NOTES ON LIE GROUPS

Abstract. Notes from a Lie groups course given at ETH Zurich, Fall 2016, unedited

Contents

1. Disclaimers

The notes recorded here are intended to *supplement* the lectures, but I have included more material here than is necessary for the course. You are *not* expected to read these notes, but I hope that providing them as a reference may be helpful. For example, although I have recorded here a fair bit of background from differential geometry, it should not be necessary for the purposes of the course.

A principal aim of the lectures is that after attending them, you should be able to do the homework problems given below their summaries.

It is very likely that the notes and exercises will contain some mistakes; any corrections would be much appreciated.

2. Summary of classes and homework assignments

2.1. **9/20: The definition of a Lie group. Objectives.** You should be able to explain the definition of "Lie group" and to prove that basic examples (e.g., orthogonal groups) are in fact Lie groups.

Summary.

(1) review of one-variable calculus and how it relates global properties of functions (e.g., monotonicity) to infinitesimal ones (e.g., positivity of derivative)
(2) review of multivariable calculus:
   (a) partial and total derivatives of a function
   (b) inverse function theorem
   (c) implicit function theorem
(3) very brief review of differential geometry:
   (a) the definition of submanifolds of open subsets of Euclidean spaces
   (b) how (in practice) to check that a subset is a submanifold
   (c) how (in practice) to compute tangent spaces of submanifolds
   (d) immersions, embeddings, submersions
(4) review of group theory:
   (a) functorial definition of "group" in terms of multiplication and inversion maps
   (b) permutation groups; Cayley's theorem
   (c) definition of topological group
(5) basic Lie-theoretic definitions:
   (a) Lie group (without recalling what a "manifold" is, other than to note that open subsets of Euclidean space and submanifolds thereof are examples of manifolds)
   (b) Lie subgroup
   (c) immersed Lie subgroup (e.g., irrational winding of the 2-torus)
   (d) the Lie algebra of a Lie group (without justifying the "algebra" in "Lie algebra")
   (e) linear Lie group
(6) how to compute Lie algebras of Lie groups in practice; examples of $\mathrm{GL}_n, \mathrm{SL}_n, O_n$

Homework 1 (Due Oct 4)**.** Write down the definitions of "Lie group" and "Lie subgroup". Using some lemmas from class, prove that $\mathrm{O}(n) := \{g \in \mathrm{GL}_n(\mathbb{R}) : gg^t = 1\}$ is a Lie subgroup of $\mathrm{GL}_n(\mathbb{R})$ of dimension $n(n-1)/2$ with Lie algebra $\mathfrak{o}(n) := \mathrm{Lie}(\mathrm{O}(n))$ given by the space $\{X \in M_n(\mathbb{R}) : X + X^t = 0\}$ of skew-symmetric matrices.

2.2. **9/22: The connected component. Objectives.** You should be able to define the "connected component" of a Lie group, explain its importance, and determine it in some basic examples (such as linear, orthogonal or unitary groups).

Summary.

(1) review of the general topological notion of connected components of a topological space, and how it specializes when the space is a manifold
(2) basics on the connected component of the identity G^0 in a Lie group G:
 (a) it is a normal Lie subgroup whose cosets are the connected components of G
 (b) G^0 (and more generally, any connected topological group) is generated by any neighborhood of the identity
(3) most of the classical groups were introduced and their number of connected components described, with some cases proved $(\mathrm{SL}_n, \mathrm{GL}_n, \mathrm{O}(n), \mathrm{SO}(n), \mathrm{U}(n))$ and others left as exercises $(\mathrm{O}(p,q), \dots)$; a large part of the class consisted of filling in and explaining the entries in the three-column table depicted in §12.2
(4) review on the matrix exponential, as in §13.1

Homework 2 (Due Oct 4)**.** Let $p, q \geq 1$ and $n := p + q$. Recall that

$$\begin{aligned}\mathrm{O}(p,q) &:= \{g \in \mathrm{GL}_n(\mathbb{R}) : Q(gv) = Q(v) \text{ for all } v \in \mathbb{R}^n\} \\ &= \{g \in \mathrm{GL}_n(\mathbb{R}) : g^t J g = J\},\end{aligned}$$

where $Q(v) := v_1^2 + \dots + v_p^2 - v_{p+1}^2 - \dots - v_n^2$ and $J := \mathrm{diag}(1, \dots, 1, -1, \dots, -1)$, and that $\mathrm{SO}(p,q) := \mathrm{SL}_n(\mathbb{R}) \cap \mathrm{O}(p,q)$. For $a \in \mathbb{R}$, set $V_a := \{v \in \mathbb{R}^n : Q(v) = a\}$. Denote by $e_1, \dots, e_n$ the standard basis vectors for $\mathbb{R}^n$.

(1) Suppose $p = q = 1$, so that $n = 2$.
 (a) Show that every matrix of the form
$$g = \begin{pmatrix} \varepsilon_1 & \\ & \varepsilon_2 \end{pmatrix} \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix} \tag{1}$$
 with $\varepsilon_1, \varepsilon_2 \in \{\pm 1\}$, $t \in \mathbb{R}$ belongs to $\mathrm{O}(1,1)$. Show that
$$\begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix} = \exp(t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}).$$
 Show that if $\varepsilon_1 = \varepsilon_2 = 1$, then $g \in \mathrm{O}(1,1)^0$.
 (b) Show that for each $v \in V_1$ with $v_1 > 0$ there is an element $g \in \mathrm{O}(1,1)^0$ so that $ge_1 = v$.
 (c) Show that every $g \in \mathrm{O}(1,1)$ with $ge_1 = e_1$ is of the form (1) with $\varepsilon_1 = 1, \varepsilon_2 = \pm 1$ and $t = 0$.
 (d) Show for $a = \pm 1$ that the space V_a has two connected components and that $\mathrm{O}(1,1)$ acts transitively on V_a. Determine the orbit of e_1 under $\mathrm{SO}(1,1)$.
 (e) Using some of the previous steps (or direct calculation), show that every element of $\mathrm{O}(1,1)$ is of the form (1) and that $\mathrm{O}(1,1)$ has four connected components.
(2) Suppose now that $p = 1, q = 2, n = 3$.
 (a) Observe (by drawing a picture, say) that V_{-1} is connected, that V_1 has two connected components, and that $e_1 \in V_1$. Denote by V_1^0 the connected component of V_1 containing e_1. Show that for each $v \in V_1^0$

there exists an $h \in \mathrm{SO}(1,2)^0$ so that $hv = e_1$. [Hint: one can reduce to part (b) of the previous exercise.]

(b) Show that the stabilizer of e_1 in $\mathrm{SO}(1,2)$ is isomorphic to $\mathrm{SO}(2)$, hence is connected.

(c) Show that any $g \in \mathrm{SO}(1,2)$ for which $ge_1 \in V_1^0$ belongs to the connected component $\mathrm{SO}(1,2)^0$.

(d) Deduce that $\mathrm{SO}(1,2)$ has two connected components.

2.3. **9/27: One-parameter subgroups and the exponential map. Objectives.** You should be able to define one-parameter subgroups and apply their basic uniqueness theorem. You should be able to define and characterize the exponential map on a Lie group in a few different ways, and be able to apply these characterizations. You should be able to apply the exponential map to relate global symmetries to infinitesimal ones (as in the example from lecture or the homework problem below). You should be able to apply the fact that the image of the exponential map contains a neighborhood of the identity, which in turn generates the connected component.

Summary.

(1) Review of the matrix exponential and its various characterizations:
  (a) as a series $\exp(X) = \sum X^n/n!$
  (b) as a limit $\exp(X) = \lim(1 + X/n)^n$, or more generally, $\exp(X) = \lim \gamma(1/n)^n$ for any curve with basepoint $\gamma(0) = 1$ and initial velocity $\gamma'(0) = X$
  (c) by requiring that for each X, the function $\Phi_X(t) = \exp(tX)$ is the unique solution to the ODE $\Phi'(t) = X\Phi(t)$ with initial condition $\Phi(0) = 1$
  (d) by requiring that for each X, the function $\Phi_X(t)$ as above is the unique smooth group homomorphism with initial velocity $\Phi'_X(0) = X$.

(2) We explained how the above generalizes to any Lie group. The key was the existence/uniqueness of one-parameter subgroups.
  (a) The uniqueness was reduced to uniqueness theorems for ODE's.
  (b) We gave a direct proof of the existence of one-parameter subgroups for GL_n, deduced it for linear Lie groups via the second characterization above, and indicated how it follows for general G by solving some ODE's and extending their solutions.

(3) As a basic application we explained how to characterize the rotation-invariant functions on $\mathbb{R}^n$ as the solutions to a finite system of homogeneous linear differential equations.

Homework 3 (Due Oct 4)**.**

(1) Use Lie's first theorem (Theorem 81) and the results of Homework 2 to show that the following are equivalent for a smooth function $f : \mathbb{R}^3 \to \mathbb{R}$:
  (a) f is constant on each connected component of $\{(x, y, z) \in \mathbb{R}^3 : z^2 - x^2 - y^2 = 1\}$.
  (b) f satisfies the differential equations

$$x\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial x} = 0,$$

$$z\frac{\partial f}{\partial x} + x\frac{\partial f}{\partial z} = 0,$$

$$z\frac{\partial f}{\partial y} + y\frac{\partial f}{\partial z} = 0$$

  on $\{(x, y, z) \in \mathbb{R}^3 : z^2 - x^2 - y^2 = 1\}$.

(2) Let G be a topological group and $H \leq G$ a subgroup with the property that there is a neighborhood U in G of the identity element so that $U \cap H = \{1\}$. Show that H is a discrete subgroup of G.

(3) Let G be a connected commutative Lie group with Lie algebra $\mathfrak{g}$. Show that $\exp : \mathfrak{g} \to G$ is a surjective homomorphism and with discrete kernel.

2.4. **9/29: The Lie algebra of a Lie group. Objectives.** You should be able to explain how the Lie bracket arises as an infinitesimal commutator of group elements. You should be able to explain the meaning of the sentence "the differential of a morphism of Lie groups is a morphism of Lie algebras"; in particular, you should be able to define all of its terms. Given a fairly explicit morphism of Lie groups (such as the representations on polynomials discussed in lecture or in the homework below), you should be able to compute the induced infinitesimal action of the Lie algebra.

Summary.

(1) tying up loose ends on application of exponential map:
   (a) connected Lie subgroups are determined by their Lie algebras
   (b) the exponential map intertwines morphisms of Lie groups with their differentials
   (c) morphisms of Lie groups with connected domain are characterized by their differentials
(2) the commutator of infinitesimal elements on the general linear group compared with the commutator bracket $[X,Y] := XY - YX$ on the matrix algebra; generalization to arbitrary Lie groups
(3) definition of Lie algebra and morphism of Lie algebra
(4) examples of Lie algebras:
   (a) $\mathrm{Lie}(G)$ for G a Lie group
   (b) $\mathrm{End}(V)$ for V a vector space
   (c) $\mathrm{Der}(A)$ for A an algebra
   (d) $\mathrm{Vect}(M) = \mathrm{Der}(C^\infty(M))$ for M a manifold
(5) proof that morphisms of Lie groups induce morphisms of Lie algebras
(6) definition of a representation of a Lie group, matrix coefficients with respect to a basis; example involving trigonometric functions and their addition law

Homework 4 (Due Oct 4)**.** Let G be the Lie group $\mathrm{SL}_2(\mathbb{C})$, $\mathfrak{g} := \mathrm{Lie}(G) = \mathfrak{sl}_2(\mathbb{C})$, and let n be a positive integer. Let $V \leq \mathbb{C}[x,y]$ be the $(n+1)$-dimensional vector space consisting of homogeneous polynomials of degree n in the variables x, y, so that a basis of V is given by the set of monomials

$$\mathcal{B} := \{x^n, x^{n-1}y, \dots, xy^{n-1}, y^n\}.$$

Let $R : G \to \mathrm{GL}(V)$ be the map given for $\phi \in V$ by

$$(R(g)\phi)(x,y) := \phi((x,y)g),$$

where $(x,y)g$ denotes matrix multiplication, so that more explicitly

$$(R(\begin{pmatrix} a & b \\ c & d \end{pmatrix})\phi)(x,y) = \phi(ax+cy, bx+dy).$$

(1) Verify that R defines a representation of G on V, hence (by a general theorem from class) that $dR : \mathfrak{g} \to \mathrm{End}(V)$ is a morphism of Lie algebras.
(2) Verify that the basis elements

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

   of $\mathfrak{g}$ satisfy $[X,Y] = H$.

(3) Compute the actions of $dR(H), dR(X), dR(Y) \in \mathrm{End}(V)$ explicitly with respect to the basis $\mathcal{B}$ of V and verify directly (without appeal to the general theorem from class) that $[dR(X), dR(Y)] = dR(H)$. [See §16.4 if the definition of $dR(X)$ is unclear.]

2.5. **10/4 (half-lecture) and 10/6: Representations of** SL(2)**. Objectives.** You should be able to analyze (finite-dimensional) representations of $\mathrm{SL}_2(\mathbb{C})$ by differentiating them to obtain representations of $\mathfrak{sl}_2(\mathbb{C})$, breaking the latter up into weight spaces, and studying how the weight spaces are permuted by raising and lowering operators.

Summary.

(1) definition of representations of Lie groups and algebras
(2) example of polynomial representations of linear Lie groups; explicit calculation of the induced representation on the Lie algebra
(3) discussion of the action of the standard basis of $\mathrm{SL}_2(\mathbb{C})$ on the \$(n+1)\$-dimensional representation W_n from Homework 4
(4) definition of invariant subspaces, irreducibility
(5) W_n is irreducible
(6) every irreducible finite-dimensional representation of $\mathrm{SL}_2(\mathbb{C})$ is isomorphic to some W_n

Homework 5 (Due Oct 11)**.** Let G be the Lie group $\mathrm{SL}_2(\mathbb{C})$. Let H, X, Y be the basis of $\mathfrak{g} := \mathrm{Lie}(G)$ as in Homework 4.

(1) For $\lambda \in \mathbb{C}$, let V_λ be the vector space with basis $(v_k)_{k \in \mathbb{Z}_{\geq 0}}$. Show that the action

$$\begin{aligned} Hv_k &= (\lambda - 2k)v_k \\ Xv_k &= k(\lambda - k + 1)v_{k-1} \quad (v_{-1} := 0) \\ Yv_k &= v_{k+1} \end{aligned}$$

defines a Lie algebra representation $\mathfrak{g} \to \mathrm{End}(V_\lambda)$. Determine the invariant subspaces of V_λ.

(2) Same question, but for the spaces U_ν ($\nu \in \mathbb{C}$) with basis $(v_k)_{k \in \mathbb{Z}}$ and action

$$\begin{aligned} Hv_k &= 2kv_k \\ Xv_k &= (\nu + k)v_{k+1} \\ Yv_k &= (\nu - k)v_{k-1}. \end{aligned}$$

(3) Let $\mathfrak{b} \leq \mathfrak{g}$ be the subspace spanned by H, X. Verify that $\mathfrak{b}$ is a Lie subalgebra. Let $\rho : \mathfrak{g} \to \mathrm{End}(V)$ be a finite-dimensional representation. Show that the following are equivalent for $v \in V$:
   (a) v is an eigenvector for every element of $\mathfrak{b}$.
   (b) v is an eigenvector of H and satisfies $Xv = 0$.

(4) Let V be a finite-dimensional representation of G. Let $v \in V$ be a nonzero element satisfying $Hv = \lambda v$ for some integer $\lambda \in \mathbb{Z}$. Define $v' \in V$ by

$$v' := \begin{cases} Y^\lambda v & \lambda \geq 0 \\ X^{-\lambda} v & \lambda \leq 0. \end{cases}$$

   (a) Verify that $Hv' = -\lambda v'$.
   (b) Prove that $v' \neq 0$. [Hint: Use the classification theorem for V proved in lecture.]

(5) (Optional) The nth Legendre polynomial P_n may be defined (perhaps up to a sign) by

$$P_n(x) := \sum_{k=0}^{n} \binom{n}{k}^2 \left(\frac{x-1}{2}\right)^{n-k} \left(\frac{x+1}{2}\right)^{k}.$$

The purpose of the exercise is to establish the formula

$$P_n(\cos\theta_1)P_n(\cos\theta_2) = \int_{\phi=-\pi}^{\pi} P_n(\cos(\theta_1)\cos(\theta_2) - \sin(\theta_2)\sin(\theta_2)\cos(\phi))\,\frac{d\phi}{2\pi}. \quad (2)$$

The geometric interpretation of the argument in the integrand is that if one fixes a point $O \in S^2$ at spherical distance θ_1 from the north pole N, then $\cos(\theta_1)\cos(\theta_2) - \sin(\theta_2)\sin(\theta_2)\cos(\phi)$ is the vertical coordinate of the point $P \in S^2$ at spherical distance θ_2 from O for which the angle between the arcs ON and OP is ϕ. [You might wish first to attempt to prove (2) directly.]

(a) Let $R : G \to \mathrm{GL}(V)$ be the $(2n+1)$-dimensional representation $V := W_{2n}$ defined in the lectures. Let $(v_k)_{k=-n..n}$ be the basis of V given by $v_k := x^{n+k}y^{n-k}$. For $i, j \in \{-n..n\}$, let $R_{ij}(g)$ denote the matrix entry of $g \in G$ with respect to this basis, i.e., the coefficient of v_i in $R(g)v_j$. For $\theta, \phi \in \mathbb{R}$, set

$$\kappa(\theta) := \begin{pmatrix} \cos(\theta/2) & i\sin(\theta/2) \\ i\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \quad \delta(\phi) := \begin{pmatrix} e^{i\phi/2} & \\ & e^{-i\phi/2} \end{pmatrix}.$$

Verify that

$$R_{00}(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \sum_{k=0}^{n} \binom{n}{k}^2 (ad)^k (cy)^{n-k}.$$

Deduce that

$$P_n(\cos(\theta)) = R_{00}(\kappa(\theta)).$$

(b) Show that for each θ_1, θ_2, ϕ there exist ϕ_1, ϕ_2, θ so that

$$\kappa(\theta_1)\delta(\phi)\kappa(\theta_2) = \delta(\phi_1)\kappa(\theta)\delta(\phi_2)$$

and moreover

$$\cos(\theta) = \cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2)\cos(\phi).$$

This can be proved directly via the geometric interpretation mentioned after (2) using the map $\mathrm{SU}(2) \to \mathrm{SO}(3)$ to be discussed next week; if one wishes to attempt an algebraic proof, it may help to note that

(i) *every* element of $\mathrm{SU}(2)$ may be decomposed as $\delta(\phi_1)\kappa(\theta)\delta(\phi_2)$,

(ii) the function

$$f : \mathrm{SU}(2) \to \mathbb{R}$$

given by

$$f(\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}) := 2|\alpha|^2 - 1$$

satisfies $f(\delta(\phi_1)g\delta(\phi_2)) = f(g)$ and $f(\kappa(\theta)) = \cos(\theta)$.

It may also help to treat first the case $\phi = 0$.

(c) Verify that $R_{kl}(\delta(\phi_1)g\delta(\phi_2)) = e^{i(-k\phi_1+l\phi_2)}$ for all relevant indices and arguments.

(d) Prove (2) by taking the $(0,0)$th matrix coefficient of the identity

$$R(g_1g_2) = R(g_1)R(g_2)$$

with

$$g_1 := \kappa(\theta_1)\delta(\phi),$$
$$g_2 := \kappa(\theta_2)$$

and integrating over ϕ. [It may be helpful to recall the Fourier inversion formula

$$\int_{\phi=-\pi}^{\pi} e^{ik\phi}\,\frac{d\phi}{2\pi} = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

for $k \in \mathbb{Z}$.]

2.6. **10/11 (half-lecture) and 10/13: The unitary trick. Objectives.** Given a real form of a complex Lie algebras, you should be able to relate representations of the two. You should be able to verify that the classical complex Lie groups have compact real forms and apply this fact to deduce their linear reductivity. You should know the definitions of Ad and ad and be able to apply the fact that they are morphisms.

Summary.

(1) introduction to and overview of the "unitary trick"
(2) defn of real form of a complex Lie algebra, comparison between representations
(3) defn of real form of a connected complex Lie group
(4) example of a representation that is not completely reducible
(5) lemma: complete reducibility is the same as invariant subspaces having invariant complements
(6) stated theorems that the following classes of groups are linearly reductive:
  (a) finite groups
  (b) (more generally) compact groups
  (c) complex connected Lie groups with a compact real form
(7) (Thursday onwards) we proved the above theorems.
(8) along the way, we proved the useful fact that any finite-dimensional representation of a compact group is unitarizable, i.e., admits an invariant inner product.
(9) we spent some time talking about examples of real forms and complexifications.
(10) we introduced Ad and ad. we related them, proved some of their basic properties, and interpreted the Jacobi identity in terms of properties of ad.

Homework 6 (Due Oct 18)**.**

(1) Prove that if $f : G \to H$ is a Lie group morphism, then $df(\mathrm{Ad}(g)X) = \mathrm{Ad}(f(g))df(X)$.
(2) Do Exercise 20.
(3) Let $G := \mathrm{SL}_2(\mathbb{C})$; it is a three-dimensional complex Lie group. Regard $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ as a three-dimensional holomorphic representation of G. Write down an explicit isomorphism between Ad and the representation $W_2 = \mathbb{C}x^2 \oplus \mathbb{C}xy \oplus \mathbb{C}y^2$ discussed in lecture.
(4) Let $\mathfrak{g}$ be a Lie algebra (the case $\mathfrak{g} = \mathrm{End}(V)$ is already interesting), let $n \geq 1$, and let

$$M = [Z_1, [Z_2, \dots, [Z_{n-1}, Z_n] \cdots] = \mathrm{ad}(Z_1)\,\mathrm{ad}(Z_2) \cdots \mathrm{ad}(Z_{n-1}) Z_n$$

be an n-fold iterated commutator of elements $Z_1, \dots, Z_n \in \mathfrak{g}$. Let $M' \in \mathfrak{g}$ be the result of formally expanding M as a sum of degree n monomials $Z_{i_1} \cdots Z_{i_n}$ and replacing each such monomial by the corresponding commutator $\mathrm{ad}(Z_{i_1}) \cdots \mathrm{ad}(Z_{i_{n-1}}) Z_{i_n}$. For example:
  (a) If $M = [X, Y]$, then we expand and set

$$M = XY - YX,$$
$$M' := [X, Y] - [Y, X]$$

and obtain $M' = 2[X, Y]$ after some simplification.

(b) If $M = [X, [Y, X]]$, then we expand and set

$$M = XYX - XXY - YXX + XYX,$$
$$M' := [X, [Y, X]] - [X, [X, Y]] - [Y, [X, X]] + [X, [Y, X]]$$

and obtain $M' = 3[X, [Y, X]]$ after some simplification.

Show that $M' = nM$. [Hint: induct on n. Use the consequence $[\mathrm{ad}(Z_{i_1}), [\mathrm{ad}(Z_{i_2}), \ldots, [\mathrm{ad}_{Z_{i_{n-1}}}, \mathrm{ad}_{Z_{i_n}}]]] = \mathrm{ad}([Z_{i_1}, [Z_{i_2}, \ldots, [Z_{i_{n-1}}, Z_{i_n}]]])$ of iterated application of the Jacobi identity in the form $\mathrm{ad}([X, Y]) = [\mathrm{ad}(X), \mathrm{ad}(Y)]$.]

2.7. **10/18 (half-lecture): The adjoint representation. Objectives.** You should be able to use the adjoint representation to describe some low-dimensional exceptional isomorphisms and to relate representations of the involved Lie groups and Lie algebras.

Summary.

(1) recap of what we've shown about representations of $\mathrm{SL}_2(\mathbb{C})$ and $\mathrm{SU}(2)$
(2) the exceptional isomorphisms $\mathrm{SL}_2(\mathbb{C})/\{\pm 1\} \cong \mathrm{SO}_3(\mathbb{C})$, $\mathrm{SU}(2)/\{\pm 1\} \cong \mathrm{SO}(3)$, $\mathrm{SL}_2(\mathbb{R})/\{\pm 1\} \cong \mathrm{SO}(1,2)^0$, (plus some generalities on quadratic spaces)

Homework 7 (Due Oct 25)**.**

(1) Write down a careful proof that the adjoint representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ of the group $G := \mathrm{SL}_2(\mathbb{R})$ induces an isomorphism of Lie groups $f : \mathrm{PSL}_2(\mathbb{R}) \xrightarrow{\cong} \mathrm{SO}(1,2)^0$. Give an explicit isomorphism of Lie algebras $df : \mathfrak{sl}_2(\mathbb{R}) \xrightarrow{\cong} \mathfrak{so}(1,2)$.
(2) Explain why the adjoint representation of $G = \mathrm{GL}_2(\mathbb{C})$ does *not* induce an isomorphism between G and $\mathrm{SO}_4(\mathbb{C})$.
(3) Let $\mathbb{H}$ denote Hamilton's quaternion algebra over $\mathbb{R}$, realized as the subalgebra of $M_2(\mathbb{C})$ given by
$$\mathbb{H} := \left\{ \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} : z, w \in \mathbb{C} \right\}.$$
Set
$$\mathbb{H}^{(1)} := \left\{ g \in \mathbb{H}^\times : \det(g) = 1 \right\}$$
and
$$\mathbb{H}_0 := \{ v \in \mathbb{H} : \mathrm{trace}(v) = 0 \}.$$
(a) Verify that $\mathbb{H}^{(1)} = \mathrm{SU}(2)$. Deduce in particular via the embedding $(z,w) \hookrightarrow \mathbb{C}^2 \hookrightarrow \mathbb{R}^4$ that $\mathrm{SU}(2)$ is diffeomorphic to the three-dimensional sphere S^3.
(b) Show that $(\mathbb{H}_0, \det)$ is a quadratic space over $\mathbb{R}$ of signature $(3,0)$.
(c) Let $\alpha : \mathbb{H}^\times \to \mathrm{GL}(\mathbb{H}_0)$ be the conjugation action $\alpha(g)(v) := gvg^{-1}$ ($g \in \mathbb{H}^\times, v \in \mathbb{H}_0$). Show that $\alpha(\mathbb{H}^{(1)}) = \mathrm{SO}(\mathbb{H}_0, \det) \cong \mathrm{SO}(3)$ and that $\{g \in \mathbb{H}^{(1)} : \alpha(g) = 1\} = \{\pm E\}$ where $E := \left(\begin{smallmatrix} 1 & \\ & 1 \end{smallmatrix}\right)$. Deduce that
$$\mathbb{H}^\times / \mathbb{R}^\times \cong \mathbb{H}^{(1)} / \{\pm 1\} \cong \mathrm{SO}(3).$$
[Use the connectedness of $\mathrm{SO}(3)$ to reduce the problem to one involving Lie algebras.]
(d) Deduce that α induces an isomorphism $\mathrm{SU}(2)/\{\pm 1\} \cong \mathrm{SO}(3)$. Compare with the proof given in class by showing that one has $\mathrm{Lie}(\mathbb{H}^{(1)}) = \mathbb{H}_0$ under the natural identification $\mathrm{Lie}(\mathbb{H}^\times) = \mathbb{H}$.

(4) (Optional) Here we understand how the map $\mathrm{SU}(2) \to \mathrm{SO}(3)$ may be defined by comparing the standard actions $\mathbb{P}^1(\mathbb{C}) \circlearrowleft \mathrm{SU}(2)$ and $S^2 \circlearrowleft \mathrm{SO}(3)$ under the identification $\mathbb{P}^1(\mathbb{C}) \cong S^2$ given by stereographic projection:
(a) Let $\mathbb{P}^1(\mathbb{C})$ be the complex projective line, that is, the set of equivalence classes $[z : w]$ of row vectors $(z,w) \in \mathbb{C}^2 - \{0\}$ under the equivalence relation $(z,w) \simeq (\lambda z, \lambda w)$ for all $\lambda \in \mathbb{C}^\times$. Verify that $\mathrm{SU}(2)$ acts on

$\mathbb{P}^1(\mathbb{C})$ via

$$[\xi : \eta] \cdot g := [a\xi + c\eta : b\xi + d\eta] \text{ for } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SU}(2).$$

(b) Let $S^2 := \{(x, y, z) : x^2 + y^2 + z^2\} \subseteq \mathbb{R}^3$ be the standard two-dimensional sphere. Let SO(3) act on S^2 in the usual way: for $v \in S^2$ and $g \in G$, $v \cdot g$ is given by matrix multiplication. Verify that an element of SO(3) is determined by its action on S^2.

(c) Let $p := (0, 0, -1) \in S^2$ denote the "south pole" and let $P := \{(u, v, 0) : u, v \in \mathbb{R}\} \subseteq \mathbb{R}^3$ denote the "equatorial plane." Let $\pi : S^2 - \{p\} \to P$ denote the result of stereographic projection from p, thus $\pi(x, y, z) = (u, v, 0)$ means that the points $(0, 0, -1)$, $(u, v, 0)$, (x, y, z) are collinear. Let $\rho : P \hookrightarrow \mathbb{P}^1(\mathbb{C})$ be the map $\rho(u, v, 0) := [u + iv : 1]$. Verify that the composition $\rho \circ \pi : S^2 - \{p\}$ extends to a homeomorphism

$$\iota : S^2 \to \mathbb{P}^1(\mathbb{C})$$

for which $\iota(p) = [1 : 0]$.

(d) Show that for each $g \in \mathrm{SU}(2)$ there is a unique $\alpha(g) \in \mathrm{SO}(3)$ so that for all $s \in S^2$, one has $\iota(s \cdot \alpha(g)) = \iota(s) \cdot g$. Show that the map $\alpha : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is a surjective morphism of Lie groups.

(e) Read about the "Hopf fibration" somewhere and understand its relevance.

2.8. **10/20 (first half): Character theory for** SL(2) **(algebraic). Objectives.** Given a (reasonably explicit) representation $\mathrm{SL}_2(\mathbb{C})$ or some closely related group, you should be able to determine its reduction into irreducibles by computing its character and multiplying by the Weyl denominator.

Summary.

(1) definitions of direct sum and tensor product of representations of Lie groups and Lie algebras
(2) characters of representations of $\mathrm{SL}_2(\mathbb{C})$ as Laurent polynomials in one variable z
(3) compatibility with direct sum and direct product
(4) the characters of the irreducibles
(5) the Weyl denominator $z - 1/z$
(6) Clebsch–Gordon decomposition

Homework 8 (Due Oct 25)**.**

(1) Verify that if $\rho_j : \mathfrak{g} \to \mathrm{End}(V_j)$ $(j = 1, 2)$ are representations of a Lie algebra, then the map $\rho_1 \otimes \rho_2 : \mathfrak{g} \to \mathrm{End}(V_1 \otimes V_2)$, defined as in class by linear extension of its definition on pure tensors $v_1 \otimes v_2 \in V_1 \otimes V_2$ by

$$((\rho_1 \otimes \rho_2)(X))(v_1 \otimes v_2) := \rho_1(X)v_1 \otimes v_2 + v_1 \otimes \rho_2(X)v_2,$$

or in abbreviated form simply by

$$X(v_1 \otimes v_2) := Xv_1 \otimes v_2 + v_1 \otimes Xv_2,$$

defines a representation of Lie algebras.

(2) Verify that the map $W_2 \oplus W_0 \to W_1 \otimes W_1$ defined in class is an isomorphism of $\mathrm{SL}_2(\mathbb{C})$-representations.

