Lie groups and Lie algebras (Winter 2024)

Contents

1. Introduction

The study of symmetries is an important theme in almost every area of mathematics. One might be dealing with finite symmetry groups (such as the symmetries of a solid) or infinite, but discrete symmetry groups (such as the symmetries of a pattern). Lie groups, by contrast, are about so-called *continuous* symmetries (more aptly called smooth symmetries), such as translational or rotational symmetries of a system. Rather then calling them smooth symmetry groups, they are named after the Norwegian mathematician Sophus Lie (1842-1899). Lie studied Lie groups mainly in the context of symmetries of differential equations. His big discovery was that the behaviour of a Lie group near the group unit is completely determined by an algebraic structure on the tangent space at the group unit, its Lie algebra. This is the starting point of a rich interplay between local and global aspects of Lie groups.

Nowadays, we are studying Lie groups in their own right – not only as symmetries of some structure. This course, after a general introduction to Lie groups and Lie algebras, will focus mainly on the theory of compact Lie groups: Their structure theory, representations, and classifications. There is a vast literature on the subject, with many excellent textbooks. Our main source is the book by Bröcker and tom Dieck, but we will also consult other references such as the book of Duistermaat and Kolk, the recent book of Kosmann-Schwarzbach, and Wikipedia.

I thank Howard Xiao for pointing out several typos and errors in these notes.

2. Definitions and first examples

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

Definition 2.1. A Lie group is a group G, equipped with a manifold structure such that the group operations

$$\mathrm{Mult}\colon G\times G\to G,\quad (g_1,g_2)\mapsto g_1g_2$$
$$\mathrm{Inv}\colon G\to G,\quad g\mapsto g^{-1}$$

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

Remark 2.2. 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*

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

[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.

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. The group $\mathrm{GL}(n, \mathbb{R})$ has two connected components distinguished by the sign of the determinant; the subgroup $\mathrm{GL}^+(n, \mathbb{R})$ of matrices with positive determinant is a Lie group too. 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 that $\mathrm{O}(n)$ is an (embedded) submanifold of $\mathrm{GL}(n, \mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{R})$. 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. (Details left as an exercise.)

Similarly, we may show that the *special linear group* $\mathrm{SL}(n, \mathbb{R})$, consisting of matrices of determinant 1, is an (embedded) submanifold of $\mathrm{GL}(n, \mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{R})$ by showing that 1 is a regular value of the determinant function.

Definition 2.3. A *matrix Lie group* is a subgroup $G \subseteq \mathrm{GL}(n, \mathbb{R})$ (for some n) which is also an embedded submanifold.

For more complicated matrix Lie groups, the direct approach of verifying a regular value property gets increasingly cumbersome. Here is another method, with the added advantage that it directly provides submanifold charts.

2.2. **Exponential charts.** Recall the exponential map for matrices:

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

This is a (entry-wise) absolutely convergent series, with infinite radius of convergence. In particular, $\exp(B)$ depends smoothly on B. Note that exp takes values in $\mathrm{GL}(n, \mathbb{R})$, and $\exp(0) = I$.

Lemma 2.4. *There exists an open ball* $U \subseteq \mathrm{Mat}_n(\mathbb{R})$ *around* 0 *such that the exponential map* exp *restricts to a diffeomorphism*

$$\exp\colon U \to \exp(U) \subseteq \mathrm{GL}(n, \mathbb{R})).$$

By open ball, we mean a set of the form $\{B |\ ||B|| < \epsilon\}$ for some $\epsilon > 0$, where $||B||$ is a matrix norm.

Proof. By the implicit function theorem, it suffices to show that the differential at 0

$$T_0 \exp\colon T_0 \mathrm{Mat}_n(\mathbb{R}) \to T_I \mathrm{Mat}_n(\mathbb{R})$$

is invertible. Since $\mathrm{Mat}_n(\mathbb{R})$ is a vector space, its tangent spaces are identified with $\mathrm{Mat}_n(\mathbb{R})$. The differential is computed as $(T_0 \exp)(B) = \frac{d}{dt}\big|_{t=0} \exp(tB) = B$; thus $T_0 \exp = \mathrm{id}_{\mathrm{Mat}_n(\mathbb{R})}$. In particular, $T_0 \exp$ is invertible. $\square$

The inverse map

$$\log\colon \exp(U) \to U \subseteq \mathrm{Mat}_n(\mathbb{R})$$

serves as a chart

$$(\exp(U), \log)$$

for $\mathrm{GL}(n,\mathbb{R})$ around I. To obtain a chart around other elements $A \in \mathrm{GL}(n,\mathbb{R})$, one uses 'left translation': Let $l_A\colon \mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$ be left multiplication by A, we obtain charts

$$(l_A(\exp(U)),\ \log \circ l_{A^{-1}})\,.$$

We shall refer to this kinds of charts as *exponential charts.* [2]

> **Fact:** For every matrix Lie group $G \subseteq \mathrm{Mat}_n(\mathbb{R})$, the exponential charts for $\mathrm{GL}(n,\mathbb{R})$ serve as submanifold charts. In particular,
>
> $$\log(G \cap \exp(U)) = \mathfrak{g} \cap U$$
>
> for some subspace $\mathfrak{g} \subseteq \mathrm{Mat}_n(\mathbb{R})$.

We postpone the proof. Let us note that the subspace $\mathfrak{g}$ is identified as

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

Indeed, $\exp(\mathfrak{g} \cap U) = G \cap \exp(U)$ implies that $T_0 \exp = \mathrm{id}_{\mathrm{Mat}_n(\mathbb{R})}$ restricts to an isomorphism from $T_0(\mathfrak{g} \cap U) = \mathfrak{g}$ to $T_I(G \cap \exp(U)) = T_I G$.

For now, we use the construction in the opposite direction: If G is a subgroup of $\mathrm{GL}(n,\mathbb{R})$ with the property that $\log(G \cap \exp(U)) = \mathfrak{g} \cap U$ for some subspace $\mathfrak{g} \subseteq \mathrm{Mat}_n(\mathbb{R})$, then we may take $(l_A(\exp(U)),\ \log \circ l_{A^{-1}})$ as submanifold charts for G. In particular, G is then a matrix Lie group. Let's see how this works for $\mathrm{O}(n)$ and $\mathrm{SL}(n,\mathbb{R})$:

Lemma 2.5. $\mathrm{O}(n)$ *is an (embedded) submanifold of* $\mathrm{GL}(n,\mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{R})$*, with the exponential charts as submanifold charts, and with*

$$\mathfrak{o}(n) = \{B \mid B + B^\top = 0\}.$$

Similarly, $\mathrm{SL}(n,\mathbb{R})$ *is an embedded submanifold of* $\mathrm{GL}(n,\mathbb{R})$*, with the exponential charts as submanifold charts, and with*

$$\mathfrak{sl}(n,\mathbb{R}) = \{B \mid \mathrm{tr}(B) = 0\}.$$

Proof. For all $B \in U \subseteq \mathrm{Mat}_n(\mathbb{R})$,

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

[2] The assumption that U is an open ball is not very important, we will need that U is 'star-shaped', i.e. invariant under scalar multiplication by $t \in [0,1]$.

Similarly,

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

where we used the identity $\det(\exp(B)) = \exp(\mathrm{tr}(B))$. $\square$

In this method, the main task is to identify a candidate for $\mathfrak{g}$. For the standard matrix Lie groups, this is usually not too difficult. For example, for $\mathrm{U}(n) = \{A|\ A^\dagger = A^{-1}\}$ one obtains

$$\mathfrak{u}(n) = \{B|\ B^\dagger = -B\};$$

by writing $B = \frac{d}{dt}|_{t=0} A_t$, and taking the t-derivative of $A_t^\dagger A_t = I$.

But if our aim is just to prove that G is a submanifold, all this work is not necessary, thanks to the following beautiful (and surprising) result of E. Cartan:

Fact: *Every (topologically) 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. For example, $\mathrm{SL}(n,\mathbb{R}) = \det^{-1}(1)$ is a submanifold simply because det is continuous.

Let us now give a few more examples of Lie groups, without detailed justifications.

Examples 2.6. (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) Given $B \in \mathrm{Mat}_n(\mathbb{R})$ let $G \subseteq \mathrm{GL}(n,\mathbb{R})$ be the invertible matrices commuting with B. This is a matrix Lie group with $\mathfrak{g}$ the set of *all* matrices commuting with B. (We may generalize by considering instead of just one B any collection of matrices B.)

(c) Let $\mathbb{H}$ 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

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

The invertible elements of $\mathbb{H}$ form a Lie group. Similarly, the unit sphere in $\mathbb{H} \cong \mathbb{R}^4$ is a Lie group.

(d) Consider a finite-dimensional associative algebra $\mathcal{A}$ over $\mathbb{R}$, with unit $1_{\mathcal{A}}$. Then the invertible elements $\mathcal{A}^\times$ are a Lie group. More generally, given $n \in \mathbb{N}$ 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 $n^2 \dim_{\mathbb{R}}(\mathcal{A})$. In particular, we have

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

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

(e) 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 $\mathrm{GL}(n,\mathbb{C}) \cap \mathrm{SL}(2n,\mathbb{R})$.)

(f) 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 $|x_1x_2| = |x_1||x_2|$, 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).

(g) 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.

(h) 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.

(i) We will show that the universal covering space of any Lie group G is a Lie group $\widetilde{G}$. The universal cover
$$\widetilde{\mathrm{SL}(2,\mathbb{R})}$$
is *not* isomorphic to a matrix Lie group. (We'll cover this in exercises.)

(j) Given a Riemannian manifold, the group of diffeomorphisms preserving the metric is a naturally Lie group (Myers-Steenrod theorem). Similarly for compact complex manifolds. These results are outside the scope of this course.

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

Definition 2.7. 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 2.8. (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.

The fact that $\mathfrak{o}(n),\ \mathfrak{sl}(n,\mathbb{R}),\ \ldots$ are closed under matrix commutator is not a coincidence. We have this phenomenon for any matrix Lie group:

Proposition 2.9. *Let $G \subseteq \mathrm{GL}(n,\mathbb{R})$ be a matrix Lie group. Then $\mathfrak{g} = T_I G$ is closed under matrix commutator, and so is a matrix Lie algebra.*

Proof. We use exponential charts. Given $B_1, B_2 \in \mathfrak{g}$, we have $tA, sB \in \mathfrak{g} \cap U$ for s,t close to 0. Hence $\exp(tA), \exp(sB) \in G \cap \exp(U)$, and also

$$\exp(tA)\exp(sB)\exp(-tA) \in G \cap \exp(U)$$

for s, t close to 0. For fixed t, this is a curve starting at $I \in G$, hence

$$\exp(tA)B\exp(-tA) = \frac{d}{ds}|_{s=0}(\exp(tA)\exp(sB)\exp(-tA)) \in \mathfrak{g}.$$

This is a curve in the subspace $\mathfrak{g}$; taking the t-derivative at $t = 0$ we obtain

$$AB - BA = \frac{d}{dt}|_{t=0}\,(\exp(tA)B\exp(-tA)) \in \mathfrak{g}$$

□

Another observation:

Proposition 2.10. *Let $G \subseteq \mathrm{GL}(n, \mathbb{R})$ be a matrix Lie group. Then $\mathfrak{g} = T_I G$ is closed under conjugation by $C \in G$:*

$$B \in \mathfrak{g} \Rightarrow CBC^{-1} \in \mathfrak{g}.$$

Proof. The conjugation action fixes I, hence its differential gives an action on $T_I\,\mathrm{GL}(n, \mathbb{R})$. This differential is computed by writing tangent vectors as $B = \frac{d}{dt}|_{t=0}A_t$ for some path $t \mapsto A_t \in G$ with $A_0 = I$

$$\frac{d}{dt}|_{t=0}CA_tC^{-1} = C(\frac{d}{dt}|_{t=0}A_t)C^{-1} = CBC^{-1},$$

so it is again just conjugation. Now, if $C \in G$ then the action on $T_I\,\mathrm{GL}(n, \mathbb{R})$ preserves the subspace $\mathfrak{g} = T_I G$. □

For example, if $A \in U(n)$ and $B \in \mathfrak{u}(n)$ then $ABA^{-1} \in \mathfrak{u}(n)$.

2.4. **Some 3-dimensional Lie groups.** 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 $\mathrm{SO}(3), \mathrm{SU}(2), \mathrm{SL}(2, \mathbb{R})$, and their topology.

The Lie group $\mathrm{SO}(3)$ 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. There is also the group $\mathrm{Sp}(1) \subseteq \mathrm{Mat}_1(\mathbb{H}) = \mathbb{H}$ of quaternions preserving the norm on $\mathbb{H}^1 = \mathbb{H}$. You may check that these are exactly the unit quaternions $S^3 \subseteq \mathbb{H}$, so we also have $\mathrm{Sp}(1) \cong S^3$ as a manifold. We may hence conjecture that

$$\mathrm{Sp}(1) \cong \mathrm{SU}(2)$$

as a Lie group. This turns out to be true: using the identification $\mathbb{H} \cong \mathbb{C}^2$, elements of $\mathbb{H} = \mathrm{Mat}_1(\mathbb{H})$ act $\mathbb{C}$-linearly on $\mathbb{C}^2$; this defines an algebra map $\mathbb{H} \to \mathrm{Mat}_2(\mathbb{C})$ taking $\mathrm{Sp}(1) \subseteq \mathbb{H}^\times$ to $\mathrm{SU}(2) \subseteq \mathrm{Mat}_2(\mathbb{C})^\times$. (Details are left as an exercise.) Let us now consider the Lie algebra

$$\mathfrak{su}(2) = T_I\,\mathrm{SU}(2) = \{B \in \mathrm{Mat}_2(\mathbb{C}) |\ B^\dagger = -B,\ \mathrm{tr}(B) = 0\}.$$

Writing

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

gives an identification $\mathfrak{su}(2) = \mathbb{R} \oplus \mathbb{C} = \mathbb{R}^3$. Note that for a matrix B of this form,

$$\det(B) = t^2 + |u|^2 = \tfrac{1}{2}||B||^2.$$

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

$$\ker(\phi) = \{I, -I\}.$$

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 local diffeomorphism, and hence is surjective (since SU(2) is compact, the image under ϕ is closed, since ϕ is a local diffeomorphism, the image is also open, since SO(3) is connected, the image is all of SO(3)). This exhibits $\mathrm{SU}(2) = S^3$ as the double cover of SO(3). In particular, $\mathrm{SO}(3) = S^3/\pm = \mathbb{R}P^3$ *as a manifold.*

Remark 2.11. The fact that SU(2) is the double cover of SO(3), and that it is isomorphic to Sp(1), is one of various instances of low-dimensional 'coincidences'. Another important example is the fact that $\mathrm{SU}(2) \times \mathrm{SU}(2)$ is a double cover of SO(4). This may be proved by looking at the action of $\mathrm{SU}(2) \times \mathrm{SU}(2)$ on $\mathbb{H} \cong \mathbb{R}^4$, given by

$$(A_1, A_2).B = A_1 B A_2^{-1}.$$

(Here $\mathbb{H}$ is viewed as a space of complex 2×2-matrices; alternatively we may view A_1, A_2 as elements of the group $\mathrm{Sp}(1) \subseteq \mathbb{H}$ of unit quaternions. (Details left as an exercise.)

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 SO(2) of rotations. Using polar decomposition of matrices, any matrix in $\mathrm{SL}(2, \mathbb{R})$ may be uniquely written as a product $A = A'A''$ where $A' \in \mathrm{SO}(2)$ while $A'' \in \mathrm{SL}(2, \mathbb{R})$ is a positive matrix. (One obtains A'' as the square root of the positive matrix $A^\top A$.) The A'''s obtained in this way are exactly the positive matrices of determinant 1. (Warning: this is not a subgroup.) It may be uniquely written as $\exp(B)$ where B is a symmetric matrix of trace 0. (The exponential map for matrices restricts to a diffeomorphism from symmetric matrices onto positive definite matrices.) In conclusion, we have

$$\mathrm{SL}(2, \mathbb{R}) = \mathrm{SO}(2) \times \{B \mid B^\top = B,\ \mathrm{tr}(B) = 0\} \cong S^1 \times \mathbb{R}^2$$

as a manifold (not as a group). This shows in particular that $\pi_1(\mathrm{SL}(2, \mathbb{R})) = \pi_1(S^1) = \mathbb{Z}$. Note that $\mathrm{SL}(2, \mathbb{R})$ has center $\{(I, -I)\}$. The quotient by this central subgroup is denoted $\mathrm{PSL}(2, \mathbb{R})$. Topologically, it is again just $S^1 \times \mathbb{R}^2$, but it is not isomorphic to $\mathrm{SL}(2, \mathbb{R})$ (since $\mathrm{PSL}(2, \mathbb{R})$ has trivial center).

2.5. **Universal covering of a Lie group.** More examples of Lie groups are obtained by taking (suitable) coverings of given Lie groups.

Remark 2.12 (A rapid review of covering spaces). 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 relation 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 $p^{-1}(x_0) \subseteq \widetilde{X}$ has a group structure given by the concatenation of paths

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

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 universal covering space is simply connected: its fundamental group is trivial.

The fundamental group acts on the covering space $\widetilde{X}$ by *deck transformations*, this action is again induced by concatenation of paths:

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

A continuous map of connected topological spaces $\Phi\colon X \to Y$ taking x_0 to the base point y_0 lifts to a continuous map $\widetilde{\Phi}\colon \widetilde{X} \to \widetilde{Y}$ of the covering spaces, by $\widetilde{\Phi}[\gamma] = [\Phi \circ \gamma]$, 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 2.13. *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 $p^{-1}(e)$ (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

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

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

Example 2.15 (Spin groups). 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'. Being double covers of compact groups, these are again compact groups.

The theorem shows in particular that the fundamental group of a connected Lie group is *abelian* (since it is a subgroup of the center of $\widetilde{G}$, and the center of a group is abelian. 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.

