_n||v_n||\ [v] = \lim_{n\to\infty} a_n\epsilon_n\ v = v. \qquad \square$$

Proof of E. Cartan's theorem. It suffices to construct a linear subspace $W \subseteq \mathfrak{g}$ and a smooth map $\phi\colon \mathfrak{g} \to G$ such that ϕ restricts to a diffeomorphism on some open neighborhood U of 0, and such that

$$\phi(U \cap W) = \phi(U) \cap H. \tag{1}$$

Indeed, these data would then give a submanifold chart around e, and by left translation one then obtains submanifold charts near arbitrary $a \in H$.

Of course, W will be T_eH, once we know that H is a submanifold. But for the time being, we only know that H is a closed subgroup. Here is the candidate for W:

$$W = \{0\} \cup \left\{\xi \in \mathfrak{g}\backslash\{0\} \,\middle|\, \exists \xi_n \neq 0\colon\ \exp(\xi_n) \in H,\quad \xi_n \to 0,\quad [\xi_n] \to [\xi]\right\}.$$

We shall prove:

(i) $\exp(W) \subseteq H$,
(ii) W is a subspace of $\mathfrak{g}$.

Proof of (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

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

Proof of (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 $u(0) = 0$. By part (i), we have that $\exp(u(t)) \in H$. Furthermore,

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

Hence

$$u(\frac{1}{n}) \to 0, \quad \exp\big(u(\frac{1}{n})\big) \in H, \quad [u(\frac{1}{n})] \to [\xi + \eta].$$

This shows $[\xi + \eta] \in W$, proving (ii).

We now prove (1). Let W' be a vector space 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 $T_0\phi$ is the identity of T_eG, there is an open neighborhood $U \subseteq \mathfrak{g}$ of 0 over which ϕ restricts to a diffeomorphism. We have that

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

We want to show that *for U sufficiently small*, this inclusion becomes an equality. Suppose not. Then, any neighborhood of $0 \in \mathfrak{g}$ contains an element $(\eta_n, \eta_n') \in W \oplus W'$ such that $\eta_n' \neq 0$ but

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

i.e. $\exp(\eta_n') \in H$. We could construct a sequence

$$\eta_n' \in W' - \{0\}\colon \quad \eta_n' \to 0,\ \exp(\eta_n') \in H.$$

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. $\square$

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 $\ker(\phi) = \phi^{-1}(e) \subseteq G$ 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 $H \subseteq G$ is a closed subgroup. Its *normalizer* $N_G(H) \subseteq G$ 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.

Remark 5.4. Given a closed subgroup of H of a Lie group G, its Lie algebra $\mathfrak{h}$ is recognized as the set of all $\xi \in \mathfrak{g}$ such that $\exp t\xi \in H$ for all $t \in \mathbb{R}$.

Theorem 5.5. *Let G, H be Lie groups, and $\phi\colon G \to H$ be a group morphism. Then ϕ is smooth if and only if it is continuous.*

Proof. Consider the graph of ϕ,

$$\mathrm{Gr}(\phi) = \{(\phi(g), g) \in H \times G |\ g \in G\}.$$

The fact that ϕ is a group morphism is equivalent to the fact that $\mathrm{Gr}(\phi)$ is a subgroup of $H \times G$. On the other hand, if ϕ is continuous then the graph is closed. By Cartan's theorem, it is thus a Lie subgroup, and in particular it is a submanifold of $H \times G$. But in turn, a map between manifolds is smooth if and only if its graph is a submanifold. (Concretely, projection to the second factor gives a diffeomorphism $\mathrm{Gr}(\phi) \cong G$, and ϕ factors as the inverse map of this diffeomorphism followed by projection to the first factor.) $\square$

As a corollary, a given topological group carries at most one smooth structure for which it is a Lie group.

6. Adjoint actions

6.1. **The adjoint representation of $\mathfrak{g}$on $\mathfrak{g}$.** Let $\mathfrak{g}$ be a Lie algebra.

Definition 6.1. A derivation of $\mathfrak{g}$ is a linear map $D\colon \mathfrak{g} \to \mathfrak{g}$ such that

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

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

Derivations are themselves a Lie algebra $\mathrm{Der}(\mathfrak{g})$, with bracket given by commutation. An example for an element of a Lie algebra derivation is the operator $\mathrm{ad}_\xi = [\xi, \cdot]$ given by bracket with ξ; the derivation property is the Jacobi identity of $\mathfrak{g}$. The map ad defines a Lie algebra morphism

$$\mathrm{ad}\colon \mathfrak{g} \to \mathrm{Der}(\mathfrak{g}).$$

The kernel of this Lie algebra morphism is the center of $\mathfrak{g}$, i.e. elements ξ with the property $[\xi, \eta] = 0$ for all η. 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 $\mathrm{Inn}(\mathfrak{g}) \subseteq \mathrm{Der}(\mathfrak{g})$ of *inner* derivations.

If $D \in \mathrm{Der}(\mathfrak{g})$ and $\xi \in \mathfrak{g}$ then

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

by the calculation

$$D(\mathrm{ad}_\xi \eta) = D([\xi, \eta]) = [D\xi, \eta] + [\xi, D\eta] = \mathrm{ad}_{D\xi} \eta + \mathrm{ad}_\xi(D\eta).$$

Hence $\mathrm{Inn}(\mathfrak{g})$ is a normal Lie subalgebra, i.e $[\mathrm{Der}(\mathfrak{g}), \mathrm{Inn}(\mathfrak{g})] \subseteq \mathrm{Inn}(\mathfrak{g})$, and the quotient space

$$\mathrm{Out}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g}) / \mathrm{Inn}(\mathfrak{g})$$

inherits a Lie algebra structure. These are the *outer automorphims.*

An *automorphism* of a Lie algebra $\mathfrak{g}$ is an invertible morphism from $\mathfrak{g}$ to itself. Thus, the group $\mathrm{Aut}(\mathfrak{g})$ of automorphisms A of a Lie algebra $\mathfrak{g}$ consists of all $A \in \mathrm{GL}(\mathfrak{g})$ such that

$$[A\xi, A\eta] = A[\xi, \eta]$$

for all $\xi, \eta \in \mathfrak{g}$, It is closed in the group $\mathrm{GL}(\mathfrak{g})$ of vector space automorphisms, hence, by Cartan's theorem, it is aLie group.

Proposition 6.2. *The Lie algebra of the group* $\mathrm{Aut}(\mathfrak{g})$ *of automorphisms is the Lie algebra* $\mathrm{Der}(\mathfrak{g})$.

Proof. The Lie algebra $\mathfrak{aut}(\mathfrak{g})$ consists of all $D \in \mathfrak{gl}(\mathfrak{g})$ with the property that $\exp(tD) \in \mathrm{Aut}(\mathfrak{g})$ for all $t \in \mathbb{R}$. Taking the t-derivative of the defining condition

$$\exp(tD)[\xi, \eta] = [\exp(tD)\xi, \exp(tD)\eta],$$

we obtain the derivation property, showing $D \in \mathrm{Der}(\mathfrak{g})$. Conversely, if $D \in \mathrm{Der}(\mathfrak{g})$ 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. $\square$

Exercise 6.3. Using similar arguments, verify that the Lie algebras 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 action of G on G.** An automorphism of a Lie group G is an invertible morphism from G to itself. The automorphisms form a group $\mathrm{Aut}(G)$. For instance, any $a \in G$ defines an 'inner' automorphism $\mathrm{Ad}_a \in \mathrm{Aut}(G)$ by conjugation:

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

Note also that $\mathrm{Ad}_{a_1 a_2} = \mathrm{Ad}_{a_1} \mathrm{Ad}_{a_2}$, 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 6.4. 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 $n \geq 3$.

Example 6.5. The group $G = \mathrm{SO}(n)$ has automorphisms given by conjugation with matrices in $B \in \mathrm{O}(n)$. More generally, if G is the identity component of a disconnected Lie group G', then any automorphism ϕ' of G' restricts to an automorphism ϕ of G, but ϕ need not be inner even if ϕ' is.

6.3. **The adjoint action of G on $\mathfrak{g}$.** The differential of the automorphism $\mathrm{Ad}_a\colon G \to G$ is a Lie algebra automorphism, denoted by the same letter: $\mathrm{Ad}_a = T_e \mathrm{Ad}_a\colon \mathfrak{g} \to \mathfrak{g}$. The resulting map

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

is called the *adjoint representation of G*.

Recall that for any Lie group morphism $\phi\colon G \to H$, the exponential map satisfies $\phi \circ \exp = \exp \circ T_e\phi$. Applying this to $H = G$ and $\phi = \mathrm{Ad}_a$ we obtain

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

Remark 6.6. If $G \subseteq \mathrm{GL}(n, \mathbb{R})$ 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.

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

Theorem 6.7. *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

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

There is a shortcut for this 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 use the representation $\rho(g) = R_g^*$ on $C^\infty(G)$, and the identity $\xi^L = \frac{\partial}{\partial s}|_{s=0}\rho(\exp(s\xi))$. We compute,

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

Here the second line used the identity $\exp(\mathrm{Ad}_a\,\eta) = \mathrm{Ad}_a\exp(\eta)$, for $a = \exp(t\xi)$ and η replaced with $s\eta$. This proves the first part. The second part is the commutativity of the diagram

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

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

Remark 6.8. 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

7.1. **Computation of $T_\xi \exp$.** We had seen that $T_0 \exp = \mathrm{id}$. More generally, one can derive a formula for the differential of the exponential map at arbitrary points $\xi \in \mathfrak{g}$. Using the identification $T_\xi \mathfrak{g} \cong \mathfrak{g}$ (since $\mathfrak{g}$ is a vector space), and using left translation to ove $T_{\exp \xi} G$ back to $T_e G = \mathfrak{g}$, this is given by 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 map*

$$T_\xi \exp \colon \mathfrak{g} = T_\xi \mathfrak{g} \to T_{\exp \xi} G$$

given by the formula,

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

Here the operator on the right hand side is defined to be the result of substituting ad_ξ for z, in the entire holomorphic function

$$\frac{1 - e^{-z}}{z} = \int_0^1 \mathrm{d}s \;\; \exp(-sz).$$

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

$$T_\xi (L_{\exp(-\xi)} \circ \exp)(\eta) = \int_0^1 \mathrm{d}s \; (\exp(-s \operatorname{ad}_\xi)\eta),$$

an equality of tangent vectors at e. In terms of the action on functions $f \in C^\infty(G)$, the left hand side is

$$\begin{aligned} T_\xi (L_{\exp(-\xi)} \circ \exp)(\eta)(f) &= \frac{\partial}{\partial t}\Big|_{t=0} f\big((L_{\exp(-\xi)} \circ \exp)(\xi + t\eta)\big) \\ &= \frac{\partial}{\partial t}\Big|_{t=0} f\big(\exp(-\xi) \; \exp(\xi + t\eta)\big) \end{aligned}$$

Letting $\rho(g) = R_g^*$ as before, we compute (as operators on functions):

$$\begin{aligned} \rho\big(\exp(-\xi)\exp(\xi + t\eta)\big) - \rho(e) &= \int_0^1 \mathrm{d}s \; \frac{\partial}{\partial s}\Big(\rho\big(\exp(-s\xi)\big)\rho\big(\exp(s(\xi + t\eta))\big)\Big) \\ &= \int_0^1 \mathrm{d}s \; \rho(\exp(-s\xi))(t\eta)^L \rho(\exp(s(\xi + t\eta)). \end{aligned}$$

Here we used the identities $\frac{\partial}{\partial s}\rho(\exp(s\zeta)) = \rho(\exp(s\zeta)) \circ \zeta^L = \zeta^L \circ \rho(\exp(s\zeta))$ for all $\zeta \in \mathfrak{g}$. Taking the t-derivative at $t = 0$, this gives

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

Applying both sides of this equation to f, and evaluating at e, we obtain

$$\frac{\partial}{\partial t}\Big|_{t=0} f\big(\exp(-\xi)\ \exp(\xi + t\eta)\big) = \Big(\int_0^1 \mathrm{d}s\ \ \exp(-s\,\mathrm{ad}_\xi)\eta\Big)(f)$$

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. $T_\xi \exp$ 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 $e^\lambda = 1$. □

7.2. **Application: Left-invariant and right-invariant vector fields in exponential coordinates.** For $\eta \in \mathfrak{g}$, we have the left-invariant vector field η^L. Letting $U \subseteq \mathfrak{g}$ be the open neighborhood of 0 on which exp is a local diffeomorphism, we may consider

$$(\exp|_U)^*\eta^L \in \mathfrak{X}(U).$$

Since U is an open subset of the vector space $\mathfrak{g}$, we may regard this vector field as a function $U \to \mathfrak{g}$. What is this function?