(3) Show that there does *not* exist a representation V of $\mathrm{SL}_2(\mathbb{C})$ whose weight spaces $V[m] := \{v \in V : Hv = mv\}$ $(m \in \mathbb{Z})$ have dimensions given by

$$\dim V[m] = \begin{cases} 1 & m \in \{-7, -6, -5, 5, 6, 7\}, \\ 0 & \text{otherwise.} \end{cases}$$

More generally, for which functions $\nu : \mathbb{Z} \to \mathbb{Z}_{\geq 0}$ does there exist a finite-dimensional representation V of $\mathrm{SL}_2(\mathbb{C})$ with $\dim V[n] = \nu(n)$ for all n? [Hint: write $V \cong \oplus W_m^{\oplus \mu(m)}$ and apply $D \cdot \mathrm{ch}(.)$ to both sides.]

(4) (Optional) Given $k \in \mathbb{Z}_{\geq 0}$ and a representation $R : G \to \mathrm{GL}(V)$ of a Lie group G, one obtains a symmetric power representation $\mathrm{Sym}^k(R) : G \to \mathrm{GL}(\mathrm{Sym}^k(V))$ on the symmetric power vector space $\mathrm{Sym}^k(V)$; see §16.9.3 or Google for some details. The purpose of this exercise is to relate the character of $\mathrm{Sym}^k(V)$ to that of V. We restrict to the case $G := \mathrm{SL}_2(\mathbb{C})$, although the arguments are somewhat more general. Let $A := \mathbb{Z}[z, z^{-1}]$ be as in lecture.

(a) Define $\sigma(V)$ to be the element of the formal power series ring $A[[T]]$ in the variable T with coefficients in A given by the formula

$$\sigma(V) := \sum_{k \in \mathbb{Z}_{\geq 0}} \mathrm{ch}(\mathrm{Sym}^k V) T^k.$$

Show that

$$\sigma(V) = \exp \sum_{k \geq 1} \frac{\Psi^k(\mathrm{ch}(V)) T^k}{k}$$

where Ψ^k is defined via the substitution $z \mapsto z^k$, i.e., by setting $\Psi^k(\chi)(z) := \chi(z^k)$ for $\chi \in A$. [Hint: the identity

$$\exp \sum_{k \geq 1} \frac{x^k}{k} = \frac{1}{1-x} = \sum_{k \geq 0} x^k$$

is relevant.]

(b) Deduce the recursion relation

$$n \operatorname{ch}(\operatorname{Sym}^k(V)) = \sum_{k=1}^{n} \Psi^k(\operatorname{ch}(V)) \operatorname{ch}(\operatorname{Sym}^{n-k}(V)).$$

Check that this is consistent with the isomorphisms $\operatorname{Sym}^n(W_1) \cong W_n$.

(c) Specialize the above relation to the case $n = 2$ to obtain

$$\operatorname{ch}(\operatorname{Sym}^2(V)) = \frac{\operatorname{ch}(V)^2 - \Psi^2(\operatorname{ch}(V))}{2}.$$

For each irreducible representation W_m of G, compute $\operatorname{ch}(\operatorname{Sym}^2(W_m))$ by the above formula and use it to derive the decomposition

$$\operatorname{Sym}^2(W_m) \cong W_{2m} \oplus W_{2m-4} \oplus \cdots = \oplus_{\substack{0 \leq j \leq 2m: \\ j \equiv 2m(4)}} W_j.$$

of $\operatorname{Sym}^2(W_m)$ into irreducibles. (It is also instructive, and not difficult, to derive this decomposition directly.) Write down an explicit isomorphism $\operatorname{Sym}^2(W_2) \cong W_4 \oplus W_0$.

2.9. **10/20 (second half): Maurer-Cartan equations; lifting morphisms of Lie algebras. Objectives.** You should know that $\operatorname{Hom}(G, H) \to \operatorname{Hom}(\mathfrak{g}, \mathfrak{h})$ is injective when G is connected and surjective when G is simply-connected and be able to give some basic counter-examples indicating the necessity of such conditions. You should be able to describe the role played by the Maurer–Cartan equations in establishing surjectivity in the simply-connected case. Given hints, you should be able to apply the Maurer–Cartan equation to related problems.

Summary.

(1) statement of main theorem on lifts of Lie algebra morphisms
(2) proof via paths and Maurer–Cartan equation

Homework 9 (Due Oct 25)**.**

(1) For a smooth scalar-valued function $f : \mathbb{R} \dashrightarrow \mathbb{R}$, the chain rule implies that

$$\frac{d}{dt} \exp(f(t)) = \exp(f(t)) f'(t). \tag{3}$$

The purpose of this exercise is to generalize the above identity as an application of a technique introduced in lecture. Let G be a real Lie group with Lie algebra $\mathfrak{g}$; the problem is already interesting when $G = \mathrm{GL}_n(\mathbb{R})$, so feel free to assume that. Prove that for a smooth function $f : \mathbb{R} \dashrightarrow \mathfrak{g}$, one has

$$\frac{d}{dt} \exp(f(t)) = \exp(f(t)) \sum_{n=1}^{\infty} \frac{(-\operatorname{ad}_{f(t)})^{n-1} f'(t)}{n!} \tag{4}$$

where we may write more explicitly

$$(-\operatorname{ad}_{f(t)})^{n-1} f'(t) = [[[f'(t), f(t)], f(t)], \dots, f(t)].$$

Observe that (4) specializes to (3) when G is abelian, so that $\operatorname{ad}_{f(t)} = 0$.

[Hint: Consider the map $g : \mathbb{R}^2 \dashrightarrow G$ given by

$$g(s,t) := \exp(sf(t)).$$

Define $\xi : \mathbb{R}^2 \dashrightarrow \mathfrak{g}$ by $\frac{\partial g}{\partial t} = g\xi$, so that $\frac{d}{dt}\exp(f(t)) = \exp(f(t))\xi(t,1)$. Apply the Maurer–Cartan equation (§19.1) to characterize ξ as the unique solution F to the differential equation

$$\frac{\partial F}{\partial s}(t,s) = f'(t) + [F(t,s), f(t)]$$

with initial condition $F(t,0) = 0$. On the other hand, verify that such a solution may be given explicitly by

$$F(t,s) := \sum_{n=1}^{\infty} s^n \frac{(-\operatorname{ad}_{f(t)})^{n-1} f'(t)}{n!}$$

and set $s := 1$ to conclude.]

2.10. **10/25 (half-lecture): universal covering group. Objectives.** You should be able to classify the Lie groups having a given Lie algebra in terms of discrete central subgroups of a simply-connected group. You should be able to describe some basic examples of covering morphisms and use them to determine the fundamental groups of some Lie groups.

Summary.

(1) The main theorem was that for any connected Lie group G there exists a simply-connected Lie group $\tilde{G}$ and a covering morphism $p : \tilde{G} \to G$ whose kernel $N = \ker(p)$ is a discrete subgroup of the center of $\tilde{G}$, with $(\tilde{G}, N)$ uniquely determined up to isomorphism. Moreover, $\pi_1(G) \cong N$.
(2) We gave several examples to which this applies. Some further examples are given on the homework.
(3) We stated without proof that every finite-dimensional Lie algebra arises from some Lie group.
(4) By combining with a result from last time, we deduced that the category of simply-connected Lie groups is equivalent to the category of finite-dimensional Lie algebras.
(5) We recalled the definition of "cover" (locally trivial fiber bundle with discrete fiber). We briefly recalled the construction of the universal cover of a connected manifold and stated its universal property (existence of unique lifts of paths). We defined the group structure on the simply-connected cover of a Lie group.
(6) We reduced the remainder of the proof of the main theorem to some lemmas left mostly as exercises.

Homework 10 (Due Nov 1)**.**

(1) Let $n \geq 1$. For the purposes of this exercise, you may use without proof that $\mathrm{SL}_n(\mathbb{C})$ and $\mathrm{SU}(n)$ are simply-connected.
    (a) Show that $\pi_1(\mathrm{PGL}_n(\mathbb{C})) \cong \mathbb{Z}/n$. [Hint: show that the natural map $p : \mathrm{SL}_n(\mathbb{C}) \to \mathrm{PGL}_n(\mathbb{C})$ is a covering morphism, determine the kernel of p, and appeal to the theorem from lecture.]
    (b) Set $\mathfrak{g} := \mathfrak{sl}_n(\mathbb{C})$. Determine the connected Lie groups G (up to isomorphism, and over either the real or complex numbers – it doesn't matter) having Lie algebra (isomorphic to) $\mathfrak{g}$, and describe their fundamental groups $\pi_1(G)$. [Hint: start by determining the center of $\mathrm{SL}_n(\mathbb{C})$.] Interpret "determine" as you wish. For instance, you should be able to answer the following question: How many isomorphism classes of connected Lie groups have Lie algebra isomorphic to $\mathfrak{sl}_{12}(\mathbb{C})$?
    (c) Same question but for $\mathfrak{g} := \mathfrak{su}(n)$.
(2) Verify that (at least one or two of) the following maps are covering morphisms of Lie groups and determine their kernels. [Hint: the lemma from lecture characterizing covering morphisms may help.]
    (a) The morphism of complex Lie groups
$$\mathbb{C} \xrightarrow{\exp} \mathbb{C}^\times.$$
    (b) The morphism of complex Lie groups
$$\mathrm{SL}_2(\mathbb{C}) \xrightarrow{\mathrm{Ad}} \mathrm{SO}(\mathfrak{sl}_2(\mathbb{C}), \det) \cong \mathrm{SO}_3(\mathbb{C}).$$

(c) The morphism of real Lie groups
$$\mathrm{SU}(2) \xrightarrow{\mathrm{Ad}} \mathrm{SO}(\mathfrak{su}(2), \det) \cong \mathrm{SO}(3).$$
(d) The morphism of real Lie groups
$$\mathrm{SL}_2(\mathbb{R}) \xrightarrow{\mathrm{Ad}} \mathrm{SO}(\mathfrak{sl}_2(\mathbb{R}), \det)^0 \cong \mathrm{SO}(1,2)^0.$$
(e) The morphism of complex Lie groups
$$\mathrm{SL}_2(\mathbb{C}) \times \mathrm{SL}_2(\mathbb{C}) \to \mathrm{SO}(M_2(\mathbb{C}), \det) \cong \mathrm{SO}_4(\mathbb{C}),$$
$$(g,h) \mapsto [x \mapsto gxh^{-1}].$$
(f) The morphism of real Lie groups
$$\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(\mathbb{H}, \det) \cong \mathrm{SO}(4),$$
$$(g,h) \mapsto [x \mapsto gxh^{-1}],$$
where $\mathbb{H} = \left\{ \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \right\} \subseteq M_2(\mathbb{C})$ denotes Hamilton's quaternion algebra as considered on previous homeworks.

(g) The morphism of *real* Lie groups
$$\mathrm{SL}_2(\mathbb{C}) \to \mathrm{SO}(V, \det)^0 \cong \mathrm{SO}(1,3)^0,$$
$$g \mapsto [x \mapsto gx\overline{g}^t],$$
where $V := \left\{X \in M_2(\mathbb{C}) : \overline{X} = X^t\right\} = \left\{ \begin{pmatrix} x & z \\ \overline{z} & y \end{pmatrix} : x, y \in \mathbb{R}, z \in \mathbb{C} \right\}$ is the space of 2×2 hermitian matrices.

2.11. **10/27: Fundamental groups of Lie groups. Objectives**. You should be able to analyze the topology of Lie groups by

- applying the homotopy exact sequence to their transitive actions, and
- via the Cartan decomposition.

Summary

(1) Description, without proof, of the fundamental groups of the classical groups; empirical observation that complex Lie groups and their compact real forms (if they exist) have the same fundamental groups
(2) Homotopy exact sequence and its consequences; proofs of some of the descriptions of fundamental groups given before (most were left as exercises). For example, we showed inductively that $\mathrm{SL}_n(\mathbb{C})$ is simply-connected.
(3) Quotient groups (abstract, topological, Lie), quotient manifolds, transitive action theorem; sketch of construction of smooth structure on the quotient
(4) Statement of Cartan decomposition; application to comparing homotopy groups, recap on the unitary trick

Homework 11 (Due Nov 1)**.**

(1) Let $p, q \geq 1$. Set $G := \mathrm{O}(p, q)$, realized as usual as a subgroup of $\mathrm{GL}_{p+q}(\mathbb{R})$. Let Θ be given by $\Theta(g) := {}^t\overline{g}^{-1}$.
   (a) Show that the subgroup $K := \{g \in G : \Theta(g) = g\}$ fixed by G identifies with $\mathrm{O}(p) \times \mathrm{O}(q)$.
   (b) Use (without proof) the Cartan decomposition (§23)
$$K \times \mathfrak{p} \cong G$$
$$(k, Y) \mapsto k \exp(Y)$$
   to show that G has four connected components.
   (c) Describe the Cartan decomposition explicitly in the special case $p := 1, q := 1$, and compare with the related problem on Homework 2.

(2) (Optional) Let
$$G := \mathrm{SL}_2(\mathbb{R}).$$
Denote by $\mathbb{H} := \{x + iy : x, y \in \mathbb{R}, y > 0\} \subseteq \mathbb{C}$ the upper half-plane. The group G acts smoothly on $\mathbb{H}$ by fractional linear transformations:
$$gz := \frac{az + b}{cz + d} \text{ if } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G, \quad z \in \mathbb{H}.$$
Denote by $\tilde{G}$ the set of all pairs (g, ϕ), where
- $g \in G$, and
- $\phi : \mathbb{H} \to \mathbb{C}$ is a holomorphic function with the property that
$$\exp(\phi(z)) = cz + d \text{ if we write } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

In other words, ϕ is a branch of $\log(cz+d)$; it is determined by g and any of its values, say $\phi(i)$. We may define on $\tilde{G}$ a smooth structure by regarding it as a submanifold of $G \times \mathbb{C}$ via the embedding $(g, \phi) \mapsto (g, \phi(i))$. We define on $\tilde{G}$ a group structure by the law
$$(g_1, \phi_1) \cdot (g_2, \phi_2) := (g_1 g_2, \phi_1^{g_2} \phi_2) \text{ where } (\phi_1^{g_2} \phi_2)(z) := \phi_1(g_2 z)\phi_2(z).$$

This group operation is then associative and smooth, and defines on $\tilde{G}$ the structure of a Lie group. The natural map

$$\pi : \tilde{G} \to G$$

given by $\pi((g, \phi)) := g$, is smooth and surjective. The group $\tilde{G}$ inherits from G an action on $\mathbb{H}$: $(g, \phi) \cdot z := gz$. The map

$$\kappa : \mathbb{R} \to \tilde{G}$$

given by

$$\kappa(\theta) := (\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}, \phi_\theta),$$

where $\phi_\theta(z)$ is the unique branch of $\log(-\sin(\theta)z + \cos(\theta))$ for which $\phi_\theta(i) = -i\theta$, is a morphism of Lie groups. [For the purposes of this exercise, all of the assertions just made may be regarded as sufficiently self-evident as not to require proof.]

(a) Write down an isomorphism $N \cong \mathbb{Z}$.

(b) Show that $\tilde{G}$ is connected. [Hint: use κ to show that $N \subseteq \tilde{G}^0$, and use that G is connected.]

(c) Let $H \leq \tilde{G}$ denote the image of κ. Show that H is the stabilizer in $\tilde{G}$ of the point $i \in \mathbb{H}$.

(d) Show that $\tilde{G}$ is simply-connected. [The homotopy exact sequence gives one way to do this; alternatively, one can find a diffeomorphism $\tilde{G} \cong \mathbb{H} \times \mathbb{R}$.]

In summary, $\tilde{G}$ is the simply-connected covering group of G, and $\pi_1(G) \cong N \cong \mathbb{Z}$.

2.12. **11/1: The Baker–Campbell–Hausdorff(–Dynkin) formula. Objectives.** You should be able to describe the BCH formula (qualitatively), specialize it to the case of \$2\$-step nilpotent groups, and apply it to derive asymptotic expansions in local exponential coordinates of products in a Lie group.

Summary.

(1) We defined what it means for a pair of Lie groups to be locally isomorphic, and explained how the lifting theorem for simply-connected Lie groups and the existence of the universal cover of a given Lie group imply that two Lie groups are locally isomorphic if and only if their Lie algebras are isomorphic.
(2) Motivated by a "local" proof of this assertion, we initiated a study of the $x * y := \log(\exp(x)\exp(y))$ for a pair of matrices x, y.
(3) We verified empirically that the first couple homogeneous components z_n in the series expansion of $x*y$ are Lie polynomials, i.e., linear combinations of iterated Lie commutators involving x and y. We stated the BCH theorem, which says that this empirical observation holds for all n.
(4) We stated Dynkin's formula and indicated briefly how it follows from the BCH theorem together with an earlier homework problem on iterated commutators.
(5) We proved the BCH theorem in its qualitative form using the homework problem on the derivative of the exponential map.

Homework 12 (Due Nov 8)**.**

(1) Let $s \in \mathbb{Z}_{\geq 0}$. A group G is said to be *s-step nilpotent* if all for all $x_1, \dots, x_{s+1} \in G$, the iterated commutator $(x_1, (x_2, \dots, (x_s, x_{s+1})))$ is the identity element. Here $(x, y) := xyx^{-1}y^{-1}$. For example, G is 1-step nilpotent if and only if it is abelian.

Similarly, a Lie algebra $\mathfrak{g}$ is said to be *s-step nilpotent* if $[x_1, [x_2, \dots, [x_s, x_{s+1}]]] = 0$ for all $x_1, \dots, x_{s+1} \in \mathfrak{g}$. We call $\mathfrak{g}$ abelian if it is 1-step nilpotent, or equivalently, if the commutator bracket vanishes identity.

(a) Verify that the Lie group $G \leq \mathrm{SL}_{s+1}(\mathbb{R})$ consisting of strictly upper-triangular matrices is s-step nilpotent.
(b) Let G be a connected Lie group with adjoint representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$. Show that $\ker(\mathrm{Ad})$ is the center of G.
(c) Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Show for $s \leq 2$ that G is s-step nilpotent if and only if $\mathfrak{g}$ is s-step nilpotent. (The same conclusion holds for all s, and can be proved similarly; the assumption $s \leq 2$ is just to simplify the homework problem.)
(d) If G is 2-step nilpotent, show that the BCH formula takes the simple form

$$x * y = x + y + \frac{1}{2}[x, y] \tag{5}$$

for small enough $x, y \in \mathfrak{g}$. Show that the quantities

$$x * (y - x), \quad y + \frac{1}{2}[x, y], \quad \frac{x}{2} * y * (\frac{-x}{2}) \tag{6}$$

coincide.

(e) Let $G \leq \mathrm{SL}_3(\mathbb{R})$ be the three-dimensional Lie group consisting of strictly upper-triangular matrices; we have seen already that it is 2-step nilpotent. Establish the formula

$$\exp(\begin{pmatrix} & x & 0 \\ & & 0 \\ & & \end{pmatrix})\exp(\begin{pmatrix} & 0 & 0 \\ & & y \\ & & \end{pmatrix}) = \exp(\begin{pmatrix} & x & xy/2 \\ & & y \\ & & \end{pmatrix})$$

in two ways:

(i) By direct calculation with power series.

(ii) By application of the BCHD formula to

$$X := \begin{pmatrix} & x & \\ & & \\ & & \end{pmatrix}, \quad Y := \begin{pmatrix} & & \\ & & y \\ & & \end{pmatrix}.$$

(2) Let G be a Lie group. Equip its Lie algebra $\mathfrak{g}$ with some norm $|.|$. Use the BCHD formula (or part of its proof) to show that for small enough $x, y \in \mathfrak{g}$,

$$x * y * (-x) = \exp(\mathrm{ad}_x)y = y + [x, y] + \frac{[x, [x, y]]}{2} + \cdots, \tag{7}$$

$$x * y = x + F_1(\mathrm{ad}_x)y + O(|y|^2), \quad F_1(z) := \frac{z}{1 - \exp(-z)} = 1 + \frac{z}{2} + \cdots, \tag{8}$$

$$x * (y - x) = F_2(\mathrm{ad}_x)y + O(|y|^2), \quad F_2(z) := \frac{\exp(z) - 1}{z} = 1 + \frac{z}{2} + \cdots. \tag{9}$$

Deduce that the quantities (6) coincide up to a possible error of size $O(|x|^2|y| + |y|^2)$. [Hint: For (8) and (9), one can specialize the BCHD formula directly. Alternatively, let $t \in \mathbb{R}$ be small and let $f(t)$ denote one of the expressions $x * ty$ or $x * (ty - x)$. Then the BCH formula in its qualitative form, together with Taylor's theorem, reduces the problem to computing $f(0)$ and $f'(0)$. For this, Homework 9 and (86) may be helpful.]

2.13. **11/3: Closed subgroups are Lie; virtual subgroups vs. Lie subalgebras. Objectives.** You should know the following trivia, and some of their basic consequences:

(1) closed subgroups of a Lie group are the same as Lie subgroups.
(2) connected immersed Lie subgroups of a given Lie group correspond to the Lie subalgebras of its Lie algebra.

Summary.

(1) I explained how the BCH formula implies directly that isomorphisms of Lie algebras lift to local isomorphisms of Lie groups, and how Lie theory is the same whether one starts with "smooth manifolds" or "analytic manifolds".
(2) I stated the theorem that closed subgroups of Lie groups are Lie subgroups, and indicated briefly how it implies that continuous homomorphisms between Lie groups are automatically smooth (hence, by BCH, analytic with respect to exponential coordinates). I then proved that theorem.
(3) I explained the correspondence between Lie subalgebras and immersed Lie subgroups and briefly mentioned some ideas of the proof.

Homework 13 (Due Nov 8)**.**

(1) Following the sketch indicated in lecture, write down a careful proof that a continuous homomorphism of Lie groups $G \to H$ is automatically smooth.

(2) Let $n \geq 1$. Denote by 1_n the $n \times n$ identity matrix. Set $J := \begin{pmatrix} & 1_n \\ -1_n & \end{pmatrix}$; it is a $2n \times 2n$ matrix. Set

$$\mathrm{Sp}_{2n}(\mathbb{C}) := \{g \in \mathrm{SL}_{2n}(\mathbb{C}) : g^t J g = J\}$$
$$\mathrm{Sp}_{2n}(\mathbb{R}) := \mathrm{Sp}_{2n}(\mathbb{C}) \cap \mathrm{SL}_{2n}(\mathbb{R})\}$$

and

$$\mathrm{Sp}(2n) := U(2n) \cap \mathrm{Sp}_{2n}(\mathbb{C}).$$

(In practice, one alternates between writing $\mathrm{Sp}(2n)$ and $\mathrm{Sp}(n)$ to mean the same thing. Beware conventions.) Check that this definition is the same as what we gave earlier using quaternions. Show that $\mathrm{Sp}(2n)$ is a compact real form of $\mathrm{Sp}_{2n}(\mathbb{C})$ and that

$$\mathfrak{sp}_{2n}(\mathbb{C}) := \mathrm{Lie}(\mathrm{Sp}_{2n}(\mathbb{C})) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in M_n(\mathbb{C}), d = -a^t, b^t = b, c^t = c \right\}.$$

Show for $\mathbf{k} = \mathbb{R}, \mathbb{C}$ and $a \in \mathrm{GL}_n(\mathbf{k})$ and $b \in M_n(\mathbf{k})$ with $b^t = b$ that $\mathrm{Sp}_{2n}(\mathbf{k})$ contains the matrices

$$\begin{pmatrix} 0_n & 1_n \\ -1_n & 0_n \end{pmatrix}, \quad \begin{pmatrix} a & \\ & {}^t a^{-1} \end{pmatrix}, \quad \begin{pmatrix} 1_n & b \\ & 1_n \end{pmatrix}.$$

Show that $\mathrm{Sp}_{2n}(\mathbb{C})$ and $\mathrm{Sp}(2n)$ are connected and simply-connected. Show that $\mathrm{Sp}_{2n}(\mathbb{R})$ is connected, that $\mathrm{Sp}_{2n}(\mathbb{R}) \cap \mathrm{U}(2n)$ is isomorphic to $\mathrm{U}(n)$, and that $\pi_1(\mathrm{Sp}_{2n}(\mathbb{R})) \cong \mathbb{Z}$. [Hint: one way is as follows. Study $\mathrm{Sp}_{2n}(\mathbb{C})$ inductively on n by considering the natural action on $\mathbb{C}^{2n} - \{0\}$, analyzing stabilizers, and using the homotopy exact sequence. Study $\mathrm{Sp}(2n)$ using the Cartan decomposition. Study $\mathrm{Sp}_{2n}(\mathbb{R})$ using either the Cartan decomposition or the homotopy exact sequence.]

2.14. **11/8: Simplicity of Lie groups and Lie algebras. Objectives.** You should be able to explain what it means for a Lie group to be "simple as a Lie group" and how this differs from being simple as an abstract group.

Summary.

(1) I defined what it means for Lie algebras and Lie groups to be simple and proved the equivalence of the following assertions concerning a connected Lie group:
  (a) It is simple (no nontrivial proper normal connected virtual Lie subgroups).
  (b) Its Lie algebra is simple (no nonzero proper ideals).
  (c) Its proper normal subgroups are discrete.

(2) I recalled the classical groups (complex forms, compact real forms, Lie algebras) and stated as a motivating goal the theorem describing when they are simple and what are the exceptional isomorphisms between them.

Homework 14 (Due Nov 15)**.**

(1) Check carefully that the following are equivalent for a connected Lie group G with Lie algebra $\mathfrak{g}$ and connected virtual Lie subgroup H with Lie algebra $\mathfrak{h}$:
  (a) H is a normal subgroup of G.
  (b) $\mathrm{Ad}(G)\mathfrak{h} \subseteq \mathfrak{h}$
  (c) $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.

  [Hint: Exercise 22 may be useful.]

(2) (a) Prove by hand that $\mathfrak{sl}_2(\mathbb{C})$ is simple.
  (b) Complete the following sentence: "a Lie algebra $\mathfrak{g}$ is simple if and only if its adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathrm{GL}(\mathfrak{g})$ is (...)." Explain then how we have already secretly proven that $\mathfrak{sl}_2(\mathbb{C})$ is simple.

2.15. **11/10: Simplicity of the special linear Lie algebra. Objectives.** You should be to analyze ideals in classical Lie algebras by decomposing them into root spaces.

Summary.

(1) We recalled briefly some facts we established long ago concerning SL(2).
(2) We recalled, with sketch of proof, the theorem from linear algebra that commuting diagonalizable operators are simultaneously diagonalizable. We then reformulated this result in terms of representations of abelian Lie algebras.
(3) We defined the set of weights of a semisimple representation of an abelian Lie algebra, and illustrated the definition in the basic case of the standard representation of the diagonal subalgebra of the matrix algebra.
(4) We defined the set of roots for the diagonal subalgebra of $\mathfrak{sl}_n(\mathbb{C})$ and described the root spaces and their commutation relations explicitly.
(5) We proved that $\mathfrak{sl}_n(\mathbb{C})$ is simple by splitting any nonzero ideal as a sum of root spaces and applying suitable commutators.

Homework 15 (Due Nov 17)**.** Set $\mathfrak{g} := \mathfrak{sp}_{2n}(\mathbb{C})$. The main purpose of this exercise is to carry out the analogue for $\mathfrak{g}$ of what was done in lecture for $\mathfrak{sl}_n(\mathbb{C})$. We include some additional computations of independent interest; they are straightforward but (I think) instructive.

(1) Recall the description of $\mathfrak{g}$ from Homework 13. Verify that $\dim(\mathfrak{g}) = 2n^2+n$. For $1 \leq j,k \leq 2n$, let $E_{jk} \in M_{2n}(\mathbb{C})$ denote the elementary matrix with 1 in the (i,j)th entry and 0 elsewhere, thus $E_{jk}e_k = e_j$ and $e_{jk}e_l = 0$ for $l \neq k$, where $e_1, \dots, e_{2n}$ denotes the standard basis of $\mathbb{C}^{2n}$. Verify that a basis for $\mathfrak{g}$ is given by elements of the following form, where $1 \leq j,k \leq n$:
   - $E_{jj} - E_{n+j,n+j}$
   - $E_{j,k} - E_{n+k,n+j}$ for $j \neq k$
   - $E_{j,n+k} + E_{k,n+j}$ for $j \leq k$
   - $E_{n+j,k} + E_{n+k,j}$ for $j \leq k$

   Write all these elements out explicitly when $n = 2$.
(2) Let $\mathfrak{h} \leq \mathfrak{g}$ denote the subalgebra of diagonal matrices. Then $\dim(\mathfrak{h}) = n$. Explicitly, $\mathfrak{h}$ has the basis $E_{jj} - E_{n+j,n+j}$ $(j = 1, \dots, n)$ and consists of of matrices of the form

$$H = \begin{pmatrix} \lambda_1(H) & 0 & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_n(H) & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_1(H) & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda_n(H) \end{pmatrix}. \tag{10}$$

A basis for the dual space $\mathfrak{h}^*$ consists of the elements $\lambda_1, \dots, \lambda_n$ defined by (10), or equivalently, by $\lambda_k(E_{jj} - E_{n+j,n+j}) := \delta_{jk}$.

Let R denote the set of *roots* for the pair $(\mathfrak{g}, \mathfrak{h})$, defined exactly as in the case of $\mathfrak{sl}_n(\mathbb{C})$ to consist of all nonzero elements $\alpha \in \mathfrak{h}^*$ for which the eigenspace

$$\mathfrak{g}^\alpha := \{X \in \mathfrak{g} : [H, X] = \alpha(H)X \text{ for all } H \in \mathfrak{h}\}$$

is nonzero. The same argument as in lecture shows that

$$\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{\alpha \in R} \mathfrak{g}^{\alpha}).$$

Show that

$$R = \{\pm(\lambda_j \pm \lambda_k) : j < k\} \cup \{\pm 2\lambda_k\},$$

where the signs $\pm$ vary independently. Verify that R spans $\mathfrak{h}^*$.

(3) For each $\alpha \in R$:
   (a) Verify that $\{n \in \mathbb{Z} : n\alpha \in R\} = \{\pm 1\}$.
   (b) Verify that $\dim(\mathfrak{g}^{\alpha}) = 1$.
   (c) Find an explicit basis element $X_\alpha \in \mathfrak{g}^{\alpha}$.
   (d) Show that there exists $Y_\alpha \in \mathfrak{g}^{-\alpha}$ so that the element $H_\alpha := [X_\alpha, Y_\alpha]$ of $\mathfrak{h}$ satisfies $\alpha(H_\alpha) = 2$; write down H_α explicitly.
   (e) Verify that for all $\alpha, \beta \in R$,

$$[\mathfrak{g}^{\alpha}, \mathfrak{g}^{\beta}] = \begin{cases} \mathfrak{g}^{\alpha+\beta} & \text{if } \alpha + \beta \in R \\ \mathbb{C}H_\alpha & \text{if } \alpha + \beta = 0 \\ 0 & \text{otherwise.} \end{cases}$$

   (f) (Optional) Show that the subspace $\mathbb{C}H_\alpha \oplus \mathbb{C}X_\alpha \oplus \mathbb{C}Y_\alpha$ of $\mathfrak{g}$ is a Lie subalgebra that is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.

(4) (Optional) Set $S := \{\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, \dots, \lambda_{n-1} - \lambda_n, 2\lambda_n\} \subseteq R$. Verify that S is a basis for $\mathfrak{h}^*$ and that every $\beta \in R$ may be written in the form $\beta = \sum_{\alpha \in S} m_\alpha \alpha$, where the m_α are integers which are either all ≥ 0 or all ≤ 0. Let C denote the set of all $\lambda \in \mathfrak{h}^*$ for which $\lambda(H_\alpha) \geq 0$ for all $\alpha \in S$. Verify that $C = \{l_1\lambda_1 + \cdots + l_n\lambda_n : l_1 \geq l_2 \geq \cdots \geq l_{n-1} \geq |l_n|\}$.

