Lie Groups

Weiyi Zhang

Mathematics Institute, University of Warwick

January 15, 2021

Contents

Chapter 1

Course description

Instructor: Weiyi Zhang
Email: weiyi.zhang@warwick.ac.uk
Webpage: http://homepages.warwick.ac.uk/staff/Weiyi.Zhang/
Lecture time (MS Teams):

Monday	3:05pm - 3:50 pm	
Thursday	9:05am - 9:50 am	from week 2

Support class:

Thursday	2:05pm - 2:50 am	from week 2

TA: Thomas Holt, Thomas.Holt@warwick.ac.uk

Reference books:

- C. Chevalley, "Theory of Lie Groups", Vol I, Princeton.
- J.J. Duistermaat, J.A.C. Kölk, "Lie Groups", Springer, 2000.
- F.W. Warner, "Foundations of Differentiable Manifolds and Lie Groups", Graduate Texts in Mathematics, 94, Springer, 1983.
- A. Kirillov, Jr., "Introduction to Lie Groups and Lie Algebras", Cambridge Studies in Advanced Mathematics, 113. Cambridge University Press, 2008.
- D. Bump, "Lie groups", Graduate Texts in Mathematics, 225. Springer-Verlag, 2004.
- W. Fulton, J. Harris, "Representation theory. A first course", Graduate Texts in Mathematics, 129. Readings in Mathematics. Springer-Verlag, New York, 1991.

Prerequisites: A knowledge of calculus of several variables including the Implicit Function and Inverse Function Theorems, as well as the existence theorem for ODEs, as well as some basic notions from topology, namely open and closed sets, continuity etc.

Knowledge of MA 3H5 Manifolds will be required. Results needed from the theory of manifolds and vector fields will be stated but not proved in the course.

Contents: The concept of continuous symmetry suggested by Sophus Lie had an enormous influence on many branches of mathematics and physics in the twentieth century. Created first as a tool in a small number of areas (e.g. PDEs) it developed into a separate theory which influences many areas of modern mathematics such as geometry, algebra, analysis, mechanics and the theory of elementary particles, to name a few.

In this module we shall introduce the classical examples of Lie groups and basic properties of the associated Lie algebra and exponential map. We will also talk about some representation theory and basic structural theory of compact Lie groups.

Chapter 2

Lie groups and Lie algebras

2.1 Lie Group

2.1.1 Smooth manifolds

To introduce the notion of Lie groups, we need to first briefly recall the definition of (smooth) manifolds.

Let M be a topological space. A local chart on M is a non-empty open set $U \subset M$ with a homeomorphism ϕ of U into an open set in $\mathbb{R}^n$.

Two charts (ϕ, U) and (ψ, V) are compatible if either $U \cap V = \emptyset$ or $U \cap V \neq \emptyset$ and $\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$ is a diffeomorphism.

An atlas on M is a collection $\{(\phi_\alpha, U_\alpha) : \alpha \in A\}$ of local charts indexed by some set A which are pairwise compatible and $M = \cup_{\alpha \in A} U_\alpha$. An atlas is called maximal if every chart compatible with all the charts of the atlas is already in the atlas. Every atlas extends to a unique maximal atlas.

A smooth manifold is a second countable Hausdorff topological space M with a choice of a maximum atlas of (smooth) atlas.

Other than proving a topological space is a manifold directly by definition, there is another way using implicit function theorem.

Theorem 2.1.1. *If $y \in \mathbb{R}^k$ is a regular value of a smooth map $F : U \to \mathbb{R}^k$ from an open set $U \subset \mathbb{R}^{n+k}$, then $F^{-1}(y) = \{x \in U : F(x) = y\}$ is a manifold of dimension n.*

We say that y is a regular value on F if for every x which satisfies $F(x) = y$, the derivative of F at x is surjective.

If M and N are manifolds then a continuous map $f : M \to N$ is smooth if for any pair of charts (U, ϕ) on M and (V, ψ) on N with $U \cap f^{-1}(V) \neq \emptyset$, the map $\psi \circ f \circ \phi^{-1}$ on $\phi(U \cap f^{-1}(V))$ is smooth (as a map between open sets in Euclidean spaces).

Say f is a diffeomorphism if f is a homeomorphism and f, f^{-1} are smooth. Two manifolds are diffeomorphic if there is a diffeomorphism between them.

If $N = \mathbb{R}$ with standard smooth structure, the smooth maps are called smooth functions on M. The set of smooth functions is denoted by $C^\infty(M)$. It is a ring with identity.

2.1.2 Lie groups

Definition 2.1.2. *A Lie group G is a smooth manifold which is also a group such that the multiplication map*

$$m : G \times G \to G, \ \ (g, h) \mapsto gh$$

and the inverse map

$$i : G \to G, \ \ g \mapsto g^{-1}$$

are smooth (i.e. C^∞) maps.

When M and N are manifolds, then their product $M \times N$ is also a manifold. Thus the definition makes sense.

Here come two remarks. First, it suffices to require that the single map $(g, h) \mapsto gh^{-1}$ is smooth. Second, we know that requiring all structures to be C^0 is enough to imply smoothness. This is Hilbert's 5th problem. In fact, the argument is much easier if we assume C^2.

Example: The following are examples of Lie groups.

1. A finite group is a Lie group of dimension zero.

2. $(\mathbb{R}^n, +)$.

3. $(\mathbb{R}^n/\mathbb{Z}^n, +)$, an n-dimensional torus. Smoothness follows because it is a local property.

4. $(\mathbb{R} \setminus \{0\}, \cdot)$, the multiplicative group of the real line.

5. $\mathrm{GL}(V)$, the automorphisms of a finite dimensional real vector space V. Choosing a basis for V gives an isomorphism of $\mathrm{GL}(V)$ and $\mathrm{GL}(n, \mathbb{R})$, the group of $n \times n$ invertible matrices with real coefficients. It is an open subset of in $M_n(\mathbb{R})$ $(= n \times n$ real matrices $\cong \mathbb{R}^{n^2})$. The multiplication is smooth since it is given by a quadratic polynomial. The inverse is smooth since the entries are polynomials divided by $\det g \neq 0$ which is smooth. It is Lie group of dimension n^2.

   For complex numbers $\mathbb{C}$ and quaternions $\mathbb{H}$, $\mathrm{GL}(n, \mathbb{C})$ and $\mathrm{GL}(n, \mathbb{H})$ (as well as $M_n(\mathbb{C}), M_n(\mathbb{H})$) are defined similarly.

6. U(n), the group of unitary matrices of rank n: *i.e.* matrices satisfying $X^* \cdot X = I_n$. Look at the map $F : M_n(\mathbb{C}) \to H_n(\mathbb{C}), F(X) = X^* \cdot X$, where $H_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) : A^* = A\}$ is the set of Hermitian matrices. $\dim H_n(\mathbb{C}) = n + \frac{n(n-1)}{2} \cdot 2 = n^2$. We check that the identity matrix I_n is a regular value of F.

$$\frac{d}{dt} F(X + tA)|_{t=0} = \frac{d}{dt}(X^* + tA^*)(X + tA) = X^*A + A^*X.$$

So the differential at X is a linear map $M_n(\mathbb{C}) \to H_n(\mathbb{C}) : A \mapsto X^*A + A^*X$. As $X \in F^{-1}(I_n) = \mathrm{U}(n)$, any Hermitian matrix B is $X^*A + A^*X$ when $A = \frac{1}{2}X \cdot B (= \frac{1}{2}(X^*)^{-1} \cdot B)$. This shows the differential is surjective and hence I_n is a regular value of F.

By Theorem 2.1.1, we know U(n) is a manifold of dimension $2n^2 - n^2 = n^2$. Multiplication and inverse are smooth since they are restrictions from that of $GL(n, \mathbb{C})$. So U(n) is a Lie group.

We have $\mathrm{U}(1) = S^1$.

7. SU(n), the group of $n \times n$ matrices with complex coefficients and satisfying $X^* \cdot X = I_n, \det X = 1$. Since $\det(X^* \cdot X) = |\det X|^2$, we know $|\det X| = 1$ for $X \in \mathrm{U}(n)$. Look at the map

$$f : M_n(\mathbb{C}) \to H_n(\mathbb{C}) \times \mathbb{R}, f(X) = (X^*X, i(\det X - \det X^*)).$$

We compute

$$\frac{d}{dt} \det(X + tA)|_{t=0} = \frac{d}{dt} \det(I_n + tX^{-1}A)|_{t=0} = \mathrm{tr}(X^{-1}A)$$

when $\det X = 1$ ($= (-1)^n$ when $\det X = -1$). So the differential of f at preimage of $(I_n, 0)$, restricting at $\det X = 1$, is a linear map

$$M_n(\mathbb{C}) \to H_n(\mathbb{C}) \times \mathbb{R} \cong \mathbb{R}^{n^2+1}.$$

$$A \mapsto (X^*A + A^*X, i \cdot \mathrm{tr}(X^*A - A^*X))$$

It is a linear algebra exercise to check this map is onto. Hence SU(n) is a manifold of dimension $n^2 - 1$ by Theorem 2.1.1.

Multiplication and inversion are smooth in the matrix entries, so $SU(n)$ is a Lie group.

One can see that

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

Write $\alpha = x_1 + ix_2, \beta = x_3 + ix_4, x_i \in \mathbb{R}$, we see that SU(2) is diffeomorphic to S^3.

In fact, the only spheres S^n which could be given the structure of Lie groups are S^1 and S^3.

In the following, we list more Lie groups without proof. The argument is similar to the above two examples.

8. Special linear group $\mathrm{SL}(n, \mathbb{R}) = \{X \in \mathrm{GL}(n, \mathbb{R}) : \det X = 1\}$ is a Lie group of dimension $n^2 - 1$.

9. $\mathrm{O}(n) = \{X \in M_n(\mathbb{R}) : X^t X = I_n\}$, the orthogonal group, is a Lie group of dimension $\frac{n(n-1)}{2}$.

10. $\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R})$, special orthogonal group. $\mathrm{SO}(n)$ is the component of the identity of $\mathrm{O}(n)$.

    $\mathrm{SO}(2) = \mathrm{U}(1) = S^1$. SO(3) is the rotation group of our 3D space.

11. $\mathrm{O}(p, q)$, $SO(p, q)$, indefinite orthogonal groups. Define non-singular bilinear form

$$B_{p,q}(x, y) = \sum_{i=1}^{p} x_i \cdot y_i - \sum_{i=p+1}^{p+q} x_i \cdot y_i = x^t I_{p,q} y.$$

    Then $\mathrm{O}(p, q) = \{g \in M_{p+q}(\mathbb{R}) : B_{p,q}(gx, gy) = B_{p,q}(x, y), \forall x, y \in \mathbb{R}^{p+q}\}$, or equivalently $g^t I_{p,q} g = I_{p,q}$. It is a Lie group of dimension $\frac{(p+q)(p+q-1)}{2}$. For $g \in \mathrm{O}(p, q)$, $(\det g)^2 = 1$.

    So $\mathrm{SO}(p, q) = \mathrm{O}(p, q) \cap \mathrm{SL}(p + q, \mathbb{R})$ is an open subgroup of $\mathrm{O}(p, q)$.

    $\mathrm{SO}(0, n) = \mathrm{SO}(n, 0) = \mathrm{SO}(n)$ and $\mathrm{SO}(1, 1) = \mathbb{R}$. Lorentz group $\mathrm{SO}(1, 3)$.

    One can similarly define $\mathrm{U}(p, q)$ and $\mathrm{SU}(p, q)$.

12. $\mathrm{Sp}(2n, \mathbb{R})$ symplectic group. Define the bilinear form $B_{2n}(x, y) = x^t J_{2n} y$ where $J_{2n} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$.

$$\mathrm{Sp}(2n, \mathbb{R}) = \{g \in M_{2n}(\mathbb{R}) : B(gx, gy) = B(x, y), \forall x, y \in \mathbb{R}^{2n}\},$$

    equivalently $g^t J_{2n} g = J_{2n}$.

    Define $F(g) = g^t J_{2n} g$, and $A_{2n}(\mathbb{R}) = \{A \in M_{2n}(\mathbb{R}) : A^t = -A\}$. Then $F : M_{2n}(\mathbb{R}) \to A_{2n}(\mathbb{R})$ and $\mathrm{Sp}(2n, \mathbb{R}) = F^{-1}(J_{2n})$. Showing J_{2n}

is a regular value would imply $\mathrm{Sp}(2n, \mathbb{R})$ is a manifold of dimension $4n^2 - \frac{2n(2n-1)}{2} = 2n^2 + n$. It is a group with matrix multiplication, so a Lie group.

13. Complex versions $\mathrm{SL}(n, \mathbb{C}), \mathrm{O}(n, \mathbb{C}), \mathrm{SO}(n, \mathbb{C}), \mathrm{O}(p, q; \mathbb{C}) = \mathrm{O}(p+q, \mathbb{C})$, $\mathrm{Sp}(2n, \mathbb{C})$. Notice for $\mathrm{Sp}(2n, \mathbb{C})$, the corresponding bilinear form is $B(x, y) = \sum_{i=1}^n (x_{n+i}y_i - x_i y_{n+i})$. If $x = (x_1, \cdots, x_{2n})$ and $y = (y_1, \cdots, y_{2n})$.

14. Compact symplectic group $\mathrm{Sp}(n) = \{X \in M_n(\mathbb{H}) : \bar{X}^t X = I\}$ is just the quaternion unitary group. It is a Lie group of dimension $4n^2 - (n + \frac{n(n-1)}{2} \cdot 4) = 2n^2 + n$.

    One can show that $\mathrm{Sp}(n) = \mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbb{C})$. The standard Hermitian form on $\mathbb{H}^n$ is $K(v', w') = \sum \bar{v}_i' w_i'$. Here the conjugation is the quaternion conjugation $\overline{a + bi + cj + dk} = a - bi - cj - dk$. Use the identification $\mathbb{C}^{2n} \cong \mathbb{H}^n$ by

    $$(z_1, \cdots, z_{2n}) \mapsto (z_1 + jz_{n+1}, \cdots, z_n + jz_{2n}),$$

    and write $v_i' = v_i + jv_{n+i}$ and similarly for w_i', we can rewrite the quaternion Hermitian form as $K(v', w') = H(v, w) - jQ(v, w)$ where H is the standard complex Hermitian form $H(v, w) = \sum_{i=1}^n \bar{v}_i w_i$ (here the conjugation is the complex one, not the quaternion one) and $Q(v, w) = \sum_{i=1}^n (v_{i+n} w_i - v_i w_{i+n})$. In the computation, notice $\bar{v}_i' = \bar{v}_i - jv_{n+i}$ where $\bar{v}_i'$ and $\bar{v}_i$ means quaternion and complex conjugation respectively, and use the relation $jz = \bar{z}j$ for any $z \in \mathbb{C}$.

    This implies $\mathrm{Sp}(n) = \mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbb{C})$. Strictly speaking, the above argument only shows $\mathrm{Sp}(n) = \mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{GL}(n, \mathbb{H})$. But we actually have $\mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbb{C}) \subset \mathrm{GL}(n, \mathbb{H})$ as $g^* g = I_{2n}$ and $g^* J_{2n} g = J_{2n}$ imply $J_{2n} g = g J_{2n}$. In fact, $\mathrm{U}(2n)$, $\mathrm{Sp}(2n, \mathbb{C})$ and $\mathrm{GL}(n, \mathbb{H})$ are called compatible triple and any element in any two belongs to the third one. Thus, $\mathrm{Sp}(n)$ is the intersection of any two of the three groups.

Exercise: Show $\mathrm{U}(n) = \mathrm{O}(2n) \cap \mathrm{Sp}(2n, \mathbb{R})$.

One notices that $\mathrm{SU}(n), \mathrm{SO}(n), \mathrm{Sp}(n)$ (and $\mathrm{U}(n)$) are compact connected Lie groups.

The following is a more direct way to show something is a Lie group. We will prove it later in the class.

Theorem 2.1.3. *Let G be a Lie group, $H \subset G$ a subgroup in the algebraic sense. If H is a closed as a subset of G, then H has a unique structure of a Lie subgroup.*

Let us now define the notion of Lie group action.

Definition 2.1.4. *A (left) action of a Lie group G on a manifold M is a smooth map $G \times M \to M$ written $(g, m) \mapsto g \cdot m$ such that $g \cdot (h \cdot m) = gh \cdot m$ and $e \cdot m = m$.*

A representation of G is a linear action on a vector space, *i.e.* a homomorphism $G \to \mathrm{GL}(V)$.

There are two important left actions of G on G itself.

$$l_g : G \to G, \;\; h \mapsto gh$$

$$r_g : G \to G, \;\; h \mapsto hg^{-1}.$$

2.2 Lie algebra

2.2.1 Lie algebra

Definition 2.2.1. *A Lie algebra is a vector space V (over a field $\mathbb{K}$) with a bilinear map (called the Lie bracket) $[\ ,\ ] : V \times V \to V$ such that for all $X, Y, Z \in V$,*

1. $[X, Y] = -[Y, X]$
2. $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ *(Jacobi identity)*

Example:.

1. Suppose $\mathcal{A}$ is an associative algebra. We can turn $\mathcal{A}$ into a Lie algebra by defining $[a, b] = ab - ba$. In particular, we may take $\mathcal{A} = \mathrm{End}(V)$ for some vector space V, *e.g.* $\mathcal{A} = M_n(\mathbb{R}), M_n(\mathbb{C})$.