2.6. **A special case of Cartan's theorem.** We shall now prove Cartan's theorem for the special case of matrix Lie groups. By definition, matrix Lie groups are embedded submanifolds of $\mathrm{GL}(n,\mathbb{R})$, hence a closed subgroup of a matrix Lie group is a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$. Hence it suffices to show:

Theorem 2.16. *Every topologically closed subgroup of* $\mathrm{GL}(n,\mathbb{R})$ *is a matrix Lie group.*

We'll set up the proof in such a way that it immediately generalizes to arbitrary Lie groups (once we have developed a notion of exponential map for abstract Lie groups). We'll need:

Proposition 2.17 (Lie product formula)**.** *For* $A, B \in \mathrm{Mat}_n(\mathbb{R})$*, we have that*

$$\exp(A+B) = \lim_{n\to\infty} \left(\exp\left(\frac{1}{n}A\right)\exp\left(\frac{1}{n}B\right)\right)^n.$$

Proof. We use the trick from calculus: A limit $n \to \infty$ may often be computed as a limit $t \to 0^+$ by letting $t = \frac{1}{n}$.

$$\begin{aligned}
\lim_{n\to\infty} \left(\exp(\frac{1}{n}A)\exp(\frac{1}{n}B)\right)^n &= \lim_{n\to\infty} \exp\left(n \log\left(\exp(\frac{1}{n}A)\exp(\frac{1}{n}B)\right)\right) \\
&= \lim_{t\to 0} \exp\left(\frac{\log(\exp(tA)\exp(tB))}{t}\right) \\
&= \exp\left(\lim_{t\to 0} \frac{\log(\exp(tA)\exp(tB))}{t}\right) \\
&= \exp\left(\frac{d}{dt}\Big|_{t=0} \log\big(\exp(tA)\exp(tB)\big)\right) \\
&= \exp(A+B).
\end{aligned}$$

(Here we used that $\log'(1) = 1$.) ☐

Proof of Theorem. Let $G \subseteq \mathrm{GL}(n,\mathbb{R})$ be a closed subgroup. It suffices to construct a submanifold chart for G near I. (Using left translation, one then obtains submanifolds charts near any given $A \in G$.) That is, we want to find an open neighborhood U of $0 \in \mathrm{Mat}_n(\mathbb{R})$, a subspace $\mathfrak{g} \subseteq \mathrm{Mat}_n(\mathbb{R})$, and a diffeomorphism

$$\phi\colon U \to \exp(U)$$

such that

$$\phi(\mathfrak{g} \cap U) = G \cap \phi(U).$$

To begin, we need a candidate for $\mathfrak{g}$. Eventually this will be $\mathfrak{g} = T_I G$, but we cannot take this as a definition since we don't know yet that G is a submanifold. Instead we define

$$\mathfrak{g} = \{B \in \mathrm{Mat}_n(\mathbb{R}) | \ \forall t\colon \ \exp(tB) \in G\}.$$

To show that this is a vector subspace, we must show that it is closed under addition and scalar multiplication. The latter is obvious from the definition. For addition, let $B, C \in \mathfrak{g}$. For all $t \in \mathbb{R}$, the Lie product formula shows

$$\exp(t(B+C)) = \lim_{n\to\infty} \left(\exp(\frac{t}{n}B)\exp(\frac{t}{n}C)\right)^n$$

a limit of a sequence of elements $\left(\exp(\frac{t}{n}B)\exp(\frac{t}{n}C)\right)^n \in G$. Since G is topologically closed, it follows that $\exp(t(B+C)) \in G$. Hence $B + C \in \mathfrak{g}$.

Next, we need chart maps. Rather than using exponential charts, it is more convenient to use a modified version:

$$\phi\colon \ \mathrm{Mat}_n(\mathbb{R}) = \mathfrak{g} \oplus \mathfrak{g}^\perp \to \mathrm{GL}(n,\mathbb{R}), \ B' + B'' \mapsto \exp(B')\exp(B'')$$

for $B' \in \mathfrak{g}, \ B'' \in \mathfrak{g}^\perp$. (Here $\perp$ is the orthogonal relative to the standard inner product on $\mathrm{Mat}_n(\mathbb{R})$.) Note that $T_0\phi = \mathrm{id}$, so that ϕ restricts to a diffeomorphism on some open ϵ-ball $B_\epsilon(0)$ around 0.

Claim. For n sufficiently large, the restriction of ϕ to $U = \frac{1}{n} B_\epsilon(0)$ satisfies $\phi(U \cap \mathfrak{g}) = \phi(U) \cap G$.

Suppose the claim is false. Then none of the subsets $U_n = \frac{1}{n} B_\epsilon(0)$ would do the job. That is, $\phi(U_n \cap \mathfrak{g})$ is a proper subset of $\phi(U_n) \cap G$. We may hence choose $B_n \in U_n$ such that

$$B_n \notin \mathfrak{g}, \quad \phi(B_n) \in G.$$

The sequence satisfies $\lim_{n\to\infty} B_n = 0$. Writing $B_n = B'_n + B''_n \in \mathfrak{g}$, we have $B_n \notin \mathfrak{g}$ if and only if $B''_n \neq 0$, and $\phi(B_n) \in G$ if and only if $\exp(B''_n) \in G$. Replacing B_n with B''_n, we may constructed a sequence

$$B_n \in \mathfrak{g}^\perp, \ B_n \neq 0, \ \exp(B_n) \in G, \quad \lim_{n\to\infty} B_n = 0.$$

Passing to a subsequence, we may assume that the elements $B_n/||B_n||$ in the unit sphere converge: Hence there exists a unit vector $B \in \mathfrak{g}^\perp$ with

$$\lim_{n\to\infty} B_n/||B_n|| = B$$

We will show $B \in \mathfrak{g}$, hence producing a contradiction. By definition, $B \in \mathfrak{g}$ if and only if for any given $t \in \mathbb{R}$, we have $\exp(tB) \in G$. Choose integers a_n with $\lim_{n\to\infty}(a_n||B_n||) = t$.[3] Then

$$\exp(B_n)^{a_n} = \exp(a_n B_n) = \exp\left(a_n ||B_n|| \, \frac{B_n}{||B_n||}\right) \to \exp(tB).$$

Since G is closed, it follows that $\exp(tB) \in G$ for all t, so $B \in \mathfrak{g}$. $\square$

[3] A calculus exercise: Given a sequence of non-zero numbers $x_n \in \mathbb{R}$ with $\lim_{n\to\infty} x_n = 0$, and any given $x \in \mathbb{R}$, prove that there are integers $a_n \in \mathbb{Z}$ with $\lim_{n\to\infty} a_n x_n = x$.

2.7. **Some exercises.**

Exercise 2.18. 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 2.19. 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 2.20. 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 2.21. 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 2.20.

Exercise 2.22. Given an explicit description of the algebra morphism $\mathbb{H} \to \mathrm{Mat}_2(\mathbb{C})$, and show that it restricts to a Lie group isomorphism $\mathrm{Sp}(1) \to \mathrm{SU}(2)$.

Exercise 2.23. 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

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

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

3. The Lie algebra of a Lie group

3.1. **Review: Tangent vectors and vector fields.** We begin with a quick reminder of some manifold theory, partly just to set up our notational conventions. Let M be a manifold, and $C^\infty(M)$ its algebra of smooth real-valued functions.

(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 $TM = \sqcup_{m\in M} T_mM$ 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 a given vector field $X \in \mathfrak{X}(M)$ 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 [4][5]

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

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

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

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

[5] 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

$$\mathfrak{g} = T_eG$$

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 $T_gL_a\colon T_gG \to T_{ag}G$; in particular

$$T_eL_a\colon \mathfrak{g} \to T_aG.$$

Taken together, they define a vector bundle isomorphism

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

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

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

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 $\mathfrak{g} = T_eG$, 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 $(T_e L_{\phi(a)})\xi' = (T_a\phi)(T_e L_a(\xi))$ 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.

Remark 3.5. Another useful observation: Recall that left trivialization gives a vector bundle isomorphism $TG \cong G \times \mathfrak{g}$, with inverse map $(g, \xi) \mapsto \xi^L|_g$. Suppose $\phi\colon G \to G'$ is a morphism of Lie group. Then the following diagram commutes:

$$\begin{tikzcd} TG \arrow[r, "T\phi"'] \arrow[d] & TG' \arrow[d] \\ G \times \mathfrak{g} \arrow[r, "\phi \times T_e\phi"'] & G' \times \mathfrak{g}' \end{tikzcd}$$

This follows since ϕ intertwines left translation: $\phi(gh) = \phi(g)\phi(h)$, i.e. $\phi \circ L_g = L_{\phi(g)} \circ \phi$, and consequently

$$T\phi \circ TL_g = TL_{\phi(g)} \circ T\phi$$

In particular, the rank of ϕ is *constant*: the rank at any $g \in G$ coincides with the rank of $T_e\phi\colon \mathfrak{g} \to \mathfrak{g}'$.

3.3. **Properties of left-invariant and right-invariant vector fields.**

Definition 3.6. A *1-parameter subgroup* of a Lie group G is a Lie group morphism

$$\gamma\colon \mathbb{R} \to G$$

The term is a bit misleading since we do not require γ to be injective.

Example 3.7. If G is a matrix Lie group (a closed subgroup of $\mathrm{GL}(n, \mathbb{R})$), the curves $\gamma(t) = \exp(tB)$ with $B \in \mathfrak{g}$ are 1-parameter subgroups. We shall see shortly that all 1-parameter subgroups of G are of this form.

Theorem 3.8. *Let G be a Lie group, with Lie algebra $\mathfrak{g} = T_eG$.*
- (a) *The left-invariant vector fields ξ^L for $\xi \in \mathfrak{g}$ are complete.*
- (b) *The unique integral curve $\gamma^\xi(t)$ of ξ^L with initial condition $\gamma^\xi(0) = e$ is a 1-parameter subgroup.*
- (c) *The flow Φ^ξ of ξ^L is given by right translations:*

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

Proof. If

$$\gamma\colon J \to G$$

(with $J \subseteq \mathbb{R}$ an open interval around 0) is any integral curve of ξ^L, then its left translates $t \mapsto a\gamma(t)$ for any $a \in G$ are again integral curves. (Indeed, L_a takes integral curves of $X = \xi^L$ to integral curves for $(L_a)_*X = \xi^L$.) In particular, for given $t_0 \in J$ 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 the flow of ξ^L. With our sign conventions, 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 3.9. *Under group inversion,* $\xi^R \sim_{\mathrm{Inv}} -\xi^L, \quad \xi^L \sim_{\mathrm{Inv}} -\xi^R$.

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.

Proposition 3.10. *The 1-parameter subgroup $\gamma^\xi(t)$ is an integral curve for ξ^R as well. The flow of ξ^R is given by left translations,*

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

We also obtain the bracket relations between left and right invariant vector fields:

Proposition 3.11. *The left-invariant and right-invariant vector fields satisfy the bracket relations,*

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

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, while the flow of ξ^R is given by left translations. Since these flows commute, the 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 $\Phi^\xi_t,\ \Phi^\eta_t$ commute. The map $t \mapsto \Phi^\xi_t \circ \Phi^\eta_t$ is the flow of $\xi^L + \eta^L$. Hence it coincides with $\Phi^{\xi+\eta}_t$. 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 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. Let us write $\exp_G, \exp_H$ for the two exponential maps. For $\xi \in \mathfrak{g}$, the map $t \mapsto \phi(\exp_G(t\xi))$ is a 1-parameter subgroup of H, with differential at 0 given by

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

It follows that $\phi(\exp_G(t\xi)) = \exp_H(t(T_e\phi)(\xi))$ with the exponential map for H. Now put $t = 1$. $\square$

We may use this to verify that our definition of exponential map for general Lie groups generalizes the one for matrix Lie groups.

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,*

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

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 $t = 1$. $\square$

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 $\mathfrak{g} = T_IG$ 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

$$\begin{aligned}\exp(t\xi)\exp(t\eta)\exp(-t\xi)\exp(-t\eta) &= I + t^2(\xi\eta - \eta\xi) + O(t^3)\\ &= \exp\big(t^2(\xi\eta-\eta\xi)+O(t^3)\big)\end{aligned}$$

(here $O(t^k)$ indicates terms of order k or higher in the Taylor expansion). This formula relates the exponential map with the matrix commutator $\xi\eta - \eta\xi$. [6] A version of this formula holds for arbitrary Lie groups, as follows. For $g \in G$, let $\rho(g) = R_g^*$ as an operator on functions $f \in C^\infty(G)$ thus

$$\rho(g)(f)|_a = f(ag).$$

Since the flow of ξ^L is by right translations, $t \mapsto R_{\exp(-t\xi)}$, we have [7]

$$\frac{d}{dt}\Big(\rho(\exp t\xi)(f)|_a\Big) = \big(\rho(\exp t\xi)\circ \xi^L\big)(f)|_a.$$

Iterating,

$$\frac{d^k}{dt^k}\Big(\rho(\exp t\xi)(f)|_a\Big) = \big(\rho(\exp t\xi)\circ (\xi^L)^k\big)(f)|_a.$$

Hence, we may write the Taylor series:

$$\rho(\exp t\xi)(f)|_a = \Big(1 + t\xi^L + \frac{t^2}{2}(\xi^L)^2\Big)(f)|_a + O(t^3)$$

Using this formula, we obtain

$$\begin{aligned}&\rho\big(\exp(t\xi)\exp(t\eta)\exp(-t\xi)\exp(-t\eta)\big)(f)|_a\\ &\qquad= \rho\big(\exp(t\xi)\big)\rho\big(\exp(t\eta)\big)\rho\big(\exp(-t\xi)\big)\rho\big(\exp(-t\eta)\big)(f)|_a\\ &\qquad= \big(1 + t^2(\xi^L\eta^L - \eta^L\xi^L)\big)(f)|_a + O(t^3)\\ &\qquad= \big(1 + t^2[\xi,\eta]^L\big)(f)|_a + O(t^3)\end{aligned}$$

where the last line used the definition of Lie bracket of vector fields, $[X,Y] = XY - YX$ (as operators on functions). We'll prove the proposition by comparing the two results. Thus assume that $G = \mathrm{GL}(n,\mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{R})$. By the matrix calculation above,

$$\exp(t\xi)\exp(t\eta)\exp(-t\xi)\exp(-t\eta) = \exp(t^2(\xi\eta + \eta\xi)).$$

---

[6] 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.

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

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

Applying the resulting operators to f and evaluating,

$$\rho\big(\exp(t\xi)\exp(t\eta)\exp(-t\xi)\exp(-t\eta)\big)(f)|_a = \Big(1 + t^2(\xi\eta - \eta\xi)^L\Big)(f)|_a + O(t^3).$$

Comparing coefficients, we obtain $[\xi,\eta]^L = (\xi\eta - \eta\xi)^L$, and finally $[\xi,\eta] = \xi\eta - \eta\xi$. ☐

Remark 4.8. This proves (again) 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. [8]

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 $U \to \exp(U)$.

Proof. We calculate, for $\xi \in \mathfrak{g}$,

$$(T_0 \exp)(\xi) = (T_0 \exp)\big(\frac{d}{dt}\Big|_{t=0} t\xi\big) = \frac{d}{dt}\Big|_{t=0} \exp(t\xi) = \xi,$$

where the last equality follows by definition of the exponential map. ☐

Theorem 4.11 (E. Cartan's theorem)**.** *Let G be a closed subgroup of a Lie group H. Then G is an embedded submanifold, and hence is a Lie subgroup.*

Having the exponential map at our disposal, it is pretty straightforward to generalize the proof for matrix Lie groups to the general case. We leave the details as an exercise.

Remark 4.12. According to the proof, the Lie algebra $\mathfrak{g}$ of G is recognized as the set of all $\xi \in \mathfrak{h}$ such that $\exp t\xi \in G$ for all $t \in \mathbb{R}$.

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 4.13.
(a) Matrix Lie groups are, by definition, closed subgroups of $\mathrm{GL}(n,\mathbb{R})$ for some $n \in \mathbb{N}$, and hence are Lie groups. (We proved this already in Theorem 2.16.)
(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.

[8] Nonetheless, some authors use that convention.

(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.

Theorem 4.14. *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. Suppose the group morphism $\phi\colon G \to H$ is continuous. Then

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

is a closed subspace. Since ϕ is a group morphism, $\mathrm{Gr}(\phi)$ is a subgroup of $H \times G$. By Cartan's theorem, it is an embedded submanifold of $H \times G$, and so is a Lie subgroup. Projection to the second factor $H \times G \to G$ restricts to a Lie group morphism

$$\psi\colon\ \mathrm{Gr}(\phi) \to G, \quad (\phi(g), g) \mapsto g$$

which furthermore is a bijection. But any morphism of Lie groups has constant rank (see Remark 3.5), hence this bijection is a diffeomorphism. [9] The map ϕ is the composition of smooth maps

$$G \xrightarrow{\psi^{-1}} \mathrm{Gr}(\phi) \hookrightarrow H \times G \xrightarrow{\mathrm{pr}_1} H$$

hence it is smooth. $\square$

[9]By the constant rank theorem, if a smooth map $\phi\colon M \to M'$ has constant rank r, its non-empty fibers are embedded submanifolds of dimension $\dim M - r$. Thus, if ϕ is a constant rank map which is also a bijection, then ϕ must be a diffeomorphism.

Exercise 4.15. Show that the exponential maps 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 4.16. 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 4.17. 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 4.18. 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.) What about the exponential map for $\mathrm{GL}^+(2,\mathbb{R})$ (invertible matrices with positive determinant)?

Exercise 4.19. Let $\exp\colon \mathfrak{g} \to G$ be the exponential map for a Lie group G. Prove the *Lie product formula*

$$\exp(\xi + \eta) = \lim_{n\to\infty} \Big(\exp(\frac{1}{n}\xi)\exp(\frac{1}{n}\eta)\Big)^n,$$

by using the same argument as for matrix Lie groups.

Exercise 4.20. Prove Cartan's theorem: If H is a Lie group, and G is a subgroup that is topologically closed, then H is an embedded submanifold (and in particular is a Lie subgroup). (Just inspect the proof for matrix Lie groups, and observe that it generalizes line-by-line.)

Exercise 4.21. Prove that a topological group carries at most one smooth structure for which it is a Lie group.

5. Automorphisms, adjoint actions

Given any category (such as the category of associative algebras, Lie algebras, Lie groups, topological spaces, etc), the automorphism group of an object of that category consists of invertible morphisms from the object to itself. We shall consider the automorphisms (and infinitesimal automorphisms) of Lie algebras and Lie groups.

5.1. **Automorphisms and derivations of Lie algebras.** If $\mathfrak{g}$ is a Lie algebra we have the group $\mathrm{Aut}(\mathfrak{g})$ of Lie algebra automorphisms. Concretely, $\mathrm{Aut}(\mathfrak{g})$ consists of all invertible linear maps $A\colon \mathfrak{g} \to \mathfrak{g}$ such that

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

for all $\xi, \eta \in \mathfrak{g}$. If $\mathfrak{g}$ is finite-dimensional, then $\mathrm{Aut}(\mathfrak{g})$ is closed in the group $\mathrm{GL}(\mathfrak{g})$ of vector space automorphisms, hence, by Cartan's theorem, it is a Lie group. To identify its Lie algebra, we introduce:

Definition 5.1. A derivation of a Lie algebra $\mathfrak{g}$ is a linear map $D\colon \mathfrak{g} \to \mathfrak{g}$ such that

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

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

The space $\mathrm{Der}(\mathfrak{g})$ of derivations is a Lie algebra , with bracket given by commutation.