By definition,

$$(T_\xi \exp)\big((\exp|_U)^*\eta^L\big)|_\xi = \eta^L|_{\exp\xi} = (T_e L_{\exp\xi})(\eta).$$

That is,

$$\big((\exp|_U)^*\eta^L\big)|_\xi = (T_\xi \exp)^{-1} \circ (T_e L_{\exp\xi})(\eta) = \frac{\mathrm{ad}_\xi}{1 - \exp(-\,\mathrm{ad}_\xi)}\ \eta.$$

This shows

Proposition 7.3. *The pullback of η^L under the exponential map is the vector field on $U \subseteq \mathfrak{g}$, given by the $\mathfrak{g}$-valued function*

$$\xi \mapsto \frac{\mathrm{ad}_\xi}{1 - \exp(-\,\mathrm{ad}_\xi)}\ \eta.$$

Similarly, the right-invariant vector field is described by the $\mathfrak{g}$-valued function

$$\xi \mapsto \frac{\mathrm{ad}_\xi}{\exp(\mathrm{ad}_\xi) - 1}\ \eta.$$

Recall that the function $z \mapsto \frac{z}{e^z-1}$ is the generating function for the Bernoulli numbers

$$\frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n = 1 - \frac{1}{2} z + \frac{1}{2!}\frac{1}{6} z^2 - \frac{1}{4!}\frac{1}{30} z^4 + \dots$$

and that $\frac{z}{1-e^{-z}} = \frac{z}{e^z-1} + z$.

7.3. **Application: The Baker-Campbell Hausdorff formula.** As another 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 $t \in [0, 1]$-dependence, and let f (as a function of t, ξ, η, for ξ, η sufficiently small) be defined by

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

We have, on the one hand,

$$(T_e L_{\exp(f)})^{-1} \frac{\partial}{\partial t} \exp(f) = (T_e L_{\exp(t\eta)})^{-1} \frac{\partial}{\partial t} \exp(t\eta) = \eta.$$

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

$$(T_e L_{\exp(f)})^{-1} \frac{\partial}{\partial t} \exp(f) = (T_e L_{\exp(f)})^{-1} (T_f \exp)(\frac{\partial f}{\partial t}) = \frac{1 - e^{-\operatorname{ad}_f}}{\operatorname{ad}_f}(\frac{\partial f}{\partial t}).$$

Hence

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

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

$$\chi(w) = \frac{\log(w)}{1 - w^{-1}},$$

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

$$\frac{\partial f}{\partial t} = \chi(e^{\operatorname{ad}_\xi} e^{t \operatorname{ad}_\eta})\eta.$$

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

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

To obtain the BCH formula, we use the series expansion of $\chi(w)$ around 1:

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

which is easily obtained from the usual power series of log. Putting $w = e^{\mathrm{ad}_\xi} e^{t\,\mathrm{ad}_\eta}$, and writing

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

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

Theorem 7.4 (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,

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

Exercise 7.5. 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, $g_1.(g_2.m) = (g_1 g_2).m$ and $e.m = m$. A smooth 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:
$$\begin{tikzcd} G \times M_1 \arrow[r] \arrow[d, "\mathrm{id}\times\Phi"] & M_1 \arrow[d, "\Phi"] \\ G \times M_2 \arrow[r] & M_2 \end{tikzcd}$$
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) There are three natural actions of G on itself:
- left multiplication, $\mathcal{A}_g = L_g$,
- right multiplication, $\mathcal{A}_g = R_{g^{-1}}$,
- conjugation (adjoint action), $\mathcal{A}_g = \mathrm{Ad}_g = L_g \circ R_{g^{-1}}$.

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 $G \subseteq G \times G$.
(c) Any G-representation $G \to \mathrm{End}(V)$ defines a G-action on V, viewed as a manifold.
(d) For any closed subgroup $H \subseteq G$, 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. 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) Some examples of actions of the orthogonal group $\mathrm{O}(n)$:
- The defining action on $\mathbb{R}^n$,
- the action on the unit sphere $S^{n-1} \subseteq \mathbb{R}^n$,
- the action on projective space $\mathbb{R}P(n-1) = S^{n-1}/\sim$,
- the action on the Grassmann manifold $\mathrm{Gr}_{\mathbb{R}}(k,n)$ of k-planes in $\mathbb{R}^n$,
- the action on the flag manifold $\mathrm{Fl}(n) \subseteq \mathrm{Gr}_{\mathbb{R}}(1,n) \times \cdots \mathrm{Gr}_{\mathbb{R}}(n-1,n)$ (consisting of sequences of subspaces $V_1 \subseteq \cdots V_{n-1} \subseteq \mathbb{R}^n$ with $\dim V_i = i$), and various types of 'partial' flag manifolds.

Except for the first example, these are all of the form G/H. (E.g, for $\mathrm{Gr}(k,n)$ one takes H to be the subgroup preserving $\mathbb{R}^k \subseteq \mathbb{R}^n$.)

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 $X \in \mathfrak{X}(M)$ defines an action of the Abelian Lie algebra $\mathbb{R}$, by $\mathbb{R} \to \mathfrak{X}(M),\ \lambda \mapsto \lambda X$.

(b) Any Lie algebra representation $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ may be viewed as a Lie algebra action $\mathfrak{g} \to \mathfrak{X}(V)$, where for $f \in C^\infty(V)$,

$$(\xi_V f)(v) = \frac{\partial}{\partial t}\Big|_{t=0} f(v - t\phi(\xi)v)$$

Using a basis e_a of V to identify $V = \mathbb{R}^n$, writing $v = \sum_a x^a\, e_a$, 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 linearity property corresponds to a linear $\mathfrak{g}$-representation.

(c) For any Lie group G, we have actions of its Lie algebra $\mathfrak{g}$ by

$$\mathcal{A}_\xi = \xi^L, \quad \mathcal{A}_\xi = -\xi^R, \quad \mathcal{A}_\xi = \xi^L - \xi^R.$$

(d) Given a closed subgroup $H \subseteq G$, 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,

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

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 $g \to \mathcal{A}_g$ on M define an action of the Lie algebra $\mathfrak{g}$ on M, given by $\xi \mapsto \mathcal{A}_\xi = \xi_M$.*

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

$$0 = [(\xi_1)_{G\times M}, (\xi_2)_{G\times M}] - [\xi_1, \xi_2]_{G\times M} \sim_\Phi [(\xi_1)_M, (\xi_2)_M] - [\xi_1, \xi_2]_M.$$

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 $\mathfrak{R} \subseteq \mathfrak{X}(M)$ of the space of vector fields, such that at any point $m \in M$, the subspace

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

is of dimension k. 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 $\mathfrak{R} \subseteq \mathfrak{X}(M)$ 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 $G \times M$, given by

$$g \cdot (a, m) = (ag^{-1}, g \cdot m).$$

The orbits of this action are exactly the fibers $\Phi^{-1}(m)$ of the action map $\Phi \colon G \times M \to M$, $(a, m) \mapsto a \cdot m$. We may think of these orbits as the leaves of a foliation, $\mathcal{L}_m = \Phi^{-1}(m)$ where

$$\mathcal{L}_m = \{(g^{-1}, g \cdot m) | \, g \in G\}.$$

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

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

Suppose now that we are given a $\mathfrak{g}$-action on M. Consider the diagonal $\mathfrak{g}$ action on $\widehat{M} = G \times M$,

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

Note that this vector field is complete, for any given ξ, since it is the sum of commuting vector fields, both of which are complete. Its flow is given by

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

Since the maps $\mathfrak{g} \to T_{(a,m)}(G \times M)$ are all injective, the generating vector fields define an integrable $\dim G$-dimensional distribution $\mathfrak{R} \subseteq \mathfrak{X}(G \times M)$. 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

$$\pi_m \colon \; \mathcal{L}_m \to G.$$

Using that any $g \in G$ can be written in the form $g = \exp(\xi_r) \cdots \exp(\xi_1)$ with $\xi_i \in \mathfrak{g}$,so $g^{-1} = R_{\exp(-\xi_r)} \cdots R_{\exp(-\xi_1)} \cdot e$, we see that π_m is *surjective* – the curve

$$\widehat{\Phi}_t^{\xi_r} \circ \cdots \circ \widehat{\Phi}_t^{\xi_1}(e, m)$$

connects (e, m) to a point of the form (g^{-1}, m'). A similar argument also shows that π_m is a covering map onto G. (Points near g^{-1} can be written as $R_{\exp(-\xi)}(g^{-1})$, and this lifts to $\Phi_1^{\xi}(g^{-1}, m')$.) Since G is simply connected by assumption, we conclude that $\pi_m \colon \mathcal{L}_m \to G$ is a diffeomorphism.

We now define the action map by $\mathcal{A}_g(m) = \mathrm{pr}_2(\pi_m^{-1}(g^{-1}))$. 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 $\mathcal{A}_{gh} = \mathcal{A}_g \circ \mathcal{A}_h$. $\square$

Remark 8.12. 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 8.13. *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. $\square$

Corollary 8.14. *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 $G \to \mathrm{GL}(V)$. $\square$

Definition 8.15. A *Lie subgroup* of a Lie group G is a subgroup $H \subseteq G$, equipped with a Lie group structure such that the inclusion is a morphism of Lie groups.

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 8.16. *Let G be a Lie group, with Lie algebra $\mathfrak{g}$. For any Lie subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ 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 $H \subseteq G$ 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. $\square$

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.

Remark 8.17. For any Lie algebra $\mathfrak{g}$ we have the adjoint representation $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$; its kernel is the center of $\mathfrak{g}$. So, for Lie algebras with trivial center the theorem of Ado is immediate.

Given such a presentation $\mathfrak{g} \subseteq \mathfrak{gl}(n, \mathbb{R})$, the lemma gives a Lie subgroup $G \subseteq \mathrm{GL}(n, \mathbb{R})$ integrating $\mathfrak{g}$. Replacing G with its universal covering, this proves (assuming Ado's theorem):

Theorem 8.18 (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 Lie's third 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 universal covering space $\widetilde{X}$ as the set of 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 space $\widetilde{X}$ has a natural topology in such a way that the map

$$p\colon \widetilde{X} \to X, \quad [\gamma] \mapsto \gamma(1)$$

is a covering map. The fiber $p^{-1}(x_0) \subseteq \widetilde{M}$ has a group structure is given by the concatenation of paths

$$(\lambda_1 * \lambda_2)(t) = \begin{cases} \lambda_1(2t) & 0 \leq t \leq \frac{1}{2}, \\ \lambda_2(2t-1) & \frac{1}{2} \leq t \leq 1 \end{cases},$$

i.e. $[\lambda_1][\lambda_2] = [\lambda_1 * \lambda_2]$ (one shows that this is well-defined). This group is denoted $\pi_1(X; x_0)$, and is called the fundamental group of X with respect to the base point x_0. The fundamental group of the covering space itself is trivial: A loop $\widetilde{\lambda}\colon [0,1] \to \widetilde{X}$ is represented by a 2-map $\sigma\colon [0,1] \times [0,1] \to X$ such that $\widetilde{\lambda}(s) = [\sigma(s,\cdot)]$, with $\sigma(s,0) = \sigma(0,t) = \sigma(1,t) = x_0$, such a map can be deformed into the constant map as indicated by the following picture. [PICTURE]

The fundamental group acts on the covering space $\widetilde{X}$ by so-called *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 $\widetilde{\Phi}[\gamma] = [\Phi \circ \gamma]$, with $\widetilde{\Psi \circ \Phi} = \widetilde{\Psi} \circ \widetilde{\Phi}$ under composition of two such maps. It restricts to a group morphism $\pi_1(X; x_0) \to \pi_1(Y; y_0)$; the map $\widetilde{\Phi}$ is equivariant with respect to the group morphism.