(5) By adapting the argument given in lecture for $\mathfrak{sl}_n(\mathbb{C})$, prove that $\mathfrak{g}$ is simple. [This is the only part of this homework that is not a straightforward computation. The same argument as in lecture shows that no nonzero ideal is contained in $\mathfrak{h}$. The key point is then to show that if an ideal contains one root space, then it contains every other root space. For this, one can certainly imitate the proof given in lecture, but it may be simpler to show that Lemma 202 applies also to $\mathfrak{sp}_{2n}(\mathbb{C})$ with $\lambda_{\max} := 2\lambda_1$; the rest of the proof then goes through unmodified.]

2.16. **11/22, 11/24: How to classify classical simple complex Lie algebras.** **Objectives.** You should be to classify classical simple complex Lie algebras by computing their Dynkin diagrams. You should be able to explain why this process is well-defined. You should develop some intuition for the following important concepts by reference to the classical examples: Cartan subalgebras, roots, simple roots, positive roots, Cartan matrix, root reflections, Weyl group, Weyl chambers and their relation with simple systems

Summary. (Tuesday.) We recalled the classical simple complex Lie algebras, explained (mechanically) how to compute their Dynkin diagrams (by reference to the handout copied in §29.4), and started explaining why the process of doing so is well-defined (i.e., depends only upon the isomorphism class of the Lie algebra); namely, we briefly discussed why Cartan subalgebras of classical simple algebras are conjugate.

(Thursday.) We started explaining why the Cartan matrix (hence the Dynkin diagram) is independent of the choice of simple root system. We divided the proof into three parts:

(1) The definition of root reflections and the observation that each root reflection stabilizes the set of roots.
(2) The interpretation of root reflections as geometric reflections with respect to the "obvious" inner product on the real span of the roots.
(3) The claim that the Weyl group, i.e., the group generated by the root reflections, acts transitively on the sets of simple systems.

We explained how these points imply that the Cartan matrix is independent of the simple system.

We did not explain why points 1,2 hold (other than by brute-force inspection of the handout, §29.4); we will return to them later. We started explaining point 3 by introducing the real span of the roots, the regular subset of that span, the Weyl chambers, the dominant Weyl chamber for a given simple system. For each classical example, we described the root reflections, the Weyl groups, the regular elements, and the dominant Weyl chambers explicitly. For the low-dimensional families $\mathrm{B}_2 = \mathrm{C}_2$, A_1 and A_2, we drew pictures of the root systems, indicating the simple roots, the irregular hyperplanes, the Weyl chambers, the dominant Weyl chamber, etc. We observed that the number of Weyl chambers is the same as the order of the Weyl group in those examples. We stated the theorem that simple systems and Weyl chambers are in natural bijection with each other, equivariantly for the action of the Weyl group, and that the Weyl group acts simply transitively on the set of Weyl chambers. We described both directions of the bijection between simple systems and Weyl chambers, without yet proving that we get well-defined bijections in this way; we explicated the bijection in the case of A_2. We explained how these facts imply that the Weyl group acts simply-transitively on the set of simple systems.

Homework 16 (Due Nov 29)**.**

(1) (Optional, but highly recommended) Carefully study the computation of the Cartan matrices N on the handout, §29.4. Check that all entries are as they should be; report to me any errors that you find. (I worked them out by hand; it might be a healthy exercise to redo them.) Verify by inspection

that the formula

$$\alpha(H_\beta) = 2\frac{(\alpha,\beta)}{(\beta,\beta)}.$$

holds (see §29.9 for details).

(2) Let $X := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathfrak{g} := \mathfrak{sl}_2(\mathbb{C})$. Show that ad_X is not semisimple (i.e., diagonalizable). Show that $\mathbb{C}X$ is a maximal abelian subalgebra of $\mathfrak{g}$ (i.e., that if $\mathfrak{h}$ is an abelian subalgebra of $\mathfrak{g}$ that contains $\mathbb{C}X$, then $\mathfrak{h} = \mathbb{C}X$.

(3) Which of the classical simple complex Lie algebras have the property that there exists an element w of the Weyl group for which $w(\alpha) = -\alpha$ for all roots α?

(4) Let $\mathfrak{g}$ be a classical simple complex Lie algebra with Weyl group W. Equip $\mathfrak{g}$ with the scalar product $(,)$ as in lecture or as in §29.9.
   (a) Suppose $\mathfrak{g} = A_n$ or D_n. Verify (by inspecting the handout (§29.4), say) that $(\alpha,\alpha) = 2$ for all roots $\alpha \in R$. Verify that W acts transitively on R.
   (b) Suppose $\mathfrak{g} = B_n$ or C_n. Observe that (α,α) takes two distinct values as α traverses R; call α *long* or *short* according as it takes the larger or the smaller of these values.
      (i) Verify (by inspecting the handout (§29.4)) that W acts transitively on the set of long roots and also on the set of short roots.
      (ii) Verify that for each root α there is a root β so that $\gamma := \alpha + \beta$ is a root and $(\alpha,\alpha) \neq (\gamma,\gamma)$.

   Remark: we may explain next week how these observations yield new proofs of the simplicity of $\mathfrak{g}$, "simpler" than the proofs we gave earlier; the key point will be to explain why the Weyl group W permutes the root spaces belonging to any ideal.

(5) Let $\mathfrak{g}$ be a classical simple complex Lie algebra, and let notation be otherwise as usual; in particular, $S \subset R$ is the "standard" simple system, R^+ the associated set of positive roots, and $(,)$ the standard inner product.
   (a) Verify that if $\alpha, \beta \in S$ are distinct simple roots, then $(\alpha,\beta) \leq 0$, or equivalently, $\alpha(H_\beta), \beta(H_\alpha) \leq 0$. [The intention here is to observe that this holds in every example; we will explain "why" next week.]
   (b) Verify that if $\alpha \in S$ is a simple root, then the root reflection s_α stabilizes the set $R^+ - \{\alpha\}$ consisting of all positive roots other than α. [This can be verified by inspection; alternatively, use the previous part of this exercise and appeal to the definition of "simple system" and the observation that $\{n \in \mathbb{Z} : n\alpha \in R\} = \{\pm 1\}$.]
   (c) Let $\rho := (1/2)\sum_{\alpha \in R^+} \alpha \in \mathfrak{h}^*$ denote the half-sum of positive roots. Compute that

$$\rho = \begin{cases} \sum_{j=1}^{n} \frac{n+1-2j}{2}\lambda_j & \mathfrak{g} = A_{n-1} \\ \sum_{j=1}^{n}(n+\frac{1}{2}-j)\lambda_j & \mathfrak{g} = B_n \\ \sum_{j=1}^{n}(n+1-j)\lambda_j & \mathfrak{g} = C_n \\ \sum_{j=1}^{n}(n-j)\lambda_j & \mathfrak{g} = D_n. \end{cases} \tag{11}$$

   Show for each $\alpha \in S$ that

$$s_\alpha\rho = \rho - \alpha, \tag{12}$$

or equivalently, that

$$\rho(H_\alpha) = 1. \tag{13}$$

[The two assertions are obviously equivalent. The first assertion (12) can be deduced from the previous part. The second assertion (13) may be compared with (11).]

(d) Verify that if $\alpha \in R^+ - S$ is a positive root that is not simple, then there exist positive roots $\beta, \gamma \in R^+$ such that $\alpha = \beta + \gamma$. [This is unrelated to the previous parts of this exercise; the intention is to observe that it holds in all examples.]

(e) Let $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$ be a regular element. Verify that $\{w \in W : w(\lambda) = \lambda\} = \{1\}$. [This is again unrelated to the previous parts of this exercise; the intent is that you verify it using the explicit description of W.]

(6) (Optional) The exceptional isomorphisms between classical complex simple Lie algebras that we have not already seen are the following:

(a) $A_3 \cong D_3$, i.e., $\mathfrak{sl}_4(\mathbb{C}) \cong \mathfrak{so}_6(\mathbb{C})$. On the group level, denote by

$$V := \Lambda^2\mathbb{C}^4 = \oplus_{i<j}\mathbb{C}e_i \wedge e_j$$

the six-dimensional vector space given by the exterior square of $\mathbb{C}^4$. Equip it with the quadratic form $Q : V \to \mathbb{C}$ defined by requiring that $v \wedge v = Q(v)e_1 \wedge e_2 \wedge e_3 \wedge e_4$. Then (V, Q) is a quadratic space (see §18.5) and the natural map $\Lambda^2 : \mathrm{SL}_4(\mathbb{C}) \to \mathrm{GL}(V)$ given by $\Lambda^2(g)(v_1 \wedge v_2) := gv_1 \wedge gv_2$ defines a covering morphism $\Lambda^2 : \mathrm{SL}_4(\mathbb{C}) \to \mathrm{SO}(V) \cong \mathrm{SO}_6(\mathbb{C})$.

(b) $C_2 \cong B_2$, i.e., $\mathfrak{sp}_4(\mathbb{C}) \cong \mathfrak{so}_5(\mathbb{C})$. [Restrict the map $\mathrm{SL}_4(\mathbb{C}) \to \mathrm{SO}_6(\mathbb{C})$ to $\mathrm{Sp}_4(\mathbb{C})$; show that its image is the stabilizer of a one-dimensional subspace L of V, hence identifies with the orthogonal group of the orthogonal complement $L^\perp$. One can read off L from the definition of $\mathrm{Sp}_4(\mathbb{C})$.]

2.17. **11/29: 11/31: Why simple Lie algebras give rise to root systems.**
Objectives. You should be able to use $\mathfrak{sl}_2$-triples to explain why simple complex Lie algebras (which contain Cartan subalgebras satisfying the expected properties) give rise to root systems.

Summary. (Tuesday.)

(1) We finished the proof from last time that the Weyl group of a classical simple complex Lie algebra acts transitively on the simple systems; this was deduced by choosing an element that maximizes the inner product of a pair of elements of the corresponding Weyl chambers and verifying that the map from simple systems to Weyl chambers is one-to-one.
(2) We stated the definition of a Cartan subalgebra of a simple complex Lie algebra and the theorem concerning the existence and uniqueness of such subalgebras; we included in this theorem also
   - the existence of a real form of the Cartan subalgebra on which the roots are real-valued, and
   - the existence of a scalar product on the ambient Lie algebra (i.e., a non-degenerate symmetric bilinear form) whose restriction to the real form of the Cartan subalgebra is positive-definite.

   We observed that the conclusion of this theorem holds for the classical families "by inspection" and gave a reference for the general case; discussing its proof would take us too far afield from (what I think are) more interesting topics to present in a first course on Lie groups.
(3) We stated the definition of a (reduced) root system: it is a finite subset of a real inner product space that satisfies some axioms that we had observed empirically last week.
(4) We stated the theorem that simple complex Lie algebras (satisfying the conclusion of the "Cartan subalgebra theorem") give rise to root systems. Our aim next time is to prove this theorem.

(Thursday.) We explained in detail how $\mathfrak{sl}_2$-triples allow us to prove that simple Lie algebras possessing Cartan subalgebras give rise to root systems that satisfy the various properties we had observed for the classical families.

Homework 17 (Due Dec 6)**.** The main purposes of the lectures this week were to demystify last week's observations concerning root systems for the classical Lie algebras and to demonstrate the power of $\mathfrak{sl}_2(\mathbb{C})$ in proving results about more general Lie algebras. This week there is one multipart exercise whose purpose is to complement that discussion by showing how one might instead use $\mathrm{SL}_2(\mathbb{C})$ to establish such properties.

Thus, let $\mathfrak{g}$ be a simple Lie algebra over $\mathbb{C}$, let $\mathfrak{h} \leq \mathfrak{g}$ be a Cartan subalgebra with real form $\mathfrak{h}_{\mathbb{R}}$ so that the set R of roots of $\mathrm{ad} : \mathfrak{h} \to \mathrm{End}(\mathfrak{g})$ satisfies $R \subseteq \mathfrak{h}_{\mathbb{R}}^* = \mathrm{Hom}_{\mathbb{R}}(\mathfrak{h}_{\mathbb{R}}, \mathbb{R}) \cong \{\lambda \in \mathfrak{h}^* : \lambda(\mathfrak{h}_{\mathbb{R}}) \subseteq \mathbb{R}\}$. Fix $\alpha \in R$. Assume given nonzero elements $X_\alpha \in \mathfrak{g}^\alpha, Y_\alpha \in \mathfrak{g}^{-\alpha}, H_\alpha \in \mathfrak{h}_{\mathbb{R}}$ satisfying the relations indicated just before the statement of Lemma 240, so that the map

$$\phi_\alpha : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{s}_\alpha := \mathbb{C}X_\alpha \oplus \mathbb{C}H_\alpha \oplus \mathbb{C}Y_\alpha \subseteq \mathfrak{g}$$

given by $X, Y, H \mapsto X_\alpha, Y_\alpha, H_\alpha$ is an isomorphism of Lie algebras. Finally, fix a complex Lie group G so that $\mathfrak{g} = \mathrm{Lie}(G)$.

Our aim here is to give alternative proofs (using Lie group methods rather than Lie algebra methods) of the first part of Lemma 242. Thus, it would be best not to invoke the statement of that lemma in the arguments to follow.

(1) Show that there is a morphism of complex Lie groups $F_\alpha : \mathrm{SL}_2(\mathbb{C}) \to G$ so that $dF_\alpha = \phi_\alpha$.

For an element t of any complex vector space on which exp is defined (e.g., t can be a complex scalar or an element of a Lie algebra over $\mathbb{C}$), introduce the abbreviation $e(t) := \exp(2\pi i t)$.

(2) Show that $e(H) = 1$.
(3) Show that $e(H_\alpha) = 1$.
(4) Let $\beta \in R$. Show for $t \in \mathbb{C}$ and $v \in \mathfrak{g}^\beta$ that $\mathrm{Ad}(e(tH_\alpha))v = e(t\beta(H_\alpha))v$.
(5) Deduce that $e(\beta(H_\alpha)) = 1$, hence that $\beta(H_\alpha) \in \mathbb{Z}$.
(6) Set $w := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C})$. Verify that $w = e^X e^{-Y} e^X$ and that $\mathrm{Ad}(w)H = -H$.
(7) Set $w_\alpha := F_\alpha(w) \in G$. Show that $\mathrm{Ad}(w_\alpha)H_\alpha = -H_\alpha$ and that $\mathrm{Ad}(w_\alpha)^2 = 1$.
(8) Suppose $H \in \mathfrak{h}$ satisfies $\alpha(H) = 0$.
   (a) Show that $[x, H] = 0$ for all $x \in \mathfrak{s}_\alpha$.
   (b) Show that $\mathrm{Ad}(g)H = H$ for all g in the image of F_α. Deduce in particular that $\mathrm{Ad}(w_\alpha)H = H$.
(9) Deduce that $\mathrm{Ad}(w_\alpha)\mathfrak{h} = \mathfrak{h}$.
(10) Recall that for $\lambda \in \mathfrak{h}^*$, we set $s_\alpha\lambda := \lambda - \lambda(H_\alpha)\alpha$. Show that for all $H \in \mathfrak{h}$, one has
$$\lambda(\mathrm{Ad}(w_\alpha)H) = (s_\alpha\lambda)(H). \tag{14}$$
(11) For any $\lambda \in \mathfrak{h}^*$, set $\mathfrak{g}^\lambda := \{v \in \mathfrak{g} : [H, v] = \lambda(H)v \text{ for all } H \in \mathfrak{h}\}$. Show that $\mathrm{Ad}(w_\alpha)$ induces a well-defined isomorphism
$$\mathrm{Ad}(w_\alpha) : \mathfrak{g}^\lambda \to \mathfrak{g}^{s_\alpha(\lambda)}. \tag{15}$$
(12) Deduce in particular that if $\beta \in R$, then $s_\alpha(\beta) \in R$.

2.18. **12/6: Complex reductive vs. compact real. Objectives.** You should be able to explain the relationships between complex simple Lie algebras, complex semisimple Lie algebras, complex reductive Lie algebras, and compact real Lie algebras.

Summary.

(1) We explained that the association constructed in previous weeks from
 (a) simple complex Lie algebras to
 (b) irreducible reduced root systems to
 (c) connected Dynkin diagrams

 is bijective, and indicated briefly how the Serre relations explain this (without proving them).

(2) We defined semisimple and reductive Lie algebras and indicated how the above bijection generalizes to them.

(3) We defined compact real Lie algebras and explained why the Lie algebra of any compact real Lie group is compact; we stated the theorem that every compact real Lie algebra arises in this way.

(4) We stated the theorem that compact real Lie algebras and complex reductive Lie algebras are in natural bijection. We gave a proof using the unitary trick and the Serre relations.

2.19. **12/8: Compact Lie groups: center, fundamental group. Objectives.** You should be able to describe the center and fundamental group of a copmact Lie group in terms of its root/weight lattices and the kernel of the exponential map.

Summary. We defined the root, weight, coroot and coweight lattices of a semisimple Lie algebra as well as the lattices of integral elements attached to a compact Lie group. We explained how the center and fundamental group are described in terms of these, and illustrate with the examples of tori and SL_n.

Homework 18 (Due Dec 13)**.** Let K be a compact Lie group with finite center. Let other notation be as in §32. Thus $\mathfrak{h}$ be a Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$. Let R be the set of roots and $S \subseteq R$ a base. Let $\mathfrak{h}_{\mathbb{R}}^*$ denote the span of R and $\mathfrak{h}_{\mathbb{R}} := \{H \in \mathfrak{h} : \alpha(H) \in \mathbb{R} \text{ for all } \alpha \in R\}$, as usual.

(1) Verify that the Weyl group W (generated by the s_α for $\alpha \in R$, as usual) acts on the root and weight lattices. Verify that the transposes of elements of the Weyl group act on the coroot and coweight lattices. One can thus form the semidirect product $W \ltimes \mathbb{Z}R^\wedge$.

(2) Let $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}} = \{H \in \mathfrak{h}_{\mathbb{R}} : \alpha(H) \notin \mathbb{Z} \text{ for all } \alpha \in R\}$.
 (a) Verify that W acts on $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$.
 (b) Verify that $\mathbb{Z}R^\wedge$ acts on $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$ (by translation).
 (c) Verify that the actions of W and $\mathbb{Z}R^\wedge$ on $\mathfrak{h}_{\mathbb{R}}$ induce an action of their semidirect product $W \ltimes \mathbb{Z}R^\wedge$ on $\mathfrak{h}_{\mathbb{R}}$, preserving $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$. Let $T \leq \mathrm{GL}(\mathfrak{h}_{\mathbb{R}})$ denote the image of $W \ltimes \mathbb{Z}R^\wedge$.
 (d) Let $n \in \mathbb{Z}$, $\alpha \in R$. Let $s_{\alpha,n} : \mathfrak{h}_{\mathbb{R}} \to \mathfrak{h}_{\mathbb{R}}$ be the reflection in the hyperplane $\alpha(H) = n$, thus $s_{\alpha,n}(H) = H - (\alpha(H) - n)H_\alpha$. Show that $s_{\alpha,n}$ belongs to T.
 (e) Show that T is generated by the $s_{\alpha,n}$.

(3) Choose an enumeration $S = \{\alpha_1, \ldots, \alpha_l\}$.
 (a) Show that the $\mathbb{Z}$-span $\mathbb{Z}R$ is has basis S in the sense that $\mathbb{Z}R = \mathbb{Z}\alpha_1 \oplus \cdots \oplus \mathbb{Z}\alpha_l$.

(b) Show that there exist unique elements $\pi_1, \ldots, \pi_l \in \mathfrak{h}_{\mathbb{R}}$ so that $H_{\alpha_i}(\pi_j) = \delta_{ij}$ for all $i, j \in \{1..l\}$. (These are called the *fundamental weights*.)
(c) Show that $(\mathbb{Z}R^\wedge)^* = \mathbb{Z}\pi_1 \oplus \cdots \oplus \mathbb{Z}\pi_l$.
(d) Show that matrix (a_{ij}) for which $\alpha_i = \sum_j a_{ij}\pi_j$ is given by the Cartan matrix.
(e) (Optional) Compute $\pi_1, \ldots, \pi_l$ for the classical families. (This can be done by hand, or by inverting the Cartan matrix.)

2.20. **12/13 Maximal tori in compact Lie groups. Objectives.** You should be able to explain the role played by maximal tori in the study of compact connected Lie groups.

Summary. (Tuesday) See §33 for details. Throughout, let K be a compact connected Lie group.

(1) We defined tori, and characterized them as compact connected abelian Lie groups.
(2) We defined maximal tori in K and characterized them as closed connected subgroups whose Lie algebras are maximal abelian subalgebras.
(3) We recorded that closed connected abelian subgroups of K are tori.
(4) As an example of the latter, we gave the connected components of closures of subgroups generated by individual elements.
(5) We indicated why the Lie algebras of maximal tori give rise to Cartan subalgebras after taking complexifications.
(6) We stated (without proof yet) the big theorem that K is the union of the conjugates of any one of its maximal tori.
(7) We derived from this the consequence that the center of K is the intersection of all (maximal) tori and indicated briefly how this implies the description given last time of the center in terms of roots.

(Thursday) See §33 for details.

(1) $\mathrm{Aut}(T)$
(2) $N(T)_0$
(3) $\Lambda(f)$
(4) Proof.
(5) T_1, T_2
(6) $Z(T) = T$

3. Selected homework solutions

- Homework 2, 2d. Here is a quick and fairly intuitive way to see that $G := \mathrm{SO}(1,2)$ has two connected components. (We've seen later in the course that this is established more efficiently using the Cartan decomposition.) We realize G as a subgroup of $\mathrm{SL}_3(\mathbb{R})$ in the evident way. It contains the subgroups

$$H_1 := \left\{ \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{O}(1,1) \right\}$$

and

$$H_2 := \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SO}(2) \right\}.$$

Note that H_2 is connected, so $H_2 \subseteq G^0$.

Let V_1 be as in the homework problem. Let V_1^+ denote the connected component containing e_1, and V_1^- the other component. (Thus $V_1^+ = V_1^0$ in the notation of the homework problem.) We now make the following observations:

(1) Since $V_1^{\pm}$ is connected and G acts on $V_1 = \cup_{\pm} V_1^{\pm}$, for each $g \in G$ we have either $gV_1^+ \subseteq V_1^+$ or $gV_1^+ \subseteq V_1^-$. It follows that G permutes the two connected components $V_1^{\pm}$ of V_1, and so the subgroup $\{g \in G : gV_1^+ \subseteq V_1^+\}$ of G has index at most 2.

(2) There exist elements $g \in G$ which map V_1^+ to V_1^-, and vice-versa. For instance, one can take $g := \mathrm{diag}(-1,-1,1)$. The subgroup $\{g \in G : gV_1^+ \subseteq V_1^+\}$ of G thus has index exactly 2.

(3) Since V_1^+ is connected, we have $G^0 \subseteq \{g \in G : gV_1^+ \subseteq V_1^+\}$.

(4) Let $v \in V_1^+$. It is of the form $v = (x,y,z)$ with $x^2 - y^2 - z^2 = 1$. Choose an element $h_2 \in H_2$ so that $h_2 v = (x, r, 0)$, where $r = \sqrt{y^2 + z^2}$. By part 1b of the same homework, we can then find $h_1 \in H_1^0$ so that $h_2 v = h_1 e_1$. Consequently $v = g e_1$ where $g := h_2^{-1} h_1 \in G^0$.

(5) Now let $g \in G$ with $gV_1^+ \subseteq V_1^+$. Then $ge_1 \in V_1^+$. By what was shown in the previous item, we can find $g_0 \in G^0$ so that $ge_1 = g_0 e_1$, thus $g \in g_0 H_2 \subseteq G^0$. Thus, $G^0 \supseteq \{g \in G : gV_1^+ \subseteq V_1^+\}$.

(6) We have seen that $G^0 = \{g \in G : gV_1^+ \subseteq V_1^+\}$ and that $\{g \in G : gV_1^+ \subseteq V_1^+\}$ has index 2 in G. Therefore G has two connected components.

One can tidy this discussion up a bit with some lemmas from lecture.

- Homework 3, part 1. Let $G := \mathrm{SO}(2,1)$. It acts on $M := V_1 := \{(x,y,z) : z^2 - x^2 - y^2 = 1\}$ as well as its connected components $V_1^{\pm}$. Let us realize M as a space of row vectors, so that G acts on M by right matrix multiplication: $m \mapsto mg$. The group G then acts on smooth functions $f : M \to \mathbb{R}$ by the formula: for $g \in G$, $m \in M$,

$$gf(m) := f(mg).$$

We saw in an earlier homework problem that G acts transitively on the connected components of M, hence f is constant on each such component if and only if

$$gf = f \text{ for all } g \in G. \tag{16}$$

Set $\mathfrak{g} := \mathrm{Lie}(G)$. We may differentiate the action of G on $C^\infty(M)$ to the action of $\mathfrak{g}$ on $C^\infty(M)$ given for $X \in \mathfrak{g}$ by

$$Xf(m) := \partial_{t=0} f(m \exp(tX)).$$

Explicitly, the Lie algebra $\mathfrak{g}$ consists of matrices $X \in M_3(\mathbb{R})$ satisfying ${}^tXJ + JX = 0$, where $J := \mathrm{diag}(1, 1, -1)$; an explicit basis for $\mathfrak{g}$ is given by the matrices

$$X_1 := \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad X_2 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, X_3 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Using that

$$\begin{aligned} (x, y, z)X_1 &= xe_2 - ye_1, \\ (x, y, z)X_2 &= ze_1 + xe_3, \\ (x, y, z)X_3 &= ze_2 + ye_3, \end{aligned}$$

we obtain

$$\begin{aligned} X_1 f(x, y, z) &= (x\partial_y - y\partial_x) f(x, y, z), \\ X_2 f(x, y, z) &= (z\partial_x + x\partial_z) f(x, y, z), \\ X_3 f(x, y, z) &= (z\partial_y + y\partial_z) f(x, y, z). \end{aligned}$$

Therefore assertion (b) in the homework problem is equivalent to saying that $X_i f = 0$ for $i = 1, 2, 3$, or equivalently, that

$$Xf = 0 \text{ for all } X \in \mathfrak{g}. \tag{17}$$

It remains only to verify that (16) and (17) are equivalent. This follows from the connectedness of G by the same argument (the "exponentiation/differentiation trick") as in §13.4.

- Homework 6, part 1. We want to show for $f : G \to H$ that $df(\mathrm{Ad}(g)X) = \mathrm{Ad}(f(g))df(X)$. The curve $t \mapsto \exp(t\,\mathrm{Ad}(g)X) = \exp(\mathrm{Ad}(g)tX)$ in G has initial velocity $\mathrm{Ad}(g)X$, hence

$$df(\mathrm{Ad}(g)X) = \partial_{t=0} f(\exp(\mathrm{Ad}(g)tX)). \tag{18}$$

In particular,

$$df(X) = \partial_{t=0} f(\exp(tX)). \tag{19}$$

By Exercise 22, we have $\exp(\mathrm{Ad}(g)tX)) = g\exp(tX)g^{-1}$; since f is a group homomorphism, it follows that

$$df(\mathrm{Ad}(g)X) = \partial_{t=0} f(g)gf(\exp(tX))f(g)^{-1} = \mathrm{Ad}(f(g))\partial_{t=0} f(\exp(tX)) \tag{20}$$

Combining (19) with (20) gives the required identity.

- Homework 6, part 2. We want to determine the complexifications of $\mathrm{SL}_n(\mathbb{H}), \mathrm{SU}(p, q)$, and $\mathrm{U}_m(\mathbb{H})$. (!!! to be written)
- Homework 8, part 3. The answer is: those functions $\nu : \mathbb{Z} \to \mathbb{Z}_{\geq 0}$ for which
  (1) $\nu(n) = 0$ for all but finitely many n,
  (2) $\nu(n) = \nu(-n)$ for all $n \in \mathbb{Z}_{\geq 0}$,
  (3) $\nu(0) \geq \nu(2) \geq \nu(4) \geq \nu(6) \geq \cdots$, and
  (4) $\nu(1) \geq \nu(3) \geq \nu(5) \geq \nu(7) \geq \cdots$.

Given such a seuqence, we can define $\mu : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ by $\mu(m) := \nu(m) - \nu(m+2)$ and set

$$V := \oplus_{m \in \mathbb{Z}_{\geq 0}} W_m^{\oplus \mu(m)}.$$

Conversely, the claimed inequalities are clearly satisfied for V of this form (by the hint suggested in the homework problem, or by directly writing out the characters).

- Homework 10, part 1: It is easy to see (by considering elementary matrices) that the center of $\mathrm{SL}_n(\mathbb{C})$ is the subgroup μ_n of scalar matrices $\mathrm{diag}(z, z, \dots, z)$ whose entries z are nth roots of unity. We have $\mathbb{Z}/n \cong \mu^n$ via the map $x \mapsto e^{2\pi i x}$.

  The inclusion $\mathrm{SL}_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ has differential given by the inclusion $\mathfrak{sl}_n(\mathbb{C}) \to \mathfrak{gl}_n(\mathbb{C})$ from the space of traceless matrices to the space of all matrices. The group $\mathrm{PGL}_n(\mathbb{C})$ is the quotient of $\mathrm{GL}_n(\mathbb{C})$ by the normal subgroup Z of scalar matrices $\mathrm{diag}(z, z, \dots, z)$ $(z \in \mathbb{C}^\times)$. Thus (by some theorem from lecture) the surjective quotient map $\mathrm{GL}_n(\mathbb{C}) \to \mathrm{PGL}_n(\mathbb{C})$ has differential given by the surjective linear map

$$\mathfrak{gl}_n(\mathbb{C}) \to \mathfrak{pgl}_n(\mathbb{C}),$$

  where $\mathfrak{pgl}_n(\mathbb{C})$ denotes the quotient of $\mathfrak{gl}_n(\mathbb{C})$ by the subgroup of diagonal scalar matrices of the form $\mathrm{diag}(Z, Z, \dots, Z)$ $(Z \in \mathbb{C})$. The composite map

$$\mathfrak{sl}_n(\mathbb{C}) \to \mathfrak{gl}_n(\mathbb{C}) \twoheadrightarrow \mathfrak{pgl}_n(\mathbb{C})$$

  is an isomorphism (indeed, one may invert it by sending an element of $\mathfrak{pgl}_n(\mathbb{C})$ to its unique traceless representative) so we may identify $\mathfrak{pgl}_n(\mathbb{C}) = \mathfrak{sl}_n(\mathbb{C})$. The map $p : \mathrm{SL}_n(\mathbb{C}) \to \mathrm{PGL}_n(\mathbb{C})$ is a morphism between connected Lie groups whose differential dp is then the "identity map" on $\mathfrak{sl}_n(\mathbb{C})$; in particular, dp is an isomorphism. Thus p is a covering morphism. We have seen that $\mathrm{SL}_n(\mathbb{C})$ is simply-connected. By the homotopy exact sequence (or the uniqueness of the discrete central subgroup "N" appearing in the theorem on the universal covering group), it follows that $\pi_1(\mathrm{PGL}_n(\mathbb{C})) \cong \ker(p) = \mu_n \cong \mathbb{Z}/n$, as required.

  An identical argument works for $G = \mathrm{SU}(n)$.