Proposition 5.2. *For any finite-dimensional Lie algebra $\mathfrak{g}$, the Lie algebra of the Lie 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. $\square$

An example of a Lie algebra derivation is the operator $\mathrm{ad}_\xi = [\xi, \cdot]$ given by bracket with a fixed Lie algebra element ξ; the derivation property is the Jacobi identity of $\mathfrak{g}$. This satisfies

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

(again by the Jacobi identity).

Definition 5.3. The *adjoint representation* of a Lie algebra is the Lie algebra morphism
$$\mathrm{ad}\colon \mathfrak{g} \to \mathrm{Der}(\mathfrak{g}), \quad \xi \mapsto \mathrm{ad}_\xi\,.$$
Elements in the image of this map are called *inner derivations*; the space of inner derivations is denoted $\mathrm{Inn}(\mathfrak{g})$.

Note that the kernel of adjoint representation is the center $\mathrm{Cent}(\mathfrak{g})$.

Lemma 5.4. *The space* $\mathrm{Inn}(\mathfrak{g}) \subseteq \mathrm{Der}(\mathfrak{g})$ *is a normal Lie subalgebra. That is,* $[\mathrm{Der}(\mathfrak{g}), \mathrm{Inn}(\mathfrak{g})] \subseteq \mathrm{Inn}(\mathfrak{g})$

Proof. If $D \in \mathrm{Der}(\mathfrak{g})$ and $\xi \in \mathfrak{g}$ then
$$[D, \mathrm{ad}_\xi] = \mathrm{ad}_{D\xi}$$
by the calculation
$$D(\mathrm{ad}_\xi \eta) = D([\xi, \eta]) = [D\xi, \eta] + [\xi, D\eta] = \mathrm{ad}_{D\xi} \eta + \mathrm{ad}_\xi(D\eta).$$
□

For any normal Lie subalgebra, the Lie bracket descends to the quotient space. (Exercise) The quotient space
$$\mathrm{Out}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g})/\,\mathrm{Inn}(\mathfrak{g})$$
is the Lie algebra of *outer derivations*.

Example 5.5. For $\mathfrak{g} = \mathfrak{so}(3)$ one can check that every Lie algebra derivation is inner. Thus $\mathrm{Out}(\mathfrak{g}) = 0$ in this case. On the other hand, if $\mathfrak{g}$ is abelian (the bracket is trivial), then the space of inner derivations is 0, and $\mathrm{Der}(\mathfrak{g}) = \mathrm{Out}(\mathfrak{g})$ are all linear maps $\mathfrak{g} \to \mathfrak{g}$.

5.2. **Automorphisms of G.** The group $\mathrm{Aut}(G)$ of automorphisms of a Lie group G is the group of invertible Lie group morphisms from G to itself. Any $a \in G$ defines an automorphism Ad_a 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.

Definition 5.6. The *adjoint action* of a Lie group G on itself is the group morphism
$$\mathrm{Ad}\colon G \to \mathrm{Aut}(G), \quad a \mapsto \mathrm{Ad}_a\,.$$
Elements in the image of this map are called *inner automorphisms*; the space of inner automorphisms is denoted $\mathrm{Inn}(G)$.

Note that the kernel of adjoint representation is the center $\mathrm{Cent}(G)$.

Lemma 5.7. *The subgroup* $\mathrm{Inn}(G) \subseteq \mathrm{Aut}(G)$ *of inner automorphisms is a normal subgroup.*

Proof. For any $\phi \in \mathrm{Aut}(G)$, and any $a \in G$,

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

□

The quotient group $\mathrm{Out}(G) = \mathrm{Aut}(G)/\,\mathrm{Inn}(G)$ is called the *outer automorphism group.*

Example 5.8. 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 5.9. 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.

Remark 5.10. Let G be a Lie group, G_0 its identity component (which is a normal subgroup), and G/G_0 its group of components. By a theorem of Hochschild, $\mathrm{Aut}(G)$ is a Lie group provided that G/G_0 is finitely generated. [10] On the other hand, if G is a countable discrete group, the group $\mathrm{Aut}(G)$ is discrete but need not be countable. An example is the group $G = \mathrm{Map}_{fin}(X, \mathbb{Z}_3)$ of finitely supported maps for a countable space X (you may take $X = \mathbb{N}$); its automorphism group contains $\mathrm{Map}(X, \mathrm{Aut}(\mathbb{Z}_3)) = \mathrm{Map}(X, \mathbb{Z}_2)$, which is not countable. (Adapted from a discussion on math stack exchange.)

5.3. **The adjoint action of G on $\mathfrak{g}$.** Suppose G is a Lie group, with Lie algebra $\mathfrak{g}$. For any Lie group automorphism $\phi \in \mathrm{Aut}(G)$, the differential at the group units is a Lie algebra automorphism $T_e\phi \in \mathrm{Aut}(\mathfrak{g})$. This differentiation map

$$\mathrm{Aut}(G) \to \mathrm{Aut}(\mathfrak{g}), \quad \phi \mapsto T_e\phi$$

is itself a (Lie) group morphism. As a special case, we may apply this to the adjoint action $G \to \mathrm{Aut}(G)$. By composition, we obtain a Lie group morphism, which is again denoted by Ad (as opposed to $T_e \mathrm{Ad}$)

$$\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathfrak{g}),\ a \mapsto \mathrm{Ad}_a = T_e \mathrm{Ad}_a .$$

It is called the *adjoint representation of the Lie group G* on $\mathfrak{g}$.

At this stage, we have constructed three kinds of adjoint actions[11]

$$\mathrm{ad}\colon \mathfrak{g} \to \mathrm{Der}(\mathfrak{g}), \quad \mathrm{Ad}\colon G \to \mathrm{Aut}(G), \quad \mathrm{Ad}\colon G \to \mathrm{Aut}(\mathfrak{g}).$$

[10] `https://www.ams.org/journals/tran/1952-072-02/S0002-9947-1952-0045735-2/S0002-9947-1952-0045735-2.pdf`

[11] We might add a forth one, by letting $\mathfrak{aut}(G)$ be the space of vector fields $X \in \mathfrak{X}(X)$ that are 'infinitesimally multiplicative' in the sense that $(X,X) \sim_{\mathrm{Mult}_G} X$. The map $\mathfrak{g} \to \mathfrak{aut}(G),\ \xi \mapsto \xi^L - \xi^R$ is a Lie algebra morphism.

Let us explore some of their properties.

Proposition 5.11. *We have*

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

for $\xi \in \mathfrak{g}$, $a \in G$.

Proof. This is just the functoriality of exp with respect to Lie group morphisms, applied to the Lie group morphism $\mathrm{Ad}_a\colon G \to G$. ☐

As a consequence, we see that 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.

Theorem 5.12. *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 on* $\mathfrak{g}$*,*

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

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

Proof. For the first part we have to show

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

There is a quick proof if G is a matrix Lie group, using $\exp(t\xi) = 1 + t\xi + O(t^2)$:

$$\mathrm{Ad}_{\exp(t\xi)} \eta = \exp(t\xi)\eta\exp(-t\xi) = \eta + t\underbrace{(\xi\eta - \eta\xi)}_{\mathrm{ad}_\xi \eta} + O(t^2).$$

The same idea works in general, once we realize everything as operators. We thus use again the representation $\rho(g) = R_g^*$ on $C^\infty(G)$. Recall

$$\rho(\exp(t\xi)) = 1 + t\xi^L + O(t^2).$$

We claim that

$$(\mathrm{Ad}_g \eta)^L = \rho(g) \circ \eta^L \circ \rho(g)^{-1} \tag{1}$$

as operators on $C^\infty(G)$. This follows by a calculation (where you should think of each line as being applied to a function $f \in C^\infty(G)$ and evaluated at some $a \in G$):

$$\begin{aligned}
(\mathrm{Ad}_g\,\eta)^L &= \frac{d}{dt}\Big|_{t=0}\left(\rho\left(\exp(t\,\mathrm{Ad}_g\,\eta)\right)\right)\\
&= \frac{d}{dt}\Big|_{t=0}\left(\rho(\mathrm{Ad}_g\exp(t\eta))\right)\\
&= \frac{d}{dt}\Big|_{t=0}\left(\rho(g)\rho(\exp(t\eta)\rho(g)^{-1}\right)\\
&= \rho(g)\circ\eta^L\circ\rho(g)^{-1}
\end{aligned}$$

With (1) in place, the calculation is essentially the same as for matrices:

$$\mathrm{Ad}_{\exp(t\xi)}\,\eta)^L = \varrho(\exp \mathrm{t}\xi)\eta^L\varrho(\exp(-t\xi)) = t(\xi^L\eta^L - \eta^L\xi^L) + O(t^2) = t[\xi,\eta]^L + O(t^2).$$

The second part is the commutativity of the diagram

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

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

Remark 5.13. 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^jCB^{n-j}, and hence must vanish if $n \geq 2N$.)

6. The differential of the exponential map

6.1. **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 move $T_{\exp\xi}G$ back to $T_eG = \mathfrak{g}$, this is given by an endomorphism of $\mathfrak{g}$.

Theorem 6.1. *Using left translation to identify $T_gG \cong \mathfrak{g}$ for all $g \in G$, The differential of the exponential map* $\exp\colon \mathfrak{g} \to G$ *at* $\xi \in \mathfrak{g}$ *is the linear map*

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

given by the operator $\phi(\mathrm{ad}_\xi)\colon \mathfrak{g} \to \mathfrak{g}$*, where* ϕ *is the entire holomorphic function*

$$\phi(z) = \frac{1-e^{-z}}{z} = \int_0^1 ds \; \exp(-sz).$$

Since the Taylor series for ϕ has infinite radius of convergence, the operator $\phi(\mathrm{ad}_\xi)$ is defined by the power series. Spelling out the left trivialization, the theorem says that the diagram

$$\begin{tikzcd} T_\xi\mathfrak{g} \arrow[r, "T_\xi \exp"] \arrow[d, "="'] & T_{\exp\xi}G \arrow[d, "\cong"', "T_{\exp\xi}L_{\exp(-\xi)}"] \\ \mathfrak{g} \arrow[r, "\phi(\mathrm{ad}_\xi)"'] & \mathfrak{g} \end{tikzcd}$$

commutes.

Proof. We want to show that for all ξ, η

$$T_\xi(L_{\exp(-\xi)} \circ \exp)(\eta) = \phi(\mathrm{ad}_\xi)\eta,$$

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

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

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

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

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

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

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

$$\frac{d}{dt}\Big|_{t=0} f\big(\exp(-\xi)\ \exp(\xi + t\eta)\big) = (\phi(\mathrm{ad}_\xi)\eta)(f)$$

as desired. $\square$

Corollary 6.2. *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 $\phi(\mathrm{ad}_\xi) = \frac{1-\exp(-\mathrm{ad}_\xi)}{\mathrm{ad}_\xi}$ is invertible, i.e. has non-zero determinant. The determinant is given in terms of the eigenvalues of ad_ξ as a product, $\prod_\lambda \phi(\lambda)$. But the zeroes of ϕ are exactly the elements of $2\pi i\mathbb{Z}\backslash\{0\}$. $\square$

6.2. **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 6.3. *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 $\frac{z}{1-e^{-z}} = \frac{z}{e^z-1} + z$.

6.3. **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, for fixed ξ, η sufficiently small) be defined by

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

Note

$$f(0) = \xi, \quad f(1) = \log(\exp \xi \exp \eta).$$

The plan is to get an equation for the t-derivative of f, and integrate.

Lemma 6.4. *f satisfies the differential equation*

$$\frac{df}{dt} = \phi(\mathrm{ad}_f)^{-1} \eta = \frac{\mathrm{ad}_f}{1 - e^{-\mathrm{ad}_f}} \eta.$$

Proof. By the formula for the differential of exp,

$$(T_e L_{\exp(f)})^{-1} \frac{d}{dt} \exp(f) = (T_e L_{\exp(f)})^{-1} (T_f \exp)(\frac{df}{dt}) = \phi(\mathrm{ad}_f)(\frac{df}{dt}).$$

On the other hand, since $\exp(f) = \exp(\xi) \exp(t\eta)$ we have

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

This proves $\phi(\mathrm{ad}_f) \frac{df}{dt} = \eta$. $\square$

Lemma 6.5. *We have*

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

with the function

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

defined on the unit disk in $\mathbb{C}$ *around* $w = 1$.

Proof. Using $\chi(w) = \phi(\log(w))^{-1}$, the previous lemma gives

$$\frac{df}{dt} = \chi(e^{\mathrm{ad}_f})\eta = \chi(e^{\mathrm{ad}_\xi} e^{t\,\mathrm{ad}_\eta})\eta.$$

The lemma follows by integrating from 0 to 1 and using $f(0) = \xi,\ f(1) = \log(\exp(\xi)\exp(\eta))$ $\square$

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 6.6 (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 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 6.7. 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.

Remark 6.8. In principle, the BCH formula allows us to reconstruct the Lie group multiplication from the Lie bracket, at least near the group unit.

7. Actions of Lie groups and Lie algebras

Groups often arise through their actions on sets. Similarly, Lie groups often arise through their actions on manifolds.

7.1. **Lie group actions.**

Definition 7.1. 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 7.2.

(a) An $\mathbb{R}$-action on M is the same thing as a (global) flow $t \mapsto \Phi_t$.
(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 $\mathbb{R}^k \subseteq \mathbb{R}^n$.)

7.2. **Lie algebra actions.**

Definition 7.3. 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 *generating 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

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

where the horizontal maps are the action maps.

Examples 7.4. (a) Consider $\mathfrak{g} = \mathbb{R}$ with the zero bracket. An Lie algebra action of $\mathbb{R}$ on M is the same thing as a vector field $X \in \mathfrak{X}(M)$, via $\mathbb{R} \to \mathfrak{X}(M),\ \lambda \mapsto \lambda X$.

(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 $f \in C^\infty(V)$,

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

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

$$\xi_V = -\sum_{ab} (\phi(\xi))_a^b \; x^a \frac{\partial}{\partial x^b}.$$

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 $\mathfrak{g}$-equivariant.

7.3. **Differentiating Lie group actions.** Informally, we expect that every Lie group action $\mathcal{A}\colon G \to \mathrm{Diff}(M)$ differentiates to a Lie algebra action $\mathcal{A}\colon \mathfrak{g} \to \mathfrak{X}(M)$, since $\mathfrak{X}(M)$ is informally the Lie algebra of $\mathrm{Diff}(M)$. Since $\mathrm{Diff}(M)$ is infinite-dimensional, a more detailed discussion is required.

Definition 7.5. 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

$$\mathcal{A}_\xi = \xi_M = \frac{d}{dt}\Big|_{t=0} \mathcal{A}^*_{\exp(-t\xi)}$$

(thinking of both sides as operators on $C^\infty(M)$).

Example 7.6. 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:

$$\xi_{M_1} \sim_\Phi \xi_{M_2}.$$

Theorem 7.7. *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. We know this fact for a special case: The generating vector fields for the action of G on itself by left multiplication are minus the right invariant vector fields, and $\xi \mapsto -\xi^R$ is indeed a Lie algebra morphism.

For the general case, note that the action map

$$\Phi\colon G \times M \to M,\ (a, m) \mapsto a.m$$

is G-equivariant, relative to the action $g.(a, m) = (ga, m)$ on $G \times M$ and the given G-action on 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

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

In turn, this implies $[(\xi_1)_M, (\xi_2)_M] - [\xi_1, \xi_2]_M = 0$. □

7.4. **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*. [12] 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 7.8 (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.

[12] We will skip this material in class, and will state Theorem 7.11 without proof.

Exercise 7.9. 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 7.10. 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 7.11. *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

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

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 7.12. 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 7.13. *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 of H on G commutes with right multiplication of G:

$$\mathcal{A}_h(ga) = \mathcal{A}_h(g)h.$$

Hence, letting

$$\psi(h) = \mathcal{A}_h(e)$$

we have that $\mathcal{A}_h(g) = \psi(h)g$. The action property now shows

$$\psi(h_1)\psi(h_2) = \mathcal{A}_{h_1}(\psi(h_2)) = \mathcal{A}_{h_1}(\mathcal{A}_{h_2}(e)) = \mathcal{A}_{h_1 h_2}(e) = \psi(h_1 h_2),$$

so that $\psi \colon H \to G$ is a Lie group morphism integrating ϕ. $\square$

As a special case, we have:

Corollary 7.14. *Let G be a connected, simply connected Lie group, with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$, and V a finiote-dimensional vector space. Then there is a 1-1 correspondence between representations of the Lie algebra $\mathfrak{g}$ on V and representations of the Lie group G on V.*

That is, every $\mathfrak{g}$-representation on V integrates uniquely to a G-representation.

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 $G \to \mathrm{GL}(V)$. □

Remark 7.15. Again, this is not true without the assumption that G is simply connected. If G is connected, then a $\mathfrak{g}$-representation on V always integrates to a representation of the universal cover $\widetilde{G} \to \mathrm{GL}(V)$. The latter descends to a G-representation if and only if the kernel of the covering map $\widetilde{G} \to G$ acts trivially. A simple example where this is not the case is the defining representation of $\mathrm{SU}(2)$ on $\mathbb{C}^2$.

The following result is about Lie subgroups. One might define a Lie subgroup of a Lie group G to be a subgroup which is also a submanifold. Indeed, some authors[13] take this as the definition. However, it is common to use the following more flexible approach.

Definition 7.16. A *Lie subgroup* of a Lie group G is a subgroup $H \subseteq G$, equipped with a Lie group structure such that the inclusion

$$i \colon H \to G$$

is a morphism of Lie groups.

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

Proposition 7.17. *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. □

[13]For example, Kirillov Jr: *An Introduction to Lie Groups and Lie Algebras*

By *Ado's theorem*, any finite-dimensional Lie algebra $\mathfrak{g}$ (over a a field of characteristic zero, in our case $\mathbb{R}$) is isomorphic to a matrix Lie algebra. We will skip the proof of this important (but relatively hard) result, since it involves a considerable amount of structure theory of Lie algebras. [14]

Remark 7.18. 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 7.19 (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.

Exercise 7.20. Show that if $H_1, H_2 \subseteq G$ are two Lie subgroups, then the intersection $H_1 \cap H_2$ is again a Lie subgroup. Hint: To define a Lie group structure on $H_1 \cap H_2$, realize it as a closed subgroup of $H_1 \times H_2$.