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. We are interested in the case of connected Lie groups G. In this case, the natural choice of base point is the group unit $x_0 = e$, and we'll write simply $\pi_1(G) = \pi_1(G; e)$. We have:

Theorem 9.1. *The universal covering space $\widetilde{G}$ of a connected Lie group G is again a Lie group, in such a way that the covering map $p\colon \widetilde{G} \to G$ is a Lie group morphism. The inclusion $\pi_1(G) = p^{-1}(\{e\}) \hookrightarrow \widetilde{G}$ is a group morphism, with image contained in 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 follows that these define a group structure on $\widetilde{G}$, in such a way that the quotient map $p\colon \widetilde{G} \to G$ is a Lie group morphism. The kernel $p^{-1}(e) \subseteq \widetilde{G}$ is a normal subgroup of $\widetilde{G}$. We claim that this group structure of $\pi^{-1}(e)$ (given by $[\lambda_1] \cdot [\lambda_2] = [\lambda_1 \lambda_2]$, using pointwise multiplication) coincides with the group structure of $\pi_1(G)$, given by concatenation. In other words, we claim that the paths

$$t \mapsto \lambda_1(t)\lambda_2(t), \quad t \mapsto (\lambda_1 * \lambda_2)(t)$$

are homotopic. To this end, let us extend the domain of definition of any loop $\lambda\colon [0,1] \to G$ to all of $\mathbb{R}$, by letting $\lambda(t) = e$ for $t \notin [0,1]$. With this convention, we have that

$$(\lambda_1 * \lambda_2)(t) = \lambda_1(2t)\lambda_2(2t-1)$$

for all $t \in \mathbb{R}$, and the desired homotopy is given by

$$\lambda_1((1+s)t)\lambda_2((1+s)t - s), \quad 0 \leq s \leq 1.$$

Hence, $\pi_1(G)$ is a discrete *normal* subgroup of $\widetilde{G}$. But if G is connected, then $\widetilde{G}$ is connected, and so the adjoint action must be trivial on $\pi_1(G)$ (since $\pi_1(G)$ is discrete). $\square$

Example 9.2. The universal covering group of the circle group $G = \mathrm{U}(1)$ is the additive group $\mathbb{R}$.

Example 9.3. $\mathrm{SU}(2)$ 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 $\mathbb{Z}_2$.

Example 9.4. 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 $\mathrm{Spin}(6) = \mathrm{SU}(4)$. 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 $\Gamma \subseteq \pi_1(G)$ 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. Basic properties of 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.

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

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

Its differential is given by

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

by the calculation

$$\frac{\partial}{\partial t}\Big|_{t=0} \det_{\mathfrak{g}}(\mathrm{Ad}_{\exp(t\xi)}) = \frac{\partial}{\partial t}\Big|_{t=0} \det_{\mathfrak{g}}(\exp(t\,\mathrm{ad}_\xi)) = \frac{\partial}{\partial t}\Big|_{t=0} \exp(t\,\mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi)) = \mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi).$$

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. A Lie group G whose modular function is trivial is called *unimodular*. In this case, the Lie algebra $\mathfrak{g}$ is unimodular, i.e. the infinitesimal moduluar character $\xi \mapsto \mathrm{tr}(\mathrm{ad}_\xi)$ is trivial. The converse holds if G is connected.

Proposition 10.1. *Compact Lie groups are unimodular.*

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 the only compact subgroups of $\mathbb{R}^\times$ are $\{-1, 1\}$ and $\{1\}$. $\square$

Remark 10.2. 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 *perfect* 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 $\chi(g) = t$.

10.2. **Volume forms and densities.** The modular function has a geometric interpretation in terms of volume forms and densities.

Definition 10.3. 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 $A \in \mathrm{GL}(E)$ and all $v_1, \ldots, v_n \in E$. 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 nowhere 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* [7]. 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

[7]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.

so is $f\mathsf{m}$. In fact, the choice of a fixed positive smooth measure m on M trivializes the density bundle, hence any density on M is then of the form $f\mathsf{m}$ with f a function. On $\mathbb{R}^n$, we have the standard 'Lebesgue measure' $|dx|$ defined by the trivialization of $T\mathbb{R}^n$; general densities on $\mathbb{R}^n$ are of the form $f|\mathrm{d}x|$ with f a function.

If $\Phi\colon M' \to M$ is a local diffeomorphism, then the pullback density $\Phi^*\mathsf{m}$ on M' is defined, similarly for volume forms.

There is an integration map, which is a linear functional on the space of densities of compact support,

$$\mathsf{m} \mapsto \int_M \mathsf{m}.$$

It is the unique linear functional such that if m supported in a chart domain $U \subseteq M$, with coordinate map $\phi\colon U \to \mathbb{R}^n$, then

$$\int_M \mathsf{m} = \int_{\mathbb{R}^n} f(x)|\mathrm{d}x|$$

where $f \in C^\infty(\mathbb{R}^n)$ is the function defined by $\mathsf{m} = \phi^*(f|\mathrm{d}x|)$. (The integration of top degree forms over oriented manifolds may be seen as a special case of the integration of densities.)

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, $M = G$. 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, the space of left-invariant top degree forms on G is 1-dimensional, and similarly for the space of left-invariant densities.

Lemma 10.4. *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

$$R_a^*\mathsf{m} = |\det(\mathrm{Ad}_{a^{-1}})|\mathsf{m}$$

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

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

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

Corollary 10.5. *On a unimodular Lie group, any left-invariant density is also right invariant. If G is unimodular and connected, the same holds true for volume forms.*

In particular, this result applies to compact Lie groups: Every left-invariant density is also right-invariant. One can normalize the left-invariant density such that $\int_G \mathsf{m} = 1$. A nonzero left-invariant measure on a Lie group G (not necessarily normalized) is often denoted $|\mathrm{d}g|$; it is referred to as a *Haar measure*.

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

$$B(\mathrm{Ad}_g\,\xi, \mathrm{Ad}_g\,\eta) = B(\xi, \eta) \tag{2}$$

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

$$B([\xi, \eta], \zeta) + B(\eta, [\xi, \zeta]) = 0. \tag{3}$$

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 (2). The converse is true if G is connected.

10.3. **Decomposition of the Lie algebra of a compact Lie group.** As an application, we obtain the following decomposition of the Lie algebra of compact Lie groups. We will use the following terminology.

Definition 10.6. (a) An *ideal* in a Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h}$ with $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$.
(b) A Lie algebra is called *simple* if it is non-abelian and does not contain non-trivial ideals.
(c) A Lie algebra is called *semi-simple* if it is a direct sum of simple ideals.

An ideal is the same thing as an invariant subspace for the adjoint representation of $\mathfrak{g}$ on itself. Note that ideals is the Lie algebra counterpart of normal subgroups; in particular, if $\mathfrak{g} = \mathrm{Lie}(G)$ then the Lie algebra of any normal subgroup is an ideal. The center of $\mathfrak{g}$ is an ideal, and the kernel of any Lie algebra morphism is an ideal. 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.

Any Lie algebra $\mathfrak{g}$ has a distinguished ideal, the so-called derived subalgebra $[\mathfrak{g}, \mathfrak{g}]$ (spanned by all Lie brackets of elements). The derived subalgebra is trivial if and only if $\mathfrak{g}$ is abelian (the bracket is zero); hence, simple Lie algebras, and more generally semi-simple ones, satisfy

$$\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}].$$

That is, semi-simple Lie algebras are *perfect*. The converse is not true: a counterexamples is the Lie algebra of 4×4-matrices of block form

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

where A, B, C are 2×2-matrices and $\mathrm{tr}(A) = \mathrm{tr}(C) = 0$.

Theorem 10.7. *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 $\mathfrak{h} \subseteq \mathfrak{g}$ is again an ideal. Indeed, $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$ implies

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

hence $[\mathfrak{g}, \mathfrak{h}^\perp] \subseteq \mathfrak{h}^\perp$. As a consequence, $\mathfrak{g}$ has an orthogonal decomposition into ideals, none of which contains a proper ideal. Hence, these summands are either simple, or 1-dimensional and abelian. Let $\mathfrak{g}_1, \ldots, \mathfrak{g}_r$ be the simple ideals, and $\mathfrak{z}$ the sum of the abelian ideals. Then $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$ is a direct sum of Lie algebras. and in particular $\mathfrak{z}$ is the center of $\mathfrak{g}$ and $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$ is the semisimple part.

For the uniqueness of the decomposition, suppose that $\mathfrak{h} \subseteq [\mathfrak{g}, \mathfrak{g}]$ is an ideal not containing any of the $\mathfrak{g}_i$'s. But then $[\mathfrak{g}_i, \mathfrak{h}] \subseteq \mathfrak{g}_i \cap \mathfrak{h} = 0$ for all i, which gives $[\mathfrak{g}, \mathfrak{h}] = \bigoplus_i [\mathfrak{g}_i, \mathfrak{h}] = 0$. Hence $\mathfrak{h} \subseteq \mathfrak{z}$. $\square$

Exercise 10.8. Show that for any Lie group G, the Lie algebra of the center of G is the center of the Lie algebra.

10.4. **Complete reducibility of representations.** Let G be a compact Lie group, and $\pi \colon G \to \mathrm{Aut}(V)$ a representation on a real vector space. Then V admits a G-invariant metric, obrained from an arbitrary given Euclidean metric $< \cdot, \cdot >'$ by averaging:

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

Given a G-invariant subspace $W \subseteq V$, the orthogonal complement $W^\perp$ is again invariant. It follows that every finite-dimensional real G-representation is a direct sum of irreducible ones.

Similarly, if V is a *complex* vector space and the representation is by complex automorphisms, we obtain an invariant *Hermitian* metric (complex inner product) by averaging. Given a G-invariant complex subspace W, its orthogonal complement $W^\perp$ 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.)

10.5. **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 $B_m \colon T_m M \times T_m M \to \mathbb{R}$ depending smoothly on m. A smooth curve $\gamma \colon J \to M$ (with $J \subseteq \mathbb{R}$ some interval) is called a *geodesic* if, for any $[t_0, t_1] \subseteq J$, 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 $t_* \subseteq J$. 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 10.9. *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 10.10. 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 10.11. *The exponential map of a compact, connected Lie group is surjective.*

Proof. Choose a bi-invariant Riemannian metric on G. Since G is compact, any two points in G are joined by a geodesic. (A length minimizing curve connecting the points is 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 $\exp(\xi) = g$. □

Remark 10.12. 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.

10.6. **The Killing form.**

Definition 10.13. The *Killing form*[8] 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 10.14. *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 $\mathrm{ad}_{\mathrm{Ad}_g(\xi)} = \mathrm{Ad}_g \circ \mathrm{ad}_\xi \circ \mathrm{Ad}_{g^{-1}}$. □

If $\mathfrak{g}$ is simple, then the Killing form κ must be nondegenerate (since otherwise the kernel of κ is a non-trivial ideal). Hence, also for semi-simple $\mathfrak{g}$ the Killing form is nondegenerate. An important result of E. Cartan says that the converse is true as well. (We will not use it in this course.)

[8]The Killing form is named after Wilhelm Killing (1847-1923). Killing's 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'.