  The connected Lie groups having Lie algebra isomorphic to $\mathfrak{sl}_n(\mathbb{C})$ are in bijection with the discrete central subgroups of the simply-connected Lie group $\mathrm{SL}_n(\mathbb{C})$ having that Lie algebra; since the center of that group is $\mu_n \cong \mathbb{Z}/n\mathbb{Z}$ and since all subgroups of the latter are uniquely of the form $d\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/(n/d)\mathbb{Z}$ for some positive divisor d of n, we obtain a bijection between the isomorphism classes of such Lie groups G and the positive divisors d of n, where $\pi_1(G) \cong \mathbb{Z}/(n/d)\mathbb{Z}$.
- Homework 12, 1a. Let $G \leq \mathrm{SL}_n(\mathbb{R})$ be the group of unipotent upper-triangular matrices; for $n = 3$, one has

$$G = \begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix}.$$

  For subgroups A, B of G, let (A, B) denote the subgroup of G generated by all commutators $(a, b) := aba^{-1}b^{-1}$ with $a \in A, b \in B$. In the special case that B is a normal subgroup of A, we may interpret $B/(A, B)$ as the *maximal quotient of B on which A acts trivially by conjugation*: if $\phi : B \to$

K is a surjective group homomorphism with the property that $\phi(aba^{-1}) = \phi(b)$ for all $a \in A, b \in B$, then ϕ factors uniquely as a composition $B \twoheadrightarrow B/(A,B) \xrightarrow{\psi} K$.

For $k \in \mathbb{Z}_{\geq 1}$, define G_k inductively by $G_1 := G$ and $G_{k+1} := (G, G_k)$. The problem is to show that $G_n = \{1\}$.

We will establish the stronger assertion that $G_k = U_k$, where

$$U_k := \{a \in G : a_{ij} = 0 \text{ if } i < j < i + k\}.$$

For example, if $n = 4$, then

$$U_1 = \begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & 1 & * \\ & & & 1 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 1 & 0 & * & * \\ & 1 & 0 & * \\ & & 1 & 0 \\ & & & 1 \end{pmatrix},$$

$$U_3 = \begin{pmatrix} 1 & 0 & 0 & * \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}, \quad U_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix} = \{1\}.$$

For notational convenience, set $U_m := \{1\}$ if $m \geq n$.

To show that $G_k = U_k$, it suffices (since $G_1 = U_1$ and $G_{k+1} = (G, G_k)$) to show that

$$(U, U_k) = U_{k+1}. \tag{21}$$

For $i, j \in \{1..n\}$ with $i < j$, let $E_{ij} \in G$ denote the "elementary matrix" that has a 1 in the (i, j)th entry and vanishes on all other off-diagonal entries. For example, if $n = 3$, then

$$E_{12} = \begin{pmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}, \quad E_{13} = \begin{pmatrix} 1 & 0 & 1 \\ & 1 & 0 \\ & & 1 \end{pmatrix}, \quad E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ & 1 & 1 \\ & & 1 \end{pmatrix}.$$

By matrix multiplication, one has the commutation relations

$$[E_{ij}, E_{jk}] = E_{ik}, \quad [E_{ij}, E_{kl}] = 0 \text{ if } j \neq k.$$

By Gaussian elimination, G is generated by the elements E_{ij} taken over all $i < j$. Similarly, U_k is generated by those E_{ij} for which $j \geq i + k$. It follows from this observation and the commutation relations that

(i) U_k is normal in U_1 for all k,
(ii) the conjugation action of U_1 on the quotient U_k/U_{k+1} is trivial, and
(iii) $(U_1, U_k) \geq U_{k+1}$ for all k.

On the other hand, $U_k/(U_1, U_k)$ is the *maximal* quotient of U_k on which U_1 acts trivially by conjugation, hence $(U_1, U_k) \leq U_{k+1}$ and therefore $(U_1, U_k) = U_{k+1}$. This completes the proof.

Remark: One can alternatively argue using the matrix logarithm/exponential. The series defining the logarithm converges everywhere on G because $g - 1$ is nilpotent for $g \in G$, hence the series is actually finite; similarly, the exponential series $\mathfrak{g} \to G$ is actually a polynomial.

Remark: One can formulate the definition of U_k more geometrically in terms of the standard complete flag $\mathbb{R}^n = V_0 \supset V_1 \supset \cdots \supset V_n = \{0\}$, where V_k denotes the span of the first k standard basis vectors $e_1, \ldots, e_k$.

Then $U_k = \{g \in \mathrm{GL}_n(\mathbb{R}) : (g-1)V_i \subseteq V_{i+k} \text{ for all } i\}$, where $V_m := \{0\}$ for $m \geq n$.

- Homework 12, 2. In what follows, x, y denote small enough elements of $\mathfrak{g}$, while g denotes an element of the group G.
  (1) Recall (from Exercise 22) the general identity $\exp(\mathrm{Ad}(g)x) = g\exp(x)g^{-1}$. Recall also (from §18.4) that $\mathrm{Ad}(\exp(x)) = \exp(\mathrm{ad}_x)$. From these identities it follows that
  $$\exp(x * y * (-x)) = \exp(x)\exp(y)\exp(-x) = \exp(\mathrm{Ad}(\exp(x))y) \tag{22}$$
  and thus
  $$x * y * (-x) = \mathrm{Ad}(\exp(x))y = \exp(\mathrm{ad}_x)y, \tag{23}$$
  as required.
  (2) Set $f(t) := x * ty$. By BCH, f is analytic near 0. We have $f(0) = x$. We have
  $$\exp(f(t)) = \exp(x)\exp(ty) \tag{24}$$
  and thus
  $$\exp(-f(0))\partial_{t=0}\exp(f(t)) = y. \tag{25}$$
  On the other hand, by Homework 9,
  $$\exp(-f(0))\partial_{t=0}\exp(f(t)) = \Psi(\mathrm{ad}_x)f'(0) \tag{26}$$
  where
  $$\Psi(z) = \sum_{n=1}^{\infty}\frac{(-z)^{n-1}}{n!} = \frac{1-\exp(-z)}{z}. \tag{27}$$
  It follows that $f'(0) = \Psi(\mathrm{ad}_x)^{-1}y$, or more verbosely, that
  $$f'(0) = \frac{\mathrm{ad}_x}{1-\exp(-\mathrm{ad}_x)}y = \sum_{n\geq 0} c_n \,\mathrm{ad}_x^n\, y = y + \frac{[x,y]}{2} + \cdots \tag{28}$$
  for some explicit coefficients c_n (Bernoulli numbers). Since f is analytic, we deduce (upon setting $t := 1$, taking x, y small enough and appealing to Taylor's theorem) that
  $$x * y = x + \frac{\mathrm{ad}_x}{1-\exp(-\mathrm{ad}_x)}y + O(|y|^2). \tag{29}$$
  (3) Take $f(t) := x * (ty - x)$. Then $\exp(f(t)) = \exp(x)\exp(ty - x)$ and $f(0) = 0$; we want to compute $f'(0)$. To that end, we compute the quantity
  $$Q := \exp(-f(0))\partial_{t=0}\exp(f(t)) = \partial_{t=0}\exp(x)\exp(ty - x). \tag{30}$$
  in two ways. First, by the rearrangement
  $$Q = \exp(x)\partial_{t=0}\exp(ty - x) \tag{31}$$
  and the formula (86), we have
  $$Q = \Psi(-\mathrm{ad}_x)y \tag{32}$$
  with Ψ as above. On the other hand, by direct application of (26) (which remains valid in this context),
  $$Q = \Psi(\mathrm{ad}_{f(0)})f'(0) = f'(0), \tag{33}$$

since $f(0) = 0$. Therefore

$$f'(0) = \Psi(-\operatorname{ad}_x)y = \frac{1-\exp(\operatorname{ad}_x)}{-\operatorname{ad}_x}y = \frac{\exp(\operatorname{ad}_x)-1}{\operatorname{ad}_x}y = y + \frac{[x,y]}{2} + \cdots . \quad (34)$$

The asymptotic formula

$$x * (y - x) = \frac{\exp(\operatorname{ad}_x)-1}{\operatorname{ad}_x}y + O(|y|)^2 \quad (35)$$

for small enough x, y follows as in the previous part of the problem.

(4) In particular, since $|\operatorname{ad}_x^n y| = O(|x|^n|y|)$, we see by Taylor's theorem that

$$\frac{x}{2} * y * \frac{-x}{2} = x + \frac{[x,y]}{2} + O(|x|^2|y| + |y|^2), \quad (36)$$

$$x * (y - x) = x + \frac{[x,y]}{2} + O(|x|^2|y| + |y|^2). \quad (37)$$

4. Some notation

4.1. Local maps.

4.1.1. *Motivation.* We shall often have occasion to consider continuous maps defined on open subsets of topological spaces for which the precise choice of domain is unimportant (other than, perhaps, in that it contains a specific point). This circumstance motivates introducing the notation and terminology to follow.

4.1.2. *Definition.* Let X, Y be topological spaces. By a *map f from X to Y*, denoted $f : X \to Y$, we shall mean a continuous function. By a *local map f from X to Y*, denoted[1]

$$f : X \dashrightarrow Y,$$

we shall mean more precisely a pair (U, f), where:

- U is an open subset of X, called the *domain of definition* or simply the *domain* of f and denoted $U =: \operatorname{dom}(f)$, and
- $f : U \to Y$ is continuous.

We say that $f : X \dashrightarrow Y$ is *defined* at a point $p \in X$ if p belongs to the domain of f.

Example 1. Let $X := Y := \mathbf{k}$. The pair (U, f), where $U := \mathbf{k}^\times$ and $f(x) := 1/x$, is a local map $f : \mathbf{k} \dashrightarrow \mathbf{k}$.

4.1.3. *Equivalence.* Two local maps $f_1, f_2 : X \dashrightarrow Y$ will be called *equivalent* if they coincide on the intersection of their domains of definition.

4.1.4. *Images.* Given a local map $f : X \dashrightarrow Y$ and a subset S of X, we denote by $f(S)$ the image of the intersection of S with the domain of f, i.e., if $U = \operatorname{dom}(f)$, then $f(S) := f(U \cap S)$.

4.1.5. *Composition.* Given local maps $f : X \dashrightarrow Y$ and $g : Y \dashrightarrow Z$, we may define their composition to be the local map

$$g \circ f : X \dashrightarrow Z$$

with $\operatorname{dom}(g \circ f) := \operatorname{dom}(f) \cap f^{-1}(\operatorname{dom}(g))$ given as usual by $(g \circ f)(x) := g(f(x))$. It can happen that $\operatorname{dom}(g \circ f) = \emptyset$.

[1] The notation is inspired by that used in algebraic geometry for rational maps.

4.1.6. *Inverses.* A local map $f : X \dashrightarrow Y$ with domain $U := \mathrm{dom}(f)$ will be called *invertible* if $f(U)$ is open and the induced map $f : U \to f(U)$ is a homeomorphism. In that case, the *inverse* of f is defined to be the local map

$$f^{-1} : Y \dashrightarrow X$$

with $\mathrm{dom}(f^{-1}) := f(U)$ given as usual by $f^{-1}(y) := x$ if $y = f(x)$.

5. Some review of calculus

To explain what will happen in this course, it will help to recall some background from calculus. Let $\mathbf{k}$ denote either of the fields $\mathbb{R}$ or $\mathbb{C}$.

5.1. **One-variable derivatives.** We say that $f : \mathbf{k} \dashrightarrow \mathbf{k}$ is *smooth* if all of its derivatives (of arbitrary order) exist at all points in the domain of definition. The first derivative $f' : \mathbf{k} \dashrightarrow \mathbf{k}$ may be characterized by the relation

$$f(p + v) = f(p) + f'(p)v + o(|v|)$$

holding for each fixed point p as $|v| \to 0$. In a single-variable calculus course, one learns to relate (apparently complicated) *global* properties of a function to simpler *local* (or *infinitesimal*) properties involving its derivatives. For example, when $\mathbf{k} = \mathbb{R}$, one learns that the following are equivalent:

- f is increasing (an ostensibly *global* property, as it requires one to check that $f(x) < f(y)$ for every pair of points with $x < y$, and such points might be quite far apart!);
- f' is positive (a *local* property, as it only requires one to check that $f'(x) > 0$ for each point x).

This test and others (concerning the second derivative, for instance) often suffice to piece together an approximate portrait of the global shape of a function from the local behavior of its derivatives at the critical points.

5.2. **Multi-variable total and partial derivatives.** We say that $f : \mathbf{k}^m \dashrightarrow \mathbf{k}^n$ is *smooth* if all of its partial derivatives exist at all points in the domain of definition. For each p in the domain of f, the total derivative $T_p f$ is a *linear* map $T_p f : \mathbf{k}^m \to \mathbf{k}^n$ that may be characterized as above by the relation

$$f(p + v) = f(p) + (T_p f)(v) + o(|v|)$$

holding for each fixed p as $|v| \to 0$. One can express $T_p f$ in matrix form

$$T_p f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_m}(p) \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1}(p) & \cdots & \frac{\partial f_n}{\partial x_m}(p) \end{pmatrix}$$

where the function f is expressed as a tuple $f = (f_1, \dots, f_n)$ of components $f_i : \mathbf{k}^m \dashrightarrow \mathbf{k}$, elements $x \in \mathbf{k}^m$ are equipped with the standard coordinates $x = (x_1, \dots, x_n)$, and the partial derivatives are characterized by

$$f_i(p + te_j) = f_i(p) + \frac{\partial f_i}{\partial x_j}(p)t + o(|t|)$$

for $t \in \mathbf{k}$ with $|t| \to 0$, where e_j denotes the standard jth basis element of $\mathbf{k}^m$ dual to the coordinate x_j; for a vector $v = (v_1, \dots, v_m) \in \mathbf{k}^m$ and a coordinate index

$i = 1, \dots, n$, one then has

$$((T_p f)v)_i = \sum_{j=1}^{m} \frac{\partial f_i}{\partial x_j}(p) v_j.$$

As in single-variable calculus, one learns various ways to relate the global behavior of a f to the local behavior of its total derivative $T_p f$ at various critical points p (Hessian test, etc). A basic case to keep in mind is that of a linear function $A : \mathbf{k}^m \to \mathbf{k}^n$, for which one has $T_p A = A$ for all $p \in \mathbf{k}^m$.

When $m = 1$, so that $f = (f_1, \dots, f_n)$ may be thought of as an n-tuple of functions $f_i : \mathbf{k} \to \mathbf{k}$, the total derivative $T_p f$ is then a linear transformation $\mathbf{k} \to \mathbf{k}^n$ and so may be identified with the vector

$$f'(p) = (f_1'(p), \dots, f_n'(p)) \in \mathbf{k}^n$$

characterized by the same relation as in the one-variable case.

5.3. **The chain rule.** Given a pair of smooth functions $f : \mathbf{k}^m \dashrightarrow \mathbf{k}^n$ and $g : \mathbf{k}^n \dashrightarrow \mathbf{k}^l$, one can form the composition $h := g \circ f : \mathbf{k}^m \dashrightarrow \mathbf{k}^l$, which is smooth. One knows the chain rule: at a point $p \in \mathbf{k}^m$ at which h is defined, the derivative of the composition h is the linear map $T_p h : \mathbf{k}^m \to \mathbf{k}^l$ given by the composition

$$T_p h = T_{f(p)} g \circ T_p f$$

of the derivatives of f, g. Expanding out in terms of matrices and using standard coordinates $x_1, \dots, x_m$ on $\mathbf{k}^m$ and $y_1, \dots, y_n$ on $\mathbf{k}^n$, the chain rule reads: for $i \in \{1..l\}$ and $k \in \{1..m\}$,

$$\frac{\partial h_i}{\partial x_k}(p) = \sum_{j=1}^{n} \frac{\partial g_i}{\partial y_j}(f(p)) \frac{\partial f_j}{\partial x_k}(p).$$

Specialized to the case $m = 1$, the chain rule reads $h'(p) = (T_{f(p)} g) f'(p)$, or expanding out further in terms in terms of components, as $h_i'(p) = \sum_{j=1}^{n} \frac{\partial h_i}{\partial y_j}(f(p)) f_j(p)$.

5.4. **Inverse function theorem.** The *inverse function theorem* is a fundamental tool for controlling the local behavior of a function f *near* a point p in terms of *linear algebraic* properties of the derivative $T_p f$.

Theorem 2. *Let $f : \mathbf{k}^n \dashrightarrow \mathbf{k}^n$ be smooth and defined at p. The following are equivalent:*

(1) *The map $T_p f : \mathbf{k}^n \to \mathbf{k}^n$ is a linear isomorphism of vector spaces, or equivalently, has nonzero Jacobian determinant* $\det(T_p f)$.
(2) *There is an open neighborhood U of p, contained in the domain of f, so that $V := f(U)$ is open and the induced map $f : U \to V$ is a diffeomorphism, i.e., admits a smooth two-sided inverse $g : V \to U$.*

5.5. **Implicit function theorem.** We state it here rather verbosely, saving a tidier formulation for the generalization to manifolds given below in §6.8.

Theorem 3. *Let n, m, d be nonnegative integers with $n = m + d$. Suppose given a point $p \in \mathbf{k}^n$ and smooth maps $f_1, \dots, f_m : \mathbf{k}^n \dashrightarrow \mathbf{k}$ defined at $p = (p_1, \dots, p_n)$ satisfying either of the following evidently equivalent properties:*

(1) *The $m \times n$ matrix of partial derivatives $\frac{\partial f_i}{\partial x_j}(p)$ ($i = 1..m, j = 1..n$) has rank m.*

(2) *The function* $f := (f_1, \dots, f_m) : \mathbf{k}^n \dashrightarrow \mathbf{k}^m$ *has the property that its total derivative* $T_p f : \mathbf{k}^n \to \mathbf{k}^m$ *at* p *is surjective.*
(3) $\dim \ker(T_p f) = d$.
(4) *The space of solutions* $(dx_1, \dots, dx_n) \in \mathbf{k}^n$ *to the system of homogeneous linear equations* $\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(p) dx_j = 0$ $(i = 1, \dots, m)$ *is* d*-dimensional.*

Then after suitably reordering the coordinates indices, one can find smooth functions $\psi_{d+1}, \dots, \psi_n : \mathbf{k}^n \dashrightarrow \mathbf{k}$ *defined at* $(p_1, \dots, p_d, f_1(p), \dots, f_m(p))$ *so for any point* $x = (x_1, \dots, x_n)$ *close enough to* p*, one has*

$$x_j = \psi_j(x_1, \dots, x_d, f_1(x), \dots, f_m(x)) \quad \text{for } j = d+1, \dots, n.$$

In particular, if we suppose moreover that $f_1(p) = \dots = f_m(p) = 0$*, then the following are equivalent for* x *close enough to* p*:*
(1) $f_1(x) = \dots = f_m(x) = 0$
(2) $x_j = \psi_j(x_1, \dots, x_d, 0, \dots, 0)$ *for* $j = d+1, \dots, n$.

Consequently the map

$$\Psi : \mathbf{k}^d \dashrightarrow \mathbf{k}^n$$

$$\Psi(x_1, \dots, x_d) := (x_1, \dots, x_d, \psi_{d+1}(x_1, \dots, x_d, 0, \dots, 0), \dots, \psi_n(x_1, \dots, x_d, 0, \dots, 0))$$

parametrizes the set $\{x : f_1(x) = \dots = f_m(x) = 0\}$ *near* p.

Proof. Suppose after suitably relabeling indices that the rightmost $m \times m$ minor of the matrix $T_p f$ is nonsingular. Consider then the map $\phi : M \dashrightarrow \mathbf{k}^n$ defined near p by the formula $\phi(x) := (x_1, \dots, x_d, f_1(x), \dots, f_m(x))$. The total derivative $T_p \phi$ is then given by the matrix

$$\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\
\frac{\partial f_1}{\partial x_1}(p) & \frac{\partial f_1}{\partial x_2}(p) & \cdots & \frac{\partial f_1}{\partial x_d}(p) & \frac{\partial f_1}{\partial x_{d+1}}(p) & \cdots & \frac{\partial f_1}{\partial x_n}(p) \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial f_m}{\partial x_1}(p) & \frac{\partial f_m}{\partial x_2}(p) & \cdots & \frac{\partial f_m}{\partial x_d}(p) & \frac{\partial f_m}{\partial x_{d+1}}(p) & \cdots & \frac{\partial f_m}{\partial x_n}(p)
\end{pmatrix}$$

By our assumption on $T_p f$, it follows that $T_p \phi$ is nonsingular. By the inverse function theorem, we can find a local inverse $\psi = (\psi_1, \dots, \psi_n) : \mathbf{k}^n \dashrightarrow \mathbf{k}^n$ to ϕ defined at $\phi(p)$. By construction, $\psi_i(x) = x_i$ for $i = 1, \dots, d$, while the ψ_i for $i > d$ satisfy the requirements of the conclusion. □

6. Some review of differential geometry

As indicated already in §1, this section is intended to be used mainly as a reference. I don't plan to use much differential geometry in the actual course.

6.1. **Charts.** Recall the notation and terminology of §4.1. By a *topological chart* on a topological space X we shall mean a pair (ϕ, n), where n is a nonnegative integer and

$$\phi : X \dashrightarrow \mathbf{k}^n$$

is an invertible local map. More verbosely, this means that a topological chart is a triple (U, ϕ, n), where n is a nonnegative integer, U is an open subset of X, and $\phi : U \to \mathbf{k}^n$ is a continuous map for which $\phi(U)$ is open and $\phi : U \to \phi(U)$ is a homeomorphism onto its image.

6.2. **Manifolds.** For us, an *n-manifold* (over the field $\mathbf{k}$) is a triple $(M, n, \mathcal{A})$, often abbreviated simply as M when the data $n, \mathcal{A}$ are understood by context, where:

(1) M is a topological space (which we assume to be Hausdorff and second-countable, hence metrizable),
(2) n is a nonnegative integer,
(3) $\mathcal{A}$ is an *maximal smooth atlas*, that is to say, a collection of topological charts $\phi : M \dashrightarrow \mathbf{k}^n$ whose domains cover M and which are *smoothly compatible* in the sense that for each pair ϕ, ψ of charts in $\mathcal{A}$, the compositions
$$\phi \circ \psi^{-1}, \psi \circ \phi^{-1} : \mathbf{k}^n \dashrightarrow \mathbf{k}^n$$
are smooth on their respective domains of definition. "Maximal" means that $\mathcal{A}$ contains every chart on M that is compatible with every chart in $\mathcal{A}$. By a *smooth chart* on M we then mean an element of $\mathcal{A}$.

A *manifold* is an n-manifold M for some n, which is called the *dimension* of M and denoted $n = \dim(M)$.

Example 4. An open subset subset M of $\mathbf{k}^n$ is a n-manifold if we take for $\mathcal{A}$ the set of all charts $\phi : M \dashrightarrow \mathbf{k}^n$ which are smooth and have smooth inverses in the ordinary sense of §5.2. In particular, $\mathbf{k}^n$ is an n-manifold. More generally, any open subset U of an n-manifold M has the natural structure of an n-manifold: if $\mathcal{A}$ is a maximal smooth atlas on M, then $\{\phi \in \mathcal{A} : \mathrm{dom}(\phi) \subseteq U\}$ is a maximal smooth atlas on U.

Remark 5.

(1) One could also work with manifolds having multiple components of varying dimension, but we will not have occasion to do so.
(2) One can show that the dimension of a manifold is determined by its topological structure, making the inclusion of n in the definition redundant, but this fact is not important for our purposes.

6.3. **Smooth maps.** Given manifolds M, N, a local map $f : M \dashrightarrow \mathbf{k}$ is *smooth* if it is smooth with respect to every pair of charts $\phi : M \dashrightarrow \mathbf{k}^m$, $\psi : N \dashrightarrow \mathbf{k}^n$ in the sense that the composition
$$\mathbf{k}^m \overset{\phi^{-1}}{\dashrightarrow} M \overset{f}{\dashrightarrow} N \overset{\psi}{\dashrightarrow} \mathbf{k}^n$$
is smooth in the sense of §5.2. When M, N are open subsets of Euclidean space, this notion generalizes the earlier one.

6.4. **Coordinate systems.** Let M be an m-manifold and let $p \in M$ be a point. By a *coordinate system for M at p*, or more simply a *coordinate system at p* or *local coordinates at p* when the ambient manifold M is clear from context, we shall mean a tuple of smooth maps $x_1, \dots, x_m : M \dashrightarrow \mathbf{k}$ arising as the components of a smooth chart $\phi = (x_1, \dots, x_m) : M \dashrightarrow \mathbf{k}^m$ on M defined at p. Such a coordinate system allows us to identify points $x \in M$ near p with tuples $x = (x_1, \dots, x_m)$ and smooth functions $f : M \dashrightarrow \mathbf{k}$ defined near p with smooth functions $f(x_1, \dots, x_m)$ of the local coordinates $x_1, \dots, x_m$. In particular, it makes sense to define partial derivatives $\partial f / \partial x_j \in \mathbf{k}$ in this setting. For example, if M is an open subset of $\mathbf{k}^m$, then the standard coordinate functions $x_1, \dots, x_m : M \to \mathbf{k}$ form a coordinate system at every point $p \in M$.

Given an m-manifold M, an n-manifold N, a smooth map $f : M \dashrightarrow N$ and a point $p \in M$ at which f is defined, one can choose local coordinates $x_1, \dots, x_m$

at p and $y_1, \dots, y_n$ at $f(p)$. With respect to such coordinates, the smooth map f identifies with a smooth map $(f_1, \dots, f_n) : \mathbf{k}^m \dashrightarrow \mathbf{k}^n$ defined at $(x_1(p), \dots, x_m(p))$ and $T_p f$ identifies with the matrix of partial derivatives $\partial f_i / \partial x_j$ as in §5.2.

6.5. **Tangent spaces.** Let M be an open subset of $\mathbf{k}^m$. A *curve on* M is a smooth map $\gamma : \mathbf{k} \dashrightarrow M$ whose domain is connected and contains 0; the point $p := \gamma(0) \in M$ is referred to as its *basepoint* and the vector $v := \gamma'(0) \in \mathbf{k}^m$ as its *initial velocity.* Given an open subset $N \subseteq \mathbf{k}^n$ and a smooth map $f : M \to N$, the composition $f \circ \gamma : \mathbf{k} \dashrightarrow N$ is then a curve on N with basepoint $f(p)$; thanks to the chain rule, its initial velocity is the image $(f \circ \gamma)'(0) = (T_p f)v$ of that of the original curve γ under the derivative of the map f at the basepoint of the original curve.

Remark 6. Observe also that for each point $p \in M$ and vector $v \in \mathbf{k}^m$ there exists a curve γ with basepoint p and initial velocity v; for example, one can set $\gamma(t) := p + tv$ and take for the domain of γ any sufficiently small neighborhood of 0.

We now recall the generalization of the above notion to an arbitrary n-manifold M:

Definition 7. Let M be a manifold. A *curve on* M is defined as above to be a smooth map $\gamma : \mathbf{k} \dashrightarrow M$ whose domain is connected and contains 0.

We continue to refer to the point $p := \gamma(0) \in M$ as the *basepoint* of γ. If one now ponders how to define "the initial velocity $\gamma'(0)$" (e.g., to which space should it belong?), one is eventually led to introduce the *tangent space* $T_p M$ to the manifold M at the point p. To motivate the definition, note first that for any smooth chart $\phi : M \dashrightarrow \mathbf{k}^n$ defined at p, the composition $\phi \circ \gamma : \mathbf{k} \dashrightarrow \mathbf{k}^n$ is (after suitably shrinking its domain so as to be connected) a curve on $\mathbf{k}^n$ with basepoint $\phi(p)$ and initial velocity

$$v_\phi := (\phi \circ \gamma)'(0) \tag{38}$$

in the Euclidean sense defined previously. Moreover, for any other smooth chart $\psi : M \dashrightarrow \mathbf{k}^n$ defined at p, we can use the chain rule in the form

$$T_0(\psi \circ \gamma) = T_0(\psi \circ \phi^{-1} \circ \phi \circ \gamma) = T_{\phi(p)}(\psi \circ \phi^{-1}) \circ T_0(\phi \circ \gamma)$$

to read off the initial velocity v_ψ of the curve $\psi \circ \gamma$ from that of $\phi \circ \gamma$:

$$v_\psi = T_{\phi(p)}(\psi \circ \phi^{-1}) v_\phi. \tag{39}$$

This observation suggests the following definition:

Definition 8. The *tangent space* $T_p M$ *to* M *at* p is the set of all tuples $v = (v_\phi)_\phi$, where ϕ traverses the set of smooth charts defined at p and the v_ϕ are elements of $\mathbf{k}^n$ satisfying the consistency condition (39).

Since the consistency condition is linear, it is clear that $T_p M$ is a vector space. We can also now define, for each curve γ on M with basepoint $p = \gamma(0)$, the initial velocity of γ to be the vector $\gamma'(0) \in T_p M$ given by $v = (v_\phi)_\phi$, where the components v_ϕ are as in (38). We have moreover the following:

Lemma 9. For any smooth chart χ on M defined at p, the map $v \mapsto v_\chi$ defines a linear isomorphism $T_p M \cong \mathbf{k}^n$. In particular, $\dim T_p M = \dim M$.

Proof. The consistency condition (39) implies that v is determined by any one of its components v_χ, so the map in question is clearly injective. Conversely, given any element $u \in \mathbf{k}^n$, we may define a tuple $v = (v_\phi)_\phi$ by the rule $v_\phi := T_{\chi(p)}(\phi \circ \chi^{-1})u$. An application of the chain rule to the composition $(\psi \circ \phi^{-1}) \circ (\phi \circ \chi^{-1})$ implies for any smooth charts ϕ, ψ defined at p that (39) is satisfied, and so v belongs to T_pM. Since $v_\chi = u$ and u was arbitrary, we deduce that the map in question is surjective. $\square$

Example 10. If M is an open subset of $\mathbf{k}^n$, then the inclusion $\phi : M \to \mathbf{k}^n$ is a smooth chart defined at all points of M, and the lemma gives a natural identification $T_pM \cong \mathbf{k}^n$ for all $p \in M$. The two senses in which we have defined the initial velocity $\gamma'(0)$ of a curve γ on M with basepoint p (first as the ordinary derivative $\frac{d}{dt}\gamma(t)|_{t=0}$, then later as a tuple $(v_\phi)_\phi$) are compatible under this identification.

Remark 11. A more customary definition is that T_pM is the space of equivalence classes $[\gamma]$ of curves γ on M with basepoint p, with two curves γ_1, γ_2 declared equivalent precisely when $(\phi \circ \gamma_1)'(0) = (\phi \circ \gamma_2)'(0)$ for all charts ϕ on M defined at p. That definition is isomorphic to the one we've used. An isomorphism from the former to the latter may be given by $[\gamma] \mapsto (v_\phi)_\phi$ with v_ϕ as in (38); this map is well-defined by (39), injective by definition of the equivalence relation defining $[\gamma]$, and surjective by Lemma 9 and Remark 6.

6.6. **Derivatives.** Given a smooth map $f : M \dashrightarrow N$ of manifolds a point $p \in M$ at which f is defined, a chart $\phi : M \dashrightarrow \mathbf{k}^m$ at p and a chart $\psi : N \dashrightarrow \mathbf{k}^n$ at $f(p)$, we can form the smooth map

$$f_{\phi\psi} := \psi \circ f \circ \phi^{-1} : \mathbf{k}^m \dashrightarrow \mathbf{k}^n$$

and consider its total derivative in the sense of §5.2, which is a linear map

$$T_{\phi(p)}(f_{\phi\psi}) : \mathbf{k}^m \dashrightarrow \mathbf{k}^n.$$

We can piece these linear maps together to form a linear map

$$T_pf : T_pM \to T_{f(p)}N,$$

called the *derivative of* f, by setting, for $v = (v_\phi)_\phi \in T_pM$,

$$(T_pf(v))_\psi := T_{\phi(p)}(f_{\phi\psi})v_\phi. \tag{40}$$

An application of the chain rule confirms that the RHS of (40) is independent of ϕ and that the the tuple $T_pf(v)$ defined componentwise above actually belongs to $T_{f(p)}N$; moreover, for a composition $L \overset{g}{\dashrightarrow} M \overset{f}{\dashrightarrow} N$ defined at a point $p \in L$, one has again

$$T_p(f \circ g) = T_{g(p)}f \circ T_pg. \tag{41}$$

Remark 12.