2. Suppose $\mathcal{A}$ is an associative algebra and $\mathrm{Der}(\mathcal{A})$ the space of derivations, *i.e.* linear maps $d : \mathcal{A} \to \mathcal{A}$ satisfying $d(ab) = a(db) + (da)b$ for all $a, b \in \mathcal{A}$. Then $\mathrm{Der}(\mathcal{A})$ is a Lie algebra under commutators, $[d_1, d_2] = d_1 d_2 - d_2 d_1$.

3. Let M be a smooth manifold, $\mathcal{A} = C^\infty(M)$. Then $\mathrm{Der}(\mathcal{A})$ is the derivations of $C^\infty(M)$, *i.e.* the Lie algebra of vector fields which will be recalled in the next subsection.

Definition 2.2.2. *A homomorphism of Lie algebra $F : \mathfrak{g} \to \mathfrak{h}$ is a K-linear map which preserves the Lie bracket, i.e. $[Fa, Fb] = F[a, b]$.*

Definition 2.2.3. *A Lie subalgebra of* $\mathfrak{g}$ *is a* K*-linear subspace* $\mathfrak{h} \subset \mathfrak{g}$ *such that* $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$.

Note that a Lie subalgebra is clearly a Lie algebra in its own right, and the inclusion $\mathfrak{h} \to \mathfrak{g}$ is a homomorphism of Lie algebras.

Definition 2.2.4. *An ideal in* $\mathfrak{g}$ *is a* K*-linear subspace* $\mathfrak{h} \subset \mathfrak{g}$ *such that* $[\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h}$.

If $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then $\mathfrak{g}/\mathfrak{h}$ has a unique Lie algebra structure so that the quotient map $\mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$ is a homomorphism of Lie algebras.

Exercise: Check the Lie bracket on $\mathfrak{g}$ induces a Lie bracket on the quotient $\mathfrak{g}/\mathfrak{h}$.

2.2.2 Tangent and cotangent spaces

To define the tangent space, we recall the notion of germ. A germ is an equivalence class of pairs (U, f_U), where U is an open neighborhood of x and $f_U : U \to \mathbb{R}$ is a smooth function. The pairs (U, f_U) and (V, f_V) are equivalent if f_U and f_V are equal on some neighborhood $W \subset U \cap V$ of x. Let $\mathcal{O}_x$ be the vector space of germs at x.

A local derivation of $\mathcal{O}_x$ is a linear map $X : \mathcal{O}_x \to \mathbb{R}$ such that

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

Any such local derivative is called a tangent vector.

Let (U, ϕ) be a local chart with coordinate functions $x_1, \cdots, x_n$ and $a_1, \cdots, a_n \in \mathbb{R}$, then

$$Xf = \sum_{i=1}^{n} a_i \frac{\partial f}{\partial x_i}|_x := \sum a_i \frac{\partial f \circ \phi^{-1}}{\partial r_i}|_{\phi(x)}$$

where r_i are standard coordinate functions on $\mathbb{R}^n$, is a local derivation. In fact, every local derivation is of this form.

Definition 2.2.5. *The tangent space* T_xM *is the set of such local derivatives. It is an* n*-dimensional real vector space if* M *is an* n*-dimensional manifold.*

Follows from the above discussion, $\{\partial_{x_1}, \cdots, \partial_{x_n}\}$ is a basis of T_xM.

Intuitively, a tangent vector is an equivalence class of paths through x: two paths are equivalent if they are tangent at x. By a path we mean a smooth map $u : (-\epsilon, \epsilon) \to M$ such that $u(0) = x$ for some $\epsilon > 0$.

Given a function, we can use the path to define a local derivation $Xf = \frac{d}{dt}f(u(t))|_{t=0} = \sum a_i \frac{\partial f}{\partial x_i}(x)$. We sometimes write this $X \in T_xM$ as $u'(0)$.

By a vector field X on M, we mean a rule that assigns to each point $x \in M$ an element $X_x \in T_xM$ and the assignment $x \mapsto X_x$ is smooth.

Proposition 2.2.6. *There is a one-to-one correspondence between vector fields on a smooth manifold M and derivations of $C^\infty(M)$.*

This implies vector fields can be restricted and patched, hence a sheaf.

Now consider two vector fields X, Y. Locally, $X = \sum a_i \frac{\partial}{\partial x_i}$ and $Y = \sum b_i \frac{\partial}{\partial x_i}$ where a_i, b_i are smooth functions. We have $[X, Y] = \sum_{i,j}(a_j \frac{\partial b_i}{\partial x_j} - b_j \frac{\partial a_i}{\partial x_j})\frac{\partial}{\partial x_i}$.

Given a smooth map $\theta : M \to N$, one can describe the map $d_x\theta : T_xM \to T_{\theta(x)}N$ (sometimes, we also write it as θ_*) by saying that it sends $v \in T_xM$ to $(\theta \circ \gamma)'(0)$ for any curve γ in M with $\gamma'(0) = v$. Or more directly, $d_x\theta(X_x)(f) = X_x(f \circ \theta)$. Two vector fields X, Y on M and N respectively are said to be θ-related if $d_x\theta(X_x) = Y_{\theta(x)}$ for all $x \in M$.

The dual space of the tangent space is called the cotangent space of M at x, and is denoted by T_x^*M. In the special case $N = \mathbb{R}$ above, we have $X_x(f) = d_xf(X_x)$. In other words, d_xf is a cotangent vector at x. In local charts, the dual basis of $\{\partial_{x_1}, \cdots, \partial_{x_n}\}$ in T_x^*M is $\{dx_1, \cdots, dx_n\}$. And we have $df = \frac{\partial f}{\partial x_1}dx_1 + \cdots + \frac{\partial f}{\partial x_n}dx_n$ at x.

With the notion of induced maps on tangent spaces, we can also define regular values. Then, we have the following general form of Theorem 2.1.1.

Theorem 2.2.7. *If $p \in N$ is a regular value of smooth map $f : M \to N$ between manifolds, then $f^{-1}(p)$ is a manifold of dimension* $\dim M - \dim N$. *Moreover,* $\ker d_qf = T_q(f^{-1}(p))$ *for any $q \in f^{-1}(p)$.*

In fact, all these tangent spaces could be bundled together to form a manifold TM, called the tangent bundle. As a set $TM = \sqcup_{x\in M}T_xM$. We have a surjective map $\pi : TM \to M$ with $\pi(T_xM) = \{x\}$ for all $x \in M$. Given a chart, $\phi : U \to V \subset \mathbb{R}^n$ of M, we have a map $\psi : TU \to V \times \mathbb{R}^n$ given by $\psi(v) = (\phi\pi(v), d_{\pi(v)}\phi(\pi(v)))$, where $x = \pi(v)$. In this way, an atlas $\{\phi_\alpha : U_\alpha \to V_\alpha\}_{\alpha\in A}$ gives rise to a smooth atlas $\{\psi_\alpha\}$ of TM. A smooth map $s : M \to TM$ such that $\pi \circ s = id$ is called a section of tangent bundle. A vector field is a smooth section of the tangent bundle.

Similarly, the cotangent spaces also form a vector bundle, called the cotangent bundle T^*M. A smooth section of the cotangent bundle is a 1-form (or sometimes called a covector). Given each $x \in M$, we can form the pth exterior power of the cotangent space. This is also a vector bundle, Λ^pT^*M. A section of it is called a p-form.

2.3 The Lie algebra of a Lie group

Definition 2.3.1. *A vector field X on a Lie group G is said to be left-invariant if $(d_g l_h)X_g = X_{l_h(g)} = X_{hg}$ for all $g, h \in G$.*

This relation will be denoted shortly by $(l_h)_*(X_g) = X_{hg}$. X is left-invariant means that X is l_h-related to itself for all $h \in G$.

Proposition 2.3.2. *The vector space of left-invariant vector fields is closed under $[,]$ and is a Lie algebra of dimension $\dim G$. If $X_e \in T_eG$, there is a unique left-invariant vector field X on G with the prescribed tangent vector at the identity.*

Proof. We first need a lemma.

Lemma 2.3.3. *If $F : M \to N$, and vector fields X, Y on M and vector fields X', Y' on N are F-related, then $[X, Y]$ is F-related to $[X', Y']$.*

Proof. Take $f \in C^\infty(N)$, then we have $X_x(f \circ F) = d_xF(X_x))(f) = X'_{F(x)}(f)$ if X and X' are F-related. It is equivalent to write

$$(X'(f) \circ F)(x) = (X'(f))(F(x)) = X'_{F(x)}(f) = X_x(f \circ F) = (X(f \circ F))(x), \forall x \in M.$$

Therefore X and X' are F-related if and only if $X'(f) \circ F = X(f \circ F)$ for all $f \in C^\infty(M)$.

Now, we have

$$\begin{aligned}([X, Y])(f \circ F) =& X(Y(f \circ F)) - Y(X(f \circ F))\\ =& X(Y'(f) \circ F) - Y(X'(f) \circ F)\\ =& X'(Y'(f)) \circ F - Y'(X'(f)) \circ F\\ =& ([X', Y'](f)) \circ F\end{aligned}$$

This means $[X, Y]$ is F-related to $[X', Y']$. □

Now since left invariant vector fields are l_h-related for all $h \in G$ by definition, the Lie bracket of two left vector fields is a left invariant vector field.

Given a tangent vector X_e at the identity e, we may define a left-invariant vector field by $X_g = (l_g)_*(X_e)$. Conversely any left-invariant vector field must satisfy this identity. So the space of left-invariant vector fields is isomorphic to the tangent space of G at e. Therefore, its dimension is $\dim G$. □

For $\xi \in T_eG$, we denote by $\tilde{\xi} = (l_g)_*\xi$ the left-invariant vector field on G whose value at e is ξ.

Definition 2.3.4. *If G is a Lie group, then its Lie algebra $\mathfrak{g}$ is the vector space T_eG with bracket $[\xi, \eta] := [\tilde{\xi}, \tilde{\eta}]_e$.*

It follows from Proposition 2.3.2, $\mathfrak{g}$ is indeed a Lie algebra.

Example:.

1. $\mathrm{GL}(n, \mathbb{R}) \subset M_n(\mathbb{R})$ is an open set so its tangent space at any $g \in \mathrm{GL}(n, \mathbb{R})$ is just $M_n(\mathbb{R})$. We will show that the Lie bracket is just the commutator $[A, B] = AB - BA$.

   To prove this, we compute $\tilde{A}$, the left invariant vector field associated with the matrix $A \in T_I \mathrm{GL}(n, \mathbb{R})$. On $\mathrm{GL}(n, \mathbb{R})$, we have global coordinate maps given by $x_{ij}(g) = g_{ij}$, the ijth entry of $g \in \mathrm{GL}(n, \mathbb{R})$. Then the corresponding tangent vector is written as $\sum A_{ij} \frac{\partial}{\partial x_{ij}}$. To determine $\tilde{A}$, we only need to evaluate $\tilde{A}_g(x_{pq})$ for any $g \in \mathrm{GL}(n, \mathbb{R})$ and $1 \le p, q \le n$. By definition, $\tilde{A}_g(x_{pq}) = \tilde{A}_I(x_{pq} \circ l_g) = A(x_{pq} \circ l_g)$. If $h \in \mathrm{GL}(n, \mathbb{R})$,

   $$(x_{pq} \circ l_g)(h) = x_{pq}(gh) = \sum_k g_{pk} h_{kq} = \sum_k g_{pk} x_{kq}(h).$$

   Hence,

   $$\tilde{A}_g(x_{pq}) = A(x_{pq} \circ l_g) = \sum_k g_{pk} A_{kq}.$$

   That is,

   $$\tilde{A}_g = \sum_{i,j,k} g_{ik} A_{kj} \frac{\partial}{\partial x_{ij}}.$$

   It follows that the Lie bracket for the vector fields is

   $$\begin{aligned}([\tilde{A}, \tilde{B}]_g) =& [\sum g_{ik} A_{kj} \frac{\partial}{\partial x_{ij}}, \sum g_{pq} B_{qr} \frac{\partial}{\partial x_{pr}}] \\ =& \sum g_{ik} A_{kj} B_{jr} \frac{\partial}{\partial x_{ir}} - \sum g_{pq} B_{qr} A_{rj} \frac{\partial}{\partial x_{pj}} \\ =& \sum g_{ik} (A_{kr} B_{rj} - B_{kr} A_{rj}) \frac{\partial}{\partial x_{ij}}.\end{aligned}$$

   In particular, its restriction to $T_I \mathrm{GL}(n, \mathbb{R})$ is $[A, B] = AB - BA$.

   This Lie algebra is called $\mathfrak{gl}(n, \mathbb{R})$.

2. All the other matrix Lie groups. Take $\mathrm{O}(n) \subset \mathrm{GL}(n, \mathbb{R})$ for example. Let $F : M_n(\mathbb{R}) \to S_n(\mathbb{R})$ by $F(g) = g^T g$. Then

$$(D_{I_n} F)(A) = \frac{d}{dt}(I_n + tA)^T (I_n + tA)|_{t=0} = A^T + A.$$

Therefore $T_{I_n}(\mathrm{O}(n)) = \ker D_{I_n} F = A_n(\mathbb{R})$, the skew-symmetric matrices.

For the Lie bracket, we can extend $A \in T_{I_n}\mathrm{O}(n)$ to $\mathrm{O}(n)$ or to $\mathrm{GL}(n, \mathbb{R})$ to get two left-invariant extensions $\tilde{A}^{\mathrm{O}(n)}$ and $\tilde{A}^{\mathrm{GL}(n,\mathbb{R})}$ and $\tilde{A}^{\mathrm{GL}(n,\mathbb{R})}|_{\mathrm{O}(n)} = \tilde{A}^{\mathrm{O}(n)}$. Hence the Lie bracket is also the commutator. This Lie algebra is denoted $\mathfrak{o}(n)$.

When $n = 3$, we can write a skew-symmetric matrix as

$$\begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}.$$

The bracket $AB - BA$ becomes the cross-product in $\mathbb{R}^3$.

Lie algebra of other matrix Lie groups could be calculated similarly as the Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R})$ or $\mathfrak{gl}(n, \mathbb{C})$.

Now suppose $\phi : G \to H$ is a Lie group homomorphism, then its differential at e gives a linear map ϕ_* from $\mathfrak{g}$ to $\mathfrak{h}$.

Theorem 2.3.5. *If $\phi : G \to H$ is a Lie group homomorphism, then the induced map $\phi_* : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism.*

Proof. Suppose $X_e, Y_e \in \mathfrak{g}$, and X, Y the corresponding left invariant vector fields. Define X', Y' be the corresponding left invariant vector fields for $\phi_* X_e, \phi_* Y_e \in \mathfrak{h}$. Since ϕ is a group homomorphism, we have $\phi \circ l_g = l_{\phi(g)} \circ \phi$.

We have

$$X'_{\phi(g)} = (l_{\phi(g)})_* \phi_* X_e = (l_{\phi(g)} \circ \phi)_* X_e = (\phi \circ l_g)_* X_e = \phi_* X_g.$$

Similarly for Y and Y'. Hence X, X' and Y, Y' are ϕ-related.

By Lemma 2.3.3, $[X, Y]$ and $[X', Y']$ are ϕ-related. In particular, this implies $\phi_*([X_e, Y_e]) = [\phi_* X_e, \phi_* Y_e]$, which is the desired relation. □

The other way to check this relation is by Proposition 2.4.6. Later we will also write ϕ_* for the induced map.

An n-dimensional manifold M is parallelizable if there is a diffeomorphism $f : M \times \mathbb{R}^n \to TM$ restricting to a linear isomorphism $\mathbb{R}^n = p \times \mathbb{R}^n \to T_pM$ for each $p \in M$. Equivalently, M is parallelizable if there exist n vector fields $X_1, \cdots, X_n$ which are linearly independent at each point in M.

When G is a Lie group, choose a basis $e_1, \cdots, e_n$ of $\mathfrak{g}$. The the left-invariant extension of e_i are linearly independent in T_gG for each $g \in G$. This proves

Proposition 2.3.6. *Any Lie group is parallelizable.*

2.4 Exponential map

Definition 2.4.1. *We say that $\gamma : I \to M$ is an integral curve of a vector field X if for any $t \in I$, $\frac{d\gamma}{dt}(t) = X_{\gamma(t)}$.*

One should understand $\frac{d\gamma}{dt}(t)$ as $\gamma_*(\frac{d}{dt})|_{\gamma(t)}$. Locally the equation $\frac{d\gamma}{dt}(t) = X(\gamma(t))$ gives rise a system of first order differential equations. By the Picard theorem, locally an integral curve always exists and is unique, and depends on the initial data smoothly.

When M is compact, then for any $p \in M$, there is a unique integral curve $\gamma_p : (-\epsilon, \epsilon) \to M$ such that $\gamma_p(0) = p$. Thus we can define a map

$$\phi^t : M \to M, \;\; p \mapsto \gamma_p(t).$$

Definition 2.4.2. *A one-parameter subgroup (1-PSG) in a Lie group G is a smooth homomorphism $\lambda : \mathbb{R} \to G$.*

A 1-PSG λ has a derivative $\lambda'(0) \in \mathfrak{g}$.

Proposition 2.4.3. *For each $\xi \in \mathfrak{g}$ there is a unique 1-PSG λ_ξ with $\lambda'_\xi(0) = \xi$.*

Proof. Take the left-invariant extension $\tilde{\xi}$. If $\lambda(t)$ is a 1-PSG with $\lambda'(0) = \xi$, then

$$\lambda'(t) = \frac{d}{ds}\lambda(t+s)|_{s=0} = \frac{d}{ds}l_{\lambda(t)}\lambda(s)|_{s=0} = d_e l_{\lambda(t)}(\lambda'(0)) = d_e l_{\lambda(t)}(\xi) = \tilde{\xi}_{\lambda(t)}.$$

Hence, $\lambda(t)$ is an integral curve for $\tilde{\xi}$. This proves the uniqueness by ODE.