Remark 10.15. It is not hard to see that if $\mathfrak{g}$ admits a non-degenerate invariant bilinear form B then $\mathfrak{g}$ is perfect: $[\mathfrak{g},\mathfrak{g}]=\mathfrak{g}$. In particular, this holds when the Killing form of $\mathfrak{g}$ is nondegenerate.

Proposition 10.16. *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 G has finite center (so that $\mathfrak{z}=0$), 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. $\xi\in\mathfrak{z}$. ☐

10.7. **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 $D=\mathrm{ad}_\xi$.

Theorem 10.17. *Suppose the Killing form of $\mathfrak{g}$ is non-degenerate (e.g., $\mathfrak{g}$ is the Lie algebra of a compact Lie group with finite center). Then any derivation of $\mathfrak{g}$ is inner. In fact,* $\mathrm{Der}(\mathfrak{g})=\mathrm{Inn}(\mathfrak{g})=\mathfrak{g}$.

Proof. Let $D\in\mathrm{Der}(\mathfrak{g})$. It defines a linear functional on $\mathfrak{g}$, given by $\eta\mapsto\mathrm{tr}(D\circ\mathrm{ad}_\eta)$. Since the Killing form is non-degenerate, this linear functional is given by some Lie algebra element ξ. That is,

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

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

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

This shows $D_0(\eta)=0$ for all η, hence $D_0=0$. This shows that every derivation of $\mathfrak{g}$ is inner. 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 $\mathrm{Aut}(\mathfrak{g})$.

Theorem 10.18. *Suppose $\mathfrak{g}$ is a finite-dimensional Lie algebra. Then the Killing form on $\mathfrak{g}$ is negative definite if and only if $\mathfrak{g}$ is the Lie algebra of a compact connected Lie group G with finite center.*

Proof. The direction $\Leftarrow$ is Proposition 10.16. For the converse, assuming that the Killing form is negative definite, let $\mathrm{Aut}(\mathfrak{g})$. Since $\mathrm{Aut}(\mathfrak{g})$ preserves the Killing form, we have

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

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

Remark 10.19. A somewhat stronger statement holds: The Lie algebra of a connected Lie group G has negative definite Killing form if and only if G is compact with finite center. This follows once we know that the universal cover $\widetilde{G}$ of a compact Lie group with finite center is again compact. Equivalently, we need to know that *for a compact connected Lie group with finite center, the fundamental group is finite.* [9] 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.

11. The maximal torus of a compact Lie group

11.1. **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.[10] 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 11.1. *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).

Proof. If $\mathfrak{g}$ is abelian, then any two left-invariant vector fields commute. Hence their flows commute, which gives

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

for all $\xi, \eta \in \mathfrak{g}$. 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 $\Gamma = \pi_1(G)$ is a discrete additive subgroup of V.

Lemma 11.2. *There are linearly independent $\gamma_1, \ldots, \gamma_k \in V$ such that*

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

[9] We may get back to this later.

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

Proof. 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, we see that any abelian Lie group is isomorphic to $\mathbb{R}^n/\mathbb{Z}^k$ for some n, k. That is:

Proposition 11.3. *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 will call

$$\Lambda = \pi_1(T) \subseteq \mathfrak{t}$$

the *integral lattice.* Thus

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

Let G be a Lie group, and $g \in G$. Then g generates an abelian subgroup

$$\{g^k |\ k \in \mathbb{Z}\}$$

of G; its closure is an abelian subgroup $H \subseteq G$. We call g a *topological generator* of G if $H = G$. Of course, this is only possible if G is abelian (but possibly disconnected).

Theorem 11.4 (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. Since T/H is a compact connected abelian Lie group, it is isomorphic to $(\mathbb{R}/\mathbb{Z})^l$ for some l.

We will show that $H \neq T$, i.e., $l > 0$, if and only if $1, u_1, \ldots, u_k \in \mathbb{R}$ are linearly *dependent* over the rationals. Note that the latter is equivalent to the existence of $a_1, \ldots, a_n$, not all zero, such that $\sum_{i=1}^k a_i u_i \in \mathbb{Z}$.

"$\Rightarrow$". If $l > 0$, 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. Its differential $T_0\phi\colon \mathbb{R}^k \to \mathbb{R}$ takes $\mathbb{Z}^k$ to $\mathbb{Z}$, hence we obtain integers $a_i = (T_0\phi)(e_i) \in \mathbb{Z}$. In terms of these integers,

$$(T_0\phi)(u) = (T_0\phi)(\sum_{i=1}^k u_i e_i) = a_i u_i.$$

But since ϕ is trivial on H, the element $t = \exp(u)$ satisfies $\phi(t) = 1$, hence $(T_0\phi)(u) \in \mathbb{Z}$. That is, $\sum_{i=1}^k a_i u_i \in \mathbb{Z}$.

"$\Leftarrow$". Conversely, given $a_i \in \mathbb{Z}$, not all zero, such that $\sum_{i=1}^k a_i u_i \in \mathbb{Z}$, define a non-trivial group homomorphism

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

by $\phi(h) = \sum_{i=1}^k a_i v_i \mod \mathbb{Z}$ for $h = exp(\sum_{i=1}^k v_i e_i)$. Then ϕ is trivial on H, but is non-trivial on T. It follows that H is a proper subgroup of T. $\square$

Remark 11.5 (Automorphisms). Given torus T, any group automorphism $\phi \in \mathrm{Aut}(T)$ induces a Lie algebra automorphism $T_0\phi \in \mathrm{Aut}(\mathfrak{t})$ 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,

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

Choose an identification $T = (\mathbb{R}/\mathbb{Z})^k$. We have

$$\mathrm{Aut}(\mathbb{Z}^k) = \mathrm{GL}(k, \mathbb{Z}) \subseteq \mathrm{Mat}_k(\mathbb{Z})$$

the group of invertible matrices A with integer coefficients whose inverse also has integer coefficients. By the formula for the inverse matrix, this is the case if and only if the determinant is ± 1:

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

The group $\mathrm{GL}(k, \mathbb{Z})$ contains the semi-direct product $(\mathbb{Z}_2)^k \rtimes S_k$, where S_k acts on $\mathbb{Z}^k$ by permutation of coordinates and $(\mathbb{Z}_2)^k$ acts by sign changes. One can show that

$$(\mathbb{Z}_2)^k \rtimes S_k \cong \mathrm{O}(k, \mathbb{Z}) = \mathrm{GL}(k, \mathbb{Z}) \cap \mathrm{O}(k),$$

the transformations preserving also the metric.

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

Theorem 11.6. *(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 $T, T' \subseteq G$ 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 11.7. The *rank* l of a compact, connected Lie group G s the dimension of a maximal torus $T \subseteq G$.

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 11.8. 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 $T' = \mathrm{SO}(2) \subseteq \mathrm{SU}(2)$. 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 $\mathrm{Ad}_a(\mathfrak{t}) = \mathfrak{t}'$.

Theorem 11.9. (Properties of maximal tori).

(a) *Every element of a Lie group is contained in some maximal torus. That is, if T is a fixed maximal torus then*

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

(b) *The intersection of all maximal tori is the center of G:*

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

(c) *If $H \subseteq G$ is a subtorus, and $g \in G$ commutes with all elements of H, then there exists a maximal torus containing H and g.*

(d) *Maximal tori are maximal abelian groups: If some $g \in G$ commutes with all elements of T then $g \in T$.*

Proof. (a) Let $g \in G$ be given. Using that fact that $\exp\colon \mathfrak{g} \to G$ is surjective, we may choose $\xi \in \mathfrak{g}$ 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.

(b) 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 $\bigcap_{a\in G}\mathrm{Ad}_a(T)\subseteq Z(G)$; the opposite inclusion is a consequence of (b) to be proved below.

(c) If $g\in H$ there is nothing to show (any maximal torus containing H will do), so assume $g\notin H$. Since g commutes with 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. Since B is compact, it can only have finitely many components; let $m\in\mathbb{N}$ be the smallest number with $g^m\in B_0$, so the components of B are 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 by part (a), the element bh is contained in some maximal torus T. Hence $B\subseteq T$.

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

□

Exercise 11.10. 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 11.11. $\dim(G/T)$ *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 $T\subseteq G$ fixes $\mathfrak{t}$, hence it restricts to a representation

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

where $\mathfrak{t}^\perp\cong\mathfrak{g}/\mathfrak{t}$ 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. □

11.3. **The Weyl group.** For any subset $S\subseteq G$ 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 $\mathrm{Ad}_g(S)\subseteq S$. Note that if $g$normalizes S then it also normalizes the closure $\overline{S}$. The normalizer $N(S)$ is a closed subgroup, hence a Lie subgroup. If H be a closed subgroup of G, then it is a is a *normal* subgroup of $N(H)$, hence the quotient $N(H)/H$ inherits a Lie 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 $T\subseteq N(T)$ on T is of course trivial, but there is a non-trivial action of the quotient $N(T)/T$.

Definition 11.12. 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 $T \subseteq N(T)$ 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 $N(T) \subseteq G$ on $\mathfrak{t}$ (since $T \subseteq N(T)$ 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 $\exp\colon \mathfrak{t} \to T$. 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. It is a general fact that if a group G acts on a manifold M, and H is a subgroup, then $N_G(H)/H$ acts on M/H. Here, we are applying this fact to the action $a \mapsto R_{a^{-1}}$ of G on itself, and te subgroup $H = T$. Explicitly, if $w \in W$ is represented by $n \in N(T)$, then
$$w.(gT) = gn^{-1}T.$$
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 11.13. 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 $G/T \cong S^2$ is the antipodal map, hence the set of maximal tori is $G/N(T) = (G/T)/W \cong \mathbb{R}P(2)$.

Example 11.14. 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 in $\mathrm{SO}(3)$ 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 11.15. *The Weyl group W of a compact, connected group G is a finite group.*

Proof. To show $N(T)_0 = T$, we only have to show the inclusion $\subseteq$. 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, $N(T)_0 \subseteq Z_G(T) = T$. $\square$

Theorem 11.16. *The action of W on T (and likewise the action on $\mathfrak{t}, \Lambda$) is faithful. That is, the map*

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

is injective.

Proof. If w acts trivially on T, then any element $g \in N(T)$ representing w lies in $Z(T) = T$, thus $w = 1$. On the other hand, w acts trivially on Λ if and only if it acts trivially on $\mathfrak{t} = \mathrm{span}_{\mathbb{R}}(\Lambda)$, if and only if it acts trivially on $T = \exp(\mathfrak{t})$. $\square$

Exercise 11.17. a) Let $\phi\colon G \to G'$ be a surjective morphism of compact connected Lie groups. Show that if $T \subseteq G$ 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: $W \cong W'$.

Remark 11.18 (The set of all maximal tori, and its tautological bundle). Let X be the set of all maximal tori of G. Over X, we have a tautological bundle of Lie groups, with fiber at x the maximal torus T_x labeled by x. Let $Q = \coprod_{x\in X} T_x$ be the disjoint union; it comes with a natural map

$$Q \to G$$

given on the fiber T_x by the obvious inclusion. The group G acts on Q, by $g \cdot (x, t) \mapsto (g \cdot x, \mathrm{Ad}_t(x))$ for $t \in T_x$; the map $Q \to G$ is G-equivariant for this action and the conjugation action on G.

Once we fix a maximal torus T, we have an identification $X = G/N(T)$, and the fiber bundle Q is realized as

$$Q = (G/T \times T)/W \to (G/T)/W = G/N(T) = X.$$

The map to G is smooth, and is a diffeomorphism on an open dense subset of Q. Note also that the bundle Q is *flat*, since we have canonical local identifications of the fibers; this also follows because it is the quotient of a trivial bundle $G/T \times T$ by the action of a discrete group.

11.4. **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 $\mathrm{Aut}(T)$.