(1) If $M \subseteq \mathbf{k}^m, N \subseteq \mathbf{k}^n$ are open and we identify $T_pM \cong \mathbf{k}^m, T_{f(p)}N \cong \mathbf{k}^n$, then the derivative $T_pf : T_pM \to T_{f(p)}N$ defined just now identifies the derivative $T_pf : \mathbf{k}^m \to \mathbf{k}^n$ as in §5.2.
(2) If M is an n-manifold, $p \in M$ is a point, and ϕ is a chart on M at p, then the map $T_\phi : T_pM \to T_{\phi(p)}\mathbf{k}^n \cong \mathbf{k}^n$ is just the projection $v \mapsto v_\phi$.

(3) If we had instead defined tangent spaces using equivalence classes of smooth curves as in Remark 11, then the definition of the derivative of f would look like: for a curve γ on M with basepoint p and equivalence class $[\gamma]$,

$$T_p f([\gamma]) := [f \circ \gamma].$$

As noted earlier in the Euclidean case, one has the following identity of elements of $T_{f(p)}N$:

$$(T_p f)\gamma'(0) = (f \circ \gamma)'(0).$$

It is occasionally notationally cumbersome to refer explicitly to the basepoint p in the total derivative $T_p f : T_p M \to T_{f(p)}N$; when that is the case, we might denote the latter as simply

$$df : T_p M \to T_{f(p)}N,$$

with the point p understood by context as the domain of any vector v to which we apply df. To illustrate, consider a composition of smooth maps of manifolds

$$\phi : K \xrightarrow{h} L \xrightarrow{g} M \xrightarrow{f} N.$$

For a point $p \in K$ and a tangent vector $v \in T_p K$, the chain rule (41) gives the slightly unwieldly formula

$$T_p\phi(v) = T_{g(h(p))} f(T_{h(p)} g(T_p h(v)))$$

which we abbreviate to simply

$$d\phi(v) = df(dg(dh(v))).$$

6.7. **Jacobians.** Given an n-dimensional vector space V, we denote by $\det(V) := \Lambda^n V$ its highest wedge power. It is a one-dimensional vector space. Given a pair V, W of n-dimensional vector spaces and a linear map $f : V \to W$, one obtains an induced map $\det(f) : \det(V) \to \det(W)$. If $W = V$, then $\det(f) : \det(V) \to \det(V)$ acts on the one-dimensional space $\det(V)$ via multiplication by the determinant of f in the sense of a first course on linear algebra.

In particular, given a manifold M, we obtain for each point $p \in M$ a one-dimensional vector space $\det T_p M$. Given a pair of manifolds M, N of the same dimension, a smooth map $f : M \dashrightarrow N$ between them, and a point p at which f is defined, we obtain a linear map $\det(T_p f) : \det(T_p M) \to \det(T_{f(p)}N)$ between one-dimensional spaces, called the *Jacobian determinant* of f. In the special case that M, N are open subsets of $\mathbf{k}^n$, we may identify $T_p M, T_{f(p)}N$ with $\mathbf{k}^n$ and the Jacobian determinant with the linear map $\det(\mathbf{k}^n) \to \det(\mathbf{k}^n)$ given by multiplication by the determinant of the Jacobian matrix describing the total derivative $T_p f : \mathbf{k}^n \to \mathbf{k}^n$. In general, one has $\det(T_p f) \neq 0$ if and only if $T_p f$ is a linear isomorphism.

6.8. **Inverse function theorem.** We record the generalization of what was given earlier in the Euclidean case.

Definition 13. A smooth map $f : M \dashrightarrow N$ between n-manifolds is said to be a *local diffeomorphism* at a point $p \in M$ if f is equivalent to an invertible local map defined at p with smooth inverse, or more verbosely, if there is an open neighborhood U of p in the domain of f so that $f(U)$ is open and the induced map $f : U \to f(U)$ is a diffeomorphism. In that case, there are coordinate systems $x_1, \dots, x_n$ on M at p and $y = (y_1, \dots, y_n)$ on N at $f(p)$ with respect to which p is the origin and so that f is given near p in these coordinates by the identity map $f : (x_1, \dots, x_n) \mapsto (x_1, \dots, x_n)$.

For example, if M is an n-manifold, then a map $f : M \dashrightarrow k^n$ is a local diffeomorphism at p if and only if it is equivalent to a smooth chart at p.

Theorem 14. *Let $f : M \to N$ be a smooth map of manifolds of the same dimension. The following are equivalent:*

(1) $T_p f : T_p M \to T_{f(p)} N$ is a linear isomorphism *of vector spaces, or equivalently, has nonzero Jacobian determinant* $\det(T_p f)$*.*
(2) f is a local diffeomorphism near p.

The problem is local, so it suffices to consider the case that M, N are open subsets $\mathbf{k}^n$, in which case the theorem reduces to special case given above in §5.4.

Here is a particularly useful consequence:

Corollary 15. *Let M be an n-manifold and $p \in M$. Let $\phi : M \dashrightarrow \mathbf{k}^n$ be a smooth map defined at p. Suppose that* $\det(T_p \phi) \neq 0$*. Then ϕ is equivalent to a smooth chart on M at p; in other words, there is a neighborhood $p \in U \subseteq M$ so that if we write $\phi|_U = (x_1, \dots, x_n)$ for some component functions $x_i : U \to \mathbf{k}$, then $x_1, \dots, x_n$ defines a coordinate system at p.*

6.9. **Local linearization of smooth maps.**

6.9.1. *Linear maps in terms of coordinates.* Suppose given an m-dimensional vector space V an n-dimensional vector W and a linear map $f : V \to W$ between them. Recall that the *rank* of f is the dimension of its image, or equivalently, the codimension of its kernel. Denote by k the rank of f. Then $k \leq m$ and $k \leq n$. One can always find bases $e_1, \dots, e_m$ of V and $\varepsilon_1, \dots, \varepsilon_n$ of W so that $f(\sum_{i=1}^m x_i e_i) = \sum_{i=1}^k x_i \varepsilon_i$; in coordinates, f takes the form $(x_1, \dots, x_m) \mapsto (x_1, \dots, x_k, 0, \dots, 0)$.

6.9.2. *The constant rank theorem.* Suppose now given an m-manifold M, an n-manifold N, a smooth map $f : M \dashrightarrow N$ and a point p in the domain of f.

Definition 16. We say that f is *linearizable* at p if there are local coordinates at p with respect to which f is given by a linear map. The *rank* of f at p is then the rank of that linear map.

Since any linear map is its own derivative, the rank of f at p is the same as the rank of $T_p f$. Denoting that rank by k (necessarily $k \leq \min(m, n)$), one can find local coordinates $x_1, \dots, x_m$ at p and $y_1, \dots, y_n$ at $f(p)$, putting both p and $f(p)$ at the origin, so that f is given in the particularly concrete form

$$f : (x_1, \dots, x_m) \mapsto (x_1, \dots, x_k, 0, \dots, 0).$$

Since a linear map has constant rank, an obvious necessary condition for f to be linearizable is the following:

Definition 17. We say that f has *constant rank* at p if the rank of $T_x f$ takes some constant value in a neighborhood of p.

In fact, the two conditions are equivalent:

Theorem 18. *f is linearizable at p if and only if f has constant rank at p.*

Proof. The interesting direction (showing that if f has constant rank, then it is linearizable) reduces to the rank theorem from multivariable calculus, whose proof is similar to that of the implicit function theorem given in §5.5. $\square$

6.9.3. *The case of maximal rank.* Given $f : M \dashrightarrow N$ as above, the function $x \mapsto \operatorname{rank}(T_x f)$ is *lower semicontinuous*, i.e., has the property that $\{x : \operatorname{rank}(T_x f) \geq k\}$ is open for all k. This is because the condition $\operatorname{rank}(T_x f) \geq k$ is detected by the nonvanishing of any $k \times k$ minor, which is an open condition. In other words, as x varies, the rank can only "jump" downwards.

The quantity $k_0 := \min(m, n)$ is the largest possible value for the rank of f at any point, i.e., one has $\operatorname{rank}(T_x f) \leq k_0$ for all x. It follows from the lower semicontinuity noted above that the set $\{x : \operatorname{rank}(T_x f) = k_0\}$ of points at which f attains its maximal rank is *open*: if f has rank k_0 at some point at some p, it automatically has rank k_0 in some small neighborhood of p. This observation motivates the utility of the following definition:

Definition 19. For $m \geq n$, we say that f is *submersive* at p if it satisfies any of the following equivalent conditions:

(1) $\operatorname{rank}(T_p f) = n$.
(2) $T_p f$ is surjective.
(3) $\dim \ker(T_p f) = m - n$.

For $m \leq n$, we say that f is *immersive* at p if it satisfies any of the following equivalent conditions:

(1) $\operatorname{rank}(T_p f) = m$.
(2) $T_p f$ is injective.
(3) $\dim \operatorname{coker}(T_p f) = n - m$, where $\operatorname{coker}(T_p f) := T_{f(p)} N / \operatorname{image}(T_p f)$.

We say that f is a submersion (resp. immersion) if it is submersive (resp. immersive) at all points p.

Theorem 20. *Suppose $f : M \dashrightarrow N$ as above is submersive at p. Then there are coordinate systems at p with respect to which f is given by a surjective linear map. For instance, there are coordinate systems $x_1, \ldots, x_m$ on M at p and $y_1, \ldots, y_n$ on N at $f(p)$, putting both p and $f(p)$ at the origin, so that f is given by $f : (x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_n)$.*

Proof. The statement is local, and reduces to that of §5.5. □

Theorem 21. *If f as above is immersive at p, then there are coordinate systems at p with respect to which f is given by an injective linear map. For instance, there are coordinates $x_1, \ldots, x_m$ on M at p and $y_1, \ldots, y_n$ on N at $f(p)$, putting both p and $f(p)$ at the origin, so that f is given by $(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_m, 0, \ldots, 0)$.*

Proof. This reduces to a local statement which can be proved as in §5.5 using the inverse function theorem. □

Corollary 22. *Let K be a k-manifold, let $g : K \dashrightarrow M$ be a continuous map, and let $f : M \dashrightarrow N$ be an immersion whose domain contains the image of g. Suppose that the composition*

$$K \overset{g}{\dashrightarrow} M \overset{f}{\dashrightarrow} N$$

is smooth. Then g is smooth.

Proof. Smoothness can be checked locally, so we may suppose that K is an open subset of $\mathbf{k}^k$. By the local description of f given the previous result, we may assume that $N = \mathbf{k}^n$ and $M = \mathbf{k}^m \cong \mathbf{k}^m \times \{0\} \subseteq \mathbf{k}^n$. We are then given a continuous map $\mathbf{k}^k \overset{g}{\dashrightarrow} \mathbf{k}^m$ with the property that the composition $\mathbf{k}^k \overset{g}{\dashrightarrow} \mathbf{k}^m \hookrightarrow \mathbf{k}^n$ is smooth (i.e., all partials exist). We want to deduce that $\mathbf{k}^k \overset{g}{\dashrightarrow} \mathbf{k}^m$ is smooth (i.e., all partials exist). What we want follows immediately from the definition. $\square$

6.10. **Submanifolds.** Submanifolds[2] are subsets of manifolds that look like vector subspaces up to a local diffeomorphism. More precisely:

Definition 23. Given an n-manifold M, a subset S of M is said to be a *d-dimensional submanifold* if for each $p \in S$, there is a smooth chart $\phi : M \dashrightarrow \mathbf{k}^n$ at p so that $\phi(S) = \phi(M) \cap \mathbf{k}^d \times \{0\} \subseteq \mathbf{k}^n$ (as defined in §4.1.4); said another way, there is a coordinate system $x_1, \ldots, x_n$ on M at p with respect to which S is given near p by the equation $x_{d+1} = \cdots = x_n = 0$.

The appropriateness of the term "submanifold" requires some justification:

Theorem 24. *Let S be a d-dimensional submanifold of an n-manifold M, regarded as a topological space with the induced topology. Then S possess a unique structure of a smooth d-manifold (i.e., a unique maximal smooth atlas) for which the inclusion map $\iota : S \to M$ is immersive.*

Proof. The uniqueness follows from Corollary 22: if $\mathcal{A}_1, \mathcal{A}_2$ are two maximal smooth atlases on S with the property that $(S, d, \mathcal{A}_1)$ and $(S, d, \mathcal{A}_2)$ are d-manifolds for which ι is immersive, then each of the inclusions $(S, d, \mathcal{A}_i) \hookrightarrow M$ is smooth, so by Corollary 22, the identity maps $(S, d, \mathcal{A}_1) \to (S, d, \mathcal{A}_2), (S, d, \mathcal{A}_2) \to (S, d, \mathcal{A}_1)$ are smooth two-sided inverses of each other; this shows that any smooth chart for $\mathcal{A}_1$ is also a smooth chart for $\mathcal{A}_2$, and vice-versa, so we may conclude by the maximality of $\mathcal{A}_1$ and $\mathcal{A}_2$ that they coincide. For the existence, we can use the local coordinates afforded by the definition of "submanifold" to define for each $p \in S$ a smooth atlas $\mathcal{A}_p$ in some neighborhood U of p for which $U \hookrightarrow M$ is an immersion; as p varies, the atlases $\mathcal{A}_p$ are compatible with one another thanks to the uniqueness assertion shown before, hence their union extends to a maximal smooth atlas on S with the required property. $\square$

By Corollary 22, we immediately obtain:

Proposition 25. *Let $f : M \to N$ be a smooth map of manifolds whose image is contained in some submanifold $S \subseteq N$. Then the induced map $f : M \to S$ is also smooth.*

Remark 26. A submanifold need not be open (think $\mathbf{k}^1 \hookrightarrow \mathbf{k}^2$) and need not be closed (think $(0,1) \hookrightarrow \mathbb{R}$), but is always *locally closed* in the following equivalent senses (as follows immediately from its local description):

(1) S is open in its closure in M.
(2) S is the intersection of a closed subset of M and an open subset of M.
(3) For each $p \in S$ there is an open neighborhood $p \in U \subseteq M$ so that $S \cap U$ is closed in U.

Exercise 1. Let S, M be manifolds and let $\iota : S \to M$ be an injective immersion with the property:

- for each $x \in M$ and each open $x \in U_1 \subseteq M$, there exists an open $x \in U \subseteq U_1 \subseteq M$ so that the open subset $\iota^{-1}(U)$ of S is connected.

Show that $\iota(S)$ is a submanifold of M and that ι is a diffeomorphism onto its image.

[2] What we call *submanifold* might normally be more verbosely called *smooth submanifold.*

6.11. **A criterion for being a submanifold.** In this section we record a handy criterion for determining when a subset is actually a submanifold. It amounts to the implicit function theorem from multivariable calculus.

Proposition 27. *Suppose given an n-manifold M and a natural number $d \leq n$. Let $S \subseteq M$ be a subset with the property that for each $p \in S$ there are $m := n - d$ smooth functions $f_1, \ldots, f_m : M \dashrightarrow \mathbf{k}$, defined at p, so that*

(1) S is given near p by the equation $f_1 = \cdots = f_m = 0$ (cf. §6.10), and
(2) $(f_1, \ldots, f_m) : M \dashrightarrow \mathbf{k}^m$ is submersive at p.

Then S is a d-dimensional submanifold of M.

Proof. This is immediate from the local description of submersive maps given in §6.9.3. □

Remark 28. The proposition is not an "if and only if." For example, consider $S := \{0\} \subseteq M := \mathbf{k}$. Clearly S is a 0-dimensional submanifold. On the other hand, one can (unwisely) define S inside M by the equation $f_1 = 0$, where $f_1(x) := x^2$. For this choice, the hypotheses of Proposition 27 fail because f_1 is not submersive at 0: its derivative $2x$ vanishes there.

6.12. **Computing tangent spaces of submanifolds.** Let M be an n-manifold and $S \subseteq M$ a d-dimensional submanifold. For each $p \in S$, the tangent space T_pS then identifies with a d-dimensional vector subspace of the n-dimensional vector space T_pM. The following is computationally helpful:

Proposition 29. *Suppose S is given near $p \in S$ by a system of smooth equations*

$$f_1 = \cdots = f_m = 0, \tag{42}$$

where $m := n - d$ and $f := (f_1, \ldots, f_m) : M \dashrightarrow \mathbf{k}^m$ is defined and submersive at p. Then $T_p(S)$ coincides with the space $\ker(T_pf)$ of solutions to the system of linear equations obtained by differentiating (42). Thus in local coordinates $x_1, \ldots, x_n$ at p,

$$T_p(S) = \left\{ (dx_1, \ldots, dx_n) \in \mathbf{k}^n : \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(p) dx_j = 0 \textit{ for } i = 1..m \right\}. \tag{43}$$

Proof. This is again immediate from the local description of submersive maps. □

6.13. **Summary of how to work with submanifolds.** Let M be an n-manifold and S a subset that one expects is a d-dimensional submanifold. Let's take a moment to explain how in practice one goes about verifying this and computing tangent spaces. First of all, the problem is local, so for each point $p \in S$, one fixes local coordinates $x_1, \ldots, x_n$ for M at p. (If M is an open subset of $\mathbf{k}^n$, then one can just use the default global coordinates.) Next, one expresses S near p as the solution set of some smooth system $f_1 = \cdots = f_m = 0$, where $m := n - d$. Next, one computes by hand the space V of solutions to the system of linear equations arising in (43). If it happens that $\dim(V) = d$, then it follows from Propositions 27 and 29 that $\dim \ker(T_pf) = d$, hence that T_pf is surjective, i.e., that f is submersive at p, hence that S is a d-dimensional submanifold and that $T_pS = V$ as subspaces of $T_pM \cong \mathbf{k}^n$.

Example 30. Let $M = \mathbb{R}^3$ and

$$S := \{(x, y, z) \in M : x^2 + y^2 + z^2 = 1\}.$$

Thus S is defined by $f = 0$, where $f(x, y, z) := x^2 + y^2 + z^2 - 1$. The derivative of f at a point $(x, y, z) \in S$ is the linear map $T_{(x,y,z)} f : \mathbb{R}^3 \to \mathbb{R}$ given by

$$T_{(x,y,z)} f(dx, dy, dz) = 2x\, dx + 2y\, dy + 2z\, dz.$$

Since at least one of x, y, z is nonzero, we see that $T_{(x,y,z)} f$ is surjective, hence that f is submersive at all points of S. Therefore S is a submanifold. Its tangent space is given at a point $(x, y, z) \in S$ by

$$T_{(x,y,z)} S = \{(dx, dy, dz) \in \mathbb{R}^3 : 2x\, dx + 2y\, dy + 2z\, dz = 0\},$$

which is a translate of (as expected) the plane tangent to S at (x, y, z) in the familiar geometric sense.

7. Some review of differential equations

The following results will be needed only briefly in the course; we record them here as a reference for completeness.

Suppose given a continuous map $f : \mathbf{k} \times \mathbf{k}^n \dashrightarrow \mathbf{k}^n$, with the first coordinate regarded as the "time" variable, the second as the "position" variable, and elements of the range as "velocities." We suppose given an initial time $t_0 \in \mathbf{k}$ and an initial position $y_0 \in \mathbf{k}^n$ for which (of course) $(t_0, y_0) \in \operatorname{dom}(f)$, and consider the existence and uniqueness problem for the linear ordinary differential equation (ODE)

$$y(t_0) = y_0, \quad y'(t) = f(t, y(t)) \text{ for all } t \in U. \tag{44}$$

Example 31. If $n = 1$ and $f(t, y) := y$ and $(t_0, y_0) := (0, 1)$, then we are considering the problem $y(0) = 1, y'(t) = y(t)$, for which it is well-known that the unique solution is the exponential map $y(t) := \exp(t) = \sum_{n=0}^{\infty} t^n/n!$.

Theorem 32 (Uniqueness)**.** *Assume that f is uniformly Lipschitz in the second variable:*

$$|f(t, y) - f(t, z)| \le C|y - z|.$$

Then for any convex open set T containing t_0, there is at most one continuously differentiable $y : T \to \mathbf{k}^n$ satisfying (44).

Proof. If $y(t), z(t)$ are two solutions to (44) defined on U, then their difference $w(t) := y(t) - z(t)$ satisfies $w(t_0) = 0$ and

$$|w'(t)| = |f(t, y(t)) - f(t, z(t))| \le C|y(t) - z(t)| = C|w(t)|.$$

Our aim is to show that the vanishing set $\Omega := \{t \in T : w(t) = 0\}$ is in fact all of T. Since Ω is nonempty (it contains t_0) and closed (w is continuous) and since T is connected, it will suffice to verify Ω is open. The mean value theorem implies that for each each $t_1, t_2 \in T$ there is some t on the line segment connecting t_1 and t_2 so that

$$|w(t_1) - w(t_2)| \le |w'(t)| \cdot |t_1 - t_2| \le C|w(t)| \cdot |t_1 - t_2|. \tag{45}$$

We apply (45) with $t_1 \in \Omega$ (so that $w(t_1) = 0$) and with t_2 in a closed ball $B \subseteq T$ with origin t_1 and radius at most $1/(2C)$ to obtain $|w(t_2)| \le (1/2)M$ with $M := \max_{t \in B} |w(t)|$. Since $t_2 \in B$ was arbitrary, it follows that $M \le (1/2)M$, hence that $M = 0$, hence that $B \subseteq \Omega$, hence that Ω is open at t_1, as required. $\square$

Theorem 33 (Existence)**.** *There exists an open ball T with origin t_0 and a continuously differentiable solution $y : T \to \mathbf{k}^n$ to (44). If f is smooth, then so is y.*

Proof. Let T_1 be a ball with origin t_0 and let Y be a ball with origin y_0 so that f is defined on $T_1 \times Y$. Let T be a ball with origin t_0 so that

$$\operatorname{radius}(T) \cdot \max_{T_1 \times Y} |f| < \operatorname{radius}(Y).$$

We will show that a solution y exists with domain T. To that end, let $\varepsilon > 0$ be small. Define as follows a function $y_\varepsilon : T \to Y \subseteq \mathbf{k}^n$:

(1) Set
$$y_\varepsilon(0) := y_0. \tag{46}$$
(2) Define y_ε on integral multiples $n\varepsilon \in T$ of ε inductively by requiring that
$$y_\varepsilon((n+1)\varepsilon) = y_\varepsilon(n\varepsilon) + \varepsilon f(n\varepsilon, y(n\varepsilon)). \tag{47}$$
This makes sense: our construction of T implies that $y(n\varepsilon) \in Y$ for all $n\varepsilon \in T$. (This is the "Euler method" for solving ODEs.)
(3) Define $y_\varepsilon(t)$ for general $t \in T$ by rounding t to the nearest integral multiple $n\varepsilon \in T$ of ε and setting $y_\varepsilon(t) := y_\varepsilon(n\varepsilon)$.

Using the elementary consequences

$$y_\varepsilon(t) \ll 1$$

and

$$|y_\varepsilon(t) - y_\varepsilon(s)| \ll |t - s| + o_{\varepsilon \to 0}(1)$$

of the construction of y_ε and arguing as in the proof of Arzeli–Ascoli, we obtain a sequence $\varepsilon_j \to 0$ and a bounded Lipschitz function $y : T \to Y$ (given by the uniform limit $y(t) = \lim_{j\to\infty} y_{\varepsilon_j}(t)$) satisfying $y(0) = y_0$ and, for all $t_1, t_2 \in T$ with $t_1 \leq t_2$,

$$\begin{aligned} y(t_2) - y(t_1) &= \lim_{j\to\infty} y_{\varepsilon_j}(t_2) - y_{\varepsilon_j}(t_1) \\ &= \lim_{j\to\infty} \sum_{\substack{n \in \mathbb{Z}: \\ n\varepsilon_j \in [t_1, t_2]}} \varepsilon_j f(n\varepsilon_j, y_{\varepsilon_j}(n\varepsilon_j)) \\ &= \lim_{\varepsilon \to 0} \sum_{\substack{n \in \mathbb{Z}: \\ n\varepsilon \in [t_1, t_2]}} \varepsilon f(n\varepsilon, y(n\varepsilon)) \\ &= \int_{t_1}^{t_2} f(t, y(t))\, dt. \end{aligned}$$

By the fundamental theorem of calculus, we conclude that y is differentiable and satisfies (44). Note finally that if f is smooth, then iterated application of the differential equation implies that y is also smooth. □

Example 34. A simple (and well-known) example illustrating the necessity of taking T sufficiently small is when $f : \mathbf{k} \times \mathbf{k}^2 \to \mathbf{k}^2$ is given by $f(t, x) := (x_1^2, x_1 x_2)$; for $y_0 = (u, v) \in \mathbf{k}^2$ with $u \neq 0$ and $t_0 := 0$, the unique solution y to (44) is given for t in a neighborhood of t_0 by

$$y(t) = \left(\frac{u}{1 - tu}, \frac{v}{1 - tu}\right),$$

which blows up as $t \to 1/u$.

We finally discuss the dependence of the solution y under "smooth deformation of parameters" in the initial condition or the differential equation.

Theorem 35 (Smooth dependence of solutions)**.** *Let* Π *be an open subset of some Euclidean space. Let*

$$f : \mathbf{k} \times \mathbf{k}^n \times \Pi \dashrightarrow \mathbf{k}^n$$
$$y_0 : \Pi \dashrightarrow \mathbf{k}^n$$

be smooth. Suppose given an initial time $t_0 \in \mathbf{k}$ *and initial parameter* $\pi_0 \in \Pi$ *so that* y_0 *is defined at* π_0 *and* f *is defined at* $(t_0, y_0(\pi_0), \pi_0)$*. Then there exist open balls* $T \subset \mathbf{k}$ *with origin* t_0 *and* $\Pi_0 \subseteq \Pi$ *with origin* π_0 *and a smooth solution* $y : T \times \Pi_0 \to \mathbf{k}^n$ *to the differential equation*

$$\frac{\partial}{\partial t} y(t, \pi) = f(t, y(t, \pi), \pi). \tag{48}$$

Proof. Arguing as in the proof of Theorem 33 and using only the continuity of f, we may choose T, Π_0 as above and a ball $Y \subseteq \mathbf{k}^n$ so that

(1) f is defined on a neighborhood of some compact set containing $T \times Y \times \Pi_0$,
(2) y_0 is defined on Π_0, and
(3) for each $\pi \in \Pi_0$, the ball in $\mathbf{k}^n$ with origin $y_0(\pi)$ and radius R, where $R := \operatorname{radius}(T) \cdot \max_{T_1 \times Y \times \Pi_0} |f|$, is contained in a compact subset of the interior of Y.

Running through the proof of Theorem 33, we obtain a function $y : Y \times \Pi \to Y$ that is smooth in the first variable and satisfies (48). We now verify that y is smooth in the second variable. By a compactness argument, it will suffice (after possibly shrinking Π_0 a bit) to verify that each $\pi_1 \in \Pi_0$ is contained in a small ball $\Pi_1 \subseteq \Pi_0$ on which y is smooth. To that end, fix $d \geq 1$ and consider for $\pi \in \Pi_1$ the Taylor series

$$y_0(\pi) = \sum_{\alpha \leq |d|} (\pi - \pi_1)^\alpha y_0^{(\alpha)}(\pi_1) + O(|\pi - \pi_1|)^{d+1},$$

$$f(t, y, \pi) = \sum_{\alpha \leq |d|} (\pi - \pi_1)^\alpha f^{(\alpha)}(t, y, \pi_1) + O(|\pi - \pi_1|)^{d+1}.$$

The errors are uniform thanks to the property (1) of f. Running through the proof of Theorem 33 and staring at (46) and (47) for a bit, we obtain an expansion

$$y(t, \pi) = \sum_{\alpha \leq |d|} (\pi - \pi_1)^\alpha y^{(\alpha)}(t, \pi_1) + O(|\pi - \pi_1|)^{d+1}.$$

Thus y is smooth in the second variable. By iterating the differential equation we conclude that y is jointly smooth in both variables. $\square$

8. Some review of group theory

8.1. **Basic definition.** Recall that a *group* is a tuple (G, m, i, e), often abbreviated simply by G, where

(1) G is a set,
(2) $m : G \times G \to G$ is a map called *multiplication* and abbreviated $xy := m(x, y)$,
(3) $i : G \to G$ is a map called *inversion* and abbreviated $x^{-1} := i(x)$, and
(4) $e \in G$ is an element called the *identity* element

and so that the usual axioms of group theory are satisfied; we do not recall them here. For example, the associativity axiom reads $m(x, m(y, z)) = m(m(x, y), z)$.

8.2. **Permutation groups.** One of the first examples of groups encountered in a basic course is the symmetric group $S(n)$. More generally, one considers for any set X the permutation group $\mathrm{Perm}(X)$, defined to consist of bijections $\sigma : X \to X$ and with the group law given by composition: $\sigma_1\sigma_2 := \sigma_1 \circ \sigma_2$. For example, $S(n) = \mathrm{Perm}(\{1, \dots, n\})$.

A particularly concrete class of groups are the subgroups of the form $G \leq \mathrm{Perm}(X)$ for some set X. Cayley's theorem asserts that *every* group is isomorphic to one of this form: indeed, one can take $X = G$ and define $G \hookrightarrow \mathrm{Perm}(G)$ via $g \mapsto [x \mapsto gx]$. Moreover, if G is finite, then one can take X to be finite.

8.3. **Topological groups.** One reason to phrase the definition in the above way is that it places the emphasis on the maps m, i. By equipping G with some additional structure and then requiring that those maps respect such structure, one obtains interesting classes of groups. For example:

Definition 36. A *topological group* is defined to be a group $G = (G, m, i, e)$ equipped with the structure of a topological space and for which the maps m, i are continuous. A *morphism of topological groups* $f : G \to H$ is a continuous group homomorphism. An *action* of a topological group G on a Hausdorff topological space X is a continuous map $\alpha : G \times X \to X$ with the property that $\alpha(e, x) = x$ and $\alpha(g_1g_2, x) = \alpha(g_1, \alpha(g_2, x))$; one typically abbreviates $gx := \alpha(g, x)$.

This definition is simple, but already fairly rich:

Exercise 2. Let X be a topological space. Let $G \subseteq X$ be a topological group that is also a subspace of X, equipped with the induced topology. Suppose there is an open $U \subseteq X$ for which $e \in U \subseteq G$, where e denotes the identity element of G. Show that G is open in X.

Exercise 3. Let G be a topological group. Let $H \leq G$ be a subgroup. Suppose that H is *locally closed* in G in the sense that there is a neighborhood $U \subseteq G$ of the identity with the property that $H \cap U$ is closed in U. Show that H is closed in G.

Exercise 4. Let G be a topological group, and $H \leq G$ an open subgroup. Show that H is closed.

Exercise 5. Let G be a connected topological group, and let U be a neighborhood of the identity. Show that U generates G.

Exercise 6. Let G be a topological group, and let $H \leq G$ be a subgroup. Equip the set G/H with the quotient topology. Show that the following are equivalent:

(1) H is closed.
(2) G/H is Hausdorff.

[Hint: If H is closed, show first that for each $g \in G - H$ there is a neighborhood U of the identity in G so that $U^{-1}gU \cap H = \emptyset$.]

Exercise 7. Let G be locally compact topological group, let $g \in G$, and let $V \subseteq G$ be a neighborhood of g. Show that there is an open neighborhood $U \subseteq G$ of e so that

(1) $\overline{U}$ is compact,
(2) $U = U^{-1}$,
(3) $U^n g \subseteq V$ for all $n \leq 100$, and
(4) $gU^n \subseteq V$ for all $n \leq 100$.