Let $\Phi^t : G \to G$ be the map induced by the integral curve of $\tilde{\xi}$. Since $\tilde{\xi}$ is left-invariant, $\Phi^t(g_1g_2) = g_1\Phi^t(g_2)$. Let $\lambda(t) = \Phi^t(1)$. Then

$$\lambda(t+s) = \Phi^{t+s}(1) = \Phi^s(\Phi^t(1)) = \Phi^s(\lambda(t)) = \lambda(t)\Phi^s(1) = \lambda(t)\lambda(s).$$

This proves the existence of λ_ξ for small t. The fact that it can be extended to any $t \in \mathbb{R}$ is obvious from $\lambda(t+s) = \lambda(t)\lambda(s)$. □

Define the *exponential map* $\exp : \mathfrak{g} \to G$ by $\exp(\xi) = \lambda_\xi(1)$. The uniqueness of 1-PSG, Proposition 2.4.3, implies that $\lambda_\xi(ct) = \lambda_{c\xi}(t)$ since $\frac{d\lambda_\xi(ct)}{dt}|_{t=0} = c\xi$. Hence, $\lambda_\xi(t) = \exp(t\xi)$. The implies that the derivative of exponential map at 0 is the identity from $\mathfrak{g}$ to $\mathfrak{g}$ and the inverse function theorem shows that the exp gives a diffeomorphism from a neighborhood of $0 \in \mathfrak{g}$ to a neighborhood of $e \in G$.

Example:.

1. When $G \subset \mathrm{GL}(n, \mathbb{R})$, then

   $$\exp(A) = e^A = 1 + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

   This is because e^{tA} is the unique 1-PSG whose derivative at $t = 0$ is A.

2. Let $G = \mathbb{R}$, so $\mathfrak{g} = \mathbb{R}$. We have $\exp(a) = a$.

3. Let $G = S^1 = \{z \in \mathbb{C} \,||z| = 1\}$, then it is identified with $U(1)$. We have $\mathfrak{g} = i\,\mathbb{R}$. The 1-PSG $\lambda_\xi(t) = e^{itx}$ and thus $\exp(ix) = e^{ix}$.

   We can also identify S^1 with $\mathbb{R}/\mathbb{Z}$. If so, item 2 gives us $\exp(x) = x$.

We can use the above Example 1 to calculate the Lie algebra of matrix groups more directly. Since they are subgroups of $\mathrm{GL}(n, \mathbb{R})$, the Lie algebras are subalgebras of $\mathfrak{gl}(n, \mathbb{R})$, by Theorem 2.3.5 and the local diffeomorphism exp.

Example:.

1. Let us consider the orthogonal group $\mathrm{O}(n)$ again. By definition, $A \in \mathfrak{o}(n)$ if and only if $(e^{tA})^T e^{tA} = I_n$, which is equivalent to saying $e^{tA^T} = e^{-tA}$. Since the exponential map is locally a diffeomorphism, we conclude that $A \in \mathfrak{o}(n)$ if and only if $A^T = -A$. Hence $\mathfrak{o}(n) = A_n$.

   Since $\mathrm{SO}(n)$ is the identity component of $\mathrm{O}(n)$, we know the Lie algebra $\mathfrak{so}(n)$ is the same as $\mathfrak{o}(n)$.

2. For the special linear group $\mathrm{SL}(n, \mathbb{R})$, and any $A \in \mathfrak{sl}(n, \mathbb{R})$, we have $\det e^A = 1$. However, $\det e^A = e^{\mathrm{tr}(A)}$. Hence, $A \in \mathfrak{sl}(n, \mathbb{R})$ if and only if $\mathrm{tr}(A) = 0$.

Proposition 2.4.4. *For any homomorphism $F : G \to H$ of Lie groups and any $\xi \in \mathfrak{g}$, we have* $\exp(F_*(\xi)) = F(\exp(\xi))$.

Proof. Both $F(\exp(t\xi))$ and $\exp(F_*(t\xi))$ are 1-PSG. They both have derivatives $F_*(\xi)$ at $t = 0$. By Proposition 2.4.3, we know $F(\exp(t\xi)) = \exp(F_*(t\xi))$, in particular $\exp(F_*(\xi)) = F(\exp(\xi))$. □

Proposition 2.4.5. *Let G, H be Lie groups. If G is connected, then any Lie group homomorphism $F : G \to H$ is uniquely determined by the map $F_* : \mathfrak{g} \to \mathfrak{h}$.*

Proof. Since $\exp(F_*(\xi)) = F(\exp(\xi))$ and the image of the exponential map contains a neighborhood of identity in G, this implies that F_* determines F in a neighborhood of identity in G.

Then the conclusion follows from the following claim

Claim: If G is connected, then any neighborhood U of identity generates G.

To show this, let $V = U \cap U^{-1}$. It is also an open neighborhood of e. It is clear that $G' = \cup_{k\geq 1} V^k$ is open. The cosets $g_1G' \cap g_2G' \neq \emptyset$ if and only if $g_1G' = g_2G'$. Therefore G is the disjoint union of open sets. Connectedness implies $G = G'$. □

2.4.1 The adjoint representation

There are other approaches to the definition of the bracket on $\mathfrak{g}$.

Lie group G acts on itself on the left by conjugation

$$\Psi : G \to \mathrm{Aut}(G),\ \ \Psi_g h = ghg^{-1}.$$

Then Ψ_g maps e to e, so we have the adjoint action $Ad_g = (\Psi_g)_* \in \mathrm{GL}(\mathfrak{g})$. In particular, we have $(Ad_g)^{-1} = Ad_{g^{-1}}$. Then the homomorphism $Ad : G \to \mathrm{GL}(\mathfrak{g})$ has a derivative at the identity. This induced map is denoted by

$$ad = Ad_* : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}).$$

This is called the infinitesimal adjoint representation.

Proposition 2.4.6. *For $X, Y \in \mathfrak{g}$, we have $ad_X Y = [X, Y]$.*

Some authors use this as an alternative definition of Lie bracket on $\mathfrak{g}$.

Proof. Let $\tilde{X}, \tilde{Y}$ be the left-invariant extensions of vectors $X, Y \in \mathfrak{g}$. Let

$f \in C^\infty(G)$.

$$\begin{aligned}
ad_X Y(f) =& \frac{d}{dt}|_{t=0}(Ad_{\exp(tX)}Y)(f) \\
=& \frac{\partial^2}{\partial t \partial u}|_{t=u=0} f(\exp(tX)\exp(uY)\exp(-tX)) \\
=& \frac{\partial^2}{\partial t \partial u}|_{t=u=0} f(\exp(tX)\exp(uY)) + \frac{\partial^2}{\partial u \partial t}|_{t=u=0} f(\exp(uY)\exp(-tX)) \\
=& (\tilde{X}\tilde{Y} - \tilde{Y}\tilde{X})f|_e \\
=& [X,Y]f
\end{aligned}$$

We use the fact of chain rule that if $F(t_1, t_2)$ is a function of two real variables,

$$\frac{d}{dt}|_{t=0}F(t,t) = \frac{\partial F}{\partial t_1}(0,0) + \frac{\partial F}{\partial t_2}(0,0).$$

□

The following is an immediate corollary of the definition.

Corollary 2.4.7. *The Lie algebra* $\mathfrak{g}$ *of an abelian Lie group* G *has trivial Lie bracket* $[X,Y] = 0$ *for any* $X, Y \in \mathfrak{g}$.

Proof. The conjugation Φ is the identity map. So the differential Ad and ad are trivial maps. □

For matrix groups, we can check it easily by expanding the matrix exponential

$$e^{tX}Ye^{-tX} = Y + t(XY - YX) + O(t^2).$$

In fact, this gives another (more conceptual) way to show the Lie bracket of $\mathfrak{gl}(n, \mathbb{R})$ is the commutator.

The Jacobi identity

$$[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$$

could take two forms

$$([ad_x, ad_y])(z) = ad_{[x,y]}(z)$$

$$ad_x[y,z] = [ad_x y, z] + [y, ad_x z]$$

The first means ad is a Lie algebra homomorphism, *i.e.* taking brackets to brackets. The second means ad is a derivation. In other words, $ad : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$.

We have a simple relation with exponential map

$$Ad_{\exp(x)} = \exp(ad_x).$$

This is easy to see by checking the derivatives of the respective curves at $t = 0$.

Given a representation $\rho : G \to \mathrm{GL}(V)$ inducing Lie algebra homomorphism $\rho_* : \mathfrak{g} \to \mathrm{End}(V)$ the map $\xi \mapsto -\operatorname{tr}(\rho_*(\xi)^2)$ is a quadratic form on $\mathfrak{g}$, invariant under the adjoint action of G. The Killing form is the quadratic form defined in this way by the adjoint representation. Namely,

$$B(u, v) = -\operatorname{tr}(ad_u \circ ad_v)$$

In general these forms could be indefinite. Cartan characterizes "semisimple" Lie algebra by the non-degeneracy (positive definite for compact groups) of the Killing form.

Proposition 2.4.8. *The Killing form on a Lie algebra satisfies*

$$B(u, v) = B(v, u), B(ad_x y, z) + B(y, ad_x z) = 0.$$

In fact, it already follows from l properties of quadratic form in the above general setting. Below is a slightly different argument.

Proof. To show $B([x, y], z) + B(y, [x, z]) = 0$. $B([x, y], z)$ is the trace of

$$ad_x ad_y ad_z - ad_y ad_x ad_z,$$

while $B(y, [x, z])$ is the trace of

$$ad_y ad_x ad_z - ad_y ad_z ad_x.$$

The sum is zero because $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ for endomorphisms X, Y of a vector space. The same fact implies $B(x, y) = B(y, x)$. □

2.4.2 Coadjoint representation

Similarly, if $\mathfrak{g}^*$ is the dual of $\mathfrak{g}$, then the coadjoint action of G on $\mathfrak{g}^*$ is

$$Ad^* : G \to \mathrm{Aut}(\mathfrak{g}^*), \;\; g \mapsto Ad^*_g,$$

where Ad^*_g is defined by

$$\langle Ad^*_g \xi, X\rangle = \langle \xi, Ad_{g^{-1}} X\rangle$$

for all $\xi \in \mathfrak{g}^*, X \in \mathfrak{g}$. Write $ad^*_X : \mathfrak{g}^* \to \mathfrak{g}^*$ for the derivative of the coadjoint action at e along the direction $X\mathfrak{g}$. Since $ad_X Y = [X, Y]$,

$$ad^*_X \xi(Y) = -\xi([X, Y]).$$

Definition 2.4.9. *For any smooth action $\tau : G \to Diff(M)$, the orbit of G through $m \in M$ is*

$$G \cdot m = \{g \cdot m | g \in G\}.$$

The stabilizer (also called the isotropic subgroup) of $m \in M$ is the subgroup

$$G_m = \{g \in G | g \cdot m = m\}.$$

In particular, for Ad and Ad^*, we have adjoint and coadjoint orbits. The orbit $O = G/G_m$ where G_m is the stabilizer of some $m \in O$. For semisimple Lie groups, the existence of non-degenerate Killing form identifies adjoint orbits with coadjoint orbits.

In general, we will see in the following that any coadjoint orbit has a natural symplectic structure.

Definition 2.4.10. *A symplectic structure on M is a non-degenerate 2-form such that $d\omega = 0$.*

We have $T_m O = T_m(G/G_m) = \mathfrak{g}/\mathfrak{g}_m$. We want to define a skew symmetric 2-form on $T_m O$. We first define a skew symmetric form

$$\omega_\xi : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}, \ \ (x, y) \mapsto \xi([x, y]).$$

We have $g \in G_\xi \Leftrightarrow Ad_g^* \xi = \xi$. Therefore

$$X \in \mathfrak{g}_\xi \Leftrightarrow ad_X^* \xi = 0.$$

Therefore, since $\xi([X, Y]) = \omega_\xi(X, Y)$, we obtain

$$\omega_\xi(X, Y) = 0, \ \forall Y \in \mathfrak{g} \Leftrightarrow X \in \mathfrak{g}_\xi.$$

The assignment $\xi \mapsto \omega_\xi$ thus gives a nondegenerate 2-form ω on O. Moreover, it satisfies $d\omega = 0$. To prove this claim, we use the Cartan formula for exterior differential

$$\begin{aligned} d\omega(X, Y, Z) =& X(\omega(Y, Z)) + Z(\omega(X, Y)) + Y(\omega(Z, X)) \\ & -(\omega([X, Y], Z) + \omega([Z, X], Y) + \omega([Y, Z], X)) \end{aligned}$$

For the first line, we apply

$$X(\omega_\xi(Y, Z)) = X(\xi([Y, Z])) = \xi([X, [Y, Z]]),$$

for the second line, we apply

$$\omega_\xi([X, Y], Z) = \xi([[X, Y], Z]),$$

as well as Jacobi identity, we know both vanish. Hence,

Theorem 2.4.11. *Any coadjoint orbit $O \subset \mathfrak{g}^*$ has a natural symplectic structure.*

This symplectic structure is sometimes called the Kirillov-Kostant-Souriau symplectic structure.

2.5 Baker-Campbell-Hausdorff formula

Through exponential map, the group law is determined by its Lie algebra. We will see how it goes in a precise manner. The following gives the first indication.

Proposition 2.5.1. *If G is a Lie group, and $X, Y \in \mathfrak{g}$ be such that $[X, Y] = 0$. Then* $\exp(X)\exp(Y) = \exp(X+Y) = \exp(Y)\exp(X)$.

Proof. By $ad_X Y = [X, Y] = 0$, we have $Ad_{\exp(X)}Y = \exp(ad_X)Y = Y$. Hence the 1-PSGs $\exp(X)\exp(tY)\exp(-X)$ and $\exp(tY)$ have the same derivative at $t = 0$: $Ad_{\exp(X)}Y = Y$. This implies $\exp(X)\exp(Y) = \exp(Y)\exp(X)$.

It implies $\exp(tX)\exp(tY)$ is a 1-PSG. Moreover, it has the same derivative, $X+Y$, as the 1-PSG $\exp(t(X+Y))$. Hence, $\exp(X)\exp(Y) = \exp(X+Y)$. □

In particular, it implies the following.

Corollary 2.5.2. *If G is an abelian Lie group and $\mathfrak{g}$ is its Lie algebra, then* $\exp : \mathfrak{g} \to G$ *is a Lie group homomorphism.*

In general, Lie bracket operation on $\mathfrak{g}$ measures the non-commutativity of the multiplication of G. In the following, we would like to find out the difference between $\exp(X)\exp(Y)$ and $\exp(X+Y)$. Without confusing, we will denote X both for an element in $\mathfrak{g}$ and its corresponding left invariant vector field. Then

$$(Xf)(a) = X_a f = \frac{d}{dt}|_{t=0} f(a\exp(tX))$$

for any $f \in C^\infty(G)$ and any $a \in G$. More generally, for any $t \in \mathbb{R}$,

$$\begin{aligned}(Xf)(a\exp(tX)) =& \frac{d}{ds}|_{s=0} f(a\exp(tX)\exp(sX)) = \frac{d}{ds}|_{s=0} f(a\exp((t+s)X)\\ =& \frac{d}{dt} f(a\exp(tX)).\end{aligned}$$

Using this and induction, one can see that for any $k \geq 0$,

$$(X^k f)(a\exp(tX)) = \frac{d^k}{dt^k}(f(a\exp(tX))).$$

In particular,

$$(X^k f)(a) = \frac{d^k}{dt^k}|_{t=0} f(a\exp(tX))$$

The formulae could be generalized to multi-variable,

$$(X_1 \cdots X_k f)(a) = \frac{\partial^k}{\partial t_1 \cdots \partial t_k}|_{t_1=\cdots=t_k=0} f(a \exp(t_1 X_1) \cdots \exp(t_k X_k)).$$

As a consequence, we have the following Taylor expansion.

Proposition 2.5.3. *If f is a smooth function on G, then for small $|t|$,*

$$f(\exp(t_1 X_1) \cdots \exp(t_k X_k)) = f(e) + \sum_i t_i X_i f(e) + \frac{1}{2}(\sum_i t_i^2 X_i^2 f(e) + 2 \sum_{i<j} t_i t_j X_i X_j f(e)) + O(t^3)$$

This formula holds for vector-valued functions as well. Then we can see how Lie brackets measures the non-commutativity for G.

Theorem 2.5.4. *Let $n \geq 1$ and $X_1, \cdots, X_n \in \mathfrak{g}$. Then for $|t|$ sufficiently small,*

$$\exp(tX_1) \cdots \exp(tX_n) = \exp(t \sum_{1 \leq i \leq n} X_i + \frac{t^2}{2} \sum_{1 \leq i < j \leq n} [X_i, X_j] + O(t^3)).$$

Proof. We apply Proposition 2.5.3 to $f(\exp(tX)) = tX$ for t small. Then $f(e) = 0$ and

$$Xf(e) = \frac{d}{dt}|_{t=0} f(\exp(tX)) = \frac{d}{dt}|_{t=0}(tX) = X, \;\; X^n f(e) = 0.$$

Since

$$\sum_i X_i^2 + 2 \sum_{i<j} X_i X_j = (X_1 + \cdots + X_n)^2 + \sum_{i<j} [X_i, X_j],$$

we have the desired formula. □

In particular,

$$\exp(tX) \exp(tY) = \exp(tX + tY + \frac{t^2}{2}[X, Y] + O(t^3))$$

for $|t|$ small. For the higher order terms, we have the following Baker-Campbell-Hausdorff formula.