Recall also that for the standard torus $(\mathbb{R}/\mathbb{Z})^l$, we have $\mathrm{Aut}((\mathbb{R}/\mathbb{Z})^l) \cong \mathrm{GL}(l,\mathbb{Z})$, and a maximal compact subgroup is

$$\mathrm{O}(l,\mathbb{Z}) = (\mathbb{Z}_2)^l \rtimes S_l \subseteq \mathrm{GL}(l,\mathbb{Z}).$$

Note however that the restriction of the metric of $\mathfrak{g}$ to the Lie algebra $\mathfrak{t}$ need not correspond to the standard metric on $\mathbb{R}^l$, hence W need not take values in $\mathrm{O}(l,\mathbb{Z})$, in general. Still, we expect W to be *conjugate* to a subgroup of $(\mathbb{Z}_2)^l \rtimes S_l$.

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 $t \in T$.

11.4.1. *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. In particular, $\mathrm{rank}(U(n)) = n$.

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

$$W \hookrightarrow S_n.$$

We claim that this map is an isomorphism. To see this, it suffices that all transpositions of adjacent elements $i, i+1$ are in the image, hence are realized as conjugation by some $n \in N(T)$. Indeed, let $n \in G$ be the matrix with

$$n_{i,i+1} = 1, \quad n_{i+1,i} = -1, \quad n_{jj} = 1 \text{ for } j \neq i, i+1,$$

and all other entries equal to zero. Conjugation by n preserves 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 $T \subseteq \mathrm{SU}(n)$, thus $\mathrm{rank}(\mathrm{SU}(n)) = n-1$, and the Weyl group is $W = S_n$, just as in the case of $\mathrm{U}(n)$.

Theorem 11.19. *The Weyl group of* $\mathrm{U}(n)$*, and also of* $\mathrm{SU}(n)$*, is the symmetric group* S_n*.*

11.4.2. *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 $T \subseteq \mathrm{SO}(2m)$. 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, and $\mathrm{rank}(\mathrm{SO}(2m)) = m$.

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 $\Gamma_m \subseteq (\mathbb{Z}_2)^m$ be the kernel of the product map $(\mathbb{Z}_2)^m \to \mathbb{Z}_2$, corresponding to an even number of sign changes.

Theorem 11.20. *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

$$g_{ij} = -g_{ji} = I, \; g_{kk} = I \quad \text{for} \quad k \neq i, j$$

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 $S_m \subseteq W$. Next, observe that the matrix

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

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 $\Gamma_n \subseteq W$.

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

$$gt(\theta_1, \theta_2, \ldots, \theta_m) = t(-\theta_1, \theta_2, \ldots, \theta_m)g.$$

As above, writing g in block form, we obtain the condition

$$R(\theta_i)g_{ij} = g_{ij}R(\theta_j)$$

for $j \geq 2$, but $R(\theta_i)g_{i1} = g_{i1}R(\theta_1)$. Taking $\theta_i = 0, \theta_j = \pi$ we see that $g_{ij} = 0$ for $i \neq j$. Thus, g must be block diagonal. Thus

$$g \in (\mathrm{O}(2) \times \cdots \times \mathrm{O}(2)) \cap \mathrm{SO}(2m).$$

From $R(\theta_i)g_{ii} = g_{ii}R(\theta_i)$ 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 $R(-\theta_1)g_{11} = g_{11}R(\theta_1)$. $\square$

Note that for $m = 2$, we have $\Gamma_2 = \mathbb{Z}_2 \subseteq \mathbb{Z}_2 \times \mathbb{Z}_2$ with the trivial action of $S_2 = \mathbb{Z}_2$. Hence $W = \mathbb{Z}_2 \times \mathbb{Z}_2$ in this case, consistent with the fact that the universal cover of $\mathrm{SO}(4)$ is $\mathrm{SU}(2) \times \mathrm{SU}(2)$.

11.4.3. *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 11.21. *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 $j(N(T')) \subseteq N(T)$, we have an inclusion of Weyl groups $W' = \Gamma_m \rtimes S_m \subseteq W$. 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 $\mathbb{Z}_2$. $\square$

11.4.4. *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

$$\mathrm{Sp}(n) = \mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbb{C}),$$

where $\mathrm{Sp}(2n, \mathbb{C})$ is the group of complex matrices satisfying $X^\top J X = J$, with

$$J = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

(see homework 1). 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

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

with $|z_i| = 1$. 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. The diagonal matrices in $\mathrm{Sp}(2n, \mathbb{C})$ are exactly those of the form $t(z_1, \ldots, z_n)$ with $z_i \notin \mathbb{C}$, and this lies in $\mathrm{U}(2n)$ exactly if all $|z_i| = 1$.

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 11.22. *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, $W \subseteq (\mathbb{Z}_2)^n \rtimes S_n$. The inclusion r defines an inclusion of the Weyl group of $\mathrm{U}(n)$, hence $S_n \subseteq W$. 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

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

where $A = \mathrm{diag}(1, 0, 1, \ldots, 1)$, $B = \mathrm{diag}(0, 1, 0, \ldots, 0)$. $\square$

11.4.5. *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.

11.4.6. *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 11.23. (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.

12. Weights and roots

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

Definition 12.1. 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 12.2. 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 12.3. Let $\Lambda \subseteq \mathfrak{t}$ 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 $2\pi i\Lambda^* \subseteq \mathfrak{t}^* \otimes \mathbb{C}$. Similarly, $X_*(T)$ is identified with $\frac{1}{2\pi i}\Lambda \subseteq \mathfrak{t} \otimes \mathbb{C}$.

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 12.4. 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 12.5. There is a natural identification of $X_*(T)$ with the fundamental group $\pi_1(T)$.

Exercise 12.6. Let

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

be a finite cover, where T, T' are tori and $\Gamma \subseteq T'$ 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 $\widehat{\Gamma} = \mathrm{Hom}(\Gamma, \mathrm{U}(1))$.

12.2. **Schur's Lemma.** Recall that a (real or complex) representation of a group G (or Lie algebra $\mathfrak{g}$) on a (real or complex) vector space V is *irreducible* if there are no invariant subspaces other than $\{0\}$ and V. The representation is called *completely reducible* if it is a direct sum of irreducible representations. We have seen that any representation of a compact Lie group is completely reducible. We will need the following simple but important fact.

Theorem 12.7 (Schur Lemma)**.** *Let G be a 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*

$$\dim(\mathrm{Hom}_G(V, V')) = \begin{cases} 1 & \text{if } V \cong V' \\ 0 & \text{otherwise} \end{cases}$$

Similar statements hold for finite-dimensional representations of Lie algebras.

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 $V' = V$, and then b) follows from a). $\square$

For any two complex G-representations V, W, one calls

$$\mathrm{Hom}_G(V, W)$$

the *space of intertwining operators* from V to W. If V is irreducible, and the representation W is completely reducible (as is automatic for G a compact Lie group), then 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.

12.3. **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 12.8. *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 any basis vector v gives an isomorphism $V \cong \mathbb{C}$. The image $\pi(T) \subseteq \mathrm{GL}(V) = \mathrm{GL}(1, \mathbb{C})$ is a compact subgroup, hence it must lie in $\mathrm{U}(1)$. Thus, π becomes a morphism $\mu\colon T \to \mathrm{U}(1)$. $\square$

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: the subspaces 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 $\Delta(V) \subseteq X^*(T)$ be the set of all weights of the T-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.

12.4. **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 $T \subseteq G$.

Proposition 12.9. *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 12.10. 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 $\varpi(t) = z$. Let $V(k)$ be the representation of $\mathrm{SU}(2)$ on the space of homogeneous polynomials of degree k on $\mathbb{C}^2$, given by

$$(g.p)(z_0, z_1) = p(g^{-1}.(z_0, z_1)).$$

The space $V(k)$ is spanned by the polynomials $p(z_0, z_1) = z_0^i z_1^{k-i}$, and since

$$(t \cdot p)(z_0, z_1) = p(z^{-1}z_0, zz_1) = z^{k-2i}\ z_0^i z_1^{k-i},$$

any such polynomial is a weight vector of weight $k - 2i$. So, the weights of the representation $V(k)$ of $\mathrm{SU}(2)$ are

$$\Delta(V(k)) = \{k\varpi,\ (k-2)\varpi, \ldots, -k\varpi\},$$

all appearing with multiplicity 1. The Weyl group $W = \mathbb{Z}_2$ acts by sign changes of weights.

Example 12.11. 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

$$\Delta(S^k\mathbb{C}^n) = \{\epsilon^{i_1} + \ldots + \epsilon^{i_k} |\ i_1 \leq \ldots \leq i_k\}.$$

(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 $\mu = \epsilon^{i_1} + \ldots + \epsilon^{i_k}$.

12.5. **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 12.12. A *root* of G is a non-zero weight for the adjoint representation on $\mathfrak{g}^{\mathbb{C}}$. The set of roots is denoted $\mathfrak{R} \subseteq X^*(T)$.

The weight spaces $\mathfrak{g}_\alpha \subseteq \mathfrak{g}^{\mathbb{C}}$ 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 12.13. 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

$$\mathfrak{R} = \{\epsilon^i - \epsilon^j |\ i \neq j\} \subseteq X^*(T).$$

Example 12.14. 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 $X_*(T) \subseteq X_*(T')$. In terms of the standard basis $\epsilon_1, \ldots, \epsilon_n$ of $X_*(T')$, the lattice $X_*(T)$ has basis

$$\epsilon_1 - \epsilon_2,\ \epsilon_2 - \epsilon_3, \ldots, \epsilon_{n-1} - \epsilon_n.$$

In terms of the dual basis $\epsilon^1, \ldots, \epsilon^n$ of $X^*(T)$, this is the annihilator of the vector $\epsilon^1 + \ldots + \epsilon^n$. Hence, the kernel of the projection map $X^*(T') \to X^*(T)$ is the 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 root system by identifying $X^*(T)$ with the orthogonal projection of $X^*(T')$ to the space

$$V = \mathrm{span}_{\mathbb{R}}\{\epsilon^1 + \ldots + \epsilon^n\}^\perp = \{a = \sum_i a_i \epsilon^i \in \mathbb{R}^n |\ \sum a_i = 0\}.$$

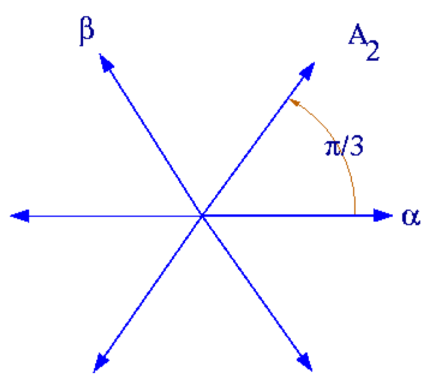
using the standard inner product on $X^*(T') \otimes_{\mathbb{Z}} \mathbb{R} = \mathbb{R}^n$. Note that the standard inner product is $W = S_n$-invariant, hence this identification respects the W-action. The projections of the ϵ^i are

$$\sigma^i = \epsilon^i - \frac{1}{n}(\epsilon^1 + \ldots + \epsilon^n), \quad i = 1, \ldots, n,$$

they generate $X^*(T) \subseteq V$. The roots are

$$\sigma^i - \sigma^j = \epsilon^i - \epsilon^j, \quad i \neq j.$$

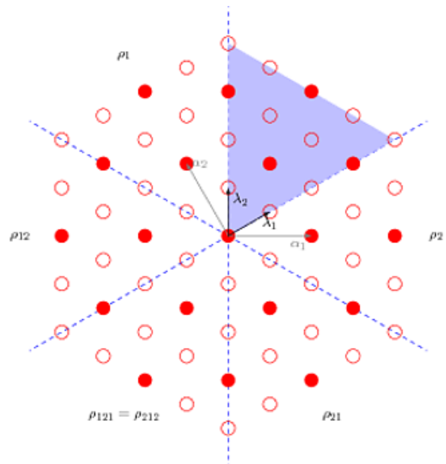
A picture of the root system of SU(3):[11]



[11]Source: wikipedia

In this picture, the so-called simple roots $\alpha = \epsilon^1 - \epsilon^2$, $\beta = \epsilon^2 - \epsilon^3$ are a basis for the lattice generated by the roots (the root lattice). Note that $\alpha + \beta = \epsilon^1 - \epsilon^3$ is also a root.