Here $U^n := \{u_1 \cdots u_n : u_1, \dots, u_n \in U\}$.

Exercise 8. Let G be a second countable topological group. Let $U \subseteq G$ be a subset with nonempty interior. Show that there is a sequence $g_n \in G$ so that $G = \cup U g_n = \cup g_n U$.

Exercise 9. Let G be a topological group, let X be a Hausdorff topological space, and suppose given a transitive action of G on X. Let U be a compact subset of G, and let $x \in X$. Show that Ux is closed.

Exercise 10. Say that a topological space X is *countable at infinity* if X can be written as a countalbe union of compact subsets.

Show that if X is locally compact and second-countable, then X is countable at infinity.

Exercise 11. Let G be a topological group, let X be a Hausdorff topological space, and suppose given a transitive action of G on X. Let $x \in X$. Show that the stabilizer $H := \{g \in G : gx = x\}$ is a closed subgroup of G. Suppose that

(1) G is locally compact and is countable at infinity, and
(2) X is locally compact.

Show that the map $\pi : G/H \to X$ given by $\pi(g) := gx$ is a homeomorphism. [It suffices to show that π is open. Use some of the previous exercises together with the following variant of the Baire category theorem: if a locally compact topological space E is a countable union of closed subsets E_n, then some E_n has nonempty interior.]

9. Some review of functional analysis

9.1. Definitions and elementary properties of operators on Hilbert spaces.

Definition 37. Recall that an *operator* on a Hilbert space (real or complex) V is a linear map $T : V \to V$. It is *bounded* if $\sup_{x \in V : |x| = 1} \|Tx\| < \infty$, *self-adjoint* if $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in V$, and *compact* if Tx_n has a convergent subsequence whenever x_n is a bounded sequence in V. An *eigenvector* of T is a nonzero element $v \in V$ for which $Tv = \lambda v$ for some scalar λ, called the *eigenvalue*.

Lemma 38. *Let T be a self-adjoint operator on a Hilbert space V.*

(1) The eigenvalues of T are real.
(2) The eigenspaces of T are orthogonal to one another.
(3) If T acts on a subspace U of V, then it acts also on the orthogonal complement $U^\perp$.

Proof.

(1) If $Tu = \lambda u$, then $\lambda \langle u, u \rangle = \langle Tu, u \rangle = \langle u, Tu \rangle = \overline{\langle Tu, u \rangle} = \overline{\lambda} \langle u, u \rangle$.
(2) If moreover $Tv = \lambda' v$, then $\lambda \langle u, v \rangle = \langle Tu, v \rangle = \langle u, Tv \rangle = \lambda' \langle u, v \rangle$, and so either $\lambda' = \lambda$ or $\langle u, v \rangle = 0$.
(3) Suppose $Tu \in U$ for all $u \in U$. Let $v \in U^\perp$. For $u \in U$, we have $Tu \in U$, and so $\langle Tv, u \rangle = \langle v, Tu \rangle = 0$. Thus $Tv \in U^\perp$.

□

9.2. **Compact self-adjoint operators on nonzero Hilbert spaces have eigenvectors.**

Theorem 39. *Let V be a nonzero Hilbert space. Let T be a compact self-adjoint operator on V. Then T has an eigenvector.*

The basic idea of the proof can seen most transparently when V is a finite-dimensional real Hilbert space: if x is an element of the unit sphere in V at which $\langle Tx, x\rangle$ assumes a local maximum, then the first derivative test implies that for any v orthogonal to x,

$$0 = \frac{d}{d\varepsilon}\langle T(x+v), x+v)\rangle|_{\varepsilon=0} = \langle Tv, x\rangle + \langle Tx, v\rangle = 2\langle Tx, v\rangle.$$

It follows that $Tx \in (x^{\perp})^{\perp} = \mathbb{R}x$, and so x is the required eigenvector.

To adapt the argument to the infinite-dimensional case, we replace the role of differential calculus with some artful application of the parallelogram law

$$4\Re\langle Tx, y\rangle = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle.$$

One of the steps en route to the solution is of independent interest:

Lemma 40. *Let T be a self-adjoint operator on a Hilbert space V. Then*

$$\sup_{x \in V : |x|=1} |\langle Tx, x\rangle| = \sup_{x,y : |x|=|y|=1} |\langle Tx, y\rangle|.$$

Proof. Denote by M the LHS and by M' the RHS. Clearly $M \leq M'$. Conversely, for $x, y \in V$ with $|x| = |y| = 1$ and $\theta \in \mathbb{C}^{(1)}$ chosen so that $\langle Tx, \theta y\rangle$ is real, the parallelogram law applied to x and θy gives

$$4\theta\langle Tx, y\rangle = \langle T(x+\theta y), x+\theta y\rangle - \langle T(x-\theta y), x-\theta y\rangle.$$

From this it follows that $M' \leq M$. $\square$

Proof of Theorem 39. Since T is compact, it is bounded, and so the quantity

$$M := \sup_{x \in V : |x|=1} |\langle Tx, x\rangle|$$

is finite. If $M = 0$, then the self-adjointness of T and the parallelogram law applied to $x \in V$ and $y := Tx$ implies that

$$4\|Tx\|^2 = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle = 0,$$

so T is the zero operator and any nonzero element of V is an eigenvector. We turn to the remaining case $M \neq 0$. Recall from the lemma that M coincides with the operator norm $\sup_{x,y:|x|=|y|=1} |\langle Tx, y\rangle|$. There is thus a nonzero real number $\lambda = \pm M$ and a sequence of unit vectors x_n so that

$$\langle Tx_n, x_n\rangle \to \lambda,$$

$$\langle Tx_n, Tx_n\rangle \leq \lambda^2.$$

It follows then from the identity

$$\|Tx_n - \lambda x_n\|^2 = \langle Tx_n, Tx_n\rangle - 2\lambda\langle Tx_n, x_n\rangle + \lambda^2$$

that

$$Tx_n - \lambda x_n \to 0. \tag{49}$$

Since T is compact, the sequence Tx_n has a subsequential limit y. By (49), one has $|y| = |\lambda|$, hence $y \notin 0$. By applying T to (49), one obtains $Ty = \lambda y$. Thus y is the required eigenvector of T. □

Remark 41. A self-adjoint operator on a Hilbert space need not have any eigenvectors; consider $f(x) \mapsto xf(x)$ on $L^2([0,1])$. In this sense, the compactness assumption is necessary.

9.3. **Spectral theorem for compact self-adjoint operators on a Hilbert space.**

Theorem 42. *Let T be a compact self-adjoint operator on a Hilbert space V. For $\lambda \in \mathbb{R}$, denote by $V_\lambda \leq V$ the λ-eigenspace of T. Then V is the Hilbert space orthogonal direct sum*

$$V = \oplus_\lambda V_\lambda \tag{50}$$

of its kernel $V_0 = \ker(T)$ and its eigenspaces V_λ with nonzero eigenvalue λ. Moreover, for any $\varepsilon > 0$,

$$\dim(\oplus_{\lambda:|\lambda|>\varepsilon} V_\lambda) < \infty. \tag{51}$$

Proof. The orthogonal complement of $\oplus_\lambda V_\lambda$ is T-stable and contains no eigenvectors for T; by Lemma 50, it is the zero space, giving (50).

For each $\varepsilon > 0$, the space $\oplus_{\lambda:|\lambda|>\varepsilon} V_\lambda$ admits an orthonormal basis of eigenvectors for T with eigenvalues of magnitude at least ε; if that basis were to contain an infinite sequence, then the image of that sequence under T would have no convergent subsequence, contradicting the compactness of T. This establishes (51). □

9.4. **Basics on matrix coefficients.**

9.5. **Finite functions on a compact group are dense.** Let G be a compact topological group. Let μ denote the probability Haar measure on G. We may define $L^2(G)$ with respect to μ. Denote by $\mathrm{U}(L^2(G))$ the group of unitary operators on $L^2(G)$. We then have the right regular representation $R : G \to \mathrm{U}(L^2(G))$ given by

$$R(g)f(x) := f(xg)$$

as well as the left regular representation $L : G \to \mathrm{U}(L^2(G))$ given by

$$L(g)f(x) := f(g^{-1}x).$$

We may extend the latter map linearly to $L : L^1(G) \to \mathrm{End}(L^2(G))$ given for $\phi \in L^2(G)$ by

$$L(\phi)f(x) := \int_{g\in G} \phi(g)f(g^{-1}x).$$

Lemma 43. *Let $f \in L^2(G)$. If the span of the right translates of f is finite-dimensional, then so is the span of its left translates, and vice-versa.*

Proof. Let $f_1, \dots, f_n$ be an orthonormal basis for the span of the right translates of f. Then for each $g \in G$ there are complex coefficients $a_1(g), \dots, a_n(g)$ so that for all $x \in G$,

$$R(g)f(x) = f(xg) = \sum_i a_i(g) f_i(x).$$

Explicitly, we may take $a_i := \langle R(g)f, f_i \rangle$, which defines a bounded function and thus an element $a_i \in L^2(G)$. It follows that

$$L(g)f(x) = f(g^{-1}x) = \sum_i f_i(g^{-1})a_i(x),$$

thus the a_i span the space of left translates of f. $\square$

Definition 44. We say that an element $f \in L^2(G)$ is *finite* if the span of its left and right translates under G is finite-dimensional. Denote by $L^2(G)_{\text{fin}}$ the space of finite functions. (Lemma 43 says that to check that a given function is finite, it suffices to show either that its left translates or its right translates have finite span.)

The main result of this subsection is as follows.

Theorem 45. *Let G be a compact group. Then the finite elements of $L^2(G)$ are dense.*

The proof requires a couple lemmas.

Lemma 46. *Set $V := L^2(G)$. For each $v \in V$ and $\varepsilon > 0$ there exists a real-valued symmetric $\phi \in C_c(G)$ so that $\|L(\phi)v - v\| < \varepsilon$.*

Proof. By the continuity of the representation L, one has for all g in some small neighborhood U of the identity in G that $\|L(g)v - v\| < \varepsilon$. We may assume after shrinking U as necessary that $U = U^{-1}$. By Urysohn's lemma, there exists a real-valued $\phi \in C_c(U) \subseteq C_c(G)$ with $\mu(\phi) = 1$. For such a ϕ, the required estimate follows from the triangle inequality. We can easily arrange that ϕ be symmetric by averaging it with the function $x \mapsto \phi(x^{-1})$. $\square$

Lemma 47. *Let $\phi \in L^1(G) \cap L^2(G)$. Then the operator $L(\phi)$ is compact.*

Proof. We will use that

$$\int_{g_1, g_2 \in G} |\phi(g_1^{-1}g_2)|^2 < \infty \tag{52}$$

as follows from the compactness of G. (The natural context for this result is thus in the setting of operators defined by kernels in $L^2(G \times G)$; see any book on functional analysis.) For concreteness, I'll give the proof of the compactness of $T := L(\phi)$ in the special case that $V := L^2(G)$ is a separable Hilbert space; this case certainly suffices when G is a compact Lie group (and perhaps somewhat more generally). We will show that the image under T of the unit ball is precompact. Let $e_1, e_2, \dots$ be a Hilbert space basis of V. Write $A_{ij} := \langle e_i, Te_j \rangle$, so that for $v = \sum a_i e_i \in V$, one has $Tv = \sum_i b_i e_i$ where $b_i := \sum_j A_{ij} a_j$. If $\|v\| \leq 1$, then Cauchy–Schwartz implies that $|b_i|^2 \leq B_i$ where $B_i := \sum_j |A_{ij}|^2$, so the image under T of the unit ball is contained in $S := \{\sum b_i e_i : |b_i|^2 \leq B_i\}$. Since

$$\sum B_i = \sum_{i,j} |A_{ij}|^2 < \infty,$$

the set S is precompact, as required. (If $v^{(n)} = \sum b_i^{(n)} e_i$ is a sequence in S, then we may assume by a diagonlization argument that after passing to a subsequence, one has $b_i^{(n)} \to b_i$ for some scalar b_i, which obviously satisfies $|b_i|^2 \leq B_i$; it then follows easily that $v := \sum b_i e_i$ belongs to V and that $v^{(n)} \to v$.) $\square$

Proof of Theorem 45. By Lemma 46, elements of the form $L(\phi)v$ with $\phi \in L^1(G), v \in L^2(G)$, and with ϕ real-valued and symmetric, are dense in $L^2(G)$, so it suffices to approximate such elements by finite functions. Thus consider some such elements ϕ, v. The operator $T := L(\phi)$ is compact and self-adjoint. Denote by V_λ its eigenspaces. Decompose $v = \sum v_\lambda$ with $v_\lambda \in V_\lambda$. For each $\varepsilon > 0$, we then have $L(\phi)v = u_\varepsilon + O(\varepsilon)$ where $u_\varepsilon := \sum_{\lambda:|\lambda|>\varepsilon} \lambda v_\lambda$. Since T commutes with $R(G)$, the eigenspaces V_λ of T are $R(G)$-invariant; Theorem 42 implies that $\dim(\oplus_{\lambda:|\lambda|>\varepsilon} V_\lambda) < \infty$, so the right translates of u_ε have finite span, and so u_ε is a finite element of $L^2(G)$. Since it converges to Tv as $\varepsilon \to 0$, we are done. □

9.6. **Schur orthogonality relations.** We continue to assume here that G is a compact group.

Lemma 48. *Let $(\pi, V), (\rho, W)$ be finite-dimensional irreducible representations of G. Then the space*

$$\mathrm{Hom}_G(V, W) := \{\phi : V \to W | \phi(\pi(g)v) = \rho(g)\phi(v) \text{ for all } v \in V, g \in G\}$$

of G-equivariant linear maps from V to W satisfies

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

If $V \cong W$, then every nonzero element of $\mathrm{Hom}_G(V, W)$ is an isomorphism $V \to W$. In particular, $\mathrm{Hom}_G(V, V)$ is the one-dimensional space consisting of scalar operators of the form $v \mapsto cv$, $c \in \mathbb{C}$.

Proof. For $\phi \in \mathrm{Hom}_G(V, W)$, we check that the kernel of ϕ is an invariant subspace of V and the image of ϕ is an invariant subspace of W. So if V is not the zero map, then its kernel is the zero space (since V is irreducible) and its image is all of W (since W is irreducible). The conclusion follows. □

Lemma 49. *Let V, W be finite-dimensional irreducible representations of G. Let $\ell \otimes v \in V^* \otimes V$ and $w \in W$. If V is not isomorphic to W, then*

$$\frac{1}{\dim W} \int_{g \in G} \ell(gv) g^{-1} w = 0. \tag{53}$$

Otherwise, let us fix an equivariant identification $W = V$. Then

$$\frac{1}{\dim W} \int_{g \in G} \ell(gv) g^{-1} w = \ell(w) v. \tag{54}$$

Proof. Fix ℓ, w, and denote by $S : V \to W$ the the linear function of v defined by the LHS of (53). Then a quick change of variables shows that S is equivariant, so by Lemma 48, it is the zero map unless $V \cong W$; this establishes (53). We turn to (54). By Lemma 48, we know that $S : V \to V$ is a scalar operator; we wish to verify that the scalar is $\ell(w)$. To that end, it will suffice to verify that $\mathrm{trace}(S) = (\dim V)\ell(w)$, or equivalently, that

$$\sum_i \int_{g \in G} \ell(g e_i) e_i^*(g^{-1} w) = \ell(w)$$

where (e_i) is a basis of V with dual basis (e_i^*) of V^*. The integrand is independent of g (since it is independent of the choice of basis, and the basis dual to $ge_1, \ldots, ge_n$

is $e_1^* \circ g^{-1}, \dots, e_n^* \circ g^{-1}$) so we reduce to showing that

$$\sum_i \ell(e_i) e_i^*(w) = \ell(w),$$

which is immediate. $\square$

Corollary 50. *Let V_1, V_2 be finite-dimensional irreducible representations of G. Let $\ell_1 \otimes v_1 \in V_1^* \otimes V_1$ and $\ell_2 \otimes v_2 \in V_2^* \otimes V_2$ Then*

$$\int_{g \in G} \ell_1(g v_1) \ell_2(g^{-1} v_2) = \begin{cases} \dim(V_1) \ell_1(v_2) \ell_2(v_1) & V_1 \cong V_2 \\ 0 & \textit{otherwise.} \end{cases} \tag{55}$$

where $\ell_1(v_2)$ and $\ell_2(v_1)$ are defined in the first case by fixing an equivariant identification between V_1 and V_2.

9.7. **Peter–Weyl theorem.** Let G be a compact group. Let $L^2(G)_{\mathrm{fin}}$ denote the subspace of finite elements; we saw in §9.5 that it is dense in $L^2(G)$.

Theorem 51. *Let V traverse the set of isomorphism classes of finite-dimensional irreducible representation of G. Then the canonical morphism of $G \times G$-modules*

$$\mathfrak{m} : \oplus V^* \otimes V \to L^2(G)_{\mathrm{fin}}$$

given in terms of matrix coefficients by setting, for $\ell \otimes v \in V^* \otimes V$,

$$\mathfrak{m}(\ell \otimes v)(g) := \ell(gv),$$

is an isomorphism with inverse

$$\mathcal{F} : L^2(G) \to \oplus \operatorname{End}(V)$$

where $\mathcal{F} = \oplus \mathcal{F}_V$ where for $u \in V$,

$$\mathcal{F}_V(f) u := \frac{1}{\dim(V)} \int_{g \in G} f(g) g^{-1} u.$$

(The action is as in the proof of Lemma 43.)

Proof. We first check surjectivity. Let $v \in L^2(G)_{\mathrm{fin}}$. Its span under the right regular representation of G is then a finite-dimensional representation W of G. We "have seen" (the chronology of the lectures differs from that of the notes) in §16.7 that any finite-dimensional representation of a compact group is completely reducible. In particular, W is completely reducible. By decomposing W into irreducibles and v into its irreducilbe components, we reduce to verifying in the special case in which W is irreducible that v belongs to the image of $W^* \otimes W$ in $L^2(G)_{\mathrm{fin}}$. But this follows immediately from the proof of Lemma 43.

We now check that $\mathcal{F} \circ \mathfrak{m} = 1$. It suffices to show for each $V, W \in \operatorname{Irr}(G)$ and $\ell \otimes v \in V^* \otimes V$ that

$$\mathcal{F}_W(\mathfrak{m}(\ell \otimes v)) = \begin{cases} \ell \otimes v & W = V \\ 0 & W \neq V. \end{cases}$$

Thus, let $w \in W$ be given. Then

$$\mathcal{F}_W(\mathfrak{m}(\ell \otimes v)) w = \frac{1}{\dim W} \int_{g \in G} \ell(gv) g^{-1} w,$$

while $(\ell \otimes v)(w) = \ell(w) v$, so the required conclusion follows from Lemma 49. $\square$

Corollary 52. *Let G act on $L^2(G)_{\mathrm{fin}}$ by the right regular representation. Then as G-representations,*

$$L^2(G)_{\mathrm{fin}} = \oplus V^{\oplus \dim(V)}.$$

where $V^{\oplus \dim(V)}$ is the image of $V^* \otimes V$, regarded now only as a G-module rather than as a $G \times G$-module.

10. Some facts concerning invariant measures

10.1. **Definition of Haar measures.** Let G be a locally compact topological group.

Definition 53. By a *Radon measure* on G we shall mean a linear functional $\mu : C_c(G) \to \mathbb{C}$ for which $f \geq 0 \implies \mu(f) \geq 0$; thanks to the Riesz representation theorem, this definition may also be formulated in terms of countably additive functions on the Borel σ-algebra satisfying certain properties.

For $y \in G$ and $f \in C_c(G)$, define the left and right translates $L_y f, R_y f \in C_c(G)$ by setting $L[y]f(x) := f(yx), R[y]f(x) := f(xy)$.

To interpret some of the statements to follow, we "recall" that it makes sense to integrate functions taking values in a Banach space. The only spaces we'll really need in the end are finite-dimensional vector spaces (where everything should be familiar) and the Hilbert space $L^2(G)$, where one doesn't lose much by interpreting everything in a pointwise fashion. (TODO: dfdfd)

Definition 54. A *left (resp. right) Haar measure* on G is a nonzero Radon measure μ with the property $\mu(L[y]f) = \mu(f)$ (resp. $\mu(R[y]f) = \mu(f)$).

We may reformulate this definition in various ways. For example, μ is a left Haar measure if $\mu(gE) = \mu(E)$ for all $g \in G$ and all Borel subsets E of G, or in integral form, if

$$\int_{g \in G} f(hg)\, d\mu(g) = \int_{g \in G} f(g)\, d\mu(g)$$

for all $h \in G$ and all $f \in C_c(G)$.

10.2. **Existence theorem.**

Theorem 55.

1. *There exist left Haar measures and there exist right Haar measures on any locally compact group G. They need neither coincide nor be scalar multiples of one another.*
2. *Any two left (resp. right) Haar measures are positive multiples of one another.*
3. *Any left or right Haar measure μ satisfies $\mu(f) > 0$ for any nonzero nonnegative $f \in C_c(G)$.*
4. *There is a continuous homomorphism $\Delta : G \to \mathbb{R}_+^\times$ so that for any left (resp. right) Haar measure μ and $f \in C_c(G)$, one has $\mu(R[g]f) = \Delta(g)$ (resp. $\mu(L[g]f) = \Delta(g^{-1})$). (TODO: check inverse here.)*
5. *If G is compact, then $\{$left Haar measures$\} = \{$right Haar measures$\}$.*

We sketch the idea of one proof; filling in the details may be regarded as an exercise, or alternatively, looked up somewhere. For each nonnegative nonzero $\phi \in C_c(G)$ and each nonnegative $f \in C_c(G)$, denote by $[f : \phi]$ the infinum of $\sum c_i$ taken over all finite tuples of positive coefficients $c_1, \dots, c_n$ and group elements $g_1, \dots, g_n$ with the property that $f \leq \sum c_i L[g_i]\phi$. Fix also some nonzero nonnegative $f_0 \in C_c(G)$. We may then attempt to define a left Haar measure μ on G by requiring that $\mu(f_0) = 1$ and that

$$\frac{\mu(f)}{\mu(f_0)} = \lim_{\phi} \frac{[f : \phi]}{[f_0 : \phi]} \tag{56}$$

where ϕ traverses a net consisting of nonzero nonnegative elements of $C_c(G)$ with support shrinking to the identity. This turns out to work. Conversely, to establish uniqueness, it suffices to show that (56) holds for any left Haar measure μ. The key lemma is that each nonnegative $f \in C_c(G)$ may uniformly approximated by some finite sum $c_i L[g_i]\phi_\alpha$ as above with support in a fixed compact; it follows then that $\mu(f)$ is approximated by $\mu(\sum c_i L[g_i]\phi_\alpha) = \mu(\phi_\alpha)\sum c_i \approx \mu(\phi_\alpha)[f : \phi]$, giving (56).

Definition 56. A locally compact group is *unimodular* if $\{$left Haar measures$\}$ = $\{$right Haar measures$\}$, or equivalently, if $\Delta(g) = 1$ for all $g \in G$. On a unimodular group, we may speak unambiguously simply about a *Haar measure* (without specifying "left" or "right"). For example, Theorem 55 says that compact groups are unimodular.

10.3. **Unimodularity of compact groups.** Note that if G is a compact group, then the image of the continuous homomorphism $\Delta : G \to \mathbb{R}_+^\times$ is a compact subgroup of $\mathbb{R}_+^\times$; the only such subgroup is $\{1\}$, so Δ is trivial, which explains why left and right Haar measures coincide on such a group. It follows that on each compact group, there is a unique (left and right) invariant probability measure.

10.4. **Direct construction for Lie groups.** When G is a Lie group, a simpler proof may be given using differential forms. Let ω_e be a nonzero element of $\det(T_e^* G)$. Denote by ω the volume form on G whose components $\omega_g \in \det(T_g^* G)$ for $g \in G$ are given by the pullback

$$\omega_g := R[g^{-1}]_g^* \omega_e$$

under the differential $R[g^{-1}]_g : T_g G \to T_e G$. Then $L[g]^*\omega = \omega$ for all $g \in G$, so ω is left-invariant. The map $C_c(G) \ni f \mapsto \int_G f \, \omega$ then defines a left Haar measure.

10.5. **Some exercises.**

Exercise 12. Let G be a Lie group with left Haar measure dg. Let $\Delta : G \to \mathbb{R}_+^\times$ be the function

$$\Delta(g) := \det(\mathrm{Ad}(g)|\mathfrak{g}).$$

Show that $\Delta(g)\, dg$ is a right Haar measure.

Exercise 13. Determine a left and right Haar measure on the Lie group

$$\mathrm{Aff}(\mathbb{R}) := \left\{ \begin{pmatrix} * & * \\ & 1 \end{pmatrix} \right\} \leq \mathrm{GL}_2(\mathbb{R}).$$

Exercise 14. (1) Let G be a locally compact group for which $[G, G]$ is dense in G. Show that G is unimodular.
(2) Let G, B, K be locally compact groups and let $\phi : B \times K \to G$ be a morphism. Suppose that G is unimodular, K is compact, and ϕ has dense

image. Let $d_l b$ be a left Haar measure on G and let dk be a Haar measure on K. Show that

$$\mu(f) := \int_{b \in B} \int_{k \in K} f(bk)\, d_l b\, dk$$

defines a Haar measure on G.

10.6. **Construction of a Haar measure on a compact group via averaging.** Let G be a compact topological group (not necessarily a Lie group). There is a nice way to construct the unique Haar probability measure μ on G via averaging.

Definition 57. Let $f : G \to \mathbb{C}$ be a continuous function on a compact group G. Let $\mathrm{Averages}(f)$ denote the space of functions $G \to \mathbb{C}$ of the form

$$G \ni x \mapsto \sum_{i=1}^{n} c_i f(\lambda_i x) \in \mathbb{C}$$

for some $n \in \mathbb{Z}_{\geq 1}$ and $\lambda_1, \dots, \lambda_n \in G$ and some $c_1, \dots, c_n \in [0, 1]$ with $c_1 + \dots + c_n = 1$.

Lemma 58. *There exists a unique constant function in the closure (with respect to the uniform topology) of* $\mathrm{Averages}(f)$.

Proof. One should be able to prove this as follows:

(1) It suffices to consider the case that f is real-valued.
(2) The space of continuous functions on G is closed with respect to the uniform topology.
(3) For a continuous real-valued function f on G, set
$$\mathrm{osc}(f) := \max_{g \in G} f(g) - \min_{g \in G} f(g).$$
Note that f is constant if and only if $\mathrm{osc}(f) = 0$. Note also that $\mathrm{osc}(f') \leq \mathrm{osc}(f)$ for all $f' \in \mathrm{Averages}(f)$.
(4) Given a continuous real-valued function f on G that is non-constant, show that there exists $f' \in \mathrm{Averages}(f)$ so that $\mathrm{osc}(f') < \mathrm{osc}(f)$. (Use the compactness of G and hence the uniform continuity of f. If f is smaller than typical in some part of G, translate f around a bit to dampen the contribution from parts of G where f is large.)
(5) The family of functions $\mathrm{Averages}(f)$ is equicontinuous.
(6) The function $\mathrm{osc} : C(G) \to \mathbb{R}_{\geq 0}$ is continuous with respect to the uniform topology on the domain.
(7) To prove the existence part of the theorem, we can take a sequence $f_i \in \mathrm{Averages}(f)$ so that $\mathrm{osc}(f_i)$ tends to the infinum of $\mathrm{osc}(h)$ over all $h \in \mathrm{Averages}(f)$. After passing to a subsequence and appealing to Arzela–Ascoli, we get a limit h of the sequence f_i. If h is non-constant, then we can find $h' \in \mathrm{Averages}(h)$ for which $\mathrm{osc}(h') < \mathrm{osc}(h)$. But one should then be able to check that h' lies in the closure of $\mathrm{Averages}(f)$, giving the required contradiction.
(8) To get uniqueness, let c_1, c_2 be values taken by constant functions in the closure of $\mathrm{Averages}(f)$. Let c' be a value taken by some constant function in the closure of the set defined analogously to $\mathrm{Averages}(f)$, but using right translations in place of left translations. Since left and right translations

commute, it's not so hard to check that $c_i = c'$ for $i = 1, 2$, hence that $c_1 = c_2$.

□

We may then define $\mu(f) = \int_G f \, d\mu$ to be the value taken by the constant function arising in Lemma 58. It's not hard to check that this defines a positive linear functional on the space of continuous functions on G, hence defines a measure; the key point is to verify additivity, which follows from some of the assertions made above.

11. Definition and basic properties of Lie groups

11.1. **Lie groups: definition.**

Definition 59. By a *Lie group* we shall mean a group G equipped with the structure of a manifold for which the maps $m : G \times G \to G$ and $i : G \to G$ are smooth. The *Lie algebra* of G is the vector space $\mathrm{Lie}(G) := T_e(G)$, often denoted $\mathfrak{g}$, given by the tangent space at the identity; it is a vector space of dimension equal to the dimension of G. (The word "algebra" appearing in the term "Lie algebra" will be justified later.)

For Lie groups G, H, a *morphism of Lie groups* or simply a *morphism* $f : G \to H$ is a smooth group homomorphism.

For a Lie group G and a manifold X, an *action of G on X* is a smooth map $\alpha : G \times X \to X$, abbreviated $gx := \alpha(g, x)$, that satisfies the same assumptions as in Definition 36.

Exercise 15. It suffices to check that m is smooth; the smoothness of i is automatic. [Hint: apply the inverse function theorem to the map $(x, y) \mapsto (x, xy)$.]

11.2. **Basic examples.** The *additive group* $(\mathbf{k}, +)$ of the field $\mathbf{k}$ is a Lie group, since the addition map $\mathbf{k} \times \mathbf{k} \ni (x, y) \mapsto x + y \in \mathbf{k}$ has the property that all of its partial derivatives exist. Similarly, the *multiplicative group* $(\mathbf{k}^\times, \times)$ is a Lie group. A slightly more interesting example can be obtained by considering any finite-dimensional unital associative algebra A over $\mathbf{k}$. A good example to keep in mind is when A is the algebra

$$A := M_n(\mathbf{k}) := \mathrm{Mat}_{n \times n}(\mathbf{k})$$

of $n \times n$ matrices, in which case

$$A^\times = \mathrm{GL}_n(\mathbf{k})$$

is the general linear group. The algebra A is a vector space, hence a manifold. Moreover, the unit group $A^\times$ is open in A: by Exercise 2, it suffices to verify that $A^\times$ contains a neighborhood of the identity element 1, and this follows from the observation that for $x \in A$ small enough, the element $1 + x$ has inverse given by the convergent series $\sum_{n \geq 0} (-x)^n$. Since the multiplication on $A^\times$ is bilinear, it is smooth, and so $A^\times$ is a Lie group of dimension $\dim(A)$. Moreover, one can naturally identify $T_0(A) = A$ and $T_1(A^\times) = A$, where $1 \in A^\times$ denotes the identity element, as in Example 10.

By what was shown above, $\mathrm{GL}_n(\mathbf{k})$ is an n^2-dimensional Lie group with

$$T_1(\mathrm{GL}_n(\mathbf{k})) = M_n(\mathbf{k}). \tag{57}$$

11.3. **Lie subgroups: definition.** There are at least a couple different conventions concerning what a "Lie subgroup" is.

Definition 60. Given a Lie group G, we say that a subset H of G is a *Lie subgroup* if it is a subgroup and a submanifold.