Theorem 2.5.5. *For small enough $X, Y \in \mathfrak{g}$, we have*

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

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

$$\mu(X,Y) = X + Y + \sum_{m\geq 2} \mu_m(X,Y),$$

where $\mu_m(X,Y)$ is a Lie polynomial in X,Y of degree m, i.e. a formal polynomial (in general non-commutative, non-associative) of X,Y of degree m whose products are Lie brackets.

This expression is independent of the Lie algebra $\mathfrak{g}$ or the choice of X,Y.

There is a more explicit formula.

Theorem 2.5.6 (Dynkin's formula). *For X,Y small,*

$$\mu(X,Y) = X+Y+\sum_{k=1} \frac{(-1)^k}{k+1} \sum \frac{(-1)^{\sum_i (l_i+m_i)}}{l_1+\cdots+l_k+1} \frac{(ad_Y)^{l_1}}{l_1!} \circ \frac{(ad_X)^{m_1}}{m_1!} \circ \cdots \circ \frac{(ad_Y)^{l_k}}{l_k!} \circ \frac{(ad_X)^{m_k}}{m_k!}(Y)$$

To write first several terms:

$$\mu(X,Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] + [Y,[Y,X]]) + \cdots$$

We will not prove the above two theorems, since higher order terms will not be used in this course.

2.6 Lie subgroups

Let us first recall various notions of submanifolds of M, from weakest to strongest.

Immersion: S is a manifold equipped with a map $\phi : S \to M$ such that ϕ_* is injective for every $s \in S$.
Injective Immersion: S a manifold equipped with a globally injective immersion.
Embedding: An injective immersion where the intrinsic topology of S agree with the induced subspace topology.

Example 2.6.1. *Consider torus $T^2 = S^1 \times S^1$. For any coprime pair of integers (p,q),*

$$H^{p,q} := \{(e^{ipt}, e^{iqt}) | t \in \mathbb{R}\}$$

is an embedded submaniold.

For any irrational number α,

$$H^{\alpha} := \{(e^{it}, e^{i\alpha t}) | t \in \mathbb{R}\}$$

is not embedded, but is injective immersed.

Closed Embedding: An embedding in which $\phi(S) \subset M$ is closed.

Definition 2.6.2. *If G is a Lie group, a Lie subgroup is an injective immersed submanifold H with the inclusion map $f : H \to G$ a homomorphism of groups.*

So all the lines on a torus are Lie subgroups.

This implies f is a homomorphism of Lie groups. Let H be a Lie subgroup of G. The induced map $f_* : \mathfrak{h} \to \mathfrak{g}$ is injective and a homomorphism of Lie algebras and hence is an isomorphism of $\mathfrak{h}$ with a Lie subalgebra of $\mathfrak{g}$.

Conversely, any Lie subalgebra gives rise to some Lie subgroup. For that, we need the notion of integrable distribution. Details can be found in *e.g.* Warner.

A k-dimensional distribution on a manifold M is a k-dimensional subbundle $\mathcal{V} \subset TM$ which assigns to every $p \in M$ a k-dimensional vector subspace $\mathcal{V}_p$ of T_pM. $\mathcal{V}$ is called smooth if for every $p \in M$, there is a neighborhood U of p and smooth vector fields $X_1, \cdots, X_k$ such that for every $q \in U$, $X_1(q), \cdots, X_k(q)$ are a basis of $\mathcal{V}_q$.

An integral manifold for a distribution $\mathcal{V}$ is a k-dimensional submanifold $N \subset M$ such that at every point $p \in N$, we have $T_pN = \mathcal{V}_p$. This is a straightforward generalization of the notion of an integral curve for direction field in ODE theory. However, for $k > 1$, the existence of integral manifold is not automatic. We say that a distribution $\mathcal{V}$ is integrable if through each point of M there exists an integral manifold of $\mathcal{V}$. Suppose N is an integral manifold for $\mathcal{V}$ at p, and $X_p, Y_p \in \mathcal{V}_p = T_pN$. Then we have $[X_p, Y_p] \in T_pN$. The converse is also true.

Theorem 2.6.3 (Frobenius Theorem)**.** *A distribution $\mathcal{V}$ is integrable if and only if for any two vector fields $X, Y \in \mathcal{V}$, one has $[X, Y] \in \mathcal{V}$. In this situation, then through every point $p \in M$, there is a unique maximal connected integral manifold of $\mathcal{V}$.*

In general, the integral submanifold is not even an embedded submanifold but only an immersed one.

Theorem 2.6.4. *If $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$, then there is a unique connected Lie subgroup H of G with Lie algebra $\mathfrak{h}$.*

Proof. Let $X_1, \cdots, X_k$ be a basis of $\mathfrak{h} \subset \mathfrak{g}$. Let $\tilde{X}_i$ be corresponding left invariant vector fields. Since they are linearly independent at e, they are linearly independent at all $g \in G$. In other words,

$$\mathcal{V}_g = span\{\tilde{X}_1(g), \cdots, \tilde{X}_k(g)\}$$

is a k-dimensional distribution on G. Since $[X_i, X_j] \in \mathfrak{h}$ for all i, j, by Frobenius theorem, $\mathcal{V}$ is integrable. That is, there is a unique maximal connected integral manifold of $\mathcal{V}$ through e. Denote this by H.

To show that H is a subgroup, note that $\mathcal{V}$ is a left invariant distribution. So the left translation of any integral manifold is an integral manifold. Now suppose $h_1, h_2 \in H$. Since $h_1 = l_{h_1} e \in H \cap l_{h_1} H$ and H is maximal, we have $l_{h_1} H \subset H$. So in particular $h_1 h_2 = l_{h_1} h_2 \in H$. Similarly, $h_1^{-1} \in H$ since $l_{h_1^{-1}}(h_1) = e \in H$ implies $l_{h_1^{-1}} H \subset H$. It follows that H is a subgroup of G. Since the group operations on H are restriction from that of G, they are smooth. So H is a Lie group.

To show uniqueness, let K be another connected Lie subgroup with Lie algebra $\mathfrak{h}$. Then K is also an integral manifold of $\mathcal{V}$. So we have $K \subset H$. Since $T_e K = T_e H$, the inclusion has to be a local isomorphism. In other words, K coincides with H near e. Since any connected Lie group is generated by any open set containing e, we conclude that $K = H$. □

2.6.1 Closed Lie subgroups

Definition 2.6.5. *A Lie subgroup H of G is said to be a closed Lie subgroup if H is also an embedded submanifold of G.*

As in Example 2.6.1, each $H^{p,q}$ is a closed Lie subgroup. However, H^α is not.

Lemma 2.6.6. *Suppose G is a Lie group, H is a subgroup of G which is an embedded submanifold as well. Then H is closed in the sense of topology.*

Proof. Since H is an embedded submanifold, it is locally closed everywhere. In particular, one can find an open neighborhood U of e in G such that $U \cap H = U \cap \bar{H}$. Take $h \in \bar{H}$. Since $hU \cap H \neq \emptyset$, we choose an element h' from it. Hence $h^{-1}h' \in U$. For any sequence $h_n \in H$ converging to h, we know $h_n^{-1}h' \in H$ converges to $h^{-1}h'$. In other words, $h^{-1}h' \in U \cap \bar{H} = U \cap H$. So $h \in H$, which implies H is closed. □

In other words, a closed Lie subgroup must be a subgroup which is closed in the sense of topology. A remarkable theorem due to E. Cartan claims that the inverse is also true, *i.e.* any subgroup which is also a closed subset must be a Lie subgroup.

Theorem 2.6.7. *Let G be a Lie group, $H \subset G$ a subgroup in the algebraic sense. If H is closed as a subset of G, then H has a unique structure of a Lie subgroup.*

Proof. Define $\mathfrak{h} := \{X \in \mathfrak{g} | \exp tX \in H, \forall t\}$.

Lemma 2.6.8. $\mathfrak{h}$ *is a linear subspace of* $\mathfrak{g}$.

Proof. Clearly $\mathfrak{h}$ is closed under scalar multiplication. It is closed under vector addition because for any $t \in \mathbb{R}$,

$$\exp(t(X+Y)) = \lim_{n\to\infty}(\exp(\frac{t(X+Y)}{n}+O(\frac{1}{n^2}))^n = \lim_{n\to\infty}(\exp(\frac{tX}{n})\exp(\frac{tY}{n}))^n \in H.$$

□

Lemma 2.6.9. *If U is a sufficiently small open neighborhood of $0 \in \mathfrak{g}$, then* $\exp(U \cap \mathfrak{h}) = \exp(U) \cap H$.

Proof. Note that $\exp \mathfrak{h} \subset H$ by definition. Therefore, $\exp(V \cap \mathfrak{h}) \subset H \cap \exp V$ for any open neighborhood V of $0 \in \mathfrak{g}$. Let $\mathfrak{h}' \subset \mathfrak{g}$ be a linear complement, *i.e.* $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}'$. Then the map

$$\Phi : \mathfrak{h} \oplus \mathfrak{h}' \to G, \quad (X, Y) \mapsto \exp X \exp Y$$

is a local diffeomorphism of a neighborhood of $0 \in \mathfrak{g}$ into a neighborhood of $e \in G$. If the lemma were false, we could find a sequence of vectors $X_i + Y_i \in \mathfrak{h} \oplus \mathfrak{h}'$ with $Y_i \neq 0$ such that $X_i + Y_i \to 0$ and $\Phi(X_i + Y_i) \in H$. Since $\exp(X_i) \in H$, we must have $\exp(Y_i) \in H$ for all i. We let Y be a limit point of $\frac{Y_i}{|Y_i|}$. We still denote the converging subsequence by Y_i, *etc.* Let $t \in \mathbb{R}$ be fixed, we take $n_i = [\frac{t}{|Y_i|}]$ be the integer part of $\frac{t}{|Y_i|}$. Then

$$\exp(tY) = \lim_{i\to\infty} \exp(n_i Y_i) = \lim_{i\to\infty} \exp(Y_i)^{n_i} \in H.$$

Since this holds for all t, it must be the case that $Y \in \mathfrak{h}$. Hence $Y \in \mathfrak{h} \cap \mathfrak{h}' = \{0\}$. Contradict to the fact that Y is a unit vector. □

Near the origin, exp is a diffeomorphism. By the preceding lemma, e has an open neighborhood U_e such that $H \cap U_e$ is a closed embedded submanifold of U_e. By translation, every $h \in H$ has an open neighborhood U_h with the same property. Since H is closed, for any $g \in G \setminus H$, there exists an neighborhood U_g of g such that $U_g \cap H = \emptyset$. Therefore, H is a closed embedded submanifold. The multiplication and inverse are restrictions from that of G, hence smooth. □

Here is an immediate consequence.

Corollary 2.6.10. *If $\phi : G \to H$ is a Lie group homomorphism, then* $\ker \phi$ *is a closed Lie subgroup of G whose Lie algebra is* $\ker(\phi_*)$.

Proof. $\ker\phi$ is a subgroup of G which is also a closed subset. According to Cartan's theorem, $\ker\phi$ is a Lie subgroup. Its Lie algebra is given by $\{X \in \mathfrak{g} | \exp(tX) \in \ker\phi, \forall t\}$. The theorem follows since $\exp(t\phi_*(X)) = \phi(\exp(tX)) = e, \forall t$, which is equivalent to saying $\phi_*(X) = 0$. □

An important consequence is a "continuity implies smoothness" result.

Theorem 2.6.11. *Every continuous homomorphism of Lie groups is smooth.*

Proof. Let $\phi : G \to H$ be a continuous homomorphism, then $\Gamma_\phi = \{(g, \phi(g)) | g \in G\}$ is a closed subgroup, and thus a Lie subgroup of $G \times H$. The projection

$$p : \Gamma_\phi \to G \times H \to G$$

is bijective, smooth and is a Lie group homomorphism. It follows that dp is a constant rank map, and thus has to be bijective at each point. So p is a local diffeomorphism everywhere. Since p is globally invertible, p is also a global diffeomorphism. The map $\phi = pr_2 \circ p^{-1}$ is smooth. □

As a consequence, for any topological group G, there is at most one smooth structure on G to make it a Lie group. (However, it is possible that one group admits two different topologies and thus have different Lie group structures.)

2.6.2 Lie group homomorphisms

Let X be a topological space. A covering space of X is a topological space C together with a continuous surjective map $p : C \to X$, such that for every $x \in X$, there exists an open neighborhood U of x, such that $p^{-1}(U)$ (the inverse image of U under p) is a union of disjoint open sets in C, each of which is mapped homeomorphically onto U by p.

Proposition 2.6.12. *Suppose $\Phi : G \to H$ is a homomorphism of connected Lie groups such that on the level of Lie algebras, $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ is bijective, then Φ is a covering map.*

Proof. By left invariance, it suffices to check the covering property at $e \in H$. Since $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ is bijective, Φ maps a neighborhood $\mathcal{U}$ of e in G bijectively to a neighborhood $\mathcal{V}$ of e in H. Hence, Φ is surjective. Let $\Gamma = \Phi^{-1}(e) \subset G$. Then Γ is a subgroup of G. Moreover, for any $a \in \Gamma$,

$$\Phi \circ l_a(g) = \Phi(ag) = \Phi(g),$$

we have $\Phi^{-1}(\mathcal{V}) = \cup_{a\in\Gamma} l_a\mathcal{U}$. The Proposition is proved if we can show $l_{a_1}\mathcal{U} \cap l_{a_2}\mathcal{U} = \emptyset$ for $a_1 \neq a_2 \in \Gamma$. If $l_{a_1}\mathcal{U} \cap l_{a_2}\mathcal{U} \neq \emptyset$, then $l_a\mathcal{U} \cap \mathcal{U} \neq \emptyset$ for $a = a_1^{-1}a_2$.

So we have $p_1, p_2 \in \mathcal{U}$ such that $p_2 = ap_1 \in l_a\mathcal{U} \cap \mathcal{U}$. Then $\Phi(p_1) = \Phi(p_2)$. However, Φ is one-to-one on $\mathcal{U}$. So $p_1 = p_2$. It follows $a = e$ and $a_1 = a_2$. This proves Φ is a covering map. $\square$

A topological space X is called simply-connected if it is path-connected and any continuous map $f : S^1 \to X$ can be contracted to a point in the following sense: there exists a continuous map $F : D^2 \to X$ (where D^2 denotes the unit disk in $\mathbb{R}^2$) such that F restricted to S^1 is f.

An equivalent formulation is this: X is simply-connected if and only if it is path-connected, and whenever $p : [0, 1] \to X$ and $q : [0, 1] \to X$ are two paths with the same start and endpoint ($p(0) = q(0)$ and $p(1) = q(1)$), then p and q are homotopic relative $\{0, 1\}$. Intuitively, this means that p can be "continuously deformed" to get q while keeping the endpoints fixed. Hence the term simply connected: for any two given points in X, there is one and "essentially" only one path connecting them.

Example 2.6.13. *$S^n, n \neq 1$, $\mathbb{R}^n$ are simply connnected.*

Every path connected covering map of a simply-connected and path connected space is a homeomorphism. A covering of X, $\pi : \tilde{X} \to X$, is called a universal covering if $\tilde{X}$ is simply connected. In fact, $\tilde{X}$ is a set of all paths $p : [0, 1] \to X$ such that $p(0) = x_0$ modulo the equivalence relation of path-homotopy. The covering map π is defined as $\pi(p) = p(1)$. If X is simply connected, $\pi : \tilde{X} \to X$ is a homeomorphism.

When G is a Lie group, its universal covering space $\tilde{G}$ also admits a group structure and the projection $\pi : \tilde{G} \to G$ is a Lie group homomorphism. Its kernel is the fundamental group $\pi_1(G)$.

Example 2.6.14. *$\pi : \mathbb{R} \to S^1$ is the universal covering of S^1. Hence, the fundamental group $\pi_1(S^1) = \mathbb{Z}$.*

Theorem 2.6.15. *Let G, H be Lie groups with Lie algebras $\mathfrak{g}, \mathfrak{h}$ and $\phi : \mathfrak{g} \to \mathfrak{h}$ a Lie algebra homomorphism. Suppose G, H are connected and G is simply connected. Then there exists a unique Lie group homomorphism $\Phi : G \to H$ such that $\Phi_* = \phi$.*

Proof. Consider the graph of ϕ as a subalgebra of $\mathfrak{g} \oplus \mathfrak{h}$:

$$\{(X, \phi(X))|X \in \mathfrak{g}\} \subset \mathfrak{g} \oplus \mathfrak{h}.$$

By Theorem 2.6.4, there exists a Lie subgroup $\Gamma \subset G \times H$ with this Lie algebra. Let p_1, p_2 be the projection maps to G and H respectively. $(p_1)_*$ restricts to the graph of ϕ is bijective, so p_1 is a covering map. Since G is simply connected, this is an isomorphism. The composition $\Phi = p_2 \circ p_1^{-1}$ is a Lie group homomorphism with $\Phi_* = \phi$. $\square$

This could be applied to representations of Lie groups and Lie algebras.

Definition 2.6.16. *A representation of a Lie algebra* $\mathfrak{g}$ *is a vector space* V *together with a homomorphism* $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$.

Let V be a complex vector space and $H = \mathrm{GL}(V)$ in Corollary 2.6.15, we have

Proposition 2.6.17. *Let* G *be a Lie group with Lie algebra* $\mathfrak{g}$. *Every representation* $\rho : G \to GL(V)$ *defines a representation* $\rho_* : \mathfrak{g} \to \mathfrak{gl}(V)$. *If* G *is connected and simply connected, then every representation of* $\mathfrak{g}$ *can be uniquely lifted to a representation of* G.