And here is a picture of the weight lattice, and some roots: [12]



In this picture, one uses the alternative notation $\alpha_1 = \alpha$, $\alpha_2 = \beta$ for the simple roots $\lambda_1 = \sigma_1, \lambda_2 = \sigma_1 + \sigma_2$ ('the fundamental weights') are chosen as a basis of the weight lattice, indicated by the dots. The solid dots are the root lattice, i.e., the lattice spanned by the roots. α_1, α_2 are two of the roots, all other roots are obtained by reflections across the walls.

Example 12.15. Let $G = \mathrm{SO}(2m)$ with its standard maximal torus $T \cong \mathrm{U}(1)^m$ given by the block diagonal matrices

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

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 $X_{ij} = -X_{ji}^{\top}$. Conjugation

$$X \mapsto t(\theta_1, \ldots, \theta_m) X t(\theta_1, \ldots, \theta_m)^{-1}$$

changes the (i, j)-block of X as follows

$$X_{ij} \rightsquigarrow R(\theta_i) X_{ij} R(-\theta_j).$$

Recall that $R(\theta)$ has eigenvalues $e^{\pm i\theta}$. Let $v_\pm \in \mathbb{C}^2$ be corresponding eigenvectors, written as column vectors. [13] Thus $R(\theta) v_\pm = e^{\pm i\theta} v_\pm$. Taking a transpose of this equation, we get

---

[12] Source: `http://tex.stackexchange.com/questions/30301/root-systems-and-weight-lattices-with-pstricks`

[13] For instance, we may take

$$v_+ = \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad v_- = \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

$v_\pm^\top R(-\theta) = e^{\pm i\theta} v_\pm^\top$. Let $i < j$ be given. Putting

$$X_{ij} = X_{ji}^\top = v_\pm v_\pm^\top$$

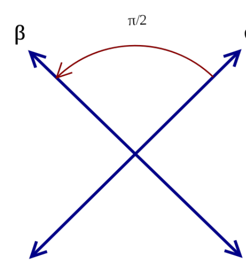
(a 2×2-matrix), we have

$$R(\theta_i) X_{ij} R(-\theta_j) = e^{\pm i\theta_i \pm i\theta^j} X_{ij}.$$

Hence, the matrix X with these entries for $X_{ij} = -X_{ji}^\top$, and all other block entries equal to zero, is a root vector 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 $\dim \mathrm{SO}(2m)/T = 2(m^2 - m)$. Below is this root system for SO(4). [14]



Example 12.16. 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

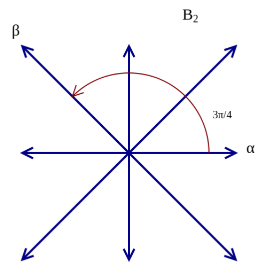
$$\mathfrak{R} = \{\pm\epsilon^i \pm \epsilon^j,\ 1 \le i < j \le m\} \cup \{\pm\epsilon^i,\ i = 1, \ldots, m\}.$$

Picture: [15]

---

[14] Source: `en.wikipedia.org/wiki/Root_system`

[15] Source: `en.wikipedia.org/wiki/Root_system`



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 12.17. Let $G = \mathrm{Sp}(n)$, viewed as $\mathrm{SU}(2n) \cap \mathrm{Spin}(2n, \mathbb{C})$, 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 $T' \subseteq \mathrm{U}(n)$ 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 $a^\top = \overline{a},\ b^\top = b$. Hence its complexification $\mathfrak{sp}(n) \otimes \mathbb{C}$ consists of complex matrices of the form

$$\xi = \begin{pmatrix} a & b \\ c & -a^\top \end{pmatrix},$$

with $b^\top = b,\ c^\top = c$. Thus

$$t\xi t^{-1} = \begin{pmatrix} ZaZ^{-1} & ZbZ \\ Z^{-1}cZ^{-1} & -Z^{-1}a^\top Z \end{pmatrix}$$

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 $i \neq j$).

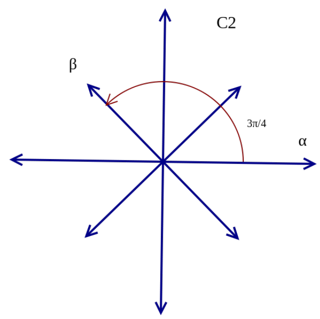
Hence we have found

$$\frac{n(n+1)}{2} + \frac{n(n+1)}{2} + (n^2 - n) = 2n^2$$

roots:

$$\mathfrak{R} = \{\pm\epsilon^i \pm \epsilon^j | 1 \le i < j \le m\} \cup \{\pm 2\epsilon^i |\ i = 1, \ldots, m\}$$

Picture: [16]



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 $\widehat{\Gamma} = \mathrm{Hom}(\Gamma, \mathrm{U}(1))$. The roots of G are identified with the roots of G' under the inclusion $X^*(T') \to X^*(T)$.

Example 12.18. 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.

To summarize some of this discussion, we have the following data for the classical groups:

	rank	name	dim	W	$X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$	$\mathfrak{R}$
A_l	$l \ge 1$	$\mathrm{SU}(l+1)$	$l^2 + 2l$	S_{l+1}	$\{a \in \mathbb{R}^{\ell+1} \mid \sum_0^\ell a_i = 0\}$	$\{\epsilon^i - \epsilon^j \mid i < j\}$
B_l	$l \ge 2$	$\mathrm{Spin}(2l+1)$	$2l^2 + l$	$(\mathbb{Z}_2)^l \rtimes S_l$	$\mathbb{R}^\ell$	$\{\pm\epsilon^i \pm \epsilon^j \mid i \ne j\} \cup \{\pm\epsilon^i\}$
C_l	$l \ge 3$	$\mathrm{Sp}(l)$	$2l^2 + l$	$(\mathbb{Z}_2)^l \rtimes S_l$	$\mathbb{R}^\ell$	$\{\pm\epsilon^i \pm \epsilon^j \mid i < j\} \cup \{\pm 2\epsilon^i\}$
D_l	$l \ge 4$	$\mathrm{Spin}(2l)$	$2l^2 - l$	$(\mathbb{Z}_2)^{l-1} \rtimes S_l$	$\mathbb{R}^\ell$	$\{\pm\epsilon^i \pm \epsilon^j \mid i < j\}$

[16] Source: `en.wikipedia.org/wiki/Root_system`

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

To obtain deeper results about root systems, we need some representation theory.

13.1. **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 13.1. 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 $\pi\colon\mathfrak{g}\to\mathrm{End}_{\mathbb{C}}(V)$. 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.

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.

If $\mathfrak{g}$ is the Lie algebra of a compact simply connected Lie group G, then every finite-dimensional $\mathfrak{g}^{\mathbb{C}}$-representation is completely reducible. Indeed, the $\mathfrak{g}^{\mathbb{C}}$-representation is in particular a $\mathfrak{g}$-representation, and the $\mathfrak{g}$-invariant complex subspaces are exactly the $\mathfrak{g}^{\mathbb{C}}$-invariant complex subspaces. But every finite-dimensional $\mathfrak{g}$-representation exponentiates to a G-representation, and we had already seen that the latter are completely reducible.

13.2. $\mathfrak{sl}(2,\mathbb{C})$**-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 $\mathfrak{sl}(2,\mathbb{C})$ given by the matrices

$$e=\begin{pmatrix}0&1\\0&0\end{pmatrix},\quad h=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad f=\begin{pmatrix}0&0\\1&0\end{pmatrix}.$$

The corresponding bracket relations read as

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

For later reference, we note that $\mathfrak{su}(2)$ is realized as the fixed point set of the conjugate-linear involution $A\mapsto -A^\dagger=-\overline{A}^\top$, that is,

$$h\mapsto -h,\ e\mapsto -f,\ f\mapsto -e.$$

Given a representation $\pi\colon\mathfrak{sl}(2,\mathbb{C})\to\mathrm{End}(V)$, we define the **Casimir operator** as follows:

$$\mathrm{Cas}=2\pi(f)\pi(e)+\frac{1}{2}\pi(h)^2+\pi(h)\in\mathrm{End}(V)$$

Lemma 13.2. *IIf π is irreducible, then* Cas *acts as a scalar.*

Proof. By Schur's Lemma, it suffices to check that this operator commutes with $\pi(h), \pi(e), \pi(f)$. For example,

$$\begin{aligned}
[\pi(e), 2\pi(f)\pi(e)] &= 2\pi(h)\pi(e), \\
\tfrac{1}{2}[\pi(e), \pi(h)^2] &= \tfrac{1}{2}[\pi(e), \pi(h)]\pi(h) + \tfrac{1}{2}\pi(h)[\pi(e), \pi(h)] \\
&= -\pi(e)\pi(h) - \pi(h)\pi(e) \\
&= -2\pi(h)\pi(e) + 2\pi(e) \\
[\pi(e), \pi(h)] &= -2\pi(e)
\end{aligned}$$

add to 0. □

The simplest non-trivial representation of $\mathfrak{sl}(2, \mathbb{C})$ is the defining representation on $\mathbb{C}^2$ (given by the matrices). Clearly, this representation is irreducible (there cannot be 1-dimensional invariant subspaces). Another picture for the defining representation is by viewing $\mathbb{C}^2$ as linear homogeneous polynomials of degree 1 in $z, w \in \mathbb{C}$. In this picture, the operators are

$$\pi(e) = z\frac{\partial}{\partial w}, \qquad \pi(h) = z\frac{\partial}{\partial z} - w\frac{\partial}{\partial w}, \qquad \pi(f) = w\frac{\partial}{\partial z}.$$

More generally, these differential operators define a representation on the space $V(k)$ of homogeneous polynomials of degree k. $V(0)$ is the trivial representation, and $V(1)$ is isomorphic to the defining representation. Introducing the basis

$$v_j = \frac{1}{(k-j)!j!} z^{k-j} w^j, \ j = 0, \ldots, k$$

for the space $V(k)$, the representation is given by the formulas

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

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

Theorem 13.3. *For all $k = 0, 1, 2, \ldots$, the representations of $\mathfrak{sl}(2, \mathbb{C})$ on $V(k)$ is irreducible. The Casimir operator acts as the scalar $\frac{1}{2}k(k+2)$ on $V(k)$.*

Proof. Suppose $W \subseteq V(k)$ is a nonzero invariant subspace. The operator $\pi(e)|_W$ has at least one eigenvector. But the formulas above show that $\pi(e)$ has a *unique* eigenvector (up to scalar) in $V(k)$, given by v_0. By iterated application of $\pi(f)$ to v_0, we obtain all basis vectors $v_0, \ldots, v_k$ (up to scalars). Hence, $W = V(k)$. To compute the action of Cas on $V(k)$, it suffices to compute its action on any vector. A convenient choice is v_k, and since $\pi(e)v_0 = 0, \ \pi(h)v_0 = k\, v_0$ the result follows. □

Theorem 13.4. *Up to isomorphism, the representations $V(k), \ k = 0, 1, \ldots$ are the unique finite-dimensional irreducible representation of $\mathfrak{sl}(2, \mathbb{C})$.*

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

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

We claim that

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

The first claim follows from the calculation, for $v \in V_{[s]}$,

$$\begin{aligned}\pi(h)\pi(e)v &= \pi([h,e])v + \pi(e)\pi(h)v \\ &= 2\pi(e)v + s\pi(e)v \\ &= (s+2)\pi(e)v.\end{aligned}$$

The second claim is proved similarly. 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 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 $j+1$. The non-zero v_j are linearly independent (since they lie in different eigenspaces for $\pi(h)$. Let k be the smallest number such that $v_{k+1} = 0$. Then $v_0, \ldots, v_k$ are a basis of V: They are linearly independent, and since their span is invariant it is all of V. Putting $j = k+1$ in the formula for $\pi(e)v_j$, we obtain $0 = (\lambda - k)v_k$, hence $\lambda = k$. ☐

Remark 13.5. 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 13.6. Show that for $\lambda \notin \mathbb{Z}_{\geq 0}$, the Verma module is irreducible.