8. Basic properties of compact Lie groups

The theory of Lie groups is particularly well-developed for the case of *compact* Lie groups. In this section we will prove the following special features of compact Lie groups G and their Lie algebras $\mathfrak{g}$:

(i) existence of a bi-invariant positive measure on G,
(ii) existence of a G-invariant inner product on $\mathfrak{g}$,
(iii) decomposition of $\mathfrak{g}$ into center and simple ideals,
(iv) complete reducibility of G-representations,
(v) surjectivity of the exponential map.

8.1. Modular function.

Let G be a Lie group, with Lie algebra $\mathfrak{g}$, and $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$ its adjoint representation.

[14] There are well-written expository articles on this topic available on the internet.

Definition 8.1. The *modular function* of a Lie group is the function

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

A Lie group G whose modular function is trivial is called *unimodular*.

Note that χ is a Lie group morphism. Let us compute the corresponding Lie algebra morphism.

Lemma 8.2. *The differential of χ at the group unit is given by*

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

Proof. We calculate

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

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. ☐

If G is unimodular, it follows that the Lie algebra morphism $\mathfrak{g} \to \mathbb{R},\ \xi \mapsto \mathrm{tr}(\mathrm{ad}_\xi)$ is trivial. The converse holds if G is connected. Generally, a Lie algebra $\mathfrak{g}$ is called unimodular if $\mathrm{tr}(\mathrm{ad}_\xi) = 0$ for all $\xi \in \mathfrak{g}$.

Proposition 8.3. *Compact Lie groups are unimodular.*

Proof. The image of the Lie group morphism $G \to \mathbb{R}_{>0},\ g \mapsto |\det_{\mathfrak{g}}(\mathrm{Ad}_g)|$ (as an image of a compact set under a continuous map) is compact. But the only compact subgroup of $\mathbb{R}_{>0}$ is the trivial subgroup $\{1\}$. ☐

Remark 8.4. There are many other examples of unimodular Lie algebras (and corresponding unimodular Lie groups).

(a) If $\mathfrak{g}$ is *perfect* i.e. $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, then any $\xi \in \mathfrak{g}$ can be written as a sum of Lie brackets $\xi = \sum_i [\eta_i, \zeta_i]$. 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. Examples of perfect Lie algebra include semi-simple Lie algebras (to be discussed later) such as $\mathfrak{sl}(n, \mathbb{R})$), but also for instance $\mathfrak{g} = \mathbb{R}^3 \rtimes \mathfrak{so}(3)$, with bracket $[(v, X), (w, Y)] = (Xw - Yv, [X, Y])$.

(b) If $\mathfrak{g}$ is *nilpotent* (i.e., ad_ξ is a nilpotent operator for all $\xi \in \mathfrak{g}$, then $\mathfrak{g}$ is unimodular. This follows since the eigenvalues of a nilpotent operator are all zero, thus $\mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}_\xi) = 0$. Examples of nilpotent Lie algebras include abelian Lie algebras, and strictly upper triangular matrices.

An example of a non-unimodular Lie group is the following

Exercise 8.5. Let G be the identity component of the affine group of the real line, consisting of transformations $x \mapsto tx + s$ with $t > 0$ and $s \in \mathbb{R}$. It may be realized as a matrices of the form

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

with $t > 0$ and $s \in \mathbb{R}$. Compute the modular function to show that G is not unimodular.

8.2. **Densities.** The modular function has a geometric interpretation in terms of densities. A density on a manifold is locally defined as an expression $f(x)|\mathrm{d}x|$ with f a smooth function, under a coordinate change $x = \Phi(y)$ it transforms as

$$\Phi^*\big(f(x)|\mathrm{d}x|\big) = |\det J(y)|\ f(\Phi(y))\,|\mathrm{d}y|$$

where $J(y) = \frac{\partial x_i}{\partial y_j}$ is the Jacobian matrix. The more formal definition is as follows.

Definition 8.6. Let E be a vector space of dimension n. A density on E is a map

$$\mathsf{m}\colon E \times \cdots \times E \to \mathbb{R}$$

satisfying

$$\mathsf{m}(Av_1, \ldots, Av_n) = |\det(A)|\,\mathsf{m}(v_1, \ldots, v_n) \tag{2}$$

for all $A \in \mathrm{GL}(E)$. The vector space of all densities is denoted $|\det|(E^*)$.

Lemma 8.7. *The space* $|\det|(E^*)$ *is 1-dimensional, and has a canonical orientation.*

Proof. A density is uniquely determined by its values on a given basis $v_1, \ldots, v_n$ of V. Indeed, if $v_1', \ldots, v_n' \in V$ are arbitrary, one has a unique linear map A such that $Av_i = v_i'$, and so the formula gives $\mathsf{m}(v_1', \ldots, v_n')$ in terms of $\mathsf{m}(v_1, \ldots, v_n)$. This shows that $\dim |\det|(E^*) = 1$. The orientation is defined by declaring that a density is positive if $\mathsf{m}(v_1, \ldots, v_n) > 0$ on some (hence, every) basis. ☐

Given an isomorphism of vector spaces $T\colon E' \to E$, one defines $T^*\colon |\det|(E^*) \to |\det|((E')^*)$ by

$$(T^*\mathsf{m})(w_1, \ldots, w_n) = \mathsf{m}(Tv_1, \ldots, Tv_n).$$

In particular, for $E' = E$ we have $A^*\mathsf{m} = |\det(A)|\,\mathsf{m}$ (by definition of a density).

Remark 8.8. Given an orientation on E, there is a canonical isomorphism between $|\det|(E^*)$ with the space $\det(E^*) = \wedge^n E^*$. The isomorphism takes a non-zero element $\Lambda \in \wedge^n E^*$ (i.e, Λ is a *volume form*) to the unique density $\mathsf{m} = |\Lambda|$ such that $\mathsf{m}(v_1, \ldots, v_n) = \Lambda(v_1, \ldots, v_n)$ whenever $v_1, \ldots, v_n$ is an oriented basis.

Remark 8.9. On $\mathbb{R}^n$, we have the standard density defined by $\mathsf{m}(e_1, \ldots, e_n) = 1$, where $e_1, \ldots, e_n$ is the standard basis. This density is typically denoted $|\mathrm{d}x|$.

Exercise 8.10. Show that
$$|\det|(E^*) \cong |\det|(E)^*$$
where $|\det|(E)$ is defined as $|\det|((E^*)^*)$.

Exercise 8.11. Prove that if $F \subseteq E$ is a subspace, with quotient space E/F, there is a canonical isomorphism
$$|\det|((E/F)) \cong |\det|(E) \otimes |\det|(F)^*.$$
More generally, for an exact sequence of vector spaces $0 \to E_1 \to E_2 \to E_3 \to \cdots \to E_r \to 0$ of vector spaces, there is a canonical isomorphism
$$|\det|(E_1) \otimes |\det|(E_2)^* \otimes |\det|(E_3) \otimes \cdots \cong \mathbb{R}$$
(with duals for the even indices).

For a manifold M, one obtains an oriented real line bundle
$$|\det|(T^*M) \to M$$
with fibers $|\det|(T^*_m M)$.

Definition 8.12. A *density* on M is a section of the density bundle $|\det|(T^*M)$. Densities on manifolds are also called *smooth measures*.

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

Remark 8.13. Being defined as sections of a vector bundle, the space of densities is a module over the algebra of functions $C^\infty(M)$: if m is a density then so is $f\mathsf{m}$. The choice of a fixed *positive* density m_0 on M trivializes $|\det|(T^*M)$, hence any density on M is of the form $\mathsf{m} = f\mathsf{m}_0$ with f a function.

On $\mathbb{R}^n$, we have the standard 'Lebesgue measure' $|\mathrm{d}x|$ 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.

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.) Under a diffeomorphism $\Phi\colon M' \to M$ of manifolds, we have

$$\int_{M'} \Phi^* \mathsf{m} = \int_M \mathsf{m}.$$

Given a G-action on M, a density is called *left-invariant* if $\mathcal{A}_g^* \mathsf{m} = \mathsf{m}$ for all $g \in G$. In particular, we can look for left-invariant measures 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 m_e can be uniquely extended to a left-invariant density. That is,

$$\Gamma\Big(|\det|(T^*G)\Big)^L \cong |\det|(\mathfrak{g}^*).$$

In particular, the space of left-invariant densities on G is 1-dimensional:

Proposition 8.14. *Let G be a Lie group, and $\chi\colon G \to \mathbb{R}_{>0}$ its modular function. If m is a left-invariant smooth density, we have*

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

for all $a \in G$.

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

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

Computing at the group unit e, we see that $\mathrm{Ad}_{a^{-1}}^* \mathsf{m}_e = |\det(\mathrm{Ad}_{a^{-1}} | \mathsf{m}_e$. ☐

Corollary 8.15. *A Lie group is unimodular if and only if the left-invariant densities are also right invariant.*

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

Corollary 8.16. *The Lie algebra $\mathfrak{g}$ of any compact G admits an* Ad*-invariant inner product B. That is,*

$$B(\mathrm{Ad}_a \xi, \mathrm{Ad}_a \eta) = B(\xi, \eta) \tag{3}$$

Proof. 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 of the inner product follows from the bi-invariance of the measure: Letting $f(g) = B'(\mathrm{Ad}_g(\xi), \mathrm{Ad}_g(\zeta))$ we have

$$\begin{aligned} B(\mathrm{Ad}_a \xi, \mathrm{Ad}_a \eta) &= \frac{1}{\mathrm{vol}(G)} \int_G B'(\mathrm{Ad}_{ga}(\xi), \mathrm{Ad}_{ga}(\zeta))\ |\mathrm{d}g|. \\ &= \frac{1}{\mathrm{vol}(G)} \int_G (R_a^* f)\ |\mathrm{d}g| \\ &= \frac{1}{\mathrm{vol}(G)} \int_G R_a^*(f|\mathrm{d}g|) \\ &= \frac{1}{\mathrm{vol}(G)} \int_G f|\mathrm{d}g| \\ &= B(\xi,\eta). \end{aligned}$$

□

Definition 8.17. A symmetric bilinear form B on a Lie algebra $\mathfrak{g}$ is called ad-invariant if

$$B([\xi,\eta],\zeta) + B(\eta,[\xi,\zeta]) = 0. \tag{4}$$

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

Exercise 8.18. Prove this last claim. Hint: begin by showing that B is invariant under the adjoint action of 1-parameter subgroups $\exp(t\xi)$.

Exercise 8.19. Prove that if $\mathfrak{g}$ admits a non-degenerate invariant bilinear form B, then $[\mathfrak{g},\mathfrak{g}]^\perp = \mathrm{Cent}(\mathfrak{g})$. Hence, if $\mathrm{Cent}(\mathfrak{g}) = 0$ then $\mathfrak{g}$ must be perfect.

8.3. **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 8.20.

(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 a $\mathfrak{g}$-invariant subspace for the adjoint representation of $\mathfrak{g}$ on itself. Note that ideals are 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.

Examples 8.21.

(a) The center of $\mathfrak{g}$ is an ideal.
(b) The kernel of any Lie algebra morphism $\phi\colon \mathfrak{g} \to \mathfrak{k}$ is an ideal (and every ideal $\mathfrak{h}$ arises in this way, since $\mathfrak{g}/\mathfrak{h}$ acquires a Lie algebra structure).
(c) 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.
(d) More generally, if $\mathfrak{h}$ is any ideal, then $[\mathfrak{g}, \mathfrak{h}]$ is another ideal, and so is $[\mathfrak{h}, \mathfrak{h}]$.
(e) Any Lie algebra $\mathfrak{g}$ has a distinguished ideal, the so-called *derived subalgebra*

$$\mathfrak{g}^{(1)} = [\mathfrak{g}, \mathfrak{g}].$$

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 8.22. *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 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. Note that $\mathfrak{z}$ 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 $\mathfrak{h} \subseteq \mathfrak{g}$ is any simple ideal, not among the list of $\mathfrak{g}_i$'s. Then $\mathfrak{g}_i \cap \mathfrak{h} = 0$ for all i (since the intersection is an ideal contained in $\mathfrak{g}_i$), and

$$[\mathfrak{g}_i, \mathfrak{h}] \subseteq \mathfrak{g}_i \cap \mathfrak{h} = 0$$

for all i, which gives $[\mathfrak{g}, \mathfrak{h}] = \bigoplus_i [\mathfrak{g}_i, \mathfrak{h}] = 0$. Hence $\mathfrak{h} \subseteq \mathfrak{z}$, which is a contradiction since $\mathfrak{h}$ is simple. $\square$

Exercise 8.23. Show that for any Lie group G, the Lie algebra of the center of G is the center of the Lie algebra.

8.4. **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 inner product, obtained from an arbitrary given (Euclidean) inner product $< \cdot, \cdot >'$ 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* 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.

Remarks 8.24. (a) One gets similar results for $\mathfrak{g}$-representations, whenever $\mathfrak{g}$ is the Lie algebra of a compact Lie group G. This follows since every finite-dimensiuonal $\mathfrak{g}$-representation integrates to a G-represntation.

(b) Suppose V is a finite-dimensional complex $\mathrm{SL}(n, \mathbb{R})$-representation. Then V is completely reducible, by the following *unitary trick*. First, the infinitesimal action of $\mathfrak{sl}(n, \mathbb{R})$ complexifies to $\mathfrak{sl}(n, \mathbb{C})$, which in turn gives an $\mathrm{SL}(n, \mathbb{C})$-representation, extending the representation of $\mathrm{SL}(n, \mathbb{R})$. But $\mathfrak{sl}(n, \mathbb{C})$ may also be regarded as the complxification of $\mathfrak{su}(n)$. The action of $\mathrm{SU}(n) \subseteq \mathrm{SL}(n, \mathbb{C})$ is completely reducible. But a complex subspace is $\mathrm{SU}(n)$-invariant if and only if it is $\mathrm{SL}(n, \mathbb{C})$-invariant, if and only if it is $\mathrm{SL}(n, \mathbb{R})$-invariant. (Q.E.D.) The same argument works more generally for all real forms of complexifications of compact Lie groups.

8.5. **The bi-invariant Riemannian metric.** Recall some material from differential geometry. Suppose M is a manifold equipped with a pseudo-Riemannian metric B. That is, B is a family of non-degenerate symmetric bilinear forms $B_m\colon T_mM \times T_mM \to \mathbb{R}$ depending smoothly on m. It is called a Riemannian metric if B is positive definite. 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 8.25. *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

$$\dot\gamma(0) = \xi, \quad \gamma(0) = e.$$

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, $u_s(t_0) = 0 = u_s(t_1)$. Using the formula for the differential of the exponential map, 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,

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

Now we compute the s-derivative at $s = 0$.

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

since $u_0 = 0,\ \dot{u}_0 = 0$. Hence, the s-derivative of $E(\gamma_s)$ at $s = 0$ is

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

□

Remark 8.26. 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 8.27. *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 8.28. 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.

8.6. **The Killing form.**

Definition 8.29. The *Killing form* 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 8.30. *The Killing form on a finite-dimensional Lie algebra $\mathfrak{g}$ is invariant under all automorphisms of $\mathfrak{g}$:*

$$\kappa(\phi(\xi), \phi(\eta)) = \kappa(\xi, \eta)$$

for all $\phi \in \mathrm{Aut}(\mathfrak{g})$, $\xi, \eta \in \mathfrak{g}$. It is hence (infinitesimally) invariant under all derivations,

$$\kappa(D(\xi), \eta) + \kappa(\xi, D(\eta)) = 0$$

for all $D \in \mathrm{Der}(\mathfrak{g})$, $\xi, \eta \in \mathfrak{g}$. In particular, κ is ad*-invariant and (when $\mathfrak{g} = \mathrm{Lie}(G)$)* Ad*-invariant.*

Proof. The invariance under automorphisms $\phi \in \mathrm{Aut}(\mathfrak{g})$ follows since

$$\mathrm{ad}_{\phi(\xi)} = \phi \circ \mathrm{ad}_\xi \circ \phi^{-1}$$

and since tr of an operator is invariant under similarity transformations. The invariance under derivations follows by the identity for $\phi_t = \exp(tD)$ by taking a t-derivative at $t = 0$. $\square$

Remark 8.31. An important theorem of É. Cartan says for any finite-dimensional real or complex Lie algebra $\mathfrak{g}$, the Killing form κ is nondegenerate if and only if $\mathfrak{g}$ is semi-simple. We won't use it in this course.

Proposition 8.32. *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

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

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

Remark 8.33. 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 É. Cartan in his Ph.D. thesis. The Cartan matrices, Cartan subalgebras, Weyl groups, root systems, Coxeter transformations etc. all appear in some form in W. Killing's work (cf. A. Borel 'Essays in the history of Lie groups and Lie algebras'.) According A. J. Coleman's paper

`https://www.math.umd.edu/~jda/744/coleman.pdf`

"*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 É. Cartan; the terminology was introduced by A. Borel. More information can be found in Helgason's article

`https://math.mit.edu/~dav/HelgasonKilling.pdf`

8.7. **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 $\xi,\eta \in \mathfrak{g}$. We may also write this as

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

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 8.34. *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 every derivation of $\mathfrak{g}$ is inner. In fact,* $\mathrm{Der}(\mathfrak{g}) = \mathrm{Inn}(\mathfrak{g}) = \mathfrak{g}$.

Proof. Every derivation $D \in \mathrm{Der}(\mathfrak{g})$ defines a linear functional on $\mathfrak{g}$,

$$\eta \mapsto \mathrm{tr}(D \circ \mathrm{ad}_\eta).$$

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

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

This shows $D_0(\eta) = 0$ for all η, hence $D_0 = 0$ and so $D = \mathrm{ad}_\xi$. 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. $\square$

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 map from the identity component of G to the identity component of $\mathrm{Aut}(\mathfrak{g})$.

Theorem 8.35. *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 8.32. For the converse, suppose that the Killing form is negative definite. 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 $\mathrm{Der}(\mathfrak{g}) = \mathrm{Inn}(\mathfrak{g}) = \mathfrak{g}$. $\square$

Remark 8.36. A 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.* (We may get back to this later.)

9. The maximal torus of a compact Lie group

A key ingredient in the theory of compact Lie groups G is the existence of a maximal torus T, unique up to the adjoint action. We'll prove this result after establishing some basic facts about abelian groups.

9.1. 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. A compact connected abelian group is called a *torus*, usually denoted by T. A Lie algebra $\mathfrak{g}$ is abelian (or commutative) if the Lie bracket is trivial, i.e. $\mathfrak{g}$ equals its center.

Proposition 9.1. *A connected Lie group G is abelian if and only if its Lie algebra $\mathfrak{g}$ is abelian. Furthermore, in this case the exponential map is a Lie group morphism, and*

$$G = \mathfrak{g}/\Gamma$$

where $\Gamma \subseteq \mathfrak{g}$ is the kernel of $\exp\colon \mathfrak{g} \to G$.