We talk about the fundamental group of a Lie group.

Lemma 2.6.18. *Let* G *be a connected topological group and* Γ *a discrete normal subgroup. Then* $\Gamma \subset Z(G)$.

Proof. Let $\gamma \in \Gamma$. Then $g \mapsto g\gamma g^{-1}$ is a continuous map $G \to \Gamma$. Since G is connected and Γ is discrete, it is constant, so $g\gamma g^{-1} = \gamma$ for all g. Therefore $\gamma \in Z(G)$. □

Proposition 2.6.19. *If* G *is a connected Lie group, then the fundamental group* $\pi_1(G)$ *is abelian.*

Proof. Let $\pi : \tilde{G} \to G$ be the universal cover. We have $\ker(\pi) = \pi_1(G)$. This is a discrete normal subgroup of $\tilde{G}$ and hence contained in $Z(\tilde{G})$ by Lemma 2.6.18. In particular, it is abelian. □

2.7 Riemannian geometry of Lie groups

A Riemannian manifold consists of a smooth manifold M and for every $x \in M$ an inner product (*i.e.* a positive definite symmetric bilinear form) $\langle , \rangle$ on the tangent space T_xM, such that for vector fields X and Y, $x \mapsto \langle X(x), Y(x) \rangle$ is a smooth function. Any smooth manifold admits a Riemannian metric. This implicitly uses the fact that a locally compact second countable Hausdorff space is paracompact, which means every open cover has a locally finite refinement.

In a system of local coordinates on the manifold M given by real functions $x_1, \cdots, x_n$, the vector fields $\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}\}$ give a basis of tangent vectors at each point of M. The metric tensor can be written in terms of the dual basis $\{dx_1, \cdots, dx_n\}$ of the cotangent bundle as $g = \sum_{i,j} g_{ij} dx_i \otimes dx_j$ where $g_{ij}(x) := \langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \rangle$.

The covariant derivative is a generalization of the directional derivative from vector calculus. More precisely, a covariant derivative ∇ assigns a tangent vector field $\nabla_X Y$ to each pair (X, Y) of vector fields X, Y, such that the following holds:

1. $\nabla_X Y$ is linear in X, *i.e.* $\nabla_{\alpha X+\beta Z} Y = \alpha \nabla_X Y + \beta \nabla_Z Y$, for functions α, β.
2. $\nabla_X Y$ is additive in Y, *i.e.* $\nabla_X (Y_1 + Y_2) = \nabla_X Y_1 + \nabla_X Y_2$.
3. $\nabla_X Y$ obeys the product rule, *i.e.* $\nabla_X (fY) = X(f)Y + f\nabla_X Y$ for functions f.

It can be written in local coordinate using Christoffel symbols. On a Riemannian manifold, there is a special one, called Levi-Civita connection, characterized by the conditions

$$\nabla_X Y - \nabla_Y X = [X, Y],$$
$$X(\langle Y, Z\rangle) = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.$$

Exercise: For Levi-Civita connection, show the Koszul formula

$$\begin{aligned} 2\langle \nabla_X Y, Z\rangle =& X(\langle Y, Z\rangle) + Y(\langle X, Z\rangle) - Z(\langle X, Y\rangle) \\ & - \langle Y, [X, Z]\rangle - \langle X, [Y, Z]\rangle - \langle Z, [Y, X]\rangle \end{aligned} \tag{2.1}$$

We now look at Riemannian metrics on a Lie group. A Riemannian metric on a Lie group G is called left-invariant if and only if

$$\langle X, Y\rangle_h = \langle (l_g)_* X, (l_g)_* Y\rangle_{gh}$$

for all $g, h \in G$ and $X, Y \in T_h G$. Similarly, we say a metric is right invariant if r_g preserve the metric. A metric that is both left-invariant and right-invariant is called bi-invariant.

Exercise: There is a bijective correspondence between left-invariant metrics on a Lie group G and inner products on the Lie algebra $\mathfrak{g}$.

Hint: If $\langle , \rangle$ is an inner product on $\mathfrak{g}$, set $\langle u, v\rangle_g = \langle (l_{g^{-1}})_* u, (l_{g^{-1}})_* v\rangle$, for all $u, v \in T_g G$.

Exercise: For a left-invariant metric $\langle , \rangle$ on G and any two left-invariant vector fields X, Y, we have $Z(\langle X, Y\rangle) = 0$ for any vector field Z.

Another exercise characterizes bi-invariant metric by adjoint action.

Lemma 2.7.1. *A left invariant metric on G is also right invariant if and only if for each $g \in G$ and $u, v \in \mathfrak{g}$,*

$$\langle Ad_g u, Ad_g v \rangle = \langle u, v \rangle. \tag{2.2}$$

When G is connected, the relation (2.2) *is equivalent to*

$$\langle [X, Y], Z \rangle = \langle X, [Y, Z] \rangle. \tag{2.3}$$

Proof. The first statement is left as an exercise.

For the second, (2.2) implies (2.3) by taking the derivative of the relation $\langle Ad_{\exp tY} X, Z \rangle = \langle X, Ad_{\exp(-tY)} Z \rangle$.

For the other direction, since a connected Lie group is generated by any neighborhood of the identity, we can assume g is sufficiently close to the identity such that $g = \exp(Y)$ for some uniquely defined Y close to zero. We have the identity $Ad_g = e^{ad_Y}$. (2.2) is same as saying $Ad_{g^{-1}} = Ad_g^*$. This is equivalent to $-ad_Y = ad_Y^*$. Hence (2.3) implies (2.2). □

Since the Killing form is Ad-invariant. Hence, if the Lie group is compact and semisimple, the Killing form provides a bi-invariant Riemannian metric. By averaging process, we can construct bi-invariant metric on any compact group. It uses Haar measure. If G is a Lie group there exists a left-invariant non-zero differential form of top degree, unique up to scaling. This could be constructed as the left invariant extension of such a differential form on the Lie algebra. This defines Haar measure. In general, when G is a locally compact Hausdorff topological group, Haar proves there exists a left-invariant regular Borel measure. It is unique up to scaling and satisfies $\mu(U) > 0$ for every nonempty open Borel set U.

Then for any $u, v \in T_e G$, and any inner product on $\mathfrak{g}$, we define

$$(u, v) := \int_G \langle Ad_g u, Ad_g v \rangle \mu(g).$$

When G is compact, $\mu(G) < \infty$. This gives an Ad-invariant inner product on $\mathfrak{g}$ and thus a bi-invariant metric on G.

In fact, we know the connected Lie group admits a bi-invariant metric if and only if it is isomorphic to the cartesian product of a compact group and an additive vector group.

Now, (2.1), Lemma 2.7.1 and the exercise before it altogether imply that for left-invariant vector fields, the Levi-Civita connection of a bi-invariant metric has the form

$$\nabla_X Y = \frac{1}{2}[X, Y]. \tag{2.4}$$

Definition 2.7.2. *A smooth curve $\gamma(t)$ is called a geodesic if $\nabla_{\gamma'(t)}\gamma'(t) = 0$.*

Since 1-PSGs are integral curves of left-invariant vector fields, we can apply (2.4) to know they are geodesics. Moreover, since geodesic equation is a 2nd order ODE, there is a unique geodesic passing through any given point and tangent to a given direction. Hence

Proposition 2.7.3. *The geodesics through $e \in G$ are the 1-PSGs.*

In the language of Riemannian geometry, it implies the exp in Lie group coincides with the notion of exponential map in Riemannian geometry when G admits a bi-invariant metric.

Corollary 2.7.4. *The exponential map of a compact connected Lie group is surjective.*

Proof. By Riemannian geometry, if M is a compact connected Riemannian manifold (see Kobayashi-Nomizu, Theorem 4.2 for example), and $x, y \in M$, then there is a geodesic $p : [0,1] \to M$ such that $p(0) = x$ and $p(1) = y$. Take $x = e$ and $y = g$. By Proposition 2.7.3, the geodesics are exponential maps $\exp(tX)$ for some $X \in \mathfrak{g}$. So $g = \exp(X)$. □

Exercise: Show the exponential map $\exp : \mathfrak{sl}(2,\mathbb{R}) \to \mathrm{SL}(2,\mathbb{R})$ is not surjective.

Hint: Which diagonal matrices are in the image?

2.8 Maximal Tori

We start with a general result on abelian Lie group.

Proposition 2.8.1. *Any connected abelian Lie group is of the form $T^r \times \mathbb{R}^k$.*

Proof. Let G be a connected abelian Lie group, then we have seen in Corollary 2.5.2 (and 2.7.4) that $\mathfrak{g} \cong \mathbb{R}^n$ and $\exp : \mathfrak{g} \to G$ is a surjective Lie group homomorphism, so G is isomorphic to $\mathfrak{g}/\ker(\exp)$.

On the other hand, $\ker(\exp)$ is a Lie subgroup of $(\mathfrak{g}, +)$, and it is discrete since exp is a local diffeomorphism near e. We can show it is a lattice, *i.e.* there exist linearly independent vectors $v_1, \cdots, v_r$ such that $\ker(\exp) = \{n_1v_1 + \cdots + n_rv_r | n_i \in \mathbb{Z}\}$ (see *e.g.* the 2016 lecture notes of algebraic number theory). Let V_1 be its span and V_2 be its complement in $\mathfrak{g}$, we have

$$G = \mathfrak{g}/\ker(\exp) = V_1/\ker(\exp) \times V_2 = T^r \times \mathbb{R}^k .$$

□

A torus is a Lie group isomorphic to $T^n = \mathbb{R}^n / \mathbb{Z}^n$. Any compact connected abelian Lie group is a torus.

A generator of a compact torus T is an element t such that the smallest closed subgroup of T containing t is T itself.

Proposition 2.8.2. *Every compact torus T has a generator. Indeed, generators are dense in T.*

Proof. Let $T = (\mathbb{R} / \mathbb{Z})^r$. Choose $(t_1, \cdots, t_r) \in \mathbb{R}^r$. The result follows from the following Kronecker theorem. □

Theorem 2.8.3. *The image of $t = (t_1, \cdots, t_r)$ in $T = (\mathbb{R} / \mathbb{Z})^r$ is a generator of T if and only if $1, t_1, \cdots, t_r$ are linearly independent over $\mathbb{Q}$.*

Proof. $1, t_1, \cdots, t_r$ are linearly dependent if and only if $t \mod \mathbb{Z}^r$ is in the kernel of a nontrivial homomorphism

$$f : T^r \to S^1, f(v_1, \cdots, v_r) = e^{2\pi i(\alpha_1 v_1 + \cdots + \alpha_r v_r)}, \alpha_j \in \mathbb{Z}.$$

If $1, t_1, \cdots, t_r$ are linearly dependent, we assume $f(t) = 1$. If f is nontrivial, this kernel is not all of T and hence is a proper closed subgroup of T, so $[t]$ cannot be a generator.

Conversely, a non-generator $[t]$ is contained in a proper closed subgroup $H \subset T$, and the quotient group is a non trivial compact connected abelian Lie group. Thus T/H is a torus $T^k, k > 0$, and $[t]$ is in the kernel of the nontrivial homomorphism

$$T \to T/H \cong T^k = S^1 \times \cdots \times S^1 \stackrel{pr_1}{\to} S^1.$$

Hence $1, t_1, \cdots, t_r$ are linearly dependent.

This completes the proof. □

Definition 2.8.4. *Let G be a compact Lie group. A subgroup $T \subset G$ is a maximal torus if T is a torus and there is no other torus T' with $T \subsetneq T' \subset G$.*

Example:

1. A maximal torus T in $\mathrm{U}(n)$ is given by the diagonal matrices

$$\mathrm{diag}(e^{i\lambda_1}, \cdots, e^{i\lambda_n}).$$

   It is known by linear algebra that any unitary matrix can be diagonalized by unitary matrices. In other words, we have

$$\mathrm{U}(n) = \cup_{g \in \mathrm{U}(n)} gTg^{-1}.$$

2. A maximal torus in $\mathrm{SU}(n)$ is given by the diagonal matrices as above, with $\sum \lambda_i = 0$. So $\mathrm{SU}(n)$ has rank $n-1$.

3. A maximal torus in $\mathrm{SO}(2n)$ is given by

$$\mathrm{SO}(2) \times \cdots \times \mathrm{SO}(2) = \begin{pmatrix} \cos t_1 & -\sin t_1 & & & & & \\ \sin t_1 & \cos t_1 & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & \cos t_n & -\sin t_n \\ & & & & & \sin t_n & \cos t_n \end{pmatrix}$$

4. A maximal torus in $\mathrm{SO}(2n+1)$ is given by

$$\mathrm{SO}(2)\times\cdots\times\mathrm{SO}(2)\times\{1\} = \begin{pmatrix} \cos t_1 & -\sin t_1 & & & & & & \\ \sin t_1 & \cos t_1 & & & & & & \\ & & \cdot & & & & & \\ & & & \cdot & & & & \\ & & & & \cdot & & & \\ & & & & & \cos t_n & -\sin t_n & \\ & & & & & \sin t_n & \cos t_n & \\ & & & & & & & 1 \end{pmatrix}$$

5. $\mathrm{Sp}(n) = \mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{U}(2n)$ consists of unitary matrices of the form

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

A maximal torus in it is the image of the maximal torus $\mathrm{U}(n)$ describe above,

$$\mathrm{diag}(e^{i\lambda_1}, \cdots, e^{i\lambda_n}, e^{-i\lambda_1}, \cdots, e^{-i\lambda_n}).$$

In all the above examples, we have $G = \cup_{g\in G} gTg^{-1}$ by linear algebra. In general, this is a deep theorem, Cartan's maximal tori theorem.

Theorem 2.8.5. *Let G be a compact connected Lie group, and T be a maximal torus. Then every maximal torus is conjugate to T, and every element of G is contained in a conjugate of T.*

Hence, we can define the rank of a group G to be the dimension of a maximal torus.

Proof. We first prove the second statement. Let $g \in G$. We will show that there exists $k \in G$ such that $g \in kTk^{-1}$.

Let t_0 be a generator of T. Since the exponential map of a compact Lie group is surjective, we have $X \in \mathfrak{g}$ and $H_0 \in \mathfrak{t}$ such that $\exp(X) = g$ and $\exp(H_0) = t_0$.

Take the Ad-invariant inner product. Let $k \in G$ be an element such that $\langle X, Ad_k H_0 \rangle$ is maximal. Let $H = Ad_k H_0$. Thus $\exp H$ generates kTk^{-1}. By this choice $\langle X, Ad_{\exp(tY)} H \rangle$ has a maximum when $t = 0$. So

$$0 = \frac{d}{dt} \langle X, Ad_{\exp(tY)} H \rangle |_{t=0} = \langle X, ad_Y H \rangle = \langle X, [Y, H] \rangle = \langle [H, X], Y \rangle$$

for all Y. Hence $[H, X] = 0$. By Proposition 2.5.1, $\exp H$ commutes with $\exp(tX)$. Since $\exp H$ generates the maximal torus kTk^{-1}, it follows that the 1-PSG $\exp(tX)$ is contained in the centralizer of kTk^{-1}. The closure of the abelian group generated by $\exp(tX)$ and kTk^{-1} is a torus containing the maximal torus kTk^{-1}. Hence, $\exp(tX) \in kTk^{-1}$. In particular, $g = \exp(X) \in kTk^{-1}$.

For the first statement, let T' be another maximal torus, and let t' be a generator. Then t' is contained in kTk^{-1} for some k, so $T' \subset kTk^{-1}$. Hence $T' = kTk^{-1}$ by maximality of T'. □

In fact, we have derived the corresponding Lie algebra version of Theoerem 2.8.5 at the same time.

Theorem 2.8.6. *Let G be a compact connected Lie group, and $\mathfrak{t}$ the Lie algebra of a maximal torus T. Then every coadjoint orbit intersects $\mathfrak{t}$.*

Corollary 2.8.7. *If G is a compact connected Lie group, then the center $Z(G) = \{g \in G | gh = hg, \forall h \in G\}$ is the intersection of all maximal tori in G.*

Proof. Suppose $g \in Z(G)$, then for any maximal torus T, there is some $h \in G$ such that $hgh^{-1} \in T$. So $g \in T$ for any maximal torus T.

Conversely suppose g lies in all maximal tori. For any $h \in G$, there is a maximal torus T such that $h \in T$. Since T is abelian, $gh = hg$. So $g \in Z(G)$. □

In general, for any subgroup $H \subset G$, the centralizer

$$Z_G(H) = \{g \in G | gh = hg, \forall h \in H\}$$

is a Lie subgroup of G.

Corollary 2.8.8. *Suppose G is a compact connected Lie group and $A \subset G$ is a connected abelian Lie subgroup. Then $Z_G(A)$ is the union of all maximal tori in G that contains A. In particular, $Z_G(A)$ is connected. When T is a maximal torus, $Z_G(T) = T$.*

Proof. Note $Z_G(A) = Z_G(\bar{A})$. So we may assume A is a torus. Suppose T is a maximal torus containing A. Then by definition $T \subset Z_G(A)$. So $Z_G(A)$ contains the union of all maximal tori in G that contains A.