Lemma 61. *Let G be a Lie group. Let H be a Lie subgroup of G. Then H is a Lie group.*

Proof. We must show that the multiplication map $\mu_H : H \times H \to H$ is smooth, and that it coincides with the restriction $\mu_G|_{H\times H} : H \times H \to G$ of the smooth multiplication map $\mu_G : G \times G \to G$. So we reduce to the following: if $f : M \to N$ is a smooth map between manifolds whose image lands in some submanifold $S \subseteq N$, then the induced map $f : M \to S$ is also smooth. This is given by Proposition 25. $\square$

Definition 62. A *linear Lie group* is a Lie subgroup G of $\mathrm{GL}_n(\mathbf{k})$ for some n. (Essentially all of our examples will be of this form.)

Definition 63. Given a Lie group G, by an *immersed Lie subgroup* we will mean a subset H of G so that there exists a pair $(\hat{H}, \iota)$, where $\hat{H}$ is a Lie group and $\iota : \hat{H} \to G$ is an injective immersion with image H. (We will see much later in the course that such a pair is essentially uniquely determined by H, at least if H is connected.)

Example 64. Let $G := (\mathbb{R}/\mathbb{Z})^2$ be the two-dimensional torus, let $H := \mathbb{R}$ be the real line, let $\alpha \in \mathbb{R} - \mathbb{Q}$ be an irrational real number, and define $\iota : H \to G$ by the formula

$$\iota(x) := (x, \alpha x).$$

Since α is irrational, the map ι is injective. It is also an immersion, since its differential is given everywhere by the column matrix

$$T_x\iota = \begin{pmatrix} 1 \\ \alpha \end{pmatrix}$$

which defines an injective linear map $\mathbf{k} \to \mathbf{k}^2$. Thus (H, ι) is an immersed Lie subgroup of G. On the other hand, $\iota(H)$ is not a Lie subgroup because it is not a submanifold: submanifolds are open in their closure (Remark 26), and $\overline{\iota(H)} = G$, but $\iota(H)$ is not open in G. On a related note, ι does not define a homeomorphism onto its image: for instance, there exist sequences $x_n \in \mathbb{R}$ with $x_n \to \infty$ for which $\iota(x_n) \to \iota(0)$.

Remark 65. An injective immersion is called an *embedding* if it defines a homeomorphism onto its image. (This is not the case in Example 64.) With this terminology, we could alternatively define a Lie subgroup to be a pair (H, ι), where H is a Lie group and ι is an embedding.

11.4. **A handy criterion for being a Lie subgroup.** Here is a very handy criterion for checking that a subgroup of a Lie group is a Lie subgroup; we shall use it in several examples.

Lemma 66. *Let G be a Lie group. Let $H \leq G$ be a subgroup that is given near the identity $e \in G$ element by a system of equations*

$$f_1 = \cdots = f_m = 0, \tag{58}$$

where the $f_i : G \dashrightarrow \mathbf{k}$ *are smooth maps defined near* e *for which* $f := (f_1, \dots, f_m) : G \dashrightarrow \mathbf{k}^m$ *is submersive, i.e., satisfies either of the equivalent conditions* $\operatorname{rank}(T_e f) = m$ *or* $\dim(V) = d$*, where* $d := \dim(G) - m$ *and* $V := \ker(T_e f)$ *is the vector space given in local coordinates* $x_1, \dots, x_n$ *at* $e \in G$ *(*$n = \dim(G)$*) by the space of solutions* $(dx_1, \dots, dx_n) \in \mathbf{k}^n$ *to the system of homogeneous linear equations*

$$\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(e) dx_j = 0 \quad (i = 1..m)$$

obtained by differentiating (58)*. Then* H *is a* d*-dimensional Lie subgroup of* G*. Moreover,* $\operatorname{Lie}(H) = T_e H$ *is equal to* V*.*

Proof. Thanks to Propositions 27, 29, we need only verify that H is a d-dimensional submanifold of G. Set $n := \dim(G)$. As in the proof of Proposition 27, there is a coordinate system $x_1, \dots, x_n$ on G at e so that H is given near e by $x_1 = \cdots = x_m = 0$. Let $p \in H$, and let $\psi : G \to G$ be the map $\psi(x) := p^{-1}x$. Since G is a Lie group, the map ψ is a diffeomorphism with $\psi(p) = e$, and so $y_1 := x_1 \circ \psi, \dots, y_n := x_n \circ \psi$ defines a coordinate system on G at p. For $g \in G$ near p, the following are then visibly equivalent:

(1) $y_1(g) = \cdots = y_m(g) = 0$
(2) $x_1(p^{-1}g) = \cdots = x_m(p^{-1}g) = 0$
(3) $p^{-1}g \in H$
(4) $g \in H$.

In relating the final two steps, we used that H is a subgroup and that $p \in H$. □

Example 67. The subgroup $H := \mathrm{SL}_n(\mathbf{k})$ of $G := \mathrm{GL}_n(\mathbf{k})$ is defined by the single equation $\det(g) = 1$. Differentiating this equation and evaluating at the identity element gives the linear equation

$$\operatorname{trace}(dg) = 0$$

in the matrix variable $dg \in M_n(\mathbf{k}) = T_e(G)$ (see (57)). Since this equation has an $n^2 - 1$-dimensional solution space, we deduce from Lemma 66 that H is a Lie subgroup of G, called the *special linear group*. Moreover,

$$T_e(\mathrm{SL}_n(\mathbf{k})) = \{dg \in M_n(\mathbf{k}) : \operatorname{trace}(dg) = 0\}$$

is the space of traceless $n \times n$ matrices.

Exercise 16. Show that the *orthogonal group*

$$O_n(\mathbf{k}) := \{g \in \mathrm{GL}_n(\mathbf{k}) : gg^t = 1\},$$

where $g \mapsto g^t$ denotes the transpose map, is defined by a system of $n(n+1)/2$ equations having full rank at the identity (i.e., satisfying the submersiveness condition). Deduce that $O_n(\mathbf{k})$ is a Lie group of dimension $n^2 - n(n+1)/2 = n(n-1)/2$.

11.5. **Lie subgroups are closed.**

Theorem 68. *Let* G *be a Lie group and* $H \leq G$ *a Lie subgroup. Then* H *is closed in* G*.*

Recalling from Remark 26 that H is locally closed in G, the proof of Theorem 68 reduces to that of the following:

Lemma 69. *A locally closed subgroup* H *of a topological group* G *is closed.*

Proof. By the continuity of the group operations in G, the closure $\overline{H}$ is itself a group. For $g \in \overline{H}$, the coset gH is then an open subset of $\overline{H}$ (using here that H is locally closed). Since H is dense in $\overline{H}$, the subsets gH and H intersect. This means that we can write $gx = y$ for some $x, y \in H$, whence $g = x^{-1}y$ belongs to H. Since g was arbitrary, we conclude as required that $\overline{H} = H$. $\square$

11.6. **Translation of tangent spaces by group elements.** Let G be a Lie group. An element $g \in G$ acts on G by the left and right multiplication maps

$$L[g] : G \to G$$
$$h \mapsto gh,$$
$$R[g] : G \to G$$
$$h \mapsto hg.$$

One has

$$L[g_1] \circ L[g_2] = L[g_1 g_2], \quad R[g_1] \circ R[g_2] = R[g_2 g_1]. \tag{59}$$

These maps are smooth, so it makes sense to differentiate them at an element $h \in G$ to obtain linear maps of tangent spaces

$$T_h L[g] : T_h G \to T_{gh}(G)$$
$$T_h R[g] : T_h G \to T_{hg}(G).$$

By the chain rule and the identities (59), these maps are in fact linear isomorphisms of the tangent spaces. It will be convenient to introduce for $X \in T_h G$ and $g \in G$ the abbreviations

$$gX := (T_h L[g])(X) \in T_{gh}(G), \quad Xg := (T_h R[g])(X) \in T_{hg}(G). \tag{60}$$

These will be used most often when $h = 1$, so that $X \in T_e G = \mathfrak{g}$. The special case worth focusing on is when G is a linear Lie group $G \leq \mathrm{GL}_n(\mathbf{k})$. In that case, we can identify the various tangent spaces $T_h(G)$ with subspaces of $M_n(\mathbf{k})$; under this identification, the quantities gX, Xg defined in (60) are given by the matrix products of $g \in \mathrm{GL}_n(\mathbf{k})$ and $X \in M_n(\mathbf{k})$.

12. The connected component

12.1. **Generalities.** Any group may be regarded as a discrete topological group, or even a discrete 0-dimensional Lie group (provided the group is countable, so as to satisfy the hypothesis of second-countability), but Lie theory has nothing interesting to say about such Lie groups; its techniques show their true strength only when the group is connected. Recall, then, that a topological space X is *connected* if it cannot be written as a disjoint union of two nonempty closed subsets, or equivalently, if every continuous map from X to a discrete topological space is constant. Any topological space X admits a unique decomposition $X = \sqcup_i X_i$ into maximal connected subsets X_i, called the *connected components* of X; since the closure of any connected set is connected, the connected components are always closed, but need not be open in general. However, if X is *locally connected*, that is to say, if each point has a connected neighborhood, then the connected components are open. In particular, any manifold is locally Euclidean, hence locally connected, and so is the disjoint union of its connected components which are in turn open submanifolds. Moreover, since manifolds are locally path-connected, we know that any connected manifold is necessarily path-connected, even by smooth paths.

In particular, the connected components of a Lie group G are submanifolds. It is customary to denote by G^0 the connected component of the identity element of G. Then G^0 has the defining property that

- if C is any connected subset of G that contains the identity element, then $C \subseteq G^0$.

We have the following:

Theorem 70. *G^0 is a normal Lie subgroup of G, and the connected components of G are precisely the cosets of G^0.*

Proof. We have already observed that G^0 is a submanifold. Let $g \in G$. Since left and right multiplication maps $x \mapsto gx, x \mapsto xg$ define homeomorphisms from G to itself, they permute the connected components. The various assertions follow easily from this:

(1) If $g \in G^0$, then gG^0 is a connected component of G containing g. Since G^0 is also a connected component of G containing g, it follows that $gG^0 = G^0$. This implies that G^0 is a subgroup.
(2) For any $g \in G$, the conjugate gG^0g^{-1} is a connected component of G that contains $gg^{-1} = 1$, hence $gG^0g^{-1} = G^0$. Thus G^0 is normal.
(3) If C is any connected component of G^0, then it is nonempty; if $g \in C$ is any element, then $g^{-1}C$ is a connected component of G^0 that contains $g^{-1}g = 1$, hence $g^{-1}C = G^0$ and so $C = gG^0$.

□

12.2. **Some examples.** In the following table, we list the number of connected components of some Lie groups. Here $\mathbf{k}$ is one of the fields $\mathbb{R}$ or $\mathbb{C}$, $n \geq 1$ and $p \geq q \geq 1$.

1 component	2 components	4 components
$\mathrm{GL}_n(\mathbb{C})$	$\mathrm{GL}_n(\mathbb{R})$	$O(p,q)$
$\mathrm{SL}_n(\mathbf{k})$	$O(n)$	
$\mathrm{SO}(n)$	$\mathrm{SO}(p,q)$	
$U(n), \mathrm{SU}(n)$ $U(p,q), \mathrm{SU}(p,q)$		

These Lie groups were defined in lecture. It was proved that $\mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbf{k})$, $\mathrm{O}(n)$, $\mathrm{SO}(n)$, $\mathrm{U}(n)$ have the indicated number of connected components; the remaining cases were left as exercises.

- For the case of $\mathrm{SL}_n(\mathbf{k})$, we argued using elementary matrices.
- The group $\mathrm{GL}_n(\mathbb{C})$ is the image of the connected domain $\mathbb{C}^\times \times \mathrm{SL}_n(\mathbb{C})$ under the continuous homomorphism $(\zeta, g) \mapsto \zeta g$, which is surjective because $\det(\zeta g) = \zeta^n$ and the nth power map on $\mathbb{C}^\times$ is surjective.
- The proof in the case of $\mathrm{SO}(n)$ was by induction on n, using that the trivial group $\mathrm{SO}(1) = \{1\}$ is connected and that $\mathrm{SO}(n)$ acts transitively on the (connected) sphere S^{n-1} with $\mathrm{SO}(n-1) \hookrightarrow \mathrm{SO}(n)$ as the stabilizer group of the point $e_n := (0, \ldots, 0, 1)$. Since several people asked for clarifications regarding this proof, I have written it down in the following subsection.

- The proof in the case of $\mathrm{U}(n)$ was to recall that every conjugacy class in $\mathrm{U}(n)$ contains a diagonal element, and that the diagonal elements

$$\begin{pmatrix} e^{i\theta_1} & & \\ & \cdots & \\ & & e^{i\theta_n} \end{pmatrix}$$

are obviously in the connected component of the identity, because (for instance) they are the values $\gamma(1)$ of the continuous maps

$$\gamma(t) = \begin{pmatrix} e^{i\theta_1 t} & & \\ & \cdots & \\ & & e^{i\theta_n t} \end{pmatrix}$$

for which $\gamma(0) = 1$ is the identity.

12.3. **Connectedness of** $\mathrm{SO}(n)$**.** We verify by induction on n that $\mathrm{SO}(n)$ is connected. The group $\mathrm{SO}(1) = \{1\}$ is trivial, hence connected. Let $n \geq 2$. The group $\mathrm{SO}(n)$ acts smoothly on the unit sphere $S^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}$. The action is transitive. In fact, the connected component already acts transitively:

Lemma 71. *Let $n \geq 2$. Then $\mathrm{SO}(n)^0$ acts transitively on S^{n-1}.*

Proof. Denote by $e_1, \dots, e_n$ the standard basis elements; they all belong to S^{n-1}. Let $v \in S^{n-1}$. We will show that there exists $g \in \mathrm{SO}(n)^0$ with $ge_n = v$.

(1) Consider first the case $n = 2$. Define $\gamma : \mathbb{R} \to \mathrm{SO}(2)$ by the formula

$$\gamma(\theta) := \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}(2).$$

Since γ is continuous and $\gamma(0) = 1$, we see that $\gamma(\theta) \in \mathrm{SO}(2)^0$ for all θ. Now let $v \in S^1$. We may then write $v = (x, y)$ where $x^2 + y^2 = 1$ and solve $x = \sin(\theta), y = \cos(\theta)$ for some θ. Then $\gamma(\theta)e_2 = (x, y)$, as required.

(2) Suppose now that $n \geq 3$. Define $\gamma : \mathbb{R} \to \mathrm{SO}(n)$ by

$$\gamma(\theta) := \begin{pmatrix} 1 & & & & \\ & \cdots & & & \\ & & 1 & & \\ & & & \cos(\theta) & \sin(\theta) \\ & & & -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

As above, $\gamma(\theta) \in \mathrm{SO}(n)^0$ for all θ. Let $v \in S^{n-1}$. If v belongs to the line spanned by e_n, then either $v = e_n$ (in which case $v = ge_n$ with $g = 1 \in \mathrm{SO}(n)^0$) or $v = -e_n$ (in which case $v = \gamma(\pi)e_n$); in either case, v is of the form $v = ge_n$ for some $g \in \mathrm{SO}(n)^0$. If v and e_n are linearly independent, we may use Gram–Schmidt to find an orthonormal basis e'_{n-1}, e_n for their span. We may then extend this to an orthonormal basis $e'_1, e'_2, \dots, e'_{n-1}, e_n$ for $\mathbb{R}^n$. We can find g in $\mathrm{O}(n)$ taking one orthonormal basis to the other, so that gv belongs to the span of e_{n-1}, e_n and $ge_n = e_n$ Then gv is of the form $(0, \dots, 0, x, y)$ with $x^2 + y^2 = 1$; we can then solve $x = \sin(\theta), y = \cos(\theta)$ as before to obtain $\gamma(\theta)e_n = gv$ and thus $g^{-1}\gamma(\theta)ge_n = v$. Since $\mathrm{SO}(n)^0 = \mathrm{O}(n)^0$ is normal in $\mathrm{O}(n)$, we have $g^{-1}\gamma(\theta)g \in \mathrm{SO}(n)^0$.

$\square$

Lemma 72. *The stabilizer group* $\mathrm{Stab}_{\mathrm{SO}(n)}(e_n)$ *is isomorphic to* $\mathrm{SO}(n-1)$*, with an isomorphism in the opposite direction given by*

$$\mathrm{SO}(n-1) \xrightarrow{\cong} \mathrm{Stab}_{\mathrm{SO}(n)}(e_n)$$

$$h \mapsto \begin{pmatrix} h & \\ & 1 \end{pmatrix}.$$

Proof. If $g \in \mathrm{SO}(n)$ fixes e_n, then the identity

$$\langle gv, e_n \rangle = \langle v, g^t e_n \rangle = \langle v, g^{-1} e_n \rangle$$

implies that g^{-1} and hence g also stabilizes the orthogonal complement $\langle e_n \rangle^{\perp} = \langle e_1, \dots, e_{n-1} \rangle$. We may thus put it in in the block-upper triangular form $g = \begin{pmatrix} h & \\ & 1 \end{pmatrix}$ for some $h \in \mathrm{GL}_{n-1}(\mathbb{R})$. The condition $gg^t = 1$ implies $hh^t = 1$, hence that $h \in \mathrm{SO}(n-1)$. □

Using the above lemmas, we now complete the inductive step in the proof that $\mathrm{SO}(n)$ is connected, i.e., that $\mathrm{SO}(n) = \mathrm{SO}(n)^0$. Let $g \in \mathrm{SO}(n)$; we wish to show that in fact $g \in \mathrm{SO}(n)^0$. Consider the element $ge_n \in S^{n-1}$. By Lemma 71, we may write $ge_n = \gamma e_n$ for some $\gamma \in \mathrm{SO}(n)^0$. We then have $\gamma^{-1} g e_n = e_n$, i.e., $\gamma^{-1} g \in \mathrm{Stab}_{\mathrm{SO}(n)}(e_n)$. By Lemma 72, we have $\gamma^{-1} g = h$ for some $h \in \mathrm{SO}(n-1) \hookrightarrow \mathrm{SO}(n)$. By the inductive hypothesis, $\mathrm{SO}(n-1)$ is connected, hence $\mathrm{SO}(n-1) \subseteq \mathrm{SO}(n)^0$. Thus $g = \gamma h$ is the product of two elements of $\mathrm{SO}(n)^0$; since the latter is known to be a group, we conclude that $g \in \mathrm{SO}(n)^0$.

Remark 73. One doesn't actually need to prove Lemma 71 to complete the inductive argument; a slightly softer way to proceed is to make use of the following general lemma:

Lemma 74. *Let* G *be a topological group that acts on a topological space* X *(i.e., we assume given an action* $\alpha : G \times X \to X$ *as in Definition 36). Assume that:*

1. *For* $x_0 \in X$*, the orbit map* $G \to X$ *given by* $g \mapsto gx_0$ *is a quotient map.*
2. *The action is transitive.*
3. X *is connected.*
4. *The stabilizer* H *in* G *of some (equivalently, any) point* $x_0 \in X$ *is connected.*

Then G *is connected.*

Proof. If not, we may find a non-constant continuous map $f : G \to D$ for some discrete topological space D (e.g., $D = \{0, 1\}$). Since H and hence any coset of H is connected, the restriction of f to any coset of H is constant, hence f induces a continuous map $G/H \to D$, where G/H is equipped with the quotient topology. Since the orbit map assumed to be a quotient map, we obtain a continuous map $X \to D$ sending gx_0 to $f(g)$. Since X is connected, this last map must be constant, hence so must the original map f. Therefore G is connected. □

The orbit maps for the action $\mathrm{SO}(n) \circlearrowright S^{n-1}$ are quotient maps because (for instance) $\mathrm{SO}(n)$ is Hausdorff and S^{n-1} is compact. Alternatively, one can appeal here to Exercise 11, which tells us that the orbits map are open maps, hence are quotient maps.

13. Basics on the exponential map

13.1. **Review of the matrix exponential.** Let $\mathbf{k}$ be either $\mathbb{R}$ of $\mathbb{C}$. Let $n \geq 1$ and

$$A := M_n(\mathbf{k}) := \mathrm{Mat}_{n\times n}(\mathbf{k}).$$

It is a finite-dimensional unital associative algebra over $\mathbf{k}$ (and the discussion to follow applies more generally to any such algebra). The operator norm $\|.\|$ on A is given by $\|x\| := \sup_{v\in\mathbf{k}^n:|v|=1}|xv|$; it satisfies the submultiplicativity property $\|xy\| \leq \|x\|\|y\|$. The unit group of A is

$$A^\times = \mathrm{GL}_n(\mathbf{k}).$$

For $x \in A$ with $\|x\| < 1$, the series $\sum_{n\geq 0} x^n$ converges (by the same proof as in the one-variable case, using the submultiplicativity). Therefore $1 - x$ has the inverse $\sum_{n\geq 0} x^n$ and hence belongs to $A^\times$ whenever $\|x\| < 1$. Therefore $A^\times$ is open in A. (One can also see this more directly.) Since A is a Euclidean space, it follows that we have natural identifications of tangent spaces

$$T_0(A) = A, \quad T_1(A^\times) = A.$$

For any $x \in A$, the series

$$\exp(x) := \sum_{n\geq 0} \frac{x^n}{n!}$$

converges. It satisfies the following properties:

(1) exp is smooth
(2) $\exp(0) = 1$
(3) One has $\exp(x + y) = \exp(x)\exp(y)$ whenever x, y commute (but not in general otherwise). In particular:
   (a) $\exp(x)\exp(-x) = \exp(0) = 1$, hence $\exp(A) \subseteq A^\times$.
   (b) $\exp((s+t)x) = \exp(sx)\exp(tx)$ for all $s, t \in \mathbf{k}$, hence the map $\mathbf{k} \ni t \mapsto \exp(tx) \in A^\times$ is morphism of Lie groups for each $x \in A$.
(4) $\frac{d}{dt}\exp(tx) = x$, hence $T_0 \exp : A \to A$ is the identity transformation. Consequently exp defines a local diffeomorphism at 0, i.e., induces a diffeomorphism $\exp : U \to V$ for some open $0 \in U \subseteq A$ and $1 \in V \subseteq A^\times$.
(5) An inverse to exp on the subset $\{1 - x : \|x\| < 1\}$ of $A^\times$ is given by the logarithm
$$\log(1 - x) := -\sum_{n\geq 1} \frac{x^n}{n}.$$
(6) For $g \in A^\times$ and $x \in A$ one has $\exp(gxg^{-1}) = g\exp(x)g^{-1}$.

For a diagonal matrix, one has

$$\exp(\begin{pmatrix} t_1 & & \\ & \cdots & \\ & & t_n \end{pmatrix}) = \begin{pmatrix} \exp(t_1) & & \\ & \cdots & \\ & & \exp(t_n) \end{pmatrix}.$$

For a basic nilpotent Jordan block N, given in the case (say) $n = 4$ by

$$N = \begin{pmatrix} & 1 & & \\ & & 1 & \\ & & & 1 \\ & & & \end{pmatrix},$$

one has $N^k = 0$ for all $k \geq n$, hence the series defining $\exp(N)$ is finite. The series defining $\log(1 + tN)$ is also finite for any $t \in \mathbf{k}$.

Exercise 17. Using the above facts and Jordan decomposition, show that $\exp : M_n(\mathbb{C}) \to \mathrm{GL}_n(\mathbb{C})$ is surjective. (The corresponding assertion over the reals is false for several reasons to be discussed in due course.)

The exponential map is very rarely injective (away from the origin); for example, $\exp(2\pi i) = 1$, and

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

Some other good examples to keep in mind are

$$\exp(t \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}) = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$$

and

$$\exp \begin{pmatrix} 0 & x & z \\ & 0 & y \\ & & 0 \end{pmatrix} = \begin{pmatrix} 1 & x & z + xy/2 \\ & 1 & y \\ & & 1 \end{pmatrix}.$$

13.2. **One-parameter subgroups.** Let G be a Lie group. Let $\mathfrak{g} := \mathrm{Lie}(G)$ denote its Lie algebra.

Definition 75. By a *one-parameter subgroup* of G, we shall mean a morphism of Lie groups $\Phi : \mathbf{k} \to G$.

Remark 76.

(1) Note the standard but slightly misleading terminology: a "one-parameter subgroup" Φ is not a subgroup; its image $\mathrm{image}(\Phi) \leq G$ a subgroup, but even if Φ is injective, the datum of Φ contains strictly more information than that of its image. For example, the one-parameter subgroups $\Phi(t) := t$ and $\Phi(t) := 2t$ in the additive group $G = (\mathbf{k}, +)$ have the same image.

(2) Note also that a one-parameter subgroup Φ of G is, in particular, a curve (in the sense of §6.5) with basepoint the identity element $e \in G$. Its initial velocity $\Phi'(0)$ is an element of $\mathfrak{g}$.

Example 77. For $v \in M_n(\mathbf{k})$, the discussion of §13.1 shows that the map $\Phi_v : \mathbf{k} \to \mathrm{GL}_n(\mathbf{k})$ given by $\Phi_v(t) := \exp(tv)$ is a one-parameter subgroup of the Lie group $\mathrm{GL}_n(\mathbf{k})$ with initial velocity $\Phi_v'(0) = v$. Moreover, the map $M_n(\mathbf{k}) \ni v \mapsto \Phi_v(1) = \exp(v) \in \mathrm{GL}_n(\mathbf{k})$ is smooth.

For general G, a one-parameter subgroup Φ satisfies the identity $\Phi(s + t) = \Phi(s)\Phi(t) = \Phi(t)\Phi(s)$. Applying $\frac{d}{ds}|_{s=0}$ to this identity gives the differential equation[3]

$$\Phi'(t) = \Phi'(0)\Phi(t) = \Phi(t)\Phi'(0). \tag{61}$$

By the initial condition $\Phi(0) = 1$ and general uniqueness theorem for ODE's (§7), it follows that Φ is determined uniquely by its initial velocity $\Phi'(0) \in \mathfrak{g}$. To put it another way, for each element $v \in \mathfrak{g}$, there is *at most one* one-parameter subgroup Φ_v of G with initial velocity v. Conversely, we will now show that such a one-parameter

[3] If the definition of products such as $\Phi'(0)\Phi(t)$ is unclear, one should either consult §11.6 or (better) assume that G is a linear Lie and interpret such products as as being given by matrix multiplication.

subgroup actually exists, and moreover, that its values $\Phi_v(t)$ vary smoothly with v. For illustration, we explain a few different ways to establish existence:

(1) In the special case $G = \mathrm{GL}_n(\mathbf{k})$, one can just take $\Phi_v(t) := \exp(tv)$.
(2) If G is a *linear* Lie group, that is to say, if it is a Lie subgroup of $\mathrm{GL}_n(\mathbf{k})$ for some n, so that $\mathfrak{g} \leq M_n(\mathbf{k})$, then it turns out that one can again take $\Phi_v(t) := \exp(tv)$ where $\exp : M_n(\mathbf{k}) \to \mathrm{GL}_n(\mathbf{k})$ is as defined above. What requires proof here is the following:

Lemma 78. *Let G be a Lie subgroup of* $\mathrm{GL}_n(\mathbf{k})$. *Then* $\exp(\mathfrak{g}) \subseteq G$.

Proof. Let $x \in \mathfrak{g}$, and let $\gamma : \mathbf{k} \dashrightarrow G$ be any curve with $\gamma'(0) = x$. (Such a curve exists, more-or-less by definition of the tangent space; see §6.5 and especially Remark 11.) Since $G \leq \mathrm{GL}_n(\mathbf{k})$ and $\gamma(t) = 1 + tv + o(t)$ as $t \to 0$, we may take for t small enough the logarithm of $\gamma(t)$, which then satisfies $\log \gamma(t) = tv + o(t)$. Taking $t := 1/n$ with $n \in \mathbb{Z}$ tending off to ∞ gives $n \log \gamma(1/n) = n(v/n + o(1/n)) = v + o(1)$. Exponentiating, one obtains $\gamma(1/n)^n = \exp(v + o(1))$. Hence $\exp(v) = \lim_{n\to\infty} \gamma(1/n)^n$. Since γ is a curve in G and G is a group, we have $\gamma(1/n)^n \in G$ for all n. Since G is closed (see §11.5), it follows that $\exp(v) \in G$, as required. $\square$

(3) For general G, we can appeal to existence theorems for ODE's (see §7) to produce a curve $\gamma : \mathbf{k} \dashrightarrow G$ satisfying

$$\gamma(0) = 1 \in G, \quad \gamma'(t) = \gamma(t)v \tag{62}$$

for all t in the domain of γ. Moreover, the values of γ vary smoothly with the initial data v. A priori, the domain of γ might be quite small, but we can now use the group structure on G, as follows, to enlarge it to all of $\mathbf{k}$. To that end, it suffices to show that each solution γ to (62) on some ball[4] B with the center the origin can be extended to a solution domain the enlarged ball $2B$ of twice the radius of B; iterating this, one eventually obtains a solution on all of $\mathbf{k}$. (Many variations on this argument are also possible.)

(a) For $s \in B$, the curve $\gamma_s(t) := \gamma(s)^{-1}\gamma(s+t)$ is defined whenever $s+t \in B$ and satisfies the same initial condition $\gamma_s(0) = 1$ and differential equation $\gamma_s'(t) = \gamma(s)^{-1}\gamma'(s+t) = \gamma(s)^{-1}\gamma(s+t)v = \gamma_s(t)v$ as $\gamma(t)$ does. By the uniqueness theorem cited above, we deduce that $\gamma_s(t) = \gamma(t)$ and hence that

$$\gamma(s+t) = \gamma(s)\gamma(t) \text{ provided that } s, t, s+t \in B. \tag{63}$$

(b) For any $s \in B$, denote by δ_s the curve in G with domain $B + s := \{t : t - s \in B\}$ given by $\delta_s(t) := \gamma(t-s)\gamma(s)$. By differentiating the condition (63) , one sees that δ_s defines a solution to (62) on its domain. For any two $s_1, s_2 \in B$, the identity (63) implies that $\delta_{s_1} = \gamma = \delta_{s_2}$ on the neighborhood $B + s_1 \cap B \cap B + s_2$ of the origin, hence by the uniqueness theorem cited above also that $\delta_{s_1} = \delta_{s_2}$ on $B + s_1 \cap B + s_2$. By the smoothness of the group operation in (63), we see that the values of δ_s still vary smoothly with v.

[4]$\mathbf{k} = \mathbb{R}$, "ball" means "interval", of course

(c) Since $\cup\{B + s : s \in B\} = 2B$, we can patch together the solutions given in the previous step to obtain a well-defined curve $\tilde{\gamma} : 2B \to G$ given for $t \in B + s$ by $\tilde{\gamma}(t) := \delta_s(t - s)$. This curve solves (62), and extends γ, as required.

To summarize the above discussion, we have the following:

Theorem 79. *Let G be a Lie group with Lie algebra $\mathfrak{g} := \mathrm{Lie}(G)$. For each $x \in \mathfrak{g}$ there exists a unique one-parameter subgroup Φ_x of G for which $\Phi_x'(0) = x$. Moreover, the map $x \mapsto \Phi_x(1)$ is smooth.*

13.3. Definition and basic properties of exponential map.

Definition 80. Let G be a Lie group with Lie algebra $\mathfrak{g}$. The *exponential map* $\exp : \mathfrak{g} \to G$ is defined by $\exp(x) := \Phi_x(1)$, where Φ_x denotes the unique one-parameter subgroup of G having initial velocity $\Phi_x'(0) = x$.