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, and so its kernel is a discrete subgroup of $\mathfrak{g}$. □

Generally, a discrete subgroup of a vector space V is called a *lattice*.

Lemma 9.2. *Every lattice $\Gamma \subseteq V$ has a basis. That is, there are linearly independent $\gamma_1, \ldots, \gamma_k \in V$ such that*

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

Proof. (After Bröcker-tom Dieck.) By induction, we may assume that the result holds for vector spaces of dimension less than $\dim V$.

Choose an inner product on V, and let $\gamma_1 \in \Gamma$ be an element of smallest length. Then

$$\mathbb{Z}\gamma_1 = \Gamma \cap \mathbb{R}\gamma_1.$$

Let $W = (\mathbb{R}\gamma_1)^\perp$, and denote by $p\colon V \to W$ the orthogonal projection. We will show that $p(\Gamma) \subseteq W$ is a lattice. Once this is established, we may use induction to choose $\gamma_2, \ldots, \gamma_k$ such that $p(\gamma_2), \ldots, p(\gamma_k)$ are a basis of $p(\Gamma)$. The elements $\gamma_1, \ldots, \gamma_k$ are then a basis of Γ.

To prove that $p(\Gamma)$ is a lattice, it suffices to show that the lengths of its nonzero elements have a positive lower bound. Given $\gamma \in \Gamma$ with $p(\gamma) \neq 0$, let $x \in \mathbb{R}\gamma_1$ be its orthogonal projection. This lies within $||\gamma_1||/2$ of the lattice $\mathbb{Z}\gamma_1$; hence there exists $n \in \mathbb{N}$ with $||n\gamma_1 - x|| \le ||\gamma_1||/2$. Hence we have

$$||p(\gamma)|| = ||\gamma - x|| \ge ||\gamma - n\gamma_1|| - ||n\gamma_1 - x|| \ge ||\gamma - n\gamma_1|| - ||\gamma_1||/2.$$

But $\gamma - n\gamma_1$ has length $\ge ||\gamma_1||$, by definition of γ_1. This shows

$$||p(\gamma)|| \ge ||\gamma_1||/2$$

giving the desired lower bound. $\square$

Extending the γ_i to a basis of V, we see that abelian Lie groups are isomorphic to $\mathbb{R}^n/\mathbb{Z}^k$ for some n, k. That is:

Proposition 9.3. *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 9.4 (Kronecker lemma). *Let $u = (u_1, \dots, u_k) \in \mathbb{R}^k$, and $t = \exp(u)$ its image in $T = (\mathbb{R}/\mathbb{Z})^k$. Then t is a topological generator if and only if $1, u_1, \dots, u_k \in \mathbb{R}$ are linearly independent over the rationals $\mathbb{Q}$. In particular, topological generators of tori exist.*

Proof. Let $T = (\mathbb{R}/\mathbb{Z})^k$, and let H be the closed subgroup (possibly disconnected) generated by $t = \exp(u)$. Let $\ell = \dim T - \dim H$, so that $T/H \cong (\mathbb{R}/\mathbb{Z})^l$. We will show $H \neq T$ if and only if $1, u_1, \dots, u_k \in \mathbb{R}$ are linearly *dependent* over the rationals. In other words:

$$\ell > 0 \quad \Leftrightarrow \quad \exists a_1, \dots, a_k \in \mathbb{Z} \quad \text{(not all zero)} \;:\; \sum_{i=1}^k a_i u_i \in \mathbb{Z}.$$

"$\Rightarrow$". Suppose $l > 0$. Choose a non-trivial group morphism $T/H \to \mathbb{R}/\mathbb{Z}$. 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 $a_i = (T_0\phi)(e_i) \in \mathbb{Z}$. Since ϕ is non-trivial, these are not all zero. In terms of these integers,

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

But since $\phi|_H$ is trivial, 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 group homomorphism

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

by

$$\phi\left(\exp\left(\sum_{i=1}^k v_i e_i\right)\right) = \sum_{i=1}^k a_i v_i \mod \mathbb{Z}.$$

Then $\phi(\exp(u)) = 1$, so $\phi|_H$ is trivial, but ϕ is non-trivial on T. It follows that H is a proper subgroup of T. □

We conclude with some comments on automorphisms of tori. (These will be relevant later in our discussion of Weyl groups.) Suppose $T = \mathfrak{t}/\Lambda$ is a torus. 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 $T = (\mathbb{R}/\mathbb{Z})^k$, thus $\Lambda = \mathbb{Z}^k$. 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 for which the inverse also has integer coefficients. For such a matrix, both $\det(A)$ and $\det(A)^{-1}$ must be integers, hence $\det(A) = \pm 1$. The converse is also true, by the formula for the inverse matrix. Thus,

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

It is known that $\mathrm{GL}(k, \mathbb{Z})$ is generated by its elementary matrices: The change-of-basis matrices corresponding to interchange of two basis vectors, adding some basis vector to another basis vector, or multiplying some basis vector by -1. We shall also be interested in the group

$$\mathrm{O}(k, \mathbb{Z}) = \mathrm{GL}(k, \mathbb{Z}) \cap \mathrm{O}(k)$$

of transformations preserving also the metric. Examples of matrices in this group are permutation matrices (these form a subgroup S_k) as well as sign changes of basis vectors (these form a subgroup $(\mathbb{Z}_2)^k$) .

Exercise 9.5. Show that, in fact,

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

9.2. **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. Given a maximal torus T, then $\mathrm{Ad}_a(T)$ is again a maximal torus, for any $a \in G$.

Theorem 9.6. *(É. Cartan) Let G be a compact, connected Lie group. Then any two maximal tori of G are conjugate.*

Proof. Given two maximal tori $T, T' \subseteq G$, pick topological generators t, t' of T, T', and choose $\xi, \xi' \in \mathfrak{g}$ with

$$\exp(\xi) = t,\ \exp(\xi') = t'.$$

Also choose an invariant inner product B on $\mathfrak{g}$. We will show that if $a \in G$ be a critical point for the function

$$g \mapsto B(\mathrm{Ad}_g(\xi), \xi')$$

(for instance, a point where it takes on the maximum), then $\mathrm{Ad}_a(T) = T'$. For $\eta \in \mathfrak{g}$, we have

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_{t=0} B(\mathrm{Ad}_{\exp(t\eta)} \mathrm{Ad}_a(\xi), \xi') \\ &= B([\eta, \mathrm{Ad}_a(\xi)], \xi') \\ &= -B(\eta, [\xi', \mathrm{Ad}_a(\xi)]). \end{aligned}$$

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

$$\mathrm{Ad}_{t'}(\mathrm{Ad}_a(t)) = \mathrm{Ad}_a(t).$$

That is, $\mathrm{Ad}_a(t)$ and t' *commute*. Since these are generators, any element in $\mathrm{Ad}_a(T)$ commutes with any element in T'. The subgroup

$$T' \mathrm{Ad}_a(T) \subseteq G$$

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 9.7. The *rank* l of a compact, connected Lie group G is 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 9.8. The group 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 $\mathrm{Ad}_a(T) = T'$.

Proposition 9.9. *Let G be a compact, connected Lie group. Then the Lie functor gives a 1-1 correspondence between maximal tori in G and maximal abelian subalgebras of $\mathfrak{g}$.*

Proof. For any abelian subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$, the closure of $\exp(\mathfrak{h})$ is a connected abelian subgroup, hence a maximal torus. If $\mathfrak{t} \subseteq \mathfrak{h}$ is maximal abelian, then the subgroup $\exp(\mathfrak{t})$ is automatically closed. (Otherwise, its closure would be a connected abelian Lie group whose Lie algebra is larger than $\mathfrak{t}$.) This shows that maximal abelian subalgebras exponentiate to maximal tori. Conversely, the Lie algebra $\mathfrak{t}$ of a maximal torus T is a maximal abelian Lie subalgebra of $\mathfrak{g}$. (For an abelian subalgebra $\mathfrak{h}$ strictly larger that $\mathfrak{t}$, the closure of $\exp(\mathfrak{h})$ would be a connected torus strictly larger than T. $\square$

Using this correspondence, Cartan's theorem says 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 9.10. (Properties of maximal tori). *Let G be a compact, connected Lie group, and $T \subseteq G$ a maximal torus.*

(a) *Every element of a Lie group is contained in some maximal torus. That is,*
$$\bigcup_{a \in G} \mathrm{Ad}_a(T) = G.$$
(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 subgroups: 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 $\exp(\xi) = g$. Let $\mathrm{Ad}_a(\mathfrak{t})$ be a maximal abelian subalgebra of $\mathfrak{g}$ containing ξ. Then $\mathrm{Ad}_a(T)$ is a maximal torus containing g.

(b) Suppose
$$c \in \bigcap_{a \in G} \mathrm{Ad}_a(T).$$

For any given a, since $c \in \mathrm{Ad}_a(T)$, we have that c commutes with all elements in $\mathrm{Ad}_a(T)$. Since $G = \bigcup_{a\in G} \mathrm{Ad}_a(T)$ it follows that c commutes with all elements of G, that is, $c \in Z(G)$. This proves $\bigcap_{a\in G} \mathrm{Ad}_a(T) \subseteq Z(G)$. Conversely, given $c \in Z(G)$, then c commutes with all elements of $\mathrm{Ad}_a(T)$, and hence lies in $\mathrm{Ad}_a(T)$ since maximal tori are maximal abelian, by part (d) 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. **Claim:** There exists $m \in \mathbb{N}$ such that

$$B \cong B_0 \times \mathbb{Z}_m$$

(direct product of Lie groups).

Once the claim is established, the proof of (c) is as follows: 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. By part (a), the element bh is contained in some maximal torus T. Hence $B \subseteq T$.

Proof of the claim: Note that each $g^k B_0$ for $k \in \mathbb{N}$ is a component of B. Since B is compact, it can only have finitely many components; let these be

$$B_0,\ gB_0,\ \ldots, g^{m-1}B_0,$$

where $m \in \mathbb{N}$ be the smallest number such that $g^m B_0 = B_0$, i.e.

$$g^m \in B_0.$$

Since B_0 is a torus, this element admits an m-th root. That is, we may write $g^m = k^m$ 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 of Lie groups.

(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 9.11. 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 9.12. $\dim(G/T)$ *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. $\square$

9.3. **The Weyl group.** It turns out that the maximal torus T of a compact, connected Lie group G captures much of the structure of the Lie group, once we include the additional data of its *Weyl group* symmetry.

Suppose as before that G is compact and connected, and let us consider the normalizer of a maximal torus T, [15]

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

By definition, it acts on T. The subgroup $T \subseteq N_G(T)$ acts trivially; hence we obtain an action of the quotient group.

Definition 9.13. 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) The action of $N_G(T)$ on $T = \mathfrak{t}/\Lambda$ descends to an action of W acts on T by Lie group automorphisms:

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

As a consequence, W also acts on the lattice Λ, as well as on the Lie algebra $\mathfrak{t}$. The latter action is induced by the adjoint representation of $N_G(T) \subseteq G$ on $\mathfrak{t}$; equivalently it is the differential of the action on T.

[15] 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 $\mathrm{Ad}_g(S) \subseteq S$. The normalizer $N(S)$ is a closed subgroup, hence a Lie subgroup. If H is 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.

(b) 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 the subgroup $H = T$, we obtain an action of

$$W \to \mathrm{Diff}(G/T).$$

Explicitly, if $w \in W$ is represented by $n \in N(T)$, then

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

Note that this action commutes with the G-action on G/T, and that it is *free*, that is, all stabilizer groups are trivial. The quotient of the W-action on G/T is $G/N_G(T)$, the space of maximal tori of G.

(c) Given any action of G on a manifold M, the Weyl group acts on the T-fixed point set M^T. In particular, for a G-representation on a vector space V, the Weyl group acts on the subspace V^T of T-fixed vectors.

Example 9.14. For $G = \mathrm{SU}(2)$, with maximal torus T consisting of diagonal matrices, we have $N_G(T) = T \cup nT$ where

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

Thus $W = N_G(T)/T = \mathbb{Z}_2$, with n descending to the non-trivial generator. One 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

$$G/N_G(T) = (G/T)/W \cong \mathbb{R}P(2).$$

Example 9.15. 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 $N_G(T)$ consist of all rotations in 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 $N_G(T)$ fixing the axis are exactly the elements of T itself. The elements in $N_G(T)$ reversing the axis are the rotations by π about any axis orthogonal to the 3-axis. Thus $W = \mathbb{Z}_2$.

Theorem 9.16. *The Weyl group W of a compact, connected group G is a finite group.*

Proof. Let $N_G(T)_0$ be the identity component of the normalizer. We have to show $N_G(T)_0 = T$. The inclusion $\supseteq$ is clear. For the other inclusion, note that the adjoint representation of $N_G(T)$ on $\mathfrak{t}$ preserves the lattice Λ. Since Λ is discrete, the identity component $N_G(T)_0$ acts trivially on Λ. It follows that $N_G(T)_0$ acts trivially on $\mathfrak{t} = \mathrm{span}_{\mathbb{R}}(\Lambda)$ and hence also on $T = \exp(\mathfrak{t})$. That is, $N_G(T)_0 \subseteq T$. $\square$

Theorem 9.17. *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. Suppose w acts trivially on T, and let $g \in N_G(T)$ be an element representing w. Then g commutes with all elements of T, hence $g \in T$ and therefore $w = 1$. On the other hand, w acts trivially on $T = \exp(\mathfrak{t})$ if and only if it acts trivially on $\mathfrak{t}$ if and only if it acts trivially on Λ. $\square$

Recall that the Lie algebra of a compact, connected Lie group G has a direct sum decomposition

$$\mathfrak{g} = \mathfrak{z} \oplus [\mathfrak{g}, \mathfrak{g}]$$

into the Lie algebra of the center and the 'semi-simple part'. The quotient $G' = G/\mathrm{Cent}(G)$ is a Lie group having Lie algebra $[\mathfrak{g}, \mathfrak{g}]$, with trivial center: $\mathrm{Cent}(G') = \{e\}$.

Proposition 9.18. *The Weyl group of a compact, connected Lie group G depends only $[\mathfrak{g}, \mathfrak{g}]$.*

Proof. The Lie algebra $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ is the Lie algebra of the compact Lie group $G' = \mathrm{Aut}(\mathfrak{g}')$. The adjoint action of G gives a surjective morphism $\phi\colon G \to G'$ with kernel $\mathrm{Cent}(G)$. Thus, we have an exact sequence

$$1 \to \mathrm{Cent}(G) \to G \to G' \to 1.$$

If $T' \subseteq G'$ is a maximal torus, then $T = \phi^{-1}(T')$ is a maximal torus in G; conversely, if $T \subseteq G$ is a maximal torus then $T' = \phi(T)$ is a maximal torus. We have the exact sequence

$$1 \to \mathrm{Cent}(G) \to T \to T' \to 1.$$

Similarly for the normalizers:

$$1 \to \mathrm{Cent}(G) \to N_G(T) \to N_{G'}(T') \to 1.$$

Taking quotients, we see obtain the isomorphism $W \cong W'$. $\square$

Remark 9.19 (The set of all maximal tori, and its tautological bundle). A brief digression, about some interesting differential geometric constructions with maximal tori. Let

$$X = \{\text{maximal tori in } G\}.$$

Over X, we have a tautological bundle of Lie groups,

$$\pi \colon Q \to X$$

with fiber at $\pi^{-1}(x)$ the maximal torus T_x labeled by x. Thus $Q = \coprod_{x \in X} T_x$ is the disjoint union; it comes with a natural surjective map $Q \to G$ given on the fiber T_x by the obvious inclusion. So, we get a diagram

$$\begin{tikzcd} Q \arrow[r] \arrow[d] & G \\ X & \end{tikzcd}$$

The group G acts transitively on the space X of maximal tori, and this action lifts to the fiber bundle $Q \to X$, by

$$g \cdot (x, t) \mapsto (g \cdot x, \mathrm{Ad}_g(t))$$

for $t \in T_x$. The map $Q \to G$ is G-equivariant for this action and the conjugation action on G.

The G-action on X is transitive, with stabilizer at a fixed maximal torus T equal to $N_G(T)$. That is, once we fix T, we have an identification $X = G/N_G(T)$. The space Q is identified with $Q = (G/T \times T)/W$ with the natural quotient map to $(G/T)/W = G/N_G(T) = X$. So, the diagram becomes

$$\begin{tikzcd} (G/T \times T)/W \arrow[r] \arrow[d] & G \\ G/N_G(T) & \end{tikzcd}$$

The map to G is given by $[(gT, t)] \mapsto gtg^{-1}$ (note that this is well-defined); it is a diffeomorphism on an open dense subset of Q (but not globally). So, we may think of it as something like a 'branched cover'. Note also that the bundle Q is canonically *flat*, since we have 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.

9.4. **Maximal tori and Weyl groups for the classical groups.** We will now describe the maximal tori and the Weyl groups for the classical groups: $\mathrm{U}(n), \mathrm{SU}(n), \mathrm{SO}(n), \mathrm{Sp}(n)$. These are all matrix Lie groups; to compute the Weyl group we take into account that the Weyl group action on T must preserve the set of eigenvalues of matrices $t \in T$ (since it comes from a conjugation action).

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

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

with $|z_i| = 1$ define a maximal torus T. To see this, take $t \in T$ to be a generator. In particular, the eigenvalues $\lambda_1, \ldots, \lambda_n$ (given as the diagonal entries) are all distinct. The only matrices commuting with t are the multiples of the identity matrix. Hence, the centralizer or T is T, as required.

If $w \in W$, then $w(t)$ is a diagonal matrix with the same set of eigenvalues as t. Thus, w acts by a permutation of the diagonal entries. (Since this is true for the generator t, it holds true for all powers of t all elements of T.) We hence obtain an injective group morphism

$$W \hookrightarrow S_n.$$

We claim that this map is an isomorphism. To see this, recall that there is a natural isomorphism between S_n and the set of permutation matrices. For $\sigma \in S_n$ one has the change-of-basis matrix $A_\sigma \in \mathrm{O}(n) \subseteq \mathrm{U}(n)$ for the corresponding permutation of standard basis vectors. The permutation matrices are unitary, and conjugation by A_σ preserves the set of diagonal matrices, with resulting action the pemutation of diagonal entries corresponding to σ. In conclusion, $W = S_n$.

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 $\mathrm{U}(n)$. (There is a small caveat here since the permutation matrices A_σ have determinant $\mathrm{sign}(\sigma) = \pm 1$. But this is not a problem, since we may replace A_σ by its product with the diagonal matrix $\mathrm{diag}(-1, 1, \ldots, 1)$ if needed.)