As explained above, any finite-dimensional $\mathfrak{sl}(2,\mathbb{C})$-representation $\pi\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V)$ is completely reducible, and hence is a direct sum of copies of the V_k's, with multiplicities n_k. There are various methods for computing the n_k's. Here are three:

Method 1: Determine the eigenspaces of the Casimir operator Cas. The eigenspace for the eigenvalue $k(k+2)/2$ is the direct sum of all irreducible sub-representations of type $V(k)$. Hence n_k is the dimension of this eigenspace, divided by $k+1$.

Method 2: For $l \in \mathbb{Z}$, let $m_l = \dim\ker(\pi(h) - l)$ 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 $V^{\mathfrak{n}}$. The multiplicity of the eigenvalue k on $V^{\mathfrak{n}}$ is then equal to n_k.

Exercise 13.7. 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 $\pi(e)$.)

Let us note the following simple consequence of the $\mathfrak{sl}(2,\mathbb{C})$-representation theory:

Lemma 13.8. *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. $\square$

14. 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 $\mathfrak{R} \subseteq X^*(T)$ of G, and of the decomposition

$$\mathfrak{g}^{\mathbb{C}} = \mathfrak{t}^{\mathbb{C}} \oplus \bigoplus_{\alpha \in \mathfrak{R}} \mathfrak{g}_\alpha.$$

14.1. **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 14.1. *For all $\alpha, \beta \in \Delta(\mathfrak{g}^{\mathbb{C}}) = \mathfrak{R} \cup \{0\}$,*

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$$

In particular, if $\alpha + \beta \notin \Delta(\mathfrak{g}^{\mathbb{C}})$ then $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = 0$, and $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{t}^{\mathbb{C}}$ 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 14.2. *The spaces $\mathfrak{g}_\alpha, \mathfrak{g}_\beta$ for $\alpha + \beta \neq 0$ 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 $B(X_\alpha, X_\beta) \neq 0$. □

14.2. **The Lie subalgebras $\mathfrak{sl}(2,\mathbb{C})_\alpha \subseteq \mathfrak{g}^{\mathbb{C}},\ \mathfrak{su}(2)_\alpha \subseteq \mathfrak{g}$.** For any weight $\mu \in X^*(T)$, seen as a group morphism $T \to \mathrm{U}(1)$, 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 $X_\alpha \in \mathfrak{g}_\alpha$ and $h \in \mathfrak{t}$. We may extend $\mathrm{d}\mu$ to a complex-linear map $\mathfrak{t}_{\mathbb{C}} \to \mathbb{C}$, with the same property.

Theorem 14.3.

(a) *For every root $\alpha \in \mathfrak{R}$, the root space $\mathfrak{g}_\alpha$ is 1-dimensional.*

(b) *The subspace $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathfrak{t}^{\mathbb{C}}$ is 1-dimensional, and contains a unique element h_α such that*

$$d\alpha(h_\alpha) = 2.$$

(c) *We may choose generators $e_\alpha \in \mathfrak{g}_\alpha,\ f_\alpha \in \mathfrak{g}_{-\alpha},\ h_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ such that*

$$[h_\alpha, e_\alpha] = 2e_\alpha, \quad [h_\alpha, f_\alpha] = -2f_\alpha, \quad [e_\alpha, f_\alpha] = h_\alpha,$$

and

$$\overline{e_\alpha} = -f_\alpha,\ \overline{f_\alpha} = -e_\alpha,\ \overline{h_\alpha} = -h_\alpha.$$

Thus $\mathfrak{sl}(2,\mathbb{C})_\alpha = \mathrm{span}_{\mathbb{C}}(e_\alpha, f_\alpha, h_\alpha) \subseteq \mathfrak{g}_{\mathbb{C}}$ is a complex Lie subalgebra isomorphic to $\mathfrak{sl}(2,\mathbb{C})$, whose fixed point set under complex conjugation is $\mathfrak{su}(2)$.

Proof. We begin with the proof of (b). Pick an invariant inner product B on $\mathfrak{g}$. Let $H_\alpha \in \mathfrak{t}^{\mathbb{C}}$ 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

$$\mathrm{d}\alpha(H_\alpha) = B(H_\alpha, H_\alpha) < 0.$$

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

$$\begin{aligned} B([e_\alpha, e_{-\alpha}], h) &= B(e_{-\alpha}, [h, e_\alpha]) \\ &= \mathrm{d}\alpha(h) B(e_{-\alpha}, e_\alpha) \\ &= B(e_{-\alpha}, e_\alpha) B(H_\alpha,\ h) \\ &= B\big(B(e_{-\alpha}, e_\alpha) H_\alpha,\ h\big). \end{aligned}$$

This shows $[e_\alpha, e_{-\alpha}] = B(e_{-\alpha}, e_\alpha) H_\alpha$, proving that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \mathrm{span}_{\mathbb{C}}(H_\alpha)$. 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 the unique element of $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ with this normalization.

We next prove (a) (and some of (c)). 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 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.

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 $\mathrm{U}(1)$) 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 14.4. *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 $\mathfrak{sl}(2, \mathbb{C})$-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 positive 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 $\mathrm{ran}(f_\alpha) \cap \ker(e_\alpha) = 0$. □

14.3. **Co-roots.** The Lie subalgebra $\mathfrak{su}(2)_\alpha \subseteq \mathfrak{g}$ 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 $\mathfrak{su}(2)_\alpha \subseteq \mathfrak{g}$. Since $\mathrm{SU}(2)$ is simply connected, with center $\mathbb{Z}_2$, the image is isomorphic either to $\mathrm{SU}(2)$ or to $\mathrm{SO}(3)$. Let $T_\alpha \subseteq \mathrm{SU}(2)_\alpha$ be the maximal torus obtained by exponentiating $\mathrm{span}_\mathbb{R}(ih_\alpha) \subseteq \mathfrak{t}$. The morphism $T_\alpha \to T$ defines an injective map of the coweight lattices,

$$X_*(T_\alpha) \to X_*(T). \tag{4}$$

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 14.5. 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 (4). The set of co-roots is denoted $\mathfrak{R}^\vee \subseteq X_*(T)$.

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 $\mu \in X^*(T)$, where the right hand side uses the pairing between weights $X^*(T) = \mathrm{Hom}(T, \mathrm{U}(1))$ and coweights $X_*(T) = \mathrm{Hom}(\mathrm{U}(1), T)$. This formula may be seen as an alternative definition of the co-root.

Remark 14.6.

(a) As for any pairing between weights and coweights, we have that $\langle \beta^\vee, \alpha \rangle \in \mathbb{Z}$ 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) By definition, the element h_α is the unique element of $[g_\alpha, \mathfrak{g}_{-\alpha}]$ satisfying $\mathrm{d}\alpha(h_\alpha) = 2$. As we have seen above, for any invariant inner product B on $\mathfrak{g}$ we have that

$$B([e_\alpha, f_\alpha], h) = B(e_\alpha, f_\alpha)\mathrm{d}\alpha(h) = B(e_\alpha, f_\alpha)\langle d\alpha, h \rangle.$$

Hence if we use B to identify $\mathfrak{t}_\mathbb{C}$ and its dual, we have that $[e_\alpha, f_\alpha]$ is a multiple of α. Thus, $\alpha^\vee$ is a multiple of α under these identifications.

(c) Recall that $X_*(T) \otimes_\mathbb{Z} \mathbb{R} = \mathrm{Hom}(i\mathbb{R}, \mathfrak{t}) \cong i\mathfrak{t} \subseteq \mathfrak{t}_\mathbb{C}$. The invariant inner product B on $\mathfrak{g}$ restricts to a W-invariant inner product on $\mathfrak{t}$, which in turn gives a nondegenerate bilinear form on $\mathfrak{t}_\mathbb{C}$. The latter is negative definite on $i\mathfrak{t}$. Let $(\cdot, \cdot)$ be obtained by a sign change, so it's 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 14.7. 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.*

14.4. **Root lengths and angles.** Choose a W-invariant inner product $(\cdot,\cdot)$ on the real vector space $E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$.

Theorem 14.8. *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 14.9. *For all roots $\alpha, \beta \in \mathfrak{R}$, the integer $\langle\alpha^\vee, \beta\rangle$ lies in the interval* $[-3, 3]$.

14.5. **Root strings.** Making further use of the $\mathfrak{sl}(2,\mathbb{C})$-representation theory, we next prove:

Theorem 14.10. *(Root strings.) Let $\alpha,\beta \in \mathfrak{R}$ be roots, with $\beta \neq \pm\alpha$. Then*

(a)
$$\langle \alpha^\vee, \beta\rangle < 0 \Rightarrow \alpha+\beta \in \mathfrak{R}.$$

(b) *There exist integers $q,p \geq 0$ such that for any integer $r \in \mathbb{Z}$,*
$$\beta + r\alpha \in \mathfrak{R} \;\Leftrightarrow\; -q \leq r \leq p.$$
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 $p+q+1$.

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) Consider
$$V = \bigoplus_{j\in\mathbb{Z}} \mathfrak{g}_{\beta+j\alpha}$$
as an $\mathfrak{sl}(2,\mathbb{C})_\alpha$-representation. The operator $\mathrm{ad}(h_\alpha)$ acts on the 1-dimensional space $\mathfrak{g}_{\beta+j\alpha}$ as $\langle \alpha^\vee, \beta\rangle + 2j$. We hence see that the eigenvalues of $\mathrm{ad}(h_\alpha)$ on V are hence either all even, or all odd, and they are all distinct (i.e., multiplicity one). But for any finite-dimensional complex $\mathfrak{sl}(2,\mathbb{C})$-representation, the number of irreducible components is the multiplicity of the eigenvalue 0 of $\mathrm{ad}(h)$, plus the multiplicity of the eigenvalue 1. This shows that V is an irreducible $\mathfrak{sl}(2,\mathbb{C})_\alpha$-representation. It is thus isomorphic to $V(k)$, where $k+1 = \dim V$.

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 $q - p = \langle \alpha^\vee, \beta\rangle$.

The last claim 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. $\square$

The set of roots $\beta + j\alpha$ with $-q \le j \le p$ is called the *α-root string through β.*

Lemma 14.11. *The length of any root string is at most* 4.

Proof. If β is such that $\beta - \alpha$ is *not* a root, we have $q = 0,\ k = p = -\langle \alpha^\vee, \beta\rangle$. By Lemma 14.9, this integer is ≤ 3. Hence, the length of any root string is at most 4. $\square$

14.6. **Weyl chambers.** Let
$$E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R}$$
be the real vector space spanned by the weight lattice. (It may be identified with $i\mathfrak{t}^* \subseteq \mathfrak{t}^*_{\mathbb{C}}$.) 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 $T_\alpha \subseteq \mathrm{SU}(2)_\alpha$ 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 14.12. *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
$$w_\alpha\, \xi = \xi - \langle \xi, \alpha\rangle \alpha^\vee, \quad \xi \in E^*.$$

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 $j_\alpha(N(T_\alpha)) \subseteq T$. We also see that w_α acts trivially on $\ker(\mathrm{d}\alpha)$, and as -1 on $\mathrm{span}(h_\alpha)$. This shows that the action of w_α on E^* is a reflection:
$$w_\alpha \xi = \xi - \langle \xi, \alpha\rangle \alpha^\vee, \quad \xi \in E^*.$$
The statement for the weight lattice follows by duality (and using $w_\alpha^2 = 1$):
$$\langle \xi, w_\alpha \mu\rangle = \langle w_\alpha \xi, \mu\rangle = \langle \xi, \mu\rangle - \langle \alpha^\vee, \mu\rangle\langle \xi, \alpha\rangle,\ \ \xi \in E^*, \mu \in E.$$
$\square$

Remark 14.13. 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*

$$H_\alpha = \mathrm{span}_{\mathbb{R}}(\alpha)^\perp \subseteq E.$$

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. Recall again that $E = \imath\mathfrak{t}$. Note that for $x = ih \in E$, the kernel $\ker(\mathrm{ad}(h)) \subseteq \mathfrak{g}^{\mathbb{C}}$ is invariant under the adjoint representation of $T \subseteq G$, hence is a sum of $\mathfrak{t}^{\mathbb{C}}$ and possibly some root spaces $\mathfrak{g}_\alpha$. But $\mathrm{ad}(h)$ acts on the root space $\mathfrak{g}_\alpha$ as a scalar $i\mathrm{d}\alpha(h) = (\alpha, x)$. This shows

$$\ker(\mathrm{ad}(h)) = \mathfrak{t}^{\mathbb{C}} \oplus \bigoplus_{\alpha\colon (\alpha,x)=0} \mathfrak{g}_\alpha.$$

In particular, $x = ih$ is regular if and only if $\ker(\mathrm{ad}(h)) = \mathfrak{t}^{\mathbb{C}}$.