Conversely, let $g \in Z_G(A)$, or equivalently $A \subset Z_G(g)$. Then the identity component $Z_G(g)^\circ$ is a compact connected Lie group, and $A \subset Z_G(g)^\circ$ since $e \in A$ and A is connected. Let T_1 be a maximal torus in $Z_G(g)^\circ$ that contains A. So by definition, $g \in Z(Z_G(g)^\circ)$. By the previous corollary, $g \in T_1$. Hence the maximal torus T_1 contains both A and g. Hence $Z_G(A)$ is contained in the union of all maximal tori in G that contains A. This completes the proof. $\square$

Corollary 2.8.9. *Let G be a compact connected Lie group and $T \subset G$ a torus. Then the centralizer $Z_G(T)$ is a closed connected Lie subgroup of G.*

Proof. By Corollary 2.8.8, $Z_G(T)$ is connected. To show it is a closed Lie subgroup, let $u \in T$ be a generator. Then $Z_G(T) = Z_G(u)$. We only need to show $Z_G(u)$ is a closed submanifold near e. Identify a neighborhood of $e \in G$ with N in $\mathfrak{g}$ by the exponential map. Since $u \exp(tX)u^{-1} = \exp(tAd_uX)$, we know $exp(tX) \in Z_G(u)$ for all t if and only if X is in the Lie subalgebra $\{X \in \mathfrak{g} | Ad_uX = X\}$. This is a closed subspace in $\mathfrak{g}$. $\square$

2.9 Weyl group

Let G be a compact connected Lie group, and $T \subset G$ a maximal torus. The normalizer of T is

$$N(T) = \{g \in G | gTg^{-1} = T\}.$$

It is a closed subgroup since if $t \in T$ is a generator, $N(T)$ is the inverse image of t under the continuous map $g \mapsto gtg^{-1}$. By definition, T is a normal subgroup of $N(T)$.

Definition 2.9.1. *The quotient group $W = N(T)/T$ is called the Weyl group of G.*

Proposition 2.9.2. *The connected component $N(T)^\circ$ of the identity in $N(T)$ is T itself. The Weyl group is a finite group.*

Proof. We first prove that the automorphism group $\mathrm{Aut}(T)$ of a torus $T = \mathbb{R}^k/\mathbb{Z}^k$ is isomorphic to $\mathrm{GL}(k,\mathbb{Z})$. First, any element of $\mathrm{GL}(k,\mathbb{Z})$ could be realized as an element in $\mathrm{Aut}(T)$. On the other hand, if $\phi \in \mathrm{Aut}(T)$, it induces an invertible linear transformation ϕ_* of the Lie algebra $\mathfrak{t}$ that commutes with the exponential map, *i.e.* $\phi \circ \exp = \exp \circ \phi_*$. It must preserve the kernel $\Lambda = \ker \exp$. Identify $\mathfrak{t} = \mathbb{R}^k$ then Λ is identified with $\mathbb{Z}^k$. This implies the matrix of $(\phi)_*$ must lie in $\mathrm{GL}(k,\mathbb{Z})$.

Since $\mathrm{GL}(k,\mathbb{Z})$ is discrete, any connected subgroup of $\mathrm{Aut}(T)$ must act trivially. Hence any element in $N(T)^\circ$ commutes with all elements in T. By Corollary 2.8.8, $N(T)^\circ \subset Z_G(T) = T$.

The quotient $N(T)/T$ is both discrete and compact and hence finite. □

Proposition 2.9.3. *Let G be a compact connected Lie group, and T a maximal torus. Then two elements $t_1, t_2 \in T$ are conjugate in G if and only if they sit on the same orbit of the Weyl group action.*

Proof. If $gt_1g^{-1} = t_2$ for $g \in G$, we have $gTg^{-1} \subset gZ_G(t_1)g^{-1} = Z_G(t_2)$. Hence, both T and gTg^{-1} are maximal tori in $Z_G(t_2)^\circ$. So there exists $h \in Z_G(t_2)^\circ$ such that $hgTg^{-1}h^{-1} = T$. It follows $hg \in N(T)$ and $hgt_1g^{-1}h^{-1} = t_2$. □

It follows that each orbit of conjugation action of G intersects the maximal torus T precisely in the orbit of the Weyl group. There are finitely many such points since $|W| < \infty$.

As we see from the proof of Proposition 2.9.2, we know the Weyl group acts on the Lie algebra $\mathfrak{t}$ of the maximal torus T preserving the lattice Λ. Therefore, the statement of Proposition 2.9.3 also holds true for Lie algebra. Precisely, suppose two elements $\xi_1, \xi_2 \in \mathfrak{t}$ are conjugate in G, then there is an element of the Weyl group mapping ξ_1 to ξ_2. Hence, we have the important fact

Corollary 2.9.4. *The adjoint orbits in G are in one-to-one correspondence with the orbits of the Weyl group acting on $\mathfrak{t}$.*

Examples:

1. For $G = U(n)$, a maximal torus is $\mathrm{diag}(e^{i\lambda_1}, \cdots, e^{i\lambda_n})$. Its normalizer $N(T)$ consists of all monomial matrices (matrices with a single nonzero entry in each row and column) in $U(n)$. One can see this by simple calculation
$$\begin{pmatrix} 0 & e^{i\theta} \\ e^{i\mu} & 0 \end{pmatrix} \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \begin{pmatrix} 0 & e^{-i\mu} \\ e^{-i\theta} & 0 \end{pmatrix} = \begin{pmatrix} y & 0 \\ 0 & x \end{pmatrix}.$$

It follows that the Weyl group $W(U(n)) = S_n$, the full symmetric group.

2. $W(\mathrm{SU}(n)) = S_n$. $|W| = n!$.

3. $W(\mathrm{SO}(2l+1)) = G(l) = (\mathbb{Z}/2\mathbb{Z})^l \rtimes S_l$, the group of permutations ϕ of the set $\{-l, \cdots, -1, 1, \cdots, l\}$ with $\phi(-k) = -\phi(k)$ for all $1 \leq k \leq l$. $|W| = 2^l l!$. These are "permutation matrices" of the diagonal 2×2 blocks, generated by direct sums of two copies of $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and diagonal 1's, and direct sums of a copy of $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ at $2i-1$ and $2i$ rows and columns and -1 in the last diagonal and other diagonals 1.

4. $W(\mathrm{SO}(2l)) = SG(l)$, the subgroup of $G(l)$ that consists of even permutations. $|W| = 2^{l-1} l!$.

5. $W(\mathrm{Sp}(n)) = G(n)$. The equality of with the same Weyl group as for $\mathrm{SO}(2n+1)$ is not a coincidence.

2.10 Examples: SO(3), SL(2, ℝ), SU(2)

We first look at SO(3). It is a 3-dimensional Lie group whose Lie algebra $\mathfrak{so}(3)$ consists of all 3×3 real anti-symmetric matrices. Take the basis

$$e_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \; e_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

One checks

$$[e_1, e_2] = e_3, \; [e_2, e_3] = e_1, \; [e_3, e_1] = e_2.$$

Hence we can calculate the Killing form

$$B(xe_1 + ye_2 + ze_3, xe_1 + ye_2 + ze_3) = 2(x^2 + y^2 + z^2).$$

For any $A \in \mathrm{SO}(3)$, the third column (row) is the cross product of the other two. Hence, a direct calculation shows that

$$A \cdot (xe_1 + ye_2 + ze_3) \cdot A^{-1} = A \cdot (xe_1 + ye_2 + ze_3) \cdot A^t = (A \cdot (x, y, z)^t) \cdot (e_1, e_2, e_3).$$

In the last part, we misuse of notations. It means a linear combination of e_i where the coefficients are the three entries of $A \cdot (x, y, z)^t$. This relation could be checked for e_1, e_2, e_3 respectively, and then by linearity.

In other words, the adjoint action of SO(3) on $\mathfrak{so}(3)$ is just the usual SO(3) action on $\mathbb{R}^3$. Each orbit of adjoint action is the sphere $x^2 + y^2 + z^2 = c$ or the origin $(0,0,0)$. The Lie subalgebra $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -t \\ 0 & t & 0 \end{pmatrix} \in \mathfrak{t}$ is the Lie algebra of the maximal torus $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}$. Its intersection with adjoint orbits is $\pm\sqrt{c}$ when $c > 0$ and 0 when $c = 0$. We notice the Weyl group is $\mathbb{Z}_2$.

For $\mathrm{SL}(2,\mathbb{R})$, one can use the similar calculation to show that the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ is three dimensional and its Killing form has signature $(2,1)$. So the adjoint action gives a homomorphism $\mathrm{SL}(2,\mathbb{R}) \to \mathrm{SO}(2,1)$. But we can see these facts in a slightly different manner. We identify $\mathfrak{sl}(2,\mathbb{R})$ with $\mathbb{R}^3$ via

$$A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \mapsto (a,b,c).$$

Nonzero traceless 2×2 matrices have the same Jordan normal form if and only if they have the equal value of the determinant $\Delta = -(a^2 + bc)$. Each orbit is either one connected component of 2-sheeted hyperboloid ($a^2 + bc = \mathit{const} > 0$), a 1-sheeted hyperboloid ($a^2 + bc = \mathit{const} < 0$), each half of the cone $a^2 + bc = 0$, or the origin $(0,0,0)$.

$\mathrm{SL}(2,\mathbb{R})$ contains two non-conjugate maximal abelian subgroups $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$ and $\mathrm{SO}(2) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$. But they are conjugate to each other within $\mathrm{SL}(2,\mathbb{C})$. Their Lie algebras are $(t,0,0)$ and $(0,-t,t)$ respectively.

The Lie algebra $\mathfrak{su}(2)$ is three dimensional and its Killing form is positive definite so the adjoint action gives a homomorphism $\mathrm{SU}(2) \to \mathrm{SO}(3)$. We can work it out more explicitly. We identify $\mathfrak{su}(2)$ with the skew Hermitian matrices $\begin{pmatrix} ix & iy - z \\ iy + z & -ix \end{pmatrix}$ of trace zero. These matrices have the same diagonalization if and only if they have the equal value of the determinant $\Delta = x^2 + y^2 + z^2$. Thus, with respect to this invariant positive definite quadratic form, the adjoint representation of SU(2) is orthogonal. Hence, we have a homomorphism $\psi : \mathrm{SU}(2) \to \mathrm{SO}(3)$. The surjectivity follows from the fact that the maximal torus $\begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}$ maps to $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2t) & -\sin(2t) \\ 0 & \sin(2t) & \cos(2t) \end{pmatrix}$ and Cartan's maximal tori theorem. Hence ψ is a covering map. The kernel is $\{\pm I\}$. Alternatively, one can check the induced Lie algebra homomorphism, *i.e.* the infinitesimal adjoint represenation, is an isomorphism. Since we have shown $\mathrm{SU}(2) = S^3$, in particular simply connected. Hence, we have $\pi_1(\mathrm{SO}(3)) = \mathbb{Z}/2\mathbb{Z}$. It can be seen from the Dirac cup trick.

In fact, SO(3) is the real projective 3-space, since in the covering map and the identification of SU(2) with S^3, antipodal points are identified. We can see it by another way. A rotation can be specified by a vector along the axis of the rotation, with magnitude giving the angle of the rotation. This identifies elements of SO(3) with points inside or on the ball of radius π centered at origin. However, antipodal points on the surface of the ball represent the same rotation, thus identified. This gives real projective 3-space. There is an implicit use of the exponential map from $\mathfrak{so}(3)$ to SO(3).

In fact, in all above examples, we can write down explicit and equivalent basis, with (inequivalent) Ad-invariant form $\mathrm{tr}(xy)$.

Chapter 3

Roots and Root system

3.1 Semisimple compact groups

A Lie algebra $\mathfrak{g}$ is simple if it has no nonzero proper ideals and if $\dim \mathfrak{g} > 1$. A Lie algebra is semisimple if it has no abelian ideals. A compact Lie group is simple/semisimple if its Lie algebra is simple/semisimple. Equivalently, a simple Lie group is a simple group, a semi-simple Lie group is one with no non-trivial abelian connected normal subgroups. Also equivalently, a semi-simple Lie algebra is isomorphic with a product of simple Lie algebras. So $U(n)$ is not semisimple since the scalar matrices in $\mathfrak{u}(n)$ form an abelian ideal.

There is an important criterion of semisimplicity of Lie algebra due to Cartan.

Theorem 3.1.1 (Cartan's criterion)**.** *Lie algebra is semisimple if and only if the Killing form is nondegenerate.*

The proof of it is based on Jordan decomposition. If we admit the Ado theorem that any Lie algebra is isomorphic to a subalgebra in $\mathfrak{gl}(V)$, then the Jordan decompositino follows from the Jordan decomposition of linear algebra: $x = x_s + x_n$ where x_s is diagnalizable and x_n is nilpotent. We will not give the proof.

A direct corollary of it is an alternative definition of semisimplicity of Lie algebra.

Corollary 3.1.2. *A Lie algabra $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g}$ can be decomposed into the direct sum*

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

where $\mathfrak{g}_i$ are simple noncommutative ideals. Such a decomposition is unique.

For Lie groups, we have

Proposition 3.1.3. *Let G be a compact connected Lie group. Then G is semisimple if and only if the center of G is finite.*

Proof. The closure of a connected abelian normal subgroup is also a connected abelian normal subgroup. Hence it is a torus which is contained in a maximal torus. But this torus is invariant under conjugation by any element. So it is contained in every maximal torus and therefore in the center. Since the Lie algebra of an abelian normal subgroup is an abelian ideal, hence G is semisimple if and only if $Z(G)$ is finite. □

3.2 Schur's lemma

Schur's lemma is an important lemma to study irreducible representations. Recall (V, π) is called a representation of G if $\pi : G \to \mathrm{GL}(V)$ is a group homomorphism.

Definition 3.2.1. *Let (V, π) be a representation of Lie group G. The representation V is irreducible (or simple) if its has no subrepresentation other than 0 and V.*

Any one dimensional representation is irreducible. The standard representation of $\mathrm{SO}(n)$ on $\mathbb{R}^n$ is irreducible.

Definition 3.2.2. *A finite dimensional representation V of G is called completely reducible (or semi-simple) if it is isomorphic to a direct sum of irreducible representations.*

Definition 3.2.3. *A representation (V, π) of G is unitary if V admits a G-invariant positive-definite Hermitian inner product, i.e. $\pi(g)$ is unitary for any $g \in G$.*

Proposition 3.2.4. *Any finite dimensional unitary representation is completely reducible.*

Proof. Let V be any reducible representation of G, and $W \subset V$ a subrepresentation. Then the space $W^\perp$ orthogonal to W is also invariant under the G-action. It follows $W^\perp$ is also a subrepresentation and $V = W \oplus W^\perp$. We continue this procedure until each component is irreducible. □

Proposition 3.2.5. *Any finite dimensional representation of a compact Lie group is completely reducible.*

Proof. We can use the averaging trick to get invariant inner product on V. Namely, let

$$(v, w) := \int_G \langle g \cdot v, g \cdot w \rangle dg$$

where dg is the Haar measure on G.

Then the conclusion follows from Proposition 3.2.4. □

The next is the Schur's lemma.

Lemma 3.2.6 (Schur's Lemma)**.** *Let V, W be irreducible complex representation of G.*

1. *If $f : V \to W$ is a G-invariant linear map, then either $f \equiv 0$ or f is invertible.*

2. *If $f_1, f_2 : V \to W$ are two G-invariant linear maps and $f_2 \neq 0$, then there exists $\lambda \in \mathbb{C}$ such that $f_1 = \lambda f_2$.*

Proof. For the first part, suppose f is not identically zero. Since $\ker(f)$ is a G-invariant subset in V, it must be $\{0\}$. So f is injective. On the other hand, $Im(f)$ is a G-invariant subspace of W, then $f(V) = W$. Hence f is invertible.

For the second part, since $f_2 \neq 0$, it is invertible. So $f = f_2^{-1} \circ f_1$ is a G-invariant linear map from V to V. For any eigenvalue λ of f, $f - \lambda Id$ is G-invariant which is not invertible. Hence $f - \lambda Id \equiv 0$ and $f_1 = \lambda f_2$. □

It immediately follows from Lemma 3.2.6 (2) that

Corollary 3.2.7. *Let V be an irreducible representation of G, then $Hom_G(V, V) = \mathbb{C} \cdot Id$.*

Conversely, we have

Lemma 3.2.8. *If (π, V) is a unitary representation of G, and $Hom_G(V, V) = \mathbb{C} \cdot Id$, then (π, V) is an irreducible representation of G.*

We leave it as an exercise.

Corollary 3.2.9. *If (π, V) is an irreducible representation of G, then for any $h \in Z(G) = \{h \in G : gh = hg, \forall g \in G\}$, $\pi(h) = \lambda \cdot Id$ for some $\lambda \in \mathbb{C}$.*

Proof. It implies $\pi(h) : V \to V$ is G-invariant. So it follows from Corollary 3.2.7. □

Corollary 3.2.10. *Any irreducible representation of an abelian Lie group is one dimensional.*

Proof. Since $G = Z(G)$, by Corollary 3.2.9, for any $g \in G$, $\pi(g)$ is a multiple of the identity map on V. It follows that any subspace of V is G-invariant. Hence $\dim V = 1$ because V is irreducible. □

Eventually, it means commutative diagonolizable matrices could be simultaneously diagonalized.

3.3 Weights and Roots

3.3.1 Weights

Suppose T is a connected compact abelian Lie group. By Corollary 2.5.2, the map $\exp : \mathfrak{t} \to T$ is a homomorphism of Lie groups. Let $L \subset \mathfrak{t}$ be its kernel. We know $L \subset \mathfrak{t}$ is closed and discrete, and $\mathfrak{t}/L = T$.