Theorem 9.20. *The Weyl group of* $\mathrm{U}(n)$*, and also of* $\mathrm{SU}(n)$*, is the symmetric group:*

$$W = S_n.$$

9.4.2. *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, write $g \in \mathrm{SO}(2m)$ in block form with 2×2-blocks $g_{ij} \in \mathrm{Mat}_2(\mathbb{R})$. For $t = t(\theta_1, \ldots, \theta_m)$ we have

$$(tgt^{-1})_{ij} = R(\theta_i) g_{ij} R(-\theta_j).$$

Thus g commutes with all $t \in T$ 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 of θ_i. 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 9.21. *The Weyl group W of $\mathrm{SO}(2m)$ is the semi-direct product $\Gamma_m \rtimes S_m$.*

Proof. The embedding $S_m \to \mathrm{O}(m)$ as permutation matrices lifts to an inclusion $S_m \to \mathrm{SO}(2m)$ (replacing every entry A_{ij} of $\mathrm{O}(n)$ by the 2×2-block $A_{ij}I_2$). This takes values in $N_G(T)$, and realizes $S_m \subseteq W$. 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 $g_{ii} = g_{jj} = K$ and $g_{\ell\ell} = I$ for $\ell \neq i, j$, lies in $N_G(T)$ 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 $g \in N_G(T)$ 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 $R(-\theta_1)g_{11} = g_{11}R(\theta_1)$. $\square$

Remark 9.22. The group $\mathrm{O}(2m)$ is compact, but disconnected. The same argument as above shows that

$$N_{\mathrm{O}(n)}(T)/T = \mathbb{Z}_2^m \rtimes S_m,$$

since now all sign changes of diagonal entries can be achieved.

Example 9.23. Consider the special case $m = 2$, so $G = \mathrm{SO}(4)$. We have $W = \Gamma_2 \rtimes S_2$. Here $S_2 = \mathbb{Z}_2$, while $\Gamma_2 = \mathbb{Z}_2 \subseteq \mathbb{Z}_2 \times \mathbb{Z}_2$ with the trivial action of $S_2 = \mathbb{Z}_2$. Hence $W = \mathbb{Z}_2 \times \mathbb{Z}_2$. This is consistent with the fact that the universal cover of $\mathrm{SO}(4)$ is $\mathrm{SU}(2) \times \mathrm{SU}(2)$.

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

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

placing the orthogonal matrix A in the upper left corner and $\det(A)$ in the lower right corner. Let T' be the standard maximal torus for $\mathrm{SO}(2m)$; 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 9.24. *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 $(\mathbb{Z}_2)^m \rtimes S_m$. Observe that if $g' \in N_{\mathrm{O}(2m)}(T')$ normalizes T' then $j(g)$ normalizes $T = j(T')$. Hence, j gives an inclusion

$$j\colon N_{\mathrm{O}(2m)}(T')) \hookrightarrow N(T).$$

Using Remark 9.22, it follows that we have an inclusion $\mathbb{Z}_2^m \rtimes S_m \subseteq W$. Hence, equality holds. $\square$

9.4.4. *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 9.25. *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 $A = \mathrm{diag}(1, 0, 1, \ldots, 1)$, $B = \mathrm{diag}(0, 1, 0, \ldots, 0)$. $\square$

9.4.5. *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.

9.4.6. *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 9.26. (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*

$$G_2,\ F_4,\ E_6,\ E_7,\ E_8$$

that are more complicated to describe.

Our main goal for the rest of this course is to explain the classification of compact Lie groups. The main idea is to take the adjoint representation $G \to \mathrm{Aut}(\mathfrak{g})$, and see how its restriction to the maximal torus breaks up into irreducibles. In order to carry out this program, we need to take a time-out and develop some representation theory.

10. Basics of representation theory

10.1. **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. Some notation and terminology: Given two G-representations V, W, we write $\mathrm{Hom}_G(V, W)$ for the space of G-equivariant linear maps $V \to W$. One calls

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

the *space of intertwining operators* from V to W. Representations V, W are isomorphic if there exists a G-equivariant isomorphism between then; in this case we write $V \cong W$.

A representation W is called *completely reducible* if it is a direct sum of irreducible representations:

$$W = W_1 \oplus \cdots \oplus W_r.$$

We have seen that any representation of a compact Lie group is completely reducible. The *multiplicity* of an irreducible representation V to appear in a (completely reducible) representation W is called its *multiplicity*. It's not quite immediate that this is well-defined, since we haven't shown that the decomposition of W is unique. To show that it's well-defined, one may use:

Lemma 10.1 (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 W is another finite-dimensional irreducible complex G-representation, then*

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

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

Proof. (a) Let λ be an eigenvalue of A. Since $\ker(A - \lambda I)$ is a G-invariant subspace, it must be all of V. Hence $A = \lambda I$.

(b) For any $A \in \mathrm{Hom}_G(V, W)$, the kernel and range of A are sub-representations. Since V, W are irreducible, it follows that $A = 0$ or A is an isomorphism. If V, W are non-isomorphic, A cannot be an isomorphism, so $A = 0$. If V, W are isomorphic, we might as well assume $W = V$, and then b) follows from a).

□

Exercise 10.2. Give a counter-example, showing that (a) does not hold for real representations.

If V is irreducible, and the representation W is completely reducible, with direct sum decomposition as above, then

$$\dim \mathrm{Hom}_G(V, W) = \sum_i \dim \mathrm{Hom}_G(V, W_i)$$

is the *multiplicity* of V in W. (In particular, the multiplicity is well-defined). The map

$$\mathrm{Hom}_G(V, W) \otimes V \to W,\ A \otimes v \mapsto A(v)$$

is injective, and its image $W_{[V]}$ is the V*-isotypical subspace* of W, i.e. the sum of all irreducible components W_i isomorphic to V. We have a canonical decomposition

$$W = \bigoplus_{V \text{ irred}} W_{[V]}$$

Each $W_{[V]}$ breaks up further into a sum of copies of V; the choice of decomposition is equivalent to a choice of basis in $\mathrm{Hom}_G(V, W)$.

Proposition 10.3. *Let T be a torus. Then the irreducible finite-dimensional complex T-representation $\pi\colon T \to \mathrm{GL}(V)$ is 1-dimensional, and are given by homomorphisms $T \to \mathrm{U}(1)$. That is, isomorphism classes of irreducible T-representations are classified according to*

$$\mathrm{Irr}(T) \cong \mathrm{Hom}(T, \mathrm{U}(1)).$$

Proof. The operators $\pi(t)$, $t \in T$ all commute, and so have a common eigenvector $v \in V$. The span of v is T-invariant, hence must be all of V is irreducible. Hence $\dim V = 1$; the action of $\pi(t)$ is multiplication by some scalar. Choose an invariant inner product on V; then $\pi(t) \in \mathrm{U}(V) = \mathrm{U}(1)$. □

If we identify $T = \mathfrak{t}/\Lambda$ and identify $\mathrm{U}(1) = S^1 = \mathbb{R}/\mathbb{Z}$, then $\mathrm{Hom}(T, \mathrm{U}(1)) = \mathrm{Hom}(\Lambda, \mathbb{Z}) = \Lambda^*$ (the dual lattice), so

$$\mathrm{Irr}(T) \cong \Lambda^*.$$

Consider next the lowest-dimensional non-abelian compact group, $G = \mathrm{SU}(2)$. (There is also SO(3), however, the representations of SO(3) may be regarded as representations of SU(2) for which $-I$ acts trivially.) The group SU(2) acts on the space $\mathbb{C}[z, w]$ of polynomials in a natural way. For any given $k = 0, 1, \ldots$, the subspace $V(k)$ of homogeneous polynomials of degree k is invariant under this action. The space $V(k)$ has basis

$$z^k,\ z^{n-1}w, \ldots, w^k$$

hence $\dim V(k) = k+1$. Let us also note that these basis vectors span invariant subspaces for the action of the standard $T \subseteq \mathrm{SU}(2)$. Namely, the action of $\mathrm{diag}(e^{i\theta}, e^{-i\theta})$ on the polynomial $z^{k-j}w^j$ is multiplication by $e^{i(k-2j)\phi}$. So, the *weights* for the T-action are

$$k, k-2, \ldots, -k.$$

Theorem 10.4. *For all $k = 0, 1, 2, \ldots$, the representation $V(k)$ is irreducible. Furthermore, $V(k)$ is the unique irreducible* SU(2)*-representations of dimension $k+1$, up to isomorphism. Hence,*

$$\mathrm{Irr}(\mathrm{SU}(2)) = \mathbb{Z}_{\geq 0}$$

We will postpone the proof to the next section. It will be convenient to phrase it in terms of representations of the complex Lie algebra $\mathfrak{sl}(2, \mathbb{C})$. This is based on the following observations.

Finite-dimensional complex representations of connected, *simply connected* Lie groups

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

are equivalent to representations of their Lie algebra

$$\pi \colon \mathfrak{g} \to \mathrm{End}(V)$$

: Given the representation of the Lie algebra, one obtains the representation of group G by exponentiation. In turn, the $\mathfrak{g}$-representation is equivalent to a representation of its complexification

$$\pi \colon \mathfrak{g}^{\mathbb{C}} \to \mathrm{End}(V);$$

here π is required to be a morphism of complex Lie algebra (in particular, is $\mathbb{C}$-linear).

For example, we obtain an equivalence of finite-dimensional complex representations of

$$\mathrm{SU}(n),\ \ \mathfrak{su}(n),\ \mathfrak{su}(n)^{\mathbb{C}} = \mathfrak{sl}(n, \mathbb{C}).$$

Remark 10.5. Since $\mathfrak{sl}(n, \mathbb{C}) = \mathfrak{sl}(n, \mathbb{R})^{\mathbb{C}}$, these are also equivalent to representations of $\mathfrak{sl}(n, \mathbb{R})$ and hence $\mathrm{SL}(n, \mathbb{R})$. 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 shall now focus on representations of SU(2), or equivalently $\mathfrak{sl}(2, \mathbb{C})$.

10.2. **$\mathfrak{sl}(2,\mathbb{C})$-representations.** We are interested in the finite-dimensional irreducible $\mathbb{C}$-linear representations of $\mathfrak{sl}(2,\mathbb{C})$. 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

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

Remark 10.6. 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,

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

Observe that this involution is an $\mathbb{R}$-linear Lie algebra automorphism of $\mathfrak{sl}(2,\mathbb{C})$.

The *quadratic Casimir element* is the expression

$$\mathrm{Cas} = 2fe + \frac{1}{2}h^2 + h$$

in the universal enveloping algebra $U(\mathfrak{sl}(2,\mathbb{C}))$: The complex unital algebra $\mathbb{C}$-linearly generated by the elements of $\mathfrak{sl}(2,\mathbb{C})$, subject to relations $\xi_1\xi_2 - \xi_2\xi_1 - [\xi_1,\xi_2]$. In other words, the enveloping algebra is defined in such a way that the algebra commutator of generators coincides with their Lie bracket. For our purposes, it is enough to work in a representation $\pi \colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V)$, and define the **Casimir operator** as follows:

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

Using the commutation relations, it may also be written

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

Lemma 10.7. *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,

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

add to 0. ∎

Note that this proof looks simpler in the enveloping algebra, where woul simply omit the π's.

The simplest non-trivial representation of $\mathfrak{sl}(2,\mathbb{C})$ (or, the corresponding Lie group $\mathrm{SL}(2,\mathbb{C})$) is the defining representation on $\mathbb{C}^2$ (given by the action 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 10.8. *For all $k = 0, 1, 2, \ldots$, the representation 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.[16] 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. $\square$

Consider now an arbitrary $\mathbb{C}$-linear representation $\pi \colon \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{End}(V)$. For $s \in \mathbb{C}$, let

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

be the eigenspace with eigenvalue s.

Lemma 10.9. *The operators $\pi(e), \pi(f)$ restrict to maps*

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

[16] In other words, the Jordan normal form of $\pi(e)$ is a single Jordan block, with zeroes on the diagonal.

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

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

hence $\pi(e)v \in V_{[s+2]}$. Similarly $\pi(f)v \in V_{[s-2]}$. $\square$

As a consequence, we see that $\pi(f)\pi(e)$ and $\pi(e)\pi(f)$ restrict to operators on $V_{[s]}$. These restrictions may be expressed in terms of the Casimir operator:

$$\pi(f)\pi(e)|_{V_{[s]}} = \frac{1}{2}\pi(\mathrm{Cas})|_{V_{[s]}} - \frac{1}{4}s(s+2),$$

$$\pi(e)\pi(f)|_{V_{[s]}} = \frac{1}{2}\pi(\mathrm{Cas})|_{V_{[s]}} - \frac{1}{4}s(s-2).$$

Note in particular that when V is irreducible (so that Cas acts as a scalar), then every nonzero $v \in V_{[s]}$ is an eigenvector of these operators.

Theorem 10.10. *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 of dimension $k+1$. Denote by $\tau \in \mathbb{C}$ the value of its Casimir operator. Since $\dim V < \infty$, we may choose an eigenvalue $\lambda \in \mathbb{C}$ of the operator $\pi(h)$, with the property that $\lambda + 2$ is not an eigenvalue. (We shall see shortly that the eigenvalues of $\pi(h)$ are in fact integers, but we don't need it yet.) Let $v_0 \in V$ be a corresponding eigenvector, and define

$$v_j = \frac{1}{j!}\pi(f)^j v_0, \ j = 0, 1, \ldots.$$

Then

$$\pi(h)v_j = (\lambda - 2j)v_j$$

by the lemma. In particular, we see that the nonzero v_j's are linearly independent, and hence there are only finitely many v_j's. We have

$$\pi(f)v_j = (j+1)v_{j+1},$$

by definition. The lemma also show that $\pi(e)v_0 = 0$ (since $V_{[\lambda+2]} = 0$ by assumption). This allows us to find the value of the Casimir element, by computing its action on v_0:

$$\tau = \tfrac{1}{2}\lambda(\lambda+2).$$

Finally, we find $\pi(e)v_j$ for $j > 0$:

$$\begin{aligned}
\pi(e)v_j &= \frac{1}{j}\pi(e)\pi(f)v_{j-1} \\
&= \frac{1}{2j}\left(\pi(\mathrm{Cas}) - \tfrac{1}{2}\pi(h)^2 + \pi(h)\right)v_{j-1} \\
&= \frac{1}{2j}\left(\tfrac{1}{2}\lambda(\lambda+2) - \tfrac{1}{2}(\lambda-2j)^2 + (\lambda-2j)\right)v_{j-1} \\
&= (\lambda - j + 1)v_{j-1}
\end{aligned}$$

This shows that the span of $v_0, v_1, \ldots$ is an invariant subspace, and so $v_0, \ldots, v_k$ is a basis of V, and $v_{k+1} = 0$. From $0 = \pi(e)v_{k+1} = (\lambda - k)v_k$, we read off that $\lambda = k$. In conclusion, we have constructed a basis $v_0, \ldots, v_k$ in which

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

These formula uniquely describe the $k+1$-dimensional representation V, up to isomorphism. On the other hand, the representation defined by these formulas really is irreducible. (E.g., if $W \subseteq V$ is an invariant subspace, we may, as before, construct a nonzero vector in W satisfying $\pi(e)w = 0$. But note that the eigenspace of $\pi(e)$ is 1-dimensional, and spanned by v_0.) □

Remark 10.11. The formulas above suggest a construction of infinite-dimensional representations of $\mathfrak{sl}(2,\mathbb{C})$. For any complex number $\lambda \in \mathbb{C}$, we obtain an infinite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ on

$$L(\lambda) = \mathrm{span}\{w_0, w_1, w_2, \ldots\},$$

by the formulas

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

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. The discussion above shows that $V(k)$ is irreducible.

Exercise 10.12. Show that for $\lambda \notin \mathbb{Z}_{\geq 0}$, the Verma module is irreducible. (This representation does not correspond to an $\mathrm{SU}(2)$-representation.)

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 $\pi(\mathrm{Cas})$. The eigenspace for the eigenvalue $k(k+2)/2$ is the direct sum of all irreducible sub-representations of type $V(k)$. (That is, it is the $V(k)$-isotypical subspace.) 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 10.13. 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 10.14. *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. ∎

11. Weights and roots

11.1. Weights and co-weights.

Let T be a torus, with Lie algebra $\mathfrak{t}$.

Definition 11.1.
- 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
$$X^*(T) \otimes_{\mathbb{Z}} \mathbb{R} = \mathrm{Hom}(\mathfrak{t}, \mathfrak{u}(1)), \quad X_*(T) \otimes_{\mathbb{Z}} \mathbb{R} = \mathrm{Hom}(\mathfrak{u}(1), \mathfrak{t})$$
respectively.
- 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 11.2. 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 11.3 (Identifications). We have avoided a choice of trivialization of $\mathfrak{u}(1)$. If we think of $\mathfrak{u}(1)$ as skew-adjoint complex matrices of rank 1, hence as $i\mathbb{R} \subseteq \mathbb{C}$, then $\mathrm{Hom}(\mathfrak{t}, \mathfrak{u}(1))$ is identified as a subspace of $\mathrm{Hom}(\mathfrak{t}, \mathbb{C}) = (\mathfrak{t}^{\mathbb{C}})^*$.

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

Sometimes, it is more convenient or more natural to absorb the $2\pi i$ factor in the definitions. This corresponds to thinking of $\mathrm{U}(1) \cong S^1 = \mathbb{R}/\mathbb{Z}$. Thus $\mathfrak{u}(1) \cong \mathbb{R}$ (with the standard lattice $\mathbb{Z}$), leading to identifications $X^*(T) \cong \Lambda^*,\quad X_*(T) \cong \Lambda$.

Exercise 11.4. 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 11.5. There is a natural identification

$$X_*(T) \cong \pi_1(T).$$

Exercise 11.6. 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 $\widehat{\Gamma} = \mathrm{Hom}(\Gamma, \mathrm{U}(1))$.

11.2. **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 11.7. *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

$$\Delta(V_1 \oplus V_2) = \Delta(V_1) \cup \Delta(V_2),$$
$$\Delta(V_1 \otimes V_2) = \Delta(V_1) + \Delta(V_2),$$
$$\Delta(V^*) = -\Delta(V).$$

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.

11.3. **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 $T \subseteq G$. For $\mu \in \Delta(V)$ denote by V_μ the corresponding T-weight space. That is,

$$\pi(t)v = \mu(t)v$$

for $v \in V_\mu,\ t \in T$.