Lemma 14.14. *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 $w = 1$. Write $x = ih$. Since $\ker(\mathrm{ad}(h)) = \mathfrak{t}^{\mathbb{C}}$, we have that $\mathfrak{t}$ is the unique maximal abelian subalgebra containing h. Equivalently, T is the unique maximal torus containing the 1-parameter subgroup S generated by the element $h \in \mathfrak{t}$. 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 14.15. This result (or rather its proof) also has the following consequence. Let $\mathfrak{g}^{\mathrm{reg}} \subseteq \mathfrak{g}$ be the set of Lie algebra elements ξ 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 14.16. For arbitrary $\xi \in \mathfrak{t}$, the stabilizer $G_\xi = \{g \in G | \operatorname{Ad}_g(\xi) = \xi\}$ 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?

Definition 14.17. 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 14.18. *The Weyl group acts freely on the set of Weyl chambers. That is, if C is a chamber and $w \in W$ with $wC \subseteq C$ then $w = 1$.*

Proof. If $wC = C$, then w preserves the interior of C. Let $x \in \operatorname{int}(C)$. Then $w^i x \in \operatorname{int}(C)$ for all $i \geq 0$. Letting k be the order of w, the element $x' := x + wx + \ldots w^{k-1}x \in \operatorname{int}(C)$ satisfies $wx' = x'$. By the previous Lemma this means $w = 1$. □

Exercise 14.19. Let C be a fixed (closed) Weyl chamber. a) Let $D \subseteq C$ one of its 'faces'. (Thus D is the intersection of C with some of the root hyperplanes.). Show that if $w(D) \subseteq D$, 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* $C, C' \subseteq E$ 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 14.20. *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 14.21. *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 14.22. *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.

15. 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 15.1. 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 15.2 (Simple roots)**.** *The set* $\Pi = \{\alpha_1, \ldots \alpha_l\}$ *of simple roots has the following properties.*

(a) Π *is a basis of the root lattice,* $\mathrm{span}_{\mathbb{Z}} \mathfrak{R} \subseteq X^*(T)$.
(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. Fix an element $x_* \in \mathrm{int}(C_+) \subseteq E$, that is, $(\alpha, x_*) > 0$ for all $\alpha \in \mathfrak{R}_+$.

1) *Proof of (c).* 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 $\alpha + \beta$ 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).

2) *Proof of linear independence of* Π. Suppose $\sum_i k_i \alpha_i = 0$ for some $k_i \in \mathbb{R}$. Let

$$\mu := \sum_{k_i > 0} k_i \alpha_i = -\sum_{k_j < 0} k_j \alpha_j. \tag{5}$$

Taking the scalar product of μ with itself, and using $(\alpha_i, \alpha_j) \leq 0$ for $i \neq j$ we obtain

$$0 \leq (\mu, \mu) = -\sum_{k_i > 0,\ k_j < 0} k_i k_j (\alpha_i, \alpha_j) \leq 0.$$

Hence $\mu = 0$. Taking the inner product with x_* we get

$$0 = \sum_{k_i > 0} k_i (\alpha_i, x_*) = -\sum_{k_j < 0} k_j (\alpha_j, x_*).$$

which is only possible if all $k_i = 0$.

3) *Proof that any $\alpha \in \mathfrak{R}_+$ can be written in the form $\alpha = \sum k_i\alpha_i$ for some $k_i \in \mathbb{Z}_{\geq 0}$.* (This will prove (b), and also finish the proof of (a).) Suppose the claim is false, and let α be a counterexample for which (α, x_*) is 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. Both $(\alpha', x_*) > 0$, $(\alpha'', x_*) > 0$ 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 $\mathbb{Z}_{\geq 0}$. Hence the same is true of α, hence α is not a counterexample. Contradiction. ☐

Corollary 15.3. *The simple co-roots $\mathfrak{R}^\vee = \{\alpha_1^\vee, \ldots, \alpha_l^\vee\}$ are a basis of the co-root lattice $\operatorname{span}_{\mathbb{Z}} \mathfrak{R}^\vee \subseteq X_*(T)$.*

Definition 15.4. 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 $\mathfrak{R} \subseteq E$).

Note that the diagonal entries of the Cartan matrix are equal to 2, the and that the off-diagonal entries are ≤ 0.

Example 15.5. Let $G = \mathrm{U}(n)$, and use the standard inner product on $E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R} = \operatorname{span}_{\mathbb{R}}(\epsilon^1, \ldots, \epsilon^n)$ 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 15.6. Let $G = \mathrm{SO}(2l+1)$. Using the standard maximal torus and the basis $X^*(T) = \operatorname{span}_{\mathbb{Z}}(\epsilon_1, \ldots, \epsilon^l)$, 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 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 SO$(2l+1)$.

A more efficient way of recording the information of a Cartan matrix is the *Dynkin diagram*[17].

Definition 15.7. The Dynkin diagram of G is a graph, with

- vertices (nodes) the simple roots,
- edges between vertices $i \neq j$ for which $(\alpha_i, \alpha_j) \neq 0$.

One gives each edge a multiplicity of $1, 2$, or 3 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 15.8. There are only four possible Dynkin diagrams with 2 nodes:

(a) a disconnected Dynkin diagram (corresponding to SU(2) $\times$ SU(2) or SO(4))
(b) a connected Dynkin diagram with an edge of multiplicity 1 (corresponding to $A_2 =$ SU(3))
(c) a connected Dynkin diagram with an edge of multiplicity 2 (corresponding to $B_2 =$ Spin(5))
(d) a connected Dynkin diagram with an edge of multiplicity 3 (corresponding to the exceptional group G_2)

Exercise 15.9. 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 15.10. *The positive Weyl chamber is 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. $\square$

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:

Theorem 15.11. *The Dynkin diagram determines the root system $\mathfrak{R}$, up to isomorphism.*

[17] Dynkin diagrams were used by E. Dynkin in his 1946 papers. Similar diagrams had previously been used by Coxeter in 1934 and Witt 1941.

Proof. The Dynkin diagram determines the set Π of simple roots, as well as their angles and relative lengths. The Weyl group W is recovered as the group generated by the simple reflections $s_i = w_{\alpha_i}$, and

$$\mathfrak{R} = W\Pi.$$

$\square$

Hence, given the Dynkin diagram one may recover the root system, the Weyl group, the Weyl chamber etc.

Example 15.12. 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 $\alpha = \mathrm{sum}_{i=1}^{l} k_i \alpha_i \in \mathfrak{R}$ (or more generally for any element of the root lattice), one defines its *height* by

$$\mathrm{ht}(\alpha) = \sum_{i=1}^{l} k_i.$$

In terms of the fundamental coweights (cf. below),

$$\mathrm{ht}(\alpha) = \sum_{i=1}^{l} \langle \varpi_i^\vee, \alpha \rangle.$$

Proposition 15.13. *For any* $\alpha \in \mathfrak{R}_+ \backslash \Pi$ *there exists* $\beta \in \mathfrak{R}_+$ *with* $\mathrm{ht}(\beta) = \mathrm{ht}(\alpha) - 1$.

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 $\alpha - \alpha_r \in \mathfrak{R}_+$. $\square$

16. 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 U(1) by the condition $[e_i, \overline{e_i}] = -h_i$, and put $f_i = -\overline{e_i}$.

Theorem 16.1. *The elements e_i, f_i, h_i generate $\mathfrak{g}^{\mathbb{C}}$. They satisfy the* Serre relations*,*

$$\begin{aligned}
&(S1) \quad [h_i, h_j] = 0,\\
&(S2) \quad [e_i, f_j] = \delta_{ij} h_i,\\
&(S3) \quad [h_i, e_j] = a_{ij} e_j,\\
&(S4) \quad [h_i, f_j] = -a_{ij} f_j,\\
&(S5) \quad \mathrm{ad}(e_i)^{1-a_{ij}}(e_j) = 0,\\
&(S6) \quad \mathrm{ad}(f_i)^{1-a_{ij}}(f_j) = 0
\end{aligned}$$

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 $\mathrm{ht}(\beta) = \mathrm{ht}(\alpha) = 1$, 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 $[\mathfrak{g}_{\alpha_i}, \mathfrak{g}_{-\alpha_j}] \subseteq \mathfrak{g}_{\alpha_i - \alpha_j} = 0$ 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. $\square$

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.

Definition 16.2. Let E be a Euclidean vector space, and $\mathfrak{R} \subseteq E \backslash \{0\}$. 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 16.3 (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*'.

17. Classification of Dynkin diagrams

There is an obvious notion of *sum* of root systems $\mathfrak{R}_1 \subseteq E_1,\ \mathfrak{R}_2 \subseteq E_2$, 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.

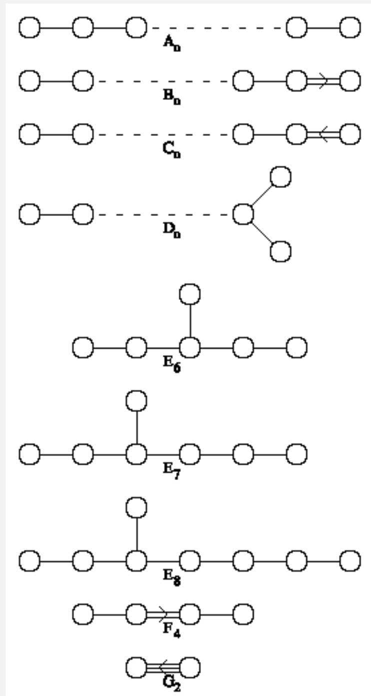
Given an abstract root system, we may as before define Weyl chambers, and the same proof as before shows that for non-orthogonal roots α, β with $||\alpha|| \geq ||\beta||$, the ratio of the root lengths is given by $||\alpha||^2/||\beta||^2 \in \{1, 2, 3\}$, and the angles in the three cases are $\pm\frac{\pi}{3}, \pm\frac{\pi}{4}, \pm\frac{\pi}{6} \mod \pi$. Hence, we may define simple roots and a Dynkin diagram as before.

Proposition 17.1. *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 $\mathfrak{R} \subseteq E$ 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 17.2. *Let $\mathfrak{R}$ be an irreducible root system. Then the Dynkin diagram is given by exactly one of the following types A_l $(l \geq 1), B_l$ $(l \geq 2), C_l$ $(l \geq 3), D_l$ $(l \geq 4)$ or E_6, E_7, E_8, F_4, G_2.* [a]



Here the subscript signifies the rank, i.e. the number of vertices of the Dynkin diagram.

[a] Picture source: `https://upload.wikimedia.org/wikipedia/commons/5/5f/ConnectedDynkinDiagrams.png`

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 17.3. *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 $\Pi' \subseteq \Pi$ 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 $\alpha_i/||\alpha_i||$, $i \le 4$ 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 $|W| = 696, 729, 600$.