By Corollary 3.2.10, all irreducible finite-dimensional complex representations of torus are one-dimensional. Therefore, all the irreducible complex representations of $T = \mathbb{R}^n/(2\pi\mathbb{Z})^n$ are given by the character group $X(T) = Hom(T, S^1)$. Explicitly,

$$(e^{i\theta_1}, \cdots, e^{i\theta_n}) \cdot z = e^{i(m_1\theta_1 + \cdots + m_n\theta_n)}z.$$

Let $\pi : G \to \mathrm{GL}(V)$ be a complex representation of G and T a maximal torus. By restriction we get a representation of T, still denote by π. Then by complete reducibility, π is a direct sum of one-dimensional representations. We know $T = \mathfrak{t}/L$, so we can lift $\pi : T \to S^1$ to a homomorphism $\mathfrak{t} \to S^1$. Using our fixed bi-invariant inner product on $\mathfrak{g}$ to identify $\mathfrak{t}$ with $\mathfrak{t}^*$, we identify $X(T)$ with a lattice $\Lambda \subset i\mathfrak{t}$

$$\Lambda = \{\lambda \in i\mathfrak{t} | \langle \lambda, L \rangle \subset 2\pi i\mathbb{Z}\}.$$

In other words, any element $\alpha \in \Lambda$ gives an irreducible representation ρ_α of T on $\mathbb{C}$ by $\rho_\alpha(\exp(X)) \cdot z = e^{\langle \alpha, X \rangle}z$ for $X \in \mathfrak{t}$.

With this understood, we have a decomposition

$$V = \oplus_{\lambda \in \Lambda} V_\lambda$$

where

$$\begin{aligned} V_\lambda :=& \{v \in V | \forall t \in T, \pi(t) \cdot v = \rho_\lambda(t) \cdot v\} \\ =& \{v \in V | \forall H \in \mathfrak{t}, \pi_*(H) \cdot v = \langle \lambda, H \rangle v\}. \end{aligned}$$

We say that λ is a weight of π if $V_\lambda \neq 0$, and $\dim V_\lambda$ is the multiplicity of the weight, and V_λ is the λ-weight space. Nonzero elements of the weight space are called weight vectors.

Example:.

1. With coordinates $i \cdot (\lambda_1, \cdots, \lambda_n)$ (*i.e.* $\text{diag}(e^{i\lambda_1}, \cdots, e^{i\lambda_n}) \in T$) on $\mathfrak{t}$ for the maximal torus $T \subset \mathrm{U}(n)$, the weights of the standard representation $\mathbb{C}^n$ are λ_i (where $\lambda_i = (0, \cdots, 1, \cdots 0)$ or when viewed as an element in the dual space $\lambda_i(i(t_1, \cdots, t_n)) = t_i$).

2. For the representation $\Lambda^2 \mathbb{C}^n$ of $\mathrm{U}(n)$ the weights are $\lambda_i + \lambda_j$, $i \neq j$. For $\Lambda^3 \mathbb{C}^n$, the weights are $\lambda_i + \lambda_j + \lambda_k$, i, j, k all different, and so on.

3. For the representation $Sym^2 \mathbb{C}^n$ of $\mathrm{U}(n)$, the weights are $\lambda_i + \lambda_j$ ($i = j$ allowed).

3.3.2 Complex representation of $\mathfrak{sl}(2, \mathbb{C})$

The simplest complex semisimple Lie algebra is $\mathfrak{sl}(2, \mathbb{C})$. As the discussion in Section 2.10, we know $\mathfrak{su}(2) \otimes \mathbb{C} = \mathfrak{sl}(2, \mathbb{R}) \otimes \mathbb{C} = \mathfrak{sl}(2, \mathbb{C})$. In fact, it plays an important role in the representation theory of general semisimple Lie algebras/groups.

Lemma 3.3.1. *The finite dimensional complex representations of* $SU(2)$, $\mathfrak{su}(2)$*,* $\mathfrak{sl}(2, \mathbb{C})$ *are one-to-one correspondent.*

Proof. Since SU(2) is simply connected, we can apply Proposition 2.6.17 to get equivalence of representations for SU(2) and $\mathfrak{su}(2)$.

The representation $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ extends to $\mathfrak{g} \otimes \mathbb{C}$ by $\rho(x + iy) = \rho(x) + i\rho(y)$. So defined ρ is complex linear and compatible with commutator. This gives equivalence of representation of $\mathfrak{su}(2)$ and $\mathfrak{sl}(2, \mathbb{C})$. □

There is a standard basis of the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$.

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

With the relations

$$[X, Y] = H, \; [H, X] = 2X, \; [H, Y] = -2Y. \tag{3.1}$$

In this case, the compact part $\mathfrak{su}(2)$ consists of the skew Hermitian matrices. Hence, $H \in i\mathfrak{t}$. We have $L = \{2\pi i n H | n \in \mathbb{Z}\}$ and $\Lambda \subset i\mathfrak{t} = \mathbb{R}$. So $\Lambda = \mathbb{Z}$.

Let (ρ, V) be an irreducible finite dimensional complex representation of $\mathfrak{s}l(2, \mathbb{C})$. The action H on V is diagonalizable, we thus have a decomposition

$$V = \oplus V_\lambda, \tag{3.2}$$

where λ runs over a collection of complex numbers (In fact, it follows from Lemma 3.3.1 that λ are integers. But we do not assume it in the following.), such that for any vector space $v \in V_\lambda$ we have $H(v) = \lambda \cdot v$. Here and later, we write $\rho(\xi)(v)$ simply by $\xi(v)$ for $\xi \in \mathfrak{sl}(2, \mathbb{C})$. Now we want to see how X, Y act on each V_λ. We apply a key computation which will be used several times later

$$\begin{aligned} H(X(v)) =& X(H(v)) + [H, X](v) \\ =& X(\lambda \cdot v) + 2X(v) \\ =& (\lambda + 2)X(v). \end{aligned}$$

That is, if v is an eigenvector for H with eigenvalue λ, then $X(v)$ is also a eigenvector with eigenvalue $\lambda + 2$. In other words, we have

$$X : V_\lambda \to V_{\lambda+2}.$$

Similar calculation shows $Y(V_\lambda) \subset V_{\lambda-2}$.

As an immediate consequence, for any irreducible representation V, the complex number that appear in the decomposition (3.2) must be congruent to one another mod 2. For any α that actually occurs, the subspace $\oplus_{j\in\mathbb{Z}} V_{\alpha+2j}$ would be invariant under $\mathfrak{sl}(2, \mathbb{C})$ and hence equal to all of V. Moreover, the V_α that appear must form an unbroken string of numbers of the form $\beta, \beta + 2, \cdots, \beta + 2k$. We denote by n the last element in this sequence.

Choose any nonzero vector $v \in V_n$. Since $V_{n+2} = \{0\}$, we must have $X(v) = 0$.

Proposition 3.3.2. *The vectors* $\{v, Y(v), Y^2(v), \cdots\}$ *span* V.

Proof. We call this space W. We show $\mathfrak{sl}(2, \mathbb{C})$ carries W into itself. Clearly, Y preserves it. Moreover, since $Y^m(v) \in V_{n-2m}$, we have $H(Y^m(v)) = (n - 2m)Y^m(v)$, so H preserves the subspace W. Thus, we only need to show $X(W) \subset W$.

First, $X(v) = 0$. Then we do the key calculation again

$$\begin{aligned} X(Y(v)) =& Y(X(v)) + [X, Y](v) \\ =& Y(0) + H(v) \\ =& nv. \end{aligned}$$

Do the same computation inductively, we have

$$\begin{aligned} X(Y^m(v)) =& (n + (n-2) + \cdots + (n - 2m + 2)) \cdot Y^{m-1}(v) \\ =& m(n - m + 1)Y^{m-1}(v). \end{aligned}$$

Hence, $W \subset V$ is a non-trivial subrepresentation. By the irreducibility of V, we have $W = V$. □

There are some immediate corollaries.

Corollary 3.3.3. *All the eigenspaces V_λ of H are one dimensional.*

Second, as we can see from the proof, V is determined by the weights, *i.e.* the collection of λ in the decomposition (3.2).

Since V is finite dimensional, we have a lower bound on the λ for which $V_\lambda \neq 0$. In other words, we must have $Y^k(v) = 0$ for sufficiently large k. If m is the smallest power of Y annihilating v, then we have

$$0 = X(Y^m(v)) = m(n-m+1)Y^{m-1}(v).$$

Hence, we have $n - m + 1 = 0$. In particular, it implies n is an integer. Moreover, the eigenvalues of H are symmetric about the origin in $\mathbb{Z}$. To summarize, we have a unique irreducible representation $V^{(n)}$ for each non-negative integer n. The representation $V^{(n)}$ is $n+1$-dimensional, with weights $n, n-2, \cdots, 2-n, -n$.

We notice two phenomena. First, the weights are symmetric about 0. This follows from the symmetry of Weyl group $\mathbb{Z}_2$. Second, irreducible representations $V^{(n)}$ are uniquely determined by the highest weight n. These two facts could be generalized for representations of a general complex semisimple Lie algebra.

The trivial 1-dimensional representation is just $V^{(0)}$. For the standard representation of $\mathfrak{sl}(2,\mathbb{C})$ on $V = \mathbb{C}^2$, if x, y are standard basis of $\mathbb{C}^2$, we have $H(x) = x, H(y) = -y$, so $V = \mathbb{C}\cdot x \oplus \mathbb{C}\cdot y = V_{-1} \oplus V_1$ is just the representation $V^{(1)}$. In general, it is straightforward to compute that the symmetric power $Sym^n(V)$ is just the representation $V^{(n)}$.

Look at the adjoint representation. As we can see from (3.1), it has weights ± 2 and 0. Hence, it is the irreducible representation $V^{(2)}$.

3.3.3 Roots

If we look at the complexification of the adjoint representation $\mathfrak{g}\otimes\mathbb{C}$ of G, the nonzero weights of V are called roots, the weight spaces are called root spaces, and weight vectors are called root vectors. We have seen that there exists inner product on the Lie algebra of a compact Lie group where ad is skew-self-adjoint. It could be extended to an inner product on $\mathfrak{g}\otimes\mathbb{C}$. When G is semisimple, we can choose it as Killing form. Explicitly, an element $\alpha \in i\mathfrak{t}$ is called a root if $\alpha \neq 0$ and there exists a nonzero $X \in \mathfrak{g}\otimes\mathbb{C}$ such that $[H, X] = \langle \alpha, H\rangle X$ for all $H \in \mathfrak{t}$. The root space is denoted by $\mathfrak{g}_\alpha$. Later, we will also sometimes write $\langle \alpha, H\rangle$ as $\alpha(H)$ for simplicity. This is a purely imaginary number.

If V is the complexification of a real representation, $V = A \otimes_{\mathbb{R}} \mathbb{C}$, then the nonzero weights come in pairs $\pm w$, with the same multiplicities (Exercise!). Hence, the roots, as elements in $\Lambda \subset i\mathfrak{t}$, come in pairs $\pm\alpha$. The weight zero space just correspond to $\mathfrak{t} \otimes \mathbb{C} =: \mathfrak{t}_{\mathbb{C}}$. If we denote the set of all roots by R, we have the direct sum decomposition

$$\mathfrak{g} \otimes \mathbb{C} = \mathfrak{t}_{\mathbb{C}} \oplus (\oplus_{\alpha \in R} \mathfrak{g}_\alpha).$$

Example:.

1. With standard coordinates $i(\lambda_1, \cdots, \lambda_n)$ on $\mathfrak{t}$ for the maximal torus $T \subset \mathrm{U}(n)$, the roots are $\lambda_i - \lambda_j$ for $i \neq j$. There are $n(n-1)$ roots and $\dim G = n^2$, rank $G = n$.

   This follows from the computation that the conjugation by an element $(\exp(i\lambda_1), \cdots, \exp(i\lambda_n)) \in T$ sends

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & e^{i(\lambda_1-\lambda_2)}a_{12} & \cdots & e^{i(\lambda_1-\lambda_n)}a_{1n} \\ e^{i(\lambda_2-\lambda_1)}a_{21} & a_{22} & \cdots & e^{i(\lambda_2-\lambda_n)}a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ e^{i(\lambda_n-\lambda_1)}a_{n1} & e^{i(\lambda_n-\lambda_2)}a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

2. For $\mathrm{SU}(n)$ everything is the same as for $\mathrm{U}(n)$, except the dimension and rank both drop by 1.

3. With the same standard coordinates, the roots of $\mathrm{SO}(2n+1)$ are $\pm\lambda_i \pm \lambda_j$ for $i \neq j$ and $\pm\lambda_i$. There are $2n(n-1) + 2n = 2n^2$ roots and $\dim G = n(2n+1)$, rank $G = n$.

4. For $\mathrm{SO}(2n)$, roots are $\pm\lambda_i \pm \lambda_j$, for $i \neq j$. There are $2n(n-1)$ roots and $\dim G = n(2n-1)$, rank $G = n$.

5. The roots of $\mathrm{Sp}(n)$ are $\lambda_i - \lambda_j$ for $i \neq j$ together with $\pm(\lambda_i + \lambda_j)$ ($i = j$ allowed). There are $n(n-1) + n(n+1) = 2n^2$ roots and $\dim G = n(2n+1)$, rank $G = n$.

Proposition 3.3.4. *For any $\alpha, \beta \in \mathfrak{t}_{\mathbb{C}}$, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. In particular, if $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, then $[X, Y] \in \mathfrak{t}_{\mathbb{C}}$. Furthermore, if X is in $\mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$, and $\alpha + \beta$ is neither zero nor a root, then $[X, Y] = 0$.*

Proof. We have Jacobi identity

$$[H, [X, Y]] = [[H, X], Y] + [X, [H, Y]].$$

Thus if $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$, we have

$$[H, [X, Y]] = [\alpha(H)X, Y] + [X, \beta(H)Y] = (\alpha + \beta)(H)[X, Y]$$

for all $H \in \mathfrak{t}_\mathbb{C}$. This shows $[X, Y] \in \mathfrak{g}_{\alpha+\beta}$. $\square$

We write $\mathfrak{g}_\mathbb{C} = \mathfrak{g} + i\mathfrak{g}$. We define the congujation with respect to $\mathfrak{g}$ by the real linear transformation $X = Y + iZ \mapsto \bar{X} = Y - iZ$, $Y, Z \in \mathfrak{g}$. For any complex number c, we have $\overline{cX} = \bar{c}\bar{X}$.

Proposition 3.3.5. *1. If $\alpha \in i\mathfrak{t}$ is a root, so is $-\alpha$. Specifically, if $X \in \mathfrak{g}_\alpha$, then $\bar{X} \in \mathfrak{g}_{-\alpha}$.*

2. If $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, then $[X, Y]$ is a multiple of α.

3. If $0 \neq X_\alpha \in \mathfrak{g}_\alpha$, then $[X_\alpha, \bar{X}_\alpha]$ is a nonzero element of $i\mathfrak{t}$, and $\alpha([X_\alpha, \bar{X}_\alpha]) \neq 0$.

4. When $\mathfrak{g}$ is semisimple, the roots span $\mathfrak{t}_\mathbb{C}$ complex linearly (and $i\mathfrak{t}$ as real vector space).

Proof. If $X \in \mathfrak{g}_\alpha$, then for all $H \in \mathfrak{t} \subset \mathfrak{g}$, we have

$$[H, \bar{X}] = \overline{[H, X]} = \overline{\alpha(H)X} = -\alpha(H)\bar{X}.$$

The last inequality is because $\alpha(H)$ is purely imaginary. Hence $\bar{X} \in \mathfrak{g}_{-\alpha}$.

For part 2, if $Z \in \mathfrak{t}_\mathbb{C}$ is orthogonal to α, then

$$\langle Z, [X, Y]\rangle = -\langle [X, Z], Y\rangle = [0, Y] = 0.$$

Then the conclusion follows from the nondegeneracy of the inner product.

For part 3, by Proposition 3.3.4, $[X_\alpha, \overline{X_\alpha}] \in \mathfrak{t}_\mathbb{C}$. It is straightforward to check $i[X_\alpha, \overline{X_\alpha}]$ is invariant under taking conjugate. Hence $[X_\alpha, \overline{X_\alpha}] \in i\mathfrak{t}$. If $[X_\alpha, \overline{X_\alpha}] = 0$, then $\ker \alpha \subset \mathfrak{t}, \Re X_\alpha, \Im X_\alpha$ span $r + 1$ dimensional abelian subspace of $i\mathfrak{g}$ over $\mathbb{R}$, contradicting to the assumption that r is the rank of $\mathfrak{g}$. We then apply part 2 to get $\alpha([X_\alpha, \bar{X}_\alpha]) \neq 0$.

For part 4, suppose that the roots do not span $i\mathfrak{t}$. Then there would be a nonzero $H \in \mathfrak{t}$ such that $\alpha(H) = 0$ for all $\alpha \in R$. Then we would have H in the center of $\mathfrak{g}$, contradicting the definition of a semisimple Lie algebra. $\square$

Proposition 3.3.6. *1. If $\alpha \in R$, then $\dim(\mathfrak{g}_\alpha) = 1$.*

2. If $\alpha, \beta \in R$ and $\alpha = c\beta, c \in \mathbb{R}$, then $c = \pm 1$.

Proof. By last proposition, we can find $X \in \mathfrak{g}_\alpha, Y \in \mathfrak{g}_{-\alpha}$, so that $H = [X,Y] \neq 0$ and $\alpha(H) \neq 0$. Adjust by scalars, they generate a subalgebra $\mathfrak{s}$ isomorphic to $\mathfrak{sl}(2,\mathbb{C})$, in particular $\alpha(iH) = 2$. Consider the adjoint action of $\mathfrak{s}$ on $V = \mathfrak{t}_\mathbb{C} \oplus (\oplus \mathfrak{g}_{k\alpha})$ sum over all nonzero complex multiple $k\alpha$ of α. From our discussion on the representation of $\mathfrak{sl}(2,\mathbb{C})$, only k's that can occur are integral multiple of $\frac{1}{2}$.