Proposition 11.8. *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 $g \in N_G(T)$ represent the Weyl group element $w \in W$. Then g^{-1} represents w^{-1}, and we have

$$g^{-1}tg = w^{-1}(t).$$

Hence, 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 11.9. Let $G = \mathrm{SU}(2)$, with its standard maximal torus $T \cong \mathrm{U}(1)$ consisting of diagonal matrices
$$t = \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix}$$
with $z \in \mathrm{U}(1)$ (i.e. $|z| = 1$). Let $\varpi \in X^*(T)$ 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
$$(t \cdot p)(z_0, z_1) = p(z^{-1} z_0, z z_1) = z^{k-2i}\ z_0^i z_1^{k-i},$$
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 11.10. Let $G = \mathrm{U}(n)$ with its standard maximal torus $T = \mathrm{U}(1)^n$ given by diagonal matrices. Let $\epsilon^i \in X^*(T) \cong \mathbb{Z}^n$ be the i-th basis vector (the map $T \to \mathrm{U}(1)$ given by 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 corresponding weight vectors are the standard basis vectors $e_1, \ldots, e_n \in \mathbb{C}^n$. The Weyl group S_n acts on the weights by permutation. 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 $e_{i_1} \wedge \cdots \wedge e_{i_k}$ 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 $\mu = \epsilon^{i_1} + \ldots + \epsilon^{i_k}$.

11.4. **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 11.11. A *root* of G is a non-zero weight for the adjoint representation on $\mathfrak{g}^{\mathbb{C}}$. The set of roots is denoted $\mathfrak{R} \subseteq X^*(T)$.

That is:
$$\mathfrak{R} = \Delta(\mathfrak{g}^{\mathbb{C}}) - \{0\},$$

a finite W-invariant subset of $X^*(T)$. 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:

$$\pi(t)v = \alpha(t)v \quad \Rightarrow \pi(t)\overline{v} = \overline{\alpha(t)}v = ((-\alpha)(t))v$$

(here the additive notation of $X^*(T)$ gets a bit inconvenient). 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

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

Example 11.12. 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

$$\pi(t)\xi = t\xi t^{-1}$$

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 11.13. For $G = \mathrm{SU}(n)$, let T be the maximal torus given by diagonal matrices. Letting T' be the maximal torus of $\mathrm{U}(n)$, again consisting of the diagonal matrices, we have an exact sequence

$$1 \to T \to T' \xrightarrow{\det} \mathrm{U}(1) \to 1$$

where the last map is given by the determinant $\mathrm{diag}(z_1, \ldots, z_n) \mapsto z_1 \cdots z_n$, hence dually a sequence of lattices

$$0 \to X^*(U(1)) \to X^*(T') \to X^*(T) \to 0.$$

The kernel of $\mathbb{Z}^n \cong X^*(T') \to X^*(T)$. The kernel of this map is spanned by the vector $\epsilon^1 + \ldots + \epsilon^n \in X^*(T')$. 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

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

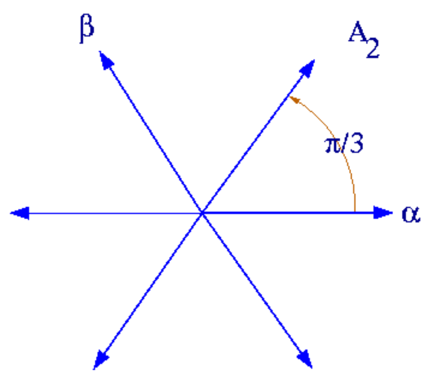
using the standard inner product on $X^*(T') \otimes_{\mathbb{Z}} \mathbb{R} = \mathbb{R}^n$. Note that the standard inner product is $W = S_n$-invariant, hence this identification respects the W-action. The projections of the ϵ^i are

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

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):[17]

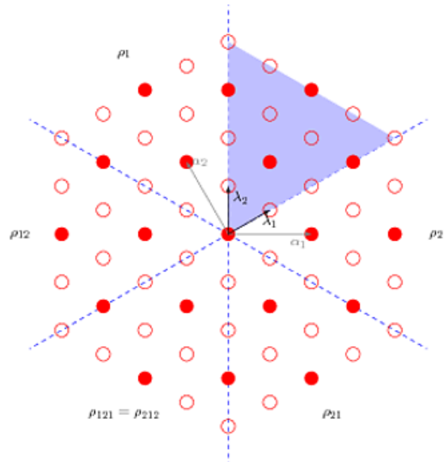


In this picture, the so-called simple roots $\alpha = \epsilon^1 - \epsilon^2$, $\beta = \epsilon^2 - \epsilon^3$ are a basis for the lattice generated by the roots (the root lattice). Note that $\alpha + \beta = \epsilon^1 - \epsilon^3$ is also a root.

And here is a picture of the weight lattice, and some roots: [18]

[17] Source: wikipedia

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



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 11.14. 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 $X_{ij} = -X_{ji}^\top \in \mathrm{Mat}_2(\mathbb{C})$. 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 $e^{\pm i\theta}$, with corresponding eigenvectors

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

Hence, the eigenvalues for the operator $A \mapsto R(\theta)AR(-\phi)$ on $\mathrm{Mat}_2(\mathbb{C})$ are given by

$$e^{i(\theta+\phi)},\ e^{i(\theta-\phi)},\ e^{i(-\theta+\phi)},\ e^{i(-\theta-\phi)}$$

with corresponding eigenvectors the matrices

$$v_+v_+^\top,\ v_+v_-^\top,\ v_-v_+^\top,\ v_-v_-^\top.$$

$$R(\theta)v_\pm = e^{\pm i\theta}v_\pm,\quad v_\pm^\top R(-\theta) = e^{\pm i\theta}v_\pm^\top.$$

Let $i < j$ be given. Putting

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

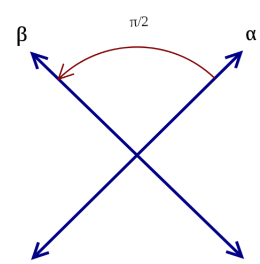
(a 2×2-matrix), we have

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

We conclude that 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 $\mathrm{SO}(4)$. [19]



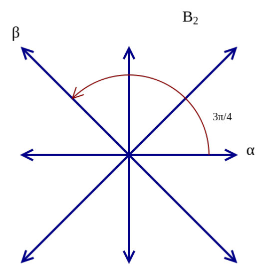
Example 11.15. Let $G = SO(2m+1)$. We write matrices in block form, corresponding to the decomposition $\mathbb{R}^{2m+1} = \mathbb{R}^2 \oplus \cdots \oplus \mathbb{R}^2 \oplus \mathbb{R}$. Thus, $X \in \mathrm{Mat}_{2m+1}(\mathbb{C})$ has 2×2-blocks X_{ij} for $i, j \le m$, a 1×1-block $X_{m+1,m+1}$, 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.

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

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: [20]



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 11.16. 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

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

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

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 $i \neq j$).

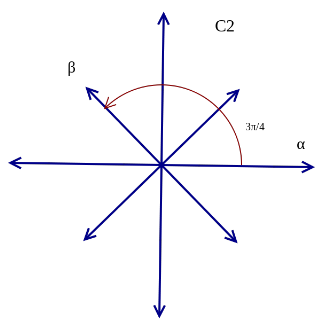
Hence we have found

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

roots:

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

Picture: [21]



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 11.17. 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 $\pi_1(\mathrm{SO}(2m)) =$

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

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

12. 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

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

12.1. **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 12.1. *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 $[X_\alpha,X_\beta]\in\mathfrak{g}_{\alpha+\beta}$. $\square$

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 12.2. *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 $B(X_\alpha, X_\beta) \neq 0$. $\square$

12.2. **The Lie subalgebras $\mathfrak{sl}(2,\mathbb{C})_\alpha$.** We shall identify

$$X^*(T) \times_{\mathbb{Z}} \mathbb{R} \cong \mathrm{Hom}(\mathfrak{t}, \mathfrak{u}(1)),$$

by the injective map taking $\mu \in \mathrm{Hom}(T, \mathrm{U}(1))$ to $T_e\mu \in \mathrm{Hom}(\mathfrak{t}, \mathfrak{u}(1))$. For simplicity, we will write μ in place of $T_e\mu$. Thus, for example,

$$\mu(\exp h) = \exp(\mu(h)), \quad h \in \mathfrak{t}$$

for $\mu \in X^*(T)$. We will regard $\mathfrak{u}(1)$ as skew-Hermitian 1×1-matrices, that is, as imaginary numbers. Hence, given a T-representation V, the elements of V_μ satisfy $\pi(t)v = \mu(t)v$ for $t \in T$, hence $\pi(h)v = \mu(h)v$ for $h \in \mathfrak{t}$.

In particular, the infinitesimal roots α satisfy

$$[h, X_\alpha] = \alpha(h) X_\alpha$$

for all $X_\alpha \in \mathfrak{g}_\alpha$ and $h \in \mathfrak{t}$.

Theorem 12.3.

(a) *For every root $\alpha \in \mathfrak{R}$, the spaces*

$$\mathfrak{g}_\alpha,\ \mathfrak{g}_{-\alpha},\ [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$$

are 1-dimensional.

(b) *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*

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

Here the choice of h_α is unique, and $\overline{h_\alpha} = -h_\alpha$. We may also arrange

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

Proof. We begin by proving that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is 1-dimensional. Pick an invariant inner product B on $\mathfrak{g}$. Let $H_\alpha \in \mathfrak{t}^{\mathbb{C}}$ be defined by

$$\alpha(h) = B(H_\alpha, h)$$

for all $h \in \mathfrak{t}^{\mathbb{C}}$. Since $\alpha(h) \in i\mathbb{R}$ for $h \in \mathfrak{t}$, we have $H_\alpha \in i\mathfrak{t}$, hence

$$\alpha(H_\alpha) = B(H_\alpha, H_\alpha) < 0.$$

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

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

This shows

$$[e_\alpha, e_{-\alpha}] = B(e_{-\alpha}, e_\alpha) H_\alpha,$$

proving that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subseteq \operatorname{span}_{\mathbb{C}}(H_\alpha)$. Actually, equality musty hold since we can choose $e_\alpha, e_{-\alpha}$ so that $B(e_\alpha, e_{-\alpha}) \neq 0$. Hence we have $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \operatorname{span}_{\mathbb{C}}(H_\alpha)$. This proves $\dim_{\mathbb{C}}[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = 1$. Let

$$h_\alpha = \frac{2}{B(H_\alpha, H_\alpha)} H_\alpha \in i\mathfrak{t},$$

a negative multiple of H_α. Then $\overline{h_\alpha} = -h_\alpha$, and

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

Clearly, h_α is the unique element of $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ with this normalization.

For fixed choice of $e_\alpha \in \mathfrak{g}_\alpha$, choose $f_\alpha \in \mathfrak{g}_{-\alpha}$ such that $B(e_\alpha, f_\alpha) \neq 0$. By the above, $[e_\alpha, f_\alpha]$ is a nonzero multiple of h_α, and we may normalize further so that

$$[e_\alpha, f_\alpha] = h_\alpha.$$

Since

$$[h_\alpha, e_\alpha] = \alpha(h_\alpha) e_\alpha = 2e_\alpha,$$

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 $\operatorname{ad}(h_\alpha)$ acts on $\mathfrak{g}_\alpha$ as the scalar $\alpha(h_\alpha) = 2$. Hence $\operatorname{ad}(f_\alpha)\colon \mathfrak{g}_\alpha \to \mathfrak{g}_0$ is injective. (See Lemma 10.14.) Since its image $\operatorname{span}_{\mathbb{C}}(h_\alpha)$ is 1-dimensional, this proves that $\mathfrak{g}_\alpha$ is 1-dimensional.

Observe next that for any choice of non-zero $e_\alpha \in \mathfrak{g}_\alpha$, the bracket

$$[e_\alpha, \overline{e_\alpha}] = B(e_\alpha, \overline{e_\alpha}) H_\alpha$$

is a positive multiple of H_α, hence a negative multiple of h_α. We hence may normalize e_α (up to a scalar in $\mathrm{U}(1)$) by requiring that $[e_\alpha, \overline{e_\alpha}] = -h_\alpha$. With the choice of $f_\alpha = -\overline{e_\alpha}$, we then have

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

confirming that $\mathfrak{sl}(2, \mathbb{C})_\alpha$ is invariant under complex conjugation, and its real part is isomorphic to $\mathfrak{su}(2)$. $\square$

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 $\mathfrak{sl}(2,\mathbb{C})$. Here h_α is uniquely determined by the condition $\alpha(h_\alpha) = 2$, while e_α (and hence $f_\alpha = -\overline{e_\alpha}$ are uniquely determined, up to complex number of absolute value 1, by the requirement $[e_\alpha, \overline{e}_\alpha] = -h_\alpha$. The fixed point set of $\mathfrak{sl}(2,\mathbb{C})_\alpha$ under complex conjugation is

$$\mathfrak{su}(2)_\alpha = \mathrm{span}_\mathbb{R}(ih_\alpha, \mathrm{Re}(e_\alpha),\ \mathrm{Im}(f_\alpha)).$$

The injective Lie algebra morphism $\mathfrak{su}(2)_\alpha \to \mathfrak{g}$ exponentiates to a Lie group morphism $\mathrm{SU}(2)_\alpha \to G$. Since $\mathrm{SU}(2)$ has center $\mathbb{Z}_2$, this morphism is either an embedding as a subgroup or a covering onto an $\mathrm{SO}(3)$ subgroup. (The two cases are distinguished by whether $\exp(\pi i h_\alpha)$ is equal to $e \in G$ or not.)

We will get a lot of information by viewing $\mathfrak{g}^\mathbb{C}$ as a representation under $\mathfrak{sl}(2,\mathbb{C})_\alpha$, for roots α. The property $[\mathfrak{g}_\gamma, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\gamma+\beta}$ for $\beta, \gamma \in \Delta(\mathfrak{g}^\mathbb{C})$ shows

$$\mathrm{ad}(e_\alpha)\colon \mathfrak{g}_\beta \to \mathfrak{g}_{\beta+\alpha},\quad \mathrm{ad}(f_\alpha)\colon \mathfrak{g}_\beta \to \mathfrak{g}_{\beta-\alpha},\quad \mathrm{ad}(h_\alpha)\colon \mathfrak{g}_\beta \mapsto \mathfrak{g}_\beta. \tag{5}$$

The scalar by which $\mathrm{ad}(h_\alpha)$ acts on $\mathfrak{g}_\beta$ is $\beta(h_\alpha)$; the $\mathfrak{sl}(2,\mathbb{C})$-representation theory tells us that this is necessarily an *integer*:

$$\beta(h_\alpha) \in \mathbb{Z}.$$

Also by $\mathfrak{sl}(2,\mathbb{C})$-representation theory, if this integer is > 0, then the map $\mathrm{ad}(f_\alpha)\colon \mathfrak{g}_\beta \to \mathfrak{g}_{\beta-\alpha}$ is *injective*. Hence, for $\beta \neq 0,\ \beta - \alpha \neq 0$ it must be an isomorphism (since both spaces are 1-dimensional), and $\mathrm{ad}(e_\alpha)\colon \mathfrak{g}_{\beta-\alpha} \to \mathfrak{g}_\beta$ is an isomorphism as well.

Theorem 12.4. *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 $\beta = t\alpha$ is a root for some $t > 1$, and take the smallest such t. Since

$$\alpha(th_\alpha) = t\alpha(h_\alpha) = 2t > 0,$$

it follows that $\mathrm{ad}(f_\alpha)\colon \mathfrak{g}_{t\alpha} \to \mathfrak{g}_{(t-1)\alpha}$ is an isomorphism. 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 an isomorphism. But this contradicts the fact that $\mathrm{ad}(e_\alpha)$ acts as zero on $\mathfrak{g}_\alpha$. (See Lemma 10.14.) ∎

12.3. **Co-roots.** Let $\mathrm{SU}(2)_\alpha \to G$ be the Lie group morphism exponentiating the inclusion $\mathfrak{su}(2)_\alpha \subseteq \mathfrak{g}$. 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 a morphism of the coweight lattices,

$$X_*(T_\alpha) \to X_*(T). \tag{6}$$

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 12.5. 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 (6). 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,

$$\alpha^\vee = h_\alpha$$

under the identification $X_*(T) \otimes_{\mathbb{Z}} \mathbb{R} = \mathrm{Hom}(\mathfrak{u}(1), \mathfrak{t}) = i\mathfrak{t}$.

Remark 12.6.
(a) As for any pairing between weights and coweights, we have that $\langle \beta^\vee, \alpha \rangle \in \mathbb{Z}$ for all $\alpha, \beta \in \mathfrak{R}$. (Actually, we already observed $\alpha(h_\beta) \in \mathbb{Z}$.) In particular,
$$\langle \alpha^\vee, \alpha \rangle = 2$$
as a consequence of the equation $\alpha(h_\alpha) = 2$.

(b) By definition, the element h_α is the unique element of $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ satisfying $\alpha(h_\alpha) = 2$. As we have seen above, for any invariant inner product B on $\mathfrak{g}$ we have that
$$B([e_\alpha, f_\alpha], h) = B(e_\alpha, f_\alpha)\alpha(h) = \langle B(e_\alpha, f_\alpha)\alpha, h \rangle.$$
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) We can be more specific. Use any W-invariant non-degenerate symmetric bilinear form $(\cdot,\cdot)$ on $i\mathfrak{t}$ to identify $V = i\mathfrak{t}$ and its dual. We have
$$h_\alpha = \alpha^\vee = c\alpha$$
for some $c \in \mathbb{R}$. Taking inner product with α, we obtain $2 = \langle \alpha, h_\alpha \rangle = c(\alpha, \alpha)$. Thus $c = 2/(\alpha, \alpha)$, and we arrive at
$$\alpha^\vee = \frac{2\alpha}{(\alpha, \alpha)}.$$
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 12.7. Recall that $\mathrm{SO}(2m)$ 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 $\mathrm{SO}(2m)$ are precisely the roots for $\mathrm{Sp}(m)$. This is an example of *Langlands duality.*

12.4. **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 12.8. *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 12.9. *For all roots $\alpha,\beta\in\mathfrak{R}$, the integer $\langle\alpha^\vee,\beta\rangle$ lies in the interval $[-3,3]$. For $\langle\alpha^\vee,\beta\rangle\neq 0$ we have*

$$|\langle\alpha^\vee,\beta\rangle|=\frac{||\beta||^2}{||\alpha||^2}.$$

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

Theorem 12.10. *(Root strings.) Let $\alpha, \beta \in \mathfrak{R}$ be roots, with $\beta \neq \pm\alpha$. Then*

(a)
$$\langle \alpha^\vee, \beta \rangle < 0 \Rightarrow \alpha + \beta \in \mathfrak{R}.$$
(Thus: If two roots form an obtuse angle then their sum is a root.)

(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 $p + q + 1$.