Now $\mathfrak{s}$ acts trivially on $\ker(\alpha) \subset \mathfrak{t}_\mathbb{C} \subset V$, and it acts irreducibly on $\mathfrak{s} \subset V$. They cover the zero weight space $\mathfrak{t}_\mathbb{C}$. Since each representation with an even weight must have a weight 0 as in Section 3.3.2, we know the only even weights occurring are 0 and ± 2 as appeared in the trivial and adjoint representations. In particular, 2α is not a root. So if we start with $\mathfrak{g}_{\frac{1}{2}\alpha}$ and $\mathfrak{g}_{-\frac{1}{2}\alpha}$ by the same argument, it would imply α is not a root. This contradiction shows $\frac{1}{2}\alpha$ is not a root, which is equivalent to saying 1 is not a weight occurring in V. But then there can be no other representations since each representation with an odd weight must have a weight 1. In other words, $V = \ker(\alpha) \oplus \mathfrak{s}$ and both statements of the proposition follow. □

The proof of the above proposition also implies the following.

Proposition 3.3.7. *For any $\alpha \in R$, a nonzero $X_\alpha \in \mathfrak{g}_\alpha$ and $\bar{X}_\alpha \in \mathfrak{g}_{-\alpha}$ generate a complex Lie subalgebra $\mathfrak{s}_\alpha \subset \mathfrak{g}_\mathbb{C}$ isomorphic to $\mathfrak{sl}(2,\mathbb{C})$. Moreover, $\mathfrak{s}_\alpha \cap \mathfrak{g}$ is isomorphic to $\mathfrak{su}(2)$.*

Proof. The first part is the construction in Proposition 3.3.6. For the second statement, we notice $iH_\alpha = i[X_\alpha, \bar{X}_\alpha]$, $X_\alpha + \bar{X}_\alpha$, and $i(X_\alpha - \bar{X}_\alpha)$ are in $\mathfrak{g}$, and generate $\mathfrak{su}(2)$. □

3.4 Root system

We define the root system.

Definition 3.4.1. *Let V be a finite-dimensional Euclidean vector space, with the standard Euclidean inner product. A root system in V is a finite set R of non-zero vectors (called roots) that satisfy the following conditions:*

1. *The roots span V.*
2. *The only scalar multiple of a root $\alpha \in R$ that belong to R are α and $-\alpha$.*
3. *For every root $\alpha \in R$, the set R is closed under reflection through the hyperplane perpendicular to α. In other words, for any root $\alpha \in R$, $s_\alpha(R) = R$ for the reflection $s_\alpha(\beta) := \beta - \frac{2(\alpha,\beta)}{(\alpha,\alpha)}\alpha$.*

4. For any two roots $\alpha, \beta \in R$, the number $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ is an integer.

The first condition is called non-degeneracy, satisfied by semisimple Lie algebras. The second condition is called reduced. Some authors omit condition 2 to define a root system, and they call ones with it reduced.

If R is a root system in V, the coroot $\alpha^\vee$ of a root α is defined by

$$\alpha^\vee := \frac{2}{(\alpha,\alpha)}\alpha.$$

The set of coroot also forms a root system in V (Exercise!), called the dual root system. By definition, $(\alpha^\vee)^\vee = \alpha$, and for reflections s, $(s\alpha)^\vee = s(\alpha^\vee)$.

Example:

For SU(2), there is only one non-trivial element of the Weyl group, represented by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. It transforms an element in $\begin{pmatrix} it & 0 \\ 0 & -it \end{pmatrix} \in \mathfrak{t}$ to $\begin{pmatrix} -it & 0 \\ 0 & it \end{pmatrix}$. In other words, it induces the reflection s_λ for the root λ, *i.e.* $s_\lambda : \lambda \mapsto -\lambda$.

Theorem 3.4.2. *If R is the set of roots associated with a compact Lie group G and its maximal torus T, then R is a root system.*

Proof. The vector space $V = i\mathfrak{t}$. By Proposition 3.3.5 (4), R spans V. By Proposition 3.3.6, the only scalar multiple of a root $\alpha \in R$ that belong to R are α and $-\alpha$.

We now show that each reflection s_α is induced by an element in the Weyl group of (G,T). Since SU(2) is simply connected, it follows from Corollary 2.6.15 that the Lie algebra inclusion $\mathfrak{su}(2) \to \mathfrak{g}$ of Proposition 3.3.7 is the differential of a homomorphism $i_\alpha : \mathrm{SU}(2) \to G$. We claim the element

$$w_\alpha = i_\alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

induces s_α. First, the adjoint map of w_α maps $\alpha \mapsto -\alpha$ as we can see in the above example. Let $T_\alpha = \{t \in T | \alpha(t) = 1\}$. We have $Lie(T_\alpha) = \mathfrak{t}_\alpha = \ker(\alpha) \subset \mathfrak{t}$. We have w_α centralizes T_α and thus fixes $i\mathfrak{t}_\alpha$. Since $\mathfrak{t}$ is generated by $\mathfrak{t}_\alpha$ and $\alpha = i_\alpha \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}$, we know we know R is preserved by w_α and the above two properties of w_α characterize s_α.

Item 4 of Definition 3.4.1 follows from a more general fact, Proposition 3.4.3. Thus R is a root system. □

Proposition 3.4.3. *Let (π, V) be a finite dimensional representation of G, and let $\lambda \in X(T)$ be a weight, i.e. $V_\lambda \neq 0$. Then $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$ for all $\alpha \in R$.*

Proof. The space $W = \oplus_{k\in\mathbb{Z}} V_{\lambda+k\alpha}$ is stable under X_α and $\bar{X}_\alpha$ by a similar argument as Proposition 3.3.4. Therefore it is invariant under $\mathfrak{s}_\alpha$. This implies it is invariant under $i_\alpha(\mathrm{SU}(2))$, in particular by w_α. It follows that the set $\{\lambda + k\alpha | k \in \mathbb{Z}\}$ is invariant under s_α. This is equivalent to saying $\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$. □

This proposition says a geometric integral element (with respect to G) is algebraic integral (with respect to $\mathfrak{g}$).

Moreover, as in section 3.3.2, the set of weights of a finite dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ is symmetric about the origin, and is an uninterrupted string of integers of the same parity. Therefore the set $\{\lambda+k\alpha|k \in \mathbb{Z}\}$ has $\{p \leq k \leq q\}$ for some $p, q \in \mathbb{Z}$, and it must be the case that

$$2p + \frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} = -(2q + \frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}),$$

i.e.

$$\frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} = -(p+q).$$

As we see from the proof, s_α fixes $t_\alpha = \ker\alpha$ and acts as -1 on other piece of $\mathfrak{t}$. Any $H \in i\mathfrak{t}$ with $\langle\alpha, H\rangle \neq 0$ for all $\alpha \in R$ is called *regular*. Hence, the regular set is the complement of a finite number of hyperplanes.

Definition 3.4.4. *A Weyl chamber in $i\mathfrak{t}$ is a connected component of the regular set.*

Choose a regular $H_0 \in i\mathfrak{t}$ and define the positive roots

$$R^+ = \{\alpha \in R | \langle\alpha, H_0\rangle > 0\}.$$

If we let $R^- = -R^+$, we have $R = R^+ \cup R^-$. It is also convex, *i.e.* if $\alpha, \beta \in R^+$ and if $\alpha+\beta \in R$ then $\alpha+\beta \in R^+$. Moreover, the Weyl chamber C^+ in which H_0 lies is characterized by

$$C^+ = \{H \in i\mathfrak{t} | \langle\alpha, H\rangle > 0, \forall\alpha \in R^+\}.$$

Hence the Weyl chambers are convex. C^+ is called the dominant Weyl chamber or the fundamental Weyl chamber.

Definition 3.4.5. *$\alpha \in R^+$ is a simple root if α cannot be expressed as $\alpha = \beta + \gamma$ with $\beta, \gamma \in R^+$.*

Let $\Delta \subset R^+$ be the set of simple roots.

We show some basic properties of simple roots.

Proposition 3.4.6. *1. For simple roots $\alpha \neq \beta$, then $\langle \alpha, \beta \rangle \leq 0$. If the equality holds, then both $\alpha \pm \beta$ are not roots.*

2. Δ is a vector space basis of the $\mathbb{R}$-linear span of R.

3. If any $\alpha \in R$ is expressed as a linear combination of simple roots, then the coefficients are integers, either all nonnegative (in which case $\alpha \in R^+$) or all non positive (in which case $\alpha \in R^-$).

Proof. Since the set of k such that $\beta + k\alpha$ is a root is a consecutive string of integers $p \leq k \leq q$, and $\frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} = -(p+q)$. So if $\langle \alpha, \beta \rangle > 0$, we have $\beta - \alpha \in R$. Interchange the roles of α, β if necessary so that $\alpha - \beta = \gamma \in R^+$. Then $\alpha = \beta + \gamma$, which is a contradiction. This is the first part of item 1.

In particular, it shows $\beta - \alpha \notin R$, *i.e.* $p \geq 0$. If $\langle \alpha, \beta \rangle = 0$, it implies $p + q = 0$. It means $p = q = 0$ and thus $\alpha + \beta \notin R$.

Now suppose that $\alpha \in R^+$. If $\alpha \notin \Delta$, we have $\alpha = \alpha' + \alpha''$. Since the terms of the decomposition $\langle \alpha, H_0 \rangle = \langle \alpha', H_0 \rangle + \langle \alpha'', H_0 \rangle$ are all positive and there are finitely many positive roots, item 2 follows.

Item 3 follows from linear independence of Δ, which will be shown in the following. If this is not true, then there is a relation for $\Delta = \{\alpha_1, \cdots, \alpha_r\}$,

$$\sum_{j \leq k} c_j \alpha_j = \sum_{j > k} c_j \alpha_j,$$

for $c_1, \cdots, c_k > 0$ and $k \geq 1$ while $c_{k+1}, \cdots, c_r \geq 0$. Take inner product gives

$$0 \leq || \sum_{j \leq k} c_j \alpha_j ||^2 = \sum_{j \leq k < i} c_j c_i \langle \alpha_j, \alpha_i \rangle \leq 0.$$

This implies $\sum_{j \leq k} c_j \alpha_j = 0$. This gives a contradiction with pairing with H_0 and notice $\alpha_j \in R^+$ and $c_j > 0$ for $j \leq k$. □

The fundamental Weyl chamber is in fact a fundamental region of Weyl group on $i\mathfrak{t}$.

Proposition 3.4.7. *The action of Weyl group on the set of chambers is simply transitive.*

Proof. To show the action is transitive, we connect two Weyl chambers by a path which crosses one root plane at a time, transversely. Then the corresponding product of reflections transforms one Weyl chamber to the other.

To show it is simply transitive, suppose w is an element of Weyl group and C is a chamber with $wC = C$. Let w have order n. and choose any point $\eta \in C$. Let $\xi = \sum_{r=0}^{n-1} w^r \eta$. By convexity of C, $\xi \in C$ which is fixed by w. Let $g \in G$ be an element whose adjoint action realizes w. So $g \in Z_G(\exp \xi)$. Since ξ is regular, we have $Z_G(\exp \xi) = T$ and so w is the identity. □

For $\xi \in \mathfrak{t}$ let W_ξ be the subgroup of the Weyl group which fixes ξ. A similar argument shows W_ξ acts transitively on the set of Weyl chambers which contain ξ.

Corollary 3.4.8. *No two distinct points in the closure of the same Weyl chamber are in the same orbit of W. Therefore the closure of C^+ is a fundamental region of the action W.*

We would like to remark without proof that there exists the following one-to-one correspondence (Borel-Weil):

- irreducible representations
- integral coadjoint orbits
- orbits of weights under Weyl group
- weights in the closure of fundamental Weyl chamber

We have seen that each adjoint orbit intersects $\mathfrak{t}$ at finitely many points which are invariant under the action of Weyl group as a set. When the intersection is on the weight lattice, this weight ξ on $\overline{C^+}$ defines a homomorphism from T to S^1 and hence a complex line bundle L_ξ over G/T when $\xi \in C^+$. Then the corresponding representation V_ξ is the space of holomorphic sections of L_ξ over G/T. When ξ is in the boundary of C^+, the corresponding line bundle over $M_\xi = G/H$ for $H \supset T$ could also be lifted to G/T. ξ is the highest weight of the representation V_ξ, which always has multiplicity one. The most difficult part of the correspondence is the construction of a representation with a given highest weight.

3.4.1 Classification

Property 4 of Definition 3.4.1 puts restrictions on the geometry of the roots. If θ is the angle between α and β, we have

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

is an integer. In particular, $4\cos^2\theta = n_{\alpha\beta}n_{\beta\alpha}$ is an integer between 0 and 4. Hence, the angles between roots could be $0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{2\pi}{3}, \frac{3\pi}{4}, \frac{5\pi}{6}$ and π. An important consequence for the classification of root system is the following.

Proposition 3.4.9. *Suppose α, β are distinct simple roots and $\langle \alpha, \beta \rangle \neq 0$. Then*

$$\angle(\alpha, \beta) = \begin{cases} 120^\circ & k = 1 \\ 135^\circ & k = 2 \\ 150^\circ & k = 3 \end{cases}$$

with $||\alpha||^2 = k||\beta||^2$ or $||\beta||^2 = k||\alpha||^2$.

In summary, we can reduce the classification of compact semisimple compact Lie groups to compact simple Lie groups. The latter could be reduced to the classification of (irreducible) root systems, which could be reduced to the classification of Dynkin diagrams of simple roots.

A Dynkin diagram of G is the graph constructed as follows

- The vertices are the simple roots.
- $\alpha, \beta \in \Delta$ are connected by k edges if $\langle \alpha, \beta \rangle \neq 0$ and $||\alpha||^2 = k||\beta||^2$ or the other direction. and are not connected if $\langle \alpha, \beta \rangle = 0$. If $k \neq 1$, the edge is directed from the longer root to the shorter.

Theorem 3.4.10. *The only possible Dynkin diagrams of connected, compact, simple Lie groups are $A_n, B_n, C_n, D_n, E_6, E_7, E_8, F_4, G_2$. Each of these does correspond to a connected, compact, simple Lie group, which is unique up to covering.*

We will not prove this. Instead, we explain why this should be true. The first part follows from some elementary analysis using inner products, *etc.* For instance, one quickly realizes that cycles are impossible, and then only one pair of nodes can have multiple edges.

In particular, the Lie groups corresponding to A_n, B_n, C_n, and D_n are $\mathrm{SU}(n+1)$, $\mathrm{SO}(2n+1)$, $\mathrm{Sp}(n)$, and $\mathrm{SO}(2n)$ respectively. The uniqueness follows because from the Dynkin diagram we can reconstruct the roots and root system. First for the $\mathfrak{sl}(2, \mathbb{C})$ corresponding to simple roots, then other roots are obtained by taking brackets.

We try to illustrate some of these exceptional collections of root systems. First for E_8. One choice of simple roots is give by the rows of the following matrix:

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix}$$

So its corresponding Lie group has rank 8 and dimension $248 = 8+112+128$, where 112 roots with integer entries obtained from $(\pm 1, \pm 1, 0, 0, 0, 0, 0, 0)$ by taking an arbitrary combination of signs and an arbitrary permutation of coordinates (roots of $\mathfrak{so}(16)$), and 128 roots with half-integer entries obtained from

$$(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$$

by taking an even number of minus signs (or, equivalently, requiring that the sum of all the eight coordinates be even).

The integral span of the E_8 root system forms a lattice in $\mathbb{R}^8$ naturally called the E_8 root lattice. This lattice is rather remarkable in that it is the only (nontrivial) even, unimodular lattice with rank less than 16. In particular, the distance between neighboring points in E_8 is $\sqrt{2}$, so we can form a packing with spheres of radius $\frac{\sqrt{2}}{2}$ and density

$$\frac{Vol(B^8_{\sqrt{2}/2})}{Vol(\mathbb{R}^8/E_8)} = \frac{\pi^4}{384}.$$

This is the optimal packing in dimension 8, in the sense the density is the largest possible one.

The root system E_7 is the set of vectors in E_8 that are perpendicular to a fixed root in E_8. The root system E_7 has 126 roots. E_6 is the subsystem of E_8 perpendicular to two suitably chosen roots of E_8. The root system E_6 has 72 roots.

The root system G_2 has rank 2 and 12 roots, which form the vertices of a hexagram. As a simply connected compact Lie group, G_2 is the automorphism $\mathrm{Aut}(\mathbb{O})$ of octonion ($\mathrm{Aut}(\mathbb{R}) \cong 1$, $\mathrm{Aut}(\mathbb{C}) \cong \mathbb{Z}_2$, $\mathrm{Aut}(\mathbb{H}) \cong \mathrm{SO}(3)$).

The root system F_4 has rank 4 and dimension 52. One choice of simple roots give by the rows of

$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

The 48 root vectors of F_4 form vertices of the 24-cell (8 vertices obtained by permuting $(\pm 1, 0, 0, 0)$ and 16 vertices of the form $(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$) and its dual (24 vertices are $(\pm 1, \pm 1, 0, 0)$). The 24-cell is one of the six 4-dimensional convex regular polytopes.

The root system also gives us some special isomorphisms, up to coverings, within the families of classical Lie groups.

- SO(3), SU(2) and Sp(1).
- SO(4), SU(2) $\times$ SU(2) and SO(3) $\times$ SO(3).
- SO(5) and Sp(2).
- SO(6) and SU(4).