(c) *If $\alpha, \beta, \alpha + \beta$ are all roots, then*
$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}.$$

Proof. We will regard $\mathfrak{g}$ as an $\mathfrak{sl}(2,\mathbb{C})_\alpha$-representation. By definition of the co-roots, we have
$$\mathrm{ad}(h_\alpha) e_\beta = \langle \alpha^\vee, \beta \rangle e_\beta$$
for $e_\beta \in \mathfrak{g}_\beta$.

(a) Suppose $\beta \neq -\alpha$ is a root with $\langle \alpha^\vee, \beta \rangle < 0$. Since $\mathrm{ad}(h_\alpha)$ acts on $\mathfrak{g}_\beta$ as a negative scalar $\langle \alpha^\vee, \beta \rangle < 0$, the $\mathfrak{sl}(2,\mathbb{C})$-representation theory shows that $\mathrm{ad}(e_\alpha) \colon \mathfrak{g}_\beta \to \mathfrak{g}_{\alpha+\beta}$ is injective. In particular, $\mathfrak{g}_{\alpha+\beta}$ is non-zero.

(b) 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 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

$$V = \bigoplus_{j=-q}^{p} \mathfrak{g}_{\beta+j\alpha}.$$

Recall that k is the eigenvalue of $\mathrm{ad}(h_\alpha)$ on $\ker(\mathrm{ad}(e_\alpha)) \cap V = \mathfrak{g}_{\beta+p\alpha}$, while $-k$ is its eigenvalue on $\mathfrak{g}_{\beta-q\alpha}$. This gives,

$$k = \langle \alpha^\vee, \beta \rangle + 2p, \quad -k = \langle \alpha^\vee, \beta \rangle - 2q.$$

Hence $k = q + p$ and $q - p = \langle \alpha^\vee, \beta \rangle$.

(c) 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 $-q \leq j \leq p$ is called the *α-root string through β.*

Lemma 12.11. *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 12.9, this integer is ≤ 3. Hence, the length of any root string is at most 4. $\square$

12.6. **The reflection group defined by the root system.** 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 $W = N_G(T)/T$ 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 12.12. *The morphism j_α takes $N(T_\alpha)$ to $N_G(T)$. 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

$$\mathfrak{t}^{\mathbb{C}} = \mathrm{span}_{\mathbb{C}}(h_\alpha) \oplus \ker(\alpha).$$

Elements $h \in \ker(\alpha)$ commute with $e_\alpha, f_\alpha, h_\alpha$, hence $[\ker(\alpha), \mathfrak{sl}(2,\mathbb{C})_\alpha] = 0$. It follows that the adjoint action of $j_\alpha(\mathrm{SU}(2)_\alpha)$ on $\mathfrak{g}^{\mathbb{C}}$ fixes the subspace $\ker(\alpha) \subseteq \mathfrak{t}^{\mathbb{C}}$. In particular, the normalizer $N(T_\alpha)$ acts trivially on $\ker(\alpha)$. On the other hand, $N(T_\alpha)$ preserves $\mathfrak{t}^{\mathbb{C}}_\alpha = \mathrm{span}_{\mathbb{C}}(h_\alpha)$ Hence, all $\mathfrak{t}^{\mathbb{C}}$ (and consequently $\mathfrak{t}$) is preserved under the action of $N(T_\alpha)$, proving that

$$j_\alpha(N(T_\alpha)) \subseteq N_G(T).$$

We also see that w_α acts trivially on $\ker(\alpha)$, 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 $w_\alpha^2 = 1$):

$$\begin{aligned}\langle \xi, w_\alpha \mu \rangle &= \langle w_\alpha \xi, \mu \rangle \\ &= \langle \xi, \mu \rangle - \langle \alpha^\vee, \mu \rangle \langle \xi, \alpha \rangle \\ &= \langle \xi, \mu - \langle \alpha^\vee, \mu \rangle \alpha \rangle,\end{aligned}$$

for $\xi \in E^*, \mu \in E$. □

Remark 12.13. 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

$$j_\alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in N_G(T).$$

Let us now use a W-invariant inner product $(\cdot,\cdot)$ 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* [22]

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

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.

Definition 12.14. An element $x \in E$ is called *regular* if it does not lie on any of the root hyperplanes, and *singular* if it does.

Let

$$E^{\mathrm{reg}} = E \backslash \bigcup_{\alpha \in \mathfrak{R}} H_\alpha, \quad E^{\mathrm{sing}} = \bigcup_{\alpha \in \mathfrak{R}} H_\alpha$$

[22]This H_α has nothing to do with the element H_α used in some earlier proof.

be the set of regular elements, respectively singular elements. For any $h \in \mathfrak{t}^{\mathbb{C}}$, the $\ker(\mathrm{ad}(h)) \subseteq \mathfrak{g}^{\mathbb{C}}$ 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 $i\alpha(h))$. This shows

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

Recall again that $E = \imath\mathfrak{t}$, this applies in particular to $x \in E$. In particular, $x \in E$ is regular if and only if $\ker(\mathrm{ad}(x)) = \mathfrak{t}^{\mathbb{C}}$.

Proposition 12.15. *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 $w \in W$ with $w(x) = x$, 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 generated by the element $h \in \mathfrak{t}$.

Let $g \in N_G(T)$ be a lift of w. Then $\mathrm{Ad}_g(h) = h$, hence Ad_g fixes 1-parameter subgroup generated by h and hence fixes the torus T. It follows that $g \in T$, proving that $w = 1$. ☐

Remark 12.16. 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 12.17. For arbitrary $\xi \in \mathfrak{t}$, the stabilizer $G_\xi = \{g \in G | \, \mathrm{Ad}_g(\xi) = \xi\}$ contains T, hence $\mathfrak{g}_\xi^{\mathbb{C}}$ is a sum of weight spaces. Which roots of G are also roots of G_ξ? What can you say about the dimension of G_ξ?

Definition 12.18. 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.

Proposition 12.19. *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 12.20. Show that for $x \in E$, the stabilizer of x under the W-action is generated by all w_α such that $x \in H_\alpha$.

Exercise 12.21. 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 12.22. *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 12.23. *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 12.24. *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.*

13. Simple roots, Dynkin diagrams

13.1. **Simple roots.** Let us fix a (closed) Weyl chamber C_+, called the *positive* or *fundamental* Weyl chamber. Then any Weyl chamber is of the form $C = wC_+$ for a uniquely determined $w \in W$. 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 $x \in \mathrm{int}(C_+)$.

Definition 13.1. 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 13.2. *The simple roots* $\Pi = \{\alpha_1, \dots \alpha_l\}$ *satisfy* $\langle \alpha_i^\vee, \alpha_j \rangle \le 0$ *for* $i \neq j$.

Proof. If α_i, α_j are distinct simple roots, then their difference $\alpha_i - \alpha_j$ is *not* a root. (Otherwise, either $\alpha_i = \alpha_j + (\alpha_i - \alpha_j)$ or $\alpha_j = \alpha_i + (\alpha_j - \alpha_i)$ would be a sum of two positive roots.)

On the other hand, we had shown (Theorem 12.10) that if two roots α, β satisfy $\langle \alpha^\vee, \beta \rangle < 0$, then their sum $\alpha + \beta$ is a root. Applying this to $\alpha = \alpha_i,\ \beta = -\alpha_j$ it follows that $\langle \alpha_i^\vee, -\alpha_j \rangle \ge 0$, hence $\langle \alpha_i^\vee, \alpha_j \rangle \le 0$. □

Proposition 13.3 (Simple roots)**.** *The set* $\Pi = \{\alpha_1, \dots \alpha_l\}$ *of simple roots has the following properties.*

(a) *The set* Π *is linearly independent.*

(b) *Every root* $\alpha \in \mathfrak{R}_+$ *may be written as a sum of simple roots (possibly with multiplicities). That is,*

$$\alpha = \sum_{i=1}^{l} k_i \alpha_i$$

for some $k_i \in \mathbb{Z}_{\ge 0}$.

Proof. Fix an element $x_* \in \mathrm{int}(C_+) \subseteq E$, that is, $(\alpha, x_*) > 0$ for all $\alpha \in \mathfrak{R}_+$.

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

$$\mu_+ = \sum_{k_i > 0} k_i \alpha_i, \quad \mu_- = \sum_{k_j < 0} k_j \alpha_j. \tag{7}$$

Then $\mu_+ = -\mu_-$. Taking the scalar product and using $(\alpha_i, \alpha_j) \le 0$ for $i \neq j$ we obtain

$$0 \le ||\mu_+||^2 = -(\mu_+, \mu_-) = -\sum_{k_i > 0,\ k_j < 0} k_i k_j (\alpha_i, \alpha_j) \le 0.$$

Hence $\mu_\pm = 0$. This shows

$$\sum_{k_i > 0} k_i \alpha_i = 0, \quad \sum_{k_j < 0} k_j \alpha_j = 0.$$

Since the α_i are strictly positive on $\mathrm{int}(C)$, the first sum must be empty – there are no $k_i > 0$. Similarly, there are no $k_i < 0$. This proves that all $k_i = 0$.

(b) We want to show that any $\alpha \in \mathfrak{R}_+$ can be written in the form $\alpha = \sum k_i \alpha_i$ for some $k_i \in \mathbb{Z}_{\geq 0}$. Suppose the claim is false, so that there exist counter-examples $\alpha \in \mathfrak{R}_+$ that cannot be written in this form. Fix $x \in \mathrm{int}(C)$, and let $\alpha \in \mathfrak{R}_+$ be a counterexample for which (α, x) takes on its minimum.

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 $(\alpha, x) > 0$. 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.

□

Let us stress the following interesting consequence of part (b): If α is a root, then its coefficients with respect to the basis of simple roots are either all ≥ 0 or are all ≤ 0.

Definition 13.4. For $\alpha \in \mathfrak{R}_+$, we define its *height* $\mathrm{ht}(\alpha) \in \mathbb{N}$ as

$$\mathrm{ht}(\alpha) = \sum_j k_j$$

where $\alpha = \sum_j k_j \alpha_j$.

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

Proposition 13.5. *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. □

In particular, the positive Weyl chamber C_+ is a simple polyhedral cone, cut out by l inequalities.

Corollary 13.6. *Every root $\alpha \in \mathfrak{R}$ is W-conjugate to a simple root α_i.*

Proof. Let C be a chamber bounded by H_α. Replacing C with $-C$ if needed, we may assume $(\alpha, x) > 0$ for $x \in \mathrm{int}(C)$. Choose $w \in W$ with $wC = C_+$. Then $H_{w\alpha}$ is one of the bounding hyperplanes for the chamber $wC = C_+$, with $w\alpha$ taking positive values on $\mathrm{int}(C_+)$. Thus $w\alpha = \pm\alpha_i$ for some simple root α_i. □

Let $s_i = w_{\alpha_i}$ be the reflection defined by the simple root α_i.

Proposition 13.7. *The Weyl group is generated by the simple reflections.*

Proof. We have to show that every $w \in W$ is a product of simple reflections. It suffices to prove this for $w = w_\alpha$. We use induction on $\mathrm{ht}(\alpha)$. Write $\alpha = \sum_j k_j \alpha_j \in \mathfrak{R}_+$. By Proposition 13.3, we have $k_j \geq 0$ for all j. Since

$$0 < (\alpha, \alpha) = \sum_j k_j (\alpha_j, \alpha),$$

there exists i with $(\alpha_i, \alpha) > 0$. That is, $\langle \alpha_i^\vee, \alpha \rangle > 0$. The corresponding reflection satisfies

$$s_i \alpha = \alpha - \langle \alpha_i^\vee, \alpha \rangle \alpha_i;$$

hence, in terms of $s_i\alpha = \sum k_j' \alpha_j$ we have $k_j' = k_j$ for $j \neq i$ while $k_i' = k_i - \langle \alpha_i^\vee, \alpha \rangle < k_i$. Hence $\mathrm{ht}(s_i\alpha) < \mathrm{ht}(\alpha)$. We have $w_{s_i\alpha} = s_i w_\alpha s_i$; by induction it is a product of simple reflections. Hence

$$w_\alpha = s_i w_{s_i\alpha} s_i$$

is a product of simple reflections. □

Exercise 13.8. Show that s_i permutes the set $\mathfrak{R}_+ - \{\alpha_i\}$. (Hint: Write $\alpha = \sum_j k_j \alpha_j$, and make use of Proposition 13.3.)

13.2. **Cartan matrix, Dynkin diagram.**

Definition 13.9. 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, and that the off-diagonal entries are ≤ 0.

Example 13.10. Let $G = \mathrm{U}(n)$, and use the standard inner product on $E = X^*(T) \otimes_{\mathbb{Z}} \mathbb{R} = \mathrm{span}_{\mathbb{R}}(\epsilon^1, \ldots, \epsilon^n)$ to identify $E \cong E^*$. Recall that $\mathrm{U}(n)$ has roots $\alpha = \epsilon^i - \epsilon^j$ for $i \neq j$. The roots coincide with the coroots, under the identification $E = E^*$.

Let $x_* = n\epsilon_1 + (n-1)\epsilon_2 + \ldots + \epsilon_n$. Then $\langle \alpha, u \rangle \neq 0$ for all roots. The positive roots are $\epsilon^i - \epsilon^j$ with $i < j$, the negative roots are those with $i > j$. The simple roots are

$$\Pi = \{\epsilon^1 - \epsilon^2,\ \epsilon^2 - \epsilon^3, \ldots,\ \epsilon^{n-1} - \epsilon^n\},$$

and are equal to the simple co-roots $\Pi^\vee$. For the Cartan matrix we obtain the $(n-1)\times(n-1)$-matrix,

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & 2 & -1 \\ \cdots & \cdots & \cdots & \cdots & -1 & 2 \end{pmatrix}$$

This is also the Cartan matrix for $\mathrm{SU}(n)$ (which has the same roots as $\mathrm{U}(n)$).

Example 13.11. Let $G = \mathrm{SO}(2l+1)$. Using the standard maximal torus and the basis $X^*(T) = \mathrm{span}_{\mathbb{Z}}(\epsilon_1, \ldots, \epsilon^l)$, we had found that the roots are

$$\mathfrak{R} = \{\pm\epsilon^i \pm \epsilon^j : i \neq j\} \cup \{\pm\epsilon^i\}.$$

Let $x_* = n\epsilon_1 + (n-1)\epsilon_2 + \ldots + \epsilon_n$. Then $(x_*, \alpha) \neq 0$ for all roots α. The positive roots are the set of all $\epsilon^i - \epsilon^j$ with $i < j$, together with all $\epsilon^i + \epsilon^j$ for $i \neq j$, together with all ϵ^i. The simple roots are

$$\Pi = \{\epsilon^1 - \epsilon^2, \epsilon^2 - \epsilon^3, \ldots, \epsilon^{l-1} - \epsilon^l, \epsilon^l\}.$$

Here is the Cartan matrix for $l = 4$:

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -2 & 2 \end{pmatrix}$$

The calculation for $\mathrm{Sp}(l)$ is similar: One has

$$\Pi = \{\epsilon^1 - \epsilon^2, \epsilon^2 - \epsilon^3, \ldots, \epsilon^{l-1} - \epsilon^l, 2\epsilon^l\},$$

with Cartan matrix the transpose of that of $\mathrm{SO}(2l+1)$.

A more efficient way of recording the information of a Cartan matrix is the *Dynkin diagram*[23].

---

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

Definition 13.12. 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

$$|\langle \alpha_j^\vee, \alpha_i \rangle| = \frac{||\alpha_i||^2}{||\alpha_j||^2}$$

(for $||\alpha_j|| < ||\alpha_i||$) 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 13.13. There are only four possible Dynkin diagrams with 2 nodes:

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

Exercise 13.14. 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 $\mathfrak{R}$.

Theorem 13.15. *The Dynkin diagram determines the root system $\mathfrak{R}$, up to isomorphism.*

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

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

□

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

Example 13.16. The Dynkin diagram of $\mathrm{SO}(5)$ 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 $(s_1 s_2)^3 = 1$.

13.3. **Serre relations.** The big question remains: How to recover the Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$ from the Dynkin diagram? It turns out that there is an explicit description in terms of generators and relations. Let G be a compact simply connected 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. Let

$$h_i \in [\mathfrak{g}_{\alpha_i}, \mathfrak{g}_{-\alpha_i}], \quad e_i \in \mathfrak{g}_{\alpha_i}, \quad f_i \in \mathfrak{g}_{-\alpha_i}$$

by standard generators of the corresponding $\mathfrak{sl}(2,\mathbb{C})$; thus $\alpha_i(h_i) = 2$, $[e_i, \overline{e_i}] = -h_i$, $f_i = -\overline{e_i}$.

Theorem 13.17 (Serre relations)**.** *The elements e_i, f_i, h_i generate $\mathfrak{g}^{\mathbb{C}}$. They satisfy the relations,*

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

where $a_{ij} = \langle \alpha_i^\vee, \alpha_j \rangle$.

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 $\alpha_j(h_i) = a_{ij}$. 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 $\mathfrak{g}^\mathbb{C}$.

13.4. **Root systems, Serre's theorem.** 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 an abstract root system, as follows.

Definition 13.18. 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 13.19 (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*'.

14. 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.

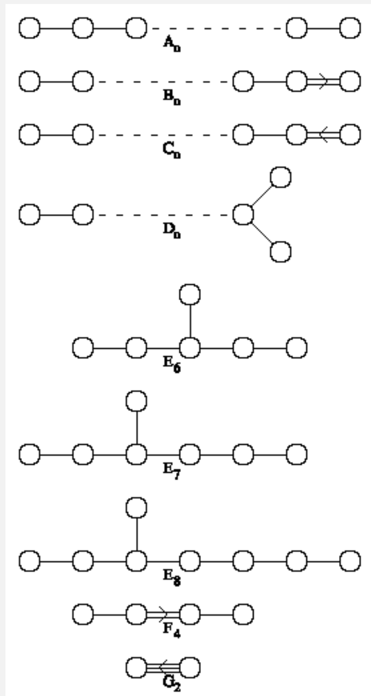
Given an abstract root system, we may as before define Weyl chambers, and the same proof as before shows that for non-orthogonal roots α, β with $||\alpha|| \geq ||\beta||$, the ratio of the root lengths is given by $||\alpha||^2/||\beta||^2 \in \{1, 2, 3\}$, and the angles in the three cases are $\pm\frac{\pi}{3},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

Proposition 14.1. *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 $\mathfrak{R}_2 = W_2\Pi_2$. $\square$

Hence, we will only consider connected Dynkin diagrams. The main theorem is as follows:

Theorem 14.2. *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 14.3. *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 $\operatorname{span}(u_1, \ldots, u_k)$.

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 $\mathfrak{R} \cap \operatorname{span}_{\mathbb{R}} \Pi'$.)

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. $\square$

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 $|W| = 696, 729, 600$.