The notation is consistent with that discussed in §13.1. We also have the following immediate consequences of §13.2:

Theorem 81 (Lie's first theorem)**.** *For a Lie group G with Lie algebra $\mathfrak{g}$, the exponential map $\exp : \mathfrak{g} \to G$ has the following properties:*

(1) $\exp : \mathfrak{g} \to G$ *is smooth*
(2) The derivative $T_0 \exp : \mathfrak{g} \to \mathfrak{g}$ is the identity transformation, or equivalently, $\frac{d}{dt} \exp(tx) = x$ for all $x \in \mathfrak{g}$.
(3) $\exp : \mathfrak{g} \to G$ is a local diffeomorphism at 0, i.e., there are open $0 \in U \subseteq \mathfrak{g}$ and $1 \in V \subseteq G$ so that $\exp : U \to V$ is a diffeomorphism.
(4) For any $X \in \mathfrak{g}$, the map $\mathbf{k} \ni t \mapsto \exp(tx)$ is the unique one-parameter subgroup of G with initial velocity x.
(5) Let γ be any curve in G with basepoint given by the identity. Set $X := \gamma'(0) \in \mathfrak{g}$. Then $\exp(X) = \lim_{n\to\infty} \gamma(1/n)^n$.

(The proof of the final assertion proceeds exactly as in the case of linear Lie groups now that we have the logarithm map on general Lie groups at our disposal.)

Here is another key consequence:

Corollary 82. $\exp(\mathfrak{g})$ *generates the connected component G^0 of G. In particular, if G is connected, then G is generated by* $\exp(\mathfrak{g})$.

Proof. We saw earlier that G^0 is generated by any neighborhood of the identity, and saw just now that exp is a local diffeomorphism at 0; in particular, $\exp(\mathfrak{g})$ contains a neighborhood of the identity in G. $\square$

Remark 83. Even if G is connected, it need not be the case that $\exp(\mathfrak{g}) = G$. For example, one can show that

$$\begin{pmatrix} -1/2 & \\ & -2 \end{pmatrix} \notin \exp(\mathfrak{sl}_2(\mathbb{R})).$$

It is a non-obvious fact (which we might conceivably see later in the course) that if G is a compact connected Lie group, then $\exp(\mathfrak{g}) = G$.

13.4. **Application to detecting invariance by a connected Lie group.** In lecture, we explained how the connectedness of $\mathrm{SO}(n)$ and Lie's first theorem imply that a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ is rotation-invariant (i.e., $f(x)$ depends only upon $|x|$) if and only if it satisfies the finite system of homogeneous linear differential equations

$$x_j \frac{\partial}{\partial x_i} f(x) = x_i \frac{\partial}{\partial x_j} f(x) \quad \text{for all } 1 \leq i < j \leq n. \tag{64}$$

This is computationally useful; for instance, it applies when f is a polynomial of large degree in many variables. This is not an earth-shaking fact, and it can probably be proved directly in various ways, but the proof we will give here illustrates in a simple way a rather fundamental technique of Lie theory.

To summarize the proof, we noted that the first condition is visibly equivalent to

$$f(gx) = f(x) \text{ for all } g \in \mathrm{SO}(n), x \in \mathbb{R}^n \tag{65}$$

while the second is visibly equivalent to

$$Xf = 0 \text{ for all } X \in \mathcal{B} \tag{66}$$

where for $X \in \mathfrak{g}$ we set

$$Xf(x) := \frac{d}{d\varepsilon} f(\exp(-\varepsilon X)x)|_{\varepsilon=0}$$

and where $\mathcal{B}$ is the basis of $\mathfrak{so}(n)$ given by

$$\mathcal{B} := \{X_{ij} : 1 \leq i < j \leq n\}$$

where

$$X_{ij} := E_{ij} - E_{ji}$$

where E_{ij} has a 1 in the (i, j) entry and 0's elsewhere; for instance, when $n = 3$, a basis for $\mathfrak{so}(3)$ is given by

$$X_{12} \begin{pmatrix} & 1 & \\ -1 & & \\ & & \end{pmatrix}, \quad X_{13} \begin{pmatrix} & & 1 \\ & & \\ -1 & & \end{pmatrix}, \quad X_{23} \begin{pmatrix} & & \\ & & 1 \\ & -1 & \end{pmatrix}.$$

To relate (64) to (66) we used that

$$X_{ij} x = x_j e_i - x_i e_j$$

and hence that

$$f(x \exp(-\varepsilon X_{ij})x) = f(x - \varepsilon(x_j e_i - x_i e_j) + O(\varepsilon^2))$$

to compute that

$$X_{ij} f(x) = -(x_j \frac{\partial}{\partial x_i} - x_i \frac{\partial}{\partial x_j}) f(x).$$

Each of the following conditions is visible equivalent to the next:

(1) The condition (65).
(2) The condition (65) but restricted to g in a generating set for G.
(3) The condition (65) for $g \in \exp(\mathfrak{g})$. (Here we use Lie's theorem and the connectedness of G.)
(4) The condition that
$$f(\exp(-tX)x)$$
be independent of t for all $x \in \mathbb{R}^n$ and all $X \in \mathfrak{g}$.

(5) The condition that
$$\frac{d}{dt} f(\exp(-tX)x) = 0 \tag{67}$$
for all $x \in \mathbb{R}^n$ and all $X \in \mathfrak{g}$.

(6) The condition that $Xf = 0$ for all $X \in \mathfrak{g}$. To relate this to the previous condition, we used the following key calculation:
$$\begin{aligned}\frac{d}{dt} f(\exp(-tX)x) &= \frac{d}{d\varepsilon} f(\exp(-(t+\varepsilon)X)x)|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} f(\exp(-\varepsilon X)\exp(-tX)x)|_{\varepsilon=0} \\ &= Xf(\exp(-tX)x).\end{aligned}$$
Thus if (67) holds, then $Xf(\exp(-tX)x) = 0$ for all t, X, x, and in particular for $t = 0$, giving $Xf(x) = 0$ and thus $Xf = 0$. Conversely, if $Xf = 0$, then in particular $Xf(\exp(-tX)x) = 0$ for all t, X, x; the above calculation then implies that (67) holds.

(7) The condition (66). (Here we used that $X \mapsto Xf$ is linear to reduce from testing all $X \in \mathfrak{g}$ to testing those in the basis $\mathcal{B}$.)

Remark 84. It's not too hard to show that $\exp : \mathfrak{so}(n) \to \mathrm{SO}(n)$ is actually surjective (in contrast to the general case mentioned in Remark 83), so we didn't *really* need Lie's theorem in the above argument, the point of which was to illustrate a technique in a simple case.

13.5. Connected Lie subgroups are determined by their Lie algebras.

Theorem 85. *Let G be a Lie group. Then connected Lie subgroups of G are classified by their Lie algebra: if H_1, H_2 are two connected Lie subgroups of G for which* $\mathrm{Lie}(H_1) = \mathrm{Lie}(H_2)$*, then $H_1 = H_2$.*

Proof. We know by Exercise 5 that H_i is generated by any neighborhood of the identity; by Theorem 81, it follows that H_i is generated by $\exp(\mathrm{Lie}(H_i))$. Since H_1, H_2 are generated by the same set, they are equal. □

13.6. The exponential map commutes with morphisms.

Theorem 86. *The exponential map commutes with Lie group morphisms: If $f : G \to H$ is a morphism of Lie groups and $x \in \mathrm{Lie}(G)$ is given, then $f(\exp(x)) = \exp(df(x))$, where $df := T_e f : \mathfrak{g} \to \mathfrak{h}$.*

For the sake of illustration, we record a few proofs.

Proof #1. It suffices to show that $f(\exp(tx)) = \exp(df(tx))$ for all t. But now both sides, viewed as functions of t, are one-parameter subgroups of H with inital velocity $df(x) = T_e f(x)$: indeed, we have
$$f(\exp(tx)) = f(1 + tx + o(t)) = 1 + t\,df(x) + o(t)$$
and similarly
$$\exp(df(tx)) = 1 + t\,df(x) + o(t).$$
By the uniqueness given in Theorem 79, we conclude. □

Proof #2. Let γ be any curve on G with basepoint $\gamma(0) = e$ and initial velocity $\gamma'(0) = x$. The curve $f \circ \gamma$ on H then has basepoint e and initial velocity $(f \circ \gamma)'(0) = df(x)$. By the final assertion of Theorem 81, we have

$$\exp(x) = \lim_{n \to \infty} \gamma(1/n)^n$$

and similarly $\exp(df(x)) = \lim_{n \to \infty} f(\gamma(1/n))^n$. Since f is a continuous group homomorphism, it follows that

$$f(\exp(x)) = f(\lim_{n \to \infty} \gamma(1/n)^n) = \lim_{n \to \infty} f(\gamma(1/n))^n = \exp(df(x)),$$

as required. □

Example 87. Set $f := \det : \mathrm{GL}_n(\mathbf{k}) \to \mathbf{k}^\times$. Since $\det(1 + \varepsilon X) = 1 + \varepsilon \operatorname{trace}(X) + O(\varepsilon^2)$, we have $df = \operatorname{trace} : M_n(\mathbf{k}) \to \mathbf{k}$. The theorem then implies for $X \in M_n(\mathbf{k})$ that

$$\det(\exp(X)) = f(\exp(X)) = \exp(df(X)) = \exp(\operatorname{trace}(X)),$$

which can also be seen directly via Jordan decomposition.

13.7. **Morphisms out of a connected Lie group are determined by their differentials.**

Theorem 88. *Let G be a connected Lie group and H any Lie group. Then morphisms $f : G \to H$ are determined by their differentials $df : \mathfrak{g} \to \mathfrak{h}$.*

Proof. Since G is connected, it is generated by a neighborhood of the identity, hence in particular by $\exp(\mathfrak{g})$. Since f is a homomorphism, it is thus determined by the quantities $\exp(X)$ for all $X \in \mathfrak{g}$. But by the result of §13.6, we have $f(\exp(X)) = \exp(df(X))$. Hence f is determined by f. □

14. Putting the "algebra" in "Lie algebra"

14.1. **The commutator of small group elements.** Let G be a Lie subgroup of $\mathrm{GL}_n(\mathbf{k})$. Consider a pair of elements $X, Y \in \mathfrak{g} := \operatorname{Lie}(G)$ and a corresponding pair of curves $\xi, \eta : \mathbf{k} \dashrightarrow G$ with basepoints $\xi(0) = \eta(0) = 1$ and initial velocities $\xi'(0) = X, \eta'(0) = Y$. For example:

(1) One can always take $\xi(s) := \exp(sX), \eta(t) := \exp(tY)$.
(2) If $G = \mathrm{GL}_n(\mathbf{k})$, one could also take $\xi(s) := 1 + sX, \eta(t) := 1 + tY$.

We can then consider, for small enough $s, t \in \mathbf{k}$, the commutator

$$\Gamma(s,t) := (\xi(s), \eta(t)) := \xi(s)\eta(t)\xi(s)^{-1}\eta(t)^{-1}.$$

From the basepoint condition we have

$$\Gamma(0,0) = \Gamma(s,0) = \Gamma(0,t) = 1,$$

so every term in the Taylor expansion of $\Gamma(s,t) - 1$ is divisible by st. We now determine the coefficient of st, or equivalently, the derivative $\partial_{s=0}\partial_{t=0}\Gamma(s,t)$. To compute this, we write

$$\xi(s)\eta(t) = \Gamma(s,t)\eta(t)\xi(s)$$

and differentiate both sides, first with respect to s at $s = 0$, giving

$$X\eta(t) = \partial_{s=0}\Gamma(s,t)\eta(t)\xi(0) + \Gamma(0,t)\eta(t)X = \partial_{s=0}\Gamma(s,t)\eta(t) + \eta(t)X$$

and then with respect to t at $t = 0$, giving after analogous simplifications that

$$XY = \partial_{s=0}\partial_{t=0}\Gamma(s,t) + YX,$$

whence

$$\Gamma(s,t) = XY - YX.$$

14.2. **The Lie bracket as an infinitesimal commutator.** For a general Lie group G, we define the commutator bracket $[,] : \mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$ by setting

$$[X,Y] := \partial_{s=0}\partial_{t=0}\Gamma(s,t),$$

with notation as in the previous section. To interpret this properly, we have

$$\partial_{t=0}\Gamma(s,t) = \partial_{t=0}\xi(s)\eta(t)\xi(s)^{-1}\eta(t)^{-1}$$

This is the initial velocity of a curve passing through the identity element of G at time $t = 0$, hence it makes sense to regard it as an element of $\mathfrak{g}$. Thus

$$s \mapsto \partial_{t=0}\Gamma(s,t)$$

defines a curve in $\mathfrak{g}$. Hence its s-derivative at $s = 0$ defines an element of $\mathfrak{g}$. Similar arguments show that $[X,Y]$ is independent of the choice of ξ, η; alternatively, one could always take $\xi(s) := \exp(sX), \eta(t) := \exp(tY)$ in the definition, but it's occasionally convenient to make other choices.

The bracket $[,]$ on the Lie algebra $\mathfrak{g}$ of a Lie group G has the following properties:

(1) $[,]$ is bilinear. This is immediate from the definition
(2) $[X,X] = 0$. This is again immediate from the definition. It follows that $[X+Y,X+Y] = [X,X]+[X,Y]+[Y,X]+[Y,Y] = [X,Y]+[Y,X]$, hence that
$$[X,Y] = -[Y,X]. \tag{68}$$
(3) It satisfies the Jacobi identity
$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \tag{69}$$
which, thanks to (68), can be put in the equivalent forms
$$[X,[Y,Z]] = [[X,Y],Z] + [Y,[X,Z]]$$
or
$$[[X,Y],Z] = [[X,Z],Y] + [X,[Y,Z]].$$
In the linear case $G \leq \mathrm{GL}_n(\mathbf{k})$, one can prove the Jacobi identity by expanding everything out using the identity $[X,Y] = XY - YX$. In general, they follow from the associativity of the group law in G in the form
$$gh = (ghg^{-1})g$$
together with some artful use of the chain rule. We do not give the details here; we promise instead that a couple more "conceptual" derivations of the Jacobi identity will be given later.

14.3. **Lie algebras.**

Definition 89. A *Lie algebra* L over $\mathbf{k}$ is a vector space equipped with a bilinear form $[,] : L \otimes L \to L$ satisfying the properties mentioned in the previous section.

Here are the basic examples:

(1) We've seen (modulo verification of the Jacobi identity in general) that for any Lie group G, what we've already called the "Lie algebra" $\mathfrak{g} := \mathrm{Lie}(G)$ of G is in fact a Lie algebra in the above sense when equipped with the commutator bracket as we've defined it.

(2) Any associative algebra A over $\mathbf{k}$ becomes a Lie algebra when equipped with the bracket $[x, y] := xy - yx$. A notable example is when $A = \mathrm{End}(V)$ for a vector space V. If $V = \mathbf{k}^n$, then of course $A = \mathrm{End}(V) = M_n(\mathbf{k})$.
(3) If L_1 is a Lie algebra and $L_2 \leq L_1$ is a vector subspace with the property that $[x, y] \in L_2$ whenever $x, y \in L_2$, then L_2 is a Lie algebra when equipped with the commutator bracket induced from that on L_1; it is then (fittingly) called a *Lie subalgebra* of L_1.
(4) Given an algebra A, the space $\mathrm{Der}(A)$ of *derivation* of A (i.e., $\mathbf{k}$ -linear maps $D : A \to A$ satisfying $D(x \cdot y) = Dx \cdot y + x \cdot Dy$) is a Lie algebra. (Exercise: check this.) It is a Lie subalgebra of $\mathrm{End}(A)$.

    When A is finite-dimensional, one can show that $\mathrm{Aut}(A)$ is a Lie group with Lie algebra $\mathrm{Der}(A)$, hence that this example is a special case of the first one. But what we've said here applies (usefully) also when A is infinite-dimensional; see below.

Every finite-dimensional example is already of the above form:

Theorem 90 (Ado)**.** *Let L be a finite-dimensional Lie algebra. Then L is isomorphic to a Lie subalgebra of* $\mathrm{End}(V)$ *for some finite-dimensional vector space V.*

The proof of this innocent-sounding theorem is not egregiously difficult, but does seem to require most of the basic structure theory of Lie algebras, and so will not be proved now. However, it may aid intuition to know up front that one can always think of any finite-dimensional Lie algebra as a Lie subalgebra of some matrix algebra.

A special case of the final example mentioned above is when $A = C^\infty(M)$ for a manifold M. In that case, it is known that

$$\mathrm{Der}(C^\infty(M)) \cong \mathrm{Vect}(M)$$

where $\mathrm{Vect}(M)$ denotes the space of vector fields on M, i.e., smooth assignments

$$X : M \to TM := \sqcup_{p \in M} T_p M$$

satisfying $X_p := X(p) \in T_p M$ for all $p \in M$. Such a vector field induces a derivation by the rule: for $f \in C^\infty(M)$, the image $Xf \in C^\infty(M)$ under $X \in \mathrm{Vect}(X)$ is defined to be

$$(Xf)(p) := (T_p f)(X_p),$$

i.e., "the directional derivative of f at p in the direction of the tangent vector X_p."

Remark 91. It can be instructive to check for some simple examples of linear Lie groups $G \leq \mathrm{GL}_n(\mathbf{k})$ with Lie algebra $\mathfrak{g} \leq M_n(\mathbf{k})$ that the bracket $[,]$ does in fact preserve $\mathfrak{g}$ (as it must). For $X, Y \in \mathfrak{sl}_n(\mathbf{k})$ we have $\mathrm{trace}([X, Y]) = 0$; indeed, the trace of any commutator is zero. For $X, Y \in \mathfrak{o}_n(\mathbf{k})$, so that $X + X^t = 0, Y + Y^t = 0$, we have $[X, Y]^t = (XY)^t - (YX)^t = Y^t X^t - X^t Y^t = (-Y)(-X) - (-X)(-Y) = -[X, Y]$, hence $[X, Y] \in \mathfrak{o}_n(\mathbf{k})$.

14.4. **The Lie bracket commutes with differentials of morphisms.** Let $f : G \to H$ be a morphism of Lie groups. Then for $X, Y \in \mathfrak{g}$, one has

$$\begin{aligned} df([X,Y]) &= df(\partial_{s=0}\partial_{t=0}(e^{sX}, e^{tY})) && \text{(definition of } [,]) \\ &= \partial_{s=0} df(\partial_{t=0}(e^{sX}, e^{tY})) && (df \text{ is linear}) \\ &= \partial_{s=0}\partial_{t=0} f((e^{sX}, e^{tY})) && \text{(definition of } df) \\ &= \partial_{s=0}\partial_{t=0}(f(e^{sX}), f(e^{tY})) && (f \text{ is a homomorphism}) \\ &= \partial_{s=0}\partial_{t=0}(e^{df(sX)}, e^{t\,df(sY)}) && (f \circ \exp = \exp \circ df) \\ &= \partial_{s=0}\partial_{t=0}(e^{s\,df(X)}, e^{t\,df(Y)}) && (df \text{ is linear}) \\ &= [df(X), df(Y)] && \text{(definition of } [,]). \end{aligned}$$

Thus $df : \mathfrak{g} \to \mathfrak{h}$ is a morphismo f Lie algebras.

15. How pretend that every Lie group is a matrix group and survive

(TODO: rewrite this section.) For many arguments it is convenient to assume that a Lie group G is a matrix group, i.e., embeds in $\mathrm{GL}_n(\mathbb{R})$, so that its Lie algebra embeds in $M_n(\mathbb{R})$. Then lots of stuff simplifies (in a non-serious way) because we can just regard everything as a matrix and not worry about which tangent space stuff belongs to, etc. Not every Lie group is a matrix group, but they are all close enough to being matrix groups (e.g., up to covering homomorphisms to be discussed later) that nothing really bad goes wrong if one pretends that they are. For example, it was convenient in class today to pretend that G was a matrix Lie group when discussing the proof of Maurer–Cartan equations.

However, there is a rigorous trick by which one can always treat a Lie group G as if it were a matrix group by embedding it in the space $\mathrm{GL}(C^\infty(G))$ of linear automorphisms of the (typically infinite-dimensional) vector space $C^\infty(G)$. The fact that $\mathrm{GL}(C^\infty(G))$ is not itself a Lie group doesn't matter much in practice. More precisely, one defines an injective homomorphism

$$G \hookrightarrow \mathrm{GL}(C^\infty(G))$$

as follows: we identify each $g \in G$ with the element of $\mathrm{Aut}(C^\infty(G))$ that sends a smooth function $\varphi \in C^\infty(G)$ to the new function $g\varphi \in C^\infty(G)$ given by right translation: for $x \in G$,

$$g\varphi(x) := \varphi(xg).$$

(This is an action: $(g_1 g_2)\varphi = g_1(g_2\varphi)$.) It makes sense to differentiate this action of G element-wise. We obtain in this way induces a morphism $X \hookrightarrow \mathrm{End}(C^\infty(G))$, whose image actually lies in an easily characterized subspace of $\mathrm{Der}(C^\infty(G))$; more on that later. The action of $X \in \mathfrak{g}$ on $\varphi \in C^\infty(G)$ is given by

$$X\varphi(x) := \partial_{t=0}\varphi(x\exp(tX)).$$

(This is a Lie algebra representation: $[X,Y]\varphi = XY\varphi - YX\varphi$.) In this way, one can make perfectly rigorous sense of identities such as

$$[X,Y] = XY - YX$$

or

$$\mathrm{Ad}(g)X = gXg^{-1}$$

even when G is not a matrix Lie group: for instance, the products XY in the above expression are just the compositions taking place inside $\mathrm{End}(C^\infty(G))$.

16. Something about representations, mostly SL_2

16.1. **Some preliminaries.** We have spoken so far in the course quite a bit about $\mathrm{GL}_n(\mathbf{k})$ and its Lie algebra $M_n(\mathbf{k})$. More abstractly, one can work with any finite-dimensional vector space V over $\mathbf{k}$. Then $\mathrm{GL}(V)$ is a Lie group over $\mathbf{k}$ with Lie algebra $\mathrm{End}(V)$. If $V = \mathbf{k}^n$, then $\mathrm{GL}(V) = \mathrm{GL}_n(\mathbf{k})$ and $\mathrm{End}(V) = M_n(\mathbf{k})$.

When $\mathbf{k} = \mathbb{C}$, we can regard $\mathrm{GL}(V)$ either as a complex Lie group or as a real Lie group.

16.2. **Definition.**

Definition 92. Let $\mathbf{k} = \mathbb{R}$ or $\mathbb{C}$. For us, a *representation* of a Lie group G is a pair (V, R), where

- V is a finite-dimensional vector space and
- $R : G \to \mathrm{GL}(V)$ is a morphism of Lie groups over $\mathbf{k}$.

We allow the possibility that $\mathbf{k} = \mathbb{R}$ and V is a complex vector space but regarded as a real vector space ($\mathbb{C}^n \cong \mathbb{R}^{2n}$). When we wish to be more specific, we might introduce the following terminology:

(1) Let G be a real Lie group. A *real representation* of G is a morphism of real Lie groups $R : G \to \mathrm{GL}(V)$ (i.e., an infinitely-real-differentiable group homomorphism).
(2) Let G be a real Lie group. A *complex representation* of G is a morphism of real Lie groups $R : G \to \mathrm{GL}(V)$ (i.e., an infinitely-real-differentiable group homomorphism).
(3) Let G be a complex Lie group. A *holomorphic representation* of G is a morphism of complex Lie groups $R : G \to \mathrm{GL}(V)$ (i.e., a complex-differentiable group homomorphism).

One can also regard a complex Lie group as a real Lie group and consider its representations in that sense.

A *representation* of a Lie algebra $\mathfrak{g}$ is likewise a pair (V, ρ), where V is as above and $\rho : \mathfrak{g} \to \mathrm{End}(V)$ is a morphism of Lie algebras over $\mathbf{k}$. We can also speak of real or complex representations of Lie algebras, or of holomorphic representations of complex Lie algebras.

The action of a representation $R : G \to \mathrm{GL}(V)$ is often abbreviated $gv := R(g)v$ and likewise that of $\rho : \mathfrak{g} \to \mathrm{End}(V)$ by $Xv := \rho(X)v$.

Given two representations $R_1 : G \to \mathrm{GL}(V_1)$ and $R_2 : G \to \mathrm{GL}(V_2)$, a *morphism of representations* or *equivariant map* $\Phi : V_1 \to V_2$ is a linear map that commutes with the action, i.e., so that $\Phi(R_1(g)v) = R_2(g)\Phi(v)$ for all $g \in G$, $v \in V$; one defines similarly the analogous notion for $\mathfrak{g}$-representations. An *isomorphism of representations* or *equivariant isomorphism* is a morphism with a two-sided inverse (equivalently, a bijective morphism), and two representations are said to be *isomorphic* if there is an isomorphism between them.

By what we've seen above, a representation $R : G \to \mathrm{GL}(V)$ of a Lie group G induces a representation

$$dR : \mathfrak{g} \to \mathrm{End}(V)$$

of its Lie algebra $\mathfrak{g}$, given explicitly by for $X \in \mathfrak{g}$ by

$$Xv := dR(X)v := \frac{d}{dt} R(\exp(tX))v|_{t=0}.$$

Example 93. Let $G := \mathrm{GL}_n(\mathbf{k})$. Let $V := \mathbb{C}[x_1, \dots, x_n]_{(d)}$ be the space of homogeneous polynomials of degree d in the variables $x_1, \dots, x_n$. One then has a representation $R : G \to \mathrm{GL}(V)$ sending $g \in G$ to the element $R(g) \in \mathrm{GL}(V)$ that acts on a polynomial $\phi \in V$ by the formula

$$g\phi(x) := (R(g)\phi)(x) := \phi(xg),$$

where $x = (x_1, \dots, x_n)$ is regarded as an n-tuple of variables and xg denotes the right multiplication of the matrix g against the row vector x. This is already an interesting representation. The differential $dR : \mathfrak{g} \to \mathrm{End}(V)$ is given on the standard basis elements E_{ij} of $\mathfrak{g} = \mathfrak{gl}_n(\mathbf{k})$

$$E_{ij}\phi(x) := (dR(E_{ij})\phi)(x) = x_i \frac{\partial}{\partial_{x_j}} \phi(x).$$

(To see this, note that $xE_{ij} = x_i e_j$ and thus $\phi(x(1 + \varepsilon E_{ij})) = \phi(x) + x_i \frac{\partial \phi}{\partial x_j}(x)\varepsilon + O(\varepsilon^2)$.) The same definition makes sense and similar considerations apply more generally when G is any subgroup of $\mathrm{GL}_n(\mathbf{k})$.

16.3. **Matrix multiplication.** Let $R : G \to \mathrm{GL}(V)$ be a representation of a Lie group G. Fixing a basis (v_i) for V, one can express a representation of G in matrix form

$$R(g) = (R_{ij}(g))_{i,j},$$

where $R_{ij}(g)$ denotes the coefficient of the basis element v_i in $R(g)v_j$. It's a fact of life that pretty much every special function of mathematics or physics is of the form $R_{ij}(g)$. Identities such as the consequence

$$\sum_j R_{ij}(g) R_{jk}(h) = R_{ik}(gh)$$

of the homomorphism property $R(g)R(h) = R(gh)$ can be of use. For example, let $G := \mathbb{R}, V := \mathbb{R}^2$,

$$R : G \to \mathrm{GL}_2(\mathbb{R})$$

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

Then

$$\cos(\theta+\phi) = R_{11}(\theta+\phi) = R_{11}(\theta)R_{11}(\phi) + R_{12}(\theta)R_{21}(\phi) = \cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi),$$

which makes for a nice way to remember addition laws for trigonometric functions.

16.4. **Invariant subspaces and irreducibility.**

Definition 94. Let G be a Lie group, and let $R : G \to \mathrm{GL}(V)$ be a finite-dimensional representation of G. A subspace W of V is said to be *invariant* (or *stable* or G-*invariant* or G-*stable*) if $R(g)W \subseteq W$ for all $g \in G$.

Similarly, given a representation $\rho : \mathfrak{g} \to \mathrm{End}(V)$ of a Lie algebra $\mathfrak{g}$, we say that a subspace W of V is *invariant* (or *stable* or $\mathfrak{g}$-*invariant* or $\mathfrak{g}$-*stable*) if $\rho(X)W \subseteq W$ for all $X \in \mathfrak{g}$.

We say that a representation (R, V) of a Lie group or a representation (ρ, V) of a Lie algebra is *irreducible* if $V \neq \{0\}$ and if V has no nonzero proper invariant subspaces (i.e., none other than $\{0\}$ and V). Otherwise, it is said to be *reducible*.

Exercise 18. Let G be a Lie group and $R : G \to \mathrm{GL}(V)$ an n-dimensional representation. Fix a basis of V and hence an identification $V := \mathbb{C}^n$. Let m be an integer satisfying $0 < m < n$, and let $W := \mathbb{C}^m$ regarded as a subspace of V via the standard inclusion $(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_m, 0, \ldots, 0)$. Denote by $P_m(V)$ the subgroup of $\mathrm{GL}(V)$ given by matrices of the form $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$, where a is an $m \times m$ matrix, b is an $m \times (n-m)$ matrix, and d is an $(n-m) \times (n-m)$ matrix. Show that the following are equivalent:

(1) R is reducible.
(2) There exist $0 < m < n$ and $\gamma \in \mathrm{GL}(V)$ so that $R(G) \subseteq \gamma P_m(V) \gamma^{-1}$.

Theorem 95. *Let G be a Lie group, and let $R : G \to \mathrm{GL}(V)$ be a finite-dimensional representation of G.*

(1) Any G-invariant subspace of V is also $\mathfrak{g}$-invariant.
(2) If G is connected, then any $\mathfrak{g}$-invariant subspace of V is also G-stable.
(3) If G is connected, then V is irreducible if and only if it is nonzero and contains no proper $\mathfrak{g}$-stable subspaces.

Proof. If $W \leq V$ is G-invariant, then for each $X \in \mathfrak{g}$ and $v \in W$ and $t \in \mathbb{R}$, we have

$$\frac{R(\exp(tX))v - v}{t} \in W$$

(because W is a vector space), hence upon differentiating that

$$dR(X)v \in V$$

(because W is closed). This shows that W is $\mathfrak{g}$-invariant. Conversely, if $W \leq V$ is $\mathfrak{g}$-invariant and G is connected, then

$$R(\exp(tX))v = \exp(t\, dR(X))v = \sum_{n \geq 0} \frac{t^n}{n!} dR(X)^n v \in W,$$

hence W is $\exp(\mathfrak{g})$-invariant, hence (because G is connected and thus generated by $\exp(\mathfrak{g})$) W is G-invariant. Etc. $\square$

16.5. **Polynomial representations of** $\mathrm{SL}_2(\mathbb{C})$**.** Here we specialize Example 93 to $G = \mathrm{SL}_2(\mathbb{C})$. Let W_n denote the space of homogeneous polynomials $\phi \in \mathbb{C}[x, y]$ of degree n. Then W_n is an $(n+1)$-dimensional vector space with basis given by the monomials $x^n, x^{n-1}y, \ldots, xy^{n-1}, y^n$. As motivation, one can check that for $g = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, the matrix coefficients $R_{ij}(g)$ of this representation with respect to the above basis give the classical spherical polynomials (e.g., when n is even, the coefficient of $x^n y^n$ in $R(g)x^n y^n$ is essentially the Legendre polynomial $P_{n/2}(\cos\theta)$).

By specializing the calculation of Example 93, we see that the basis elements

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

of $\mathfrak{g}$ act by

$$dR(X) = x\partial_y, \quad dR(Y) = y\partial_x, \quad dR(H) = x\partial_x - y\partial_y.$$

Their effects on the basis elements is thus given by $Xx^n = 0, Yy^n = 0$ and in all other cases by

$$Xx^{n-k}y^k = kx^{n-k+1}y^{k-1}, \quad Yx^{n-k}y^k = (n-k)x^{n-k-1}y^{k+1}, \quad Hx^{n-k}y^k = (n-2k)x^{n-k}y^k.$$

In lecture, we drew a picture in which the basis elements $y^n, xy^{n-1} \ldots, x^n y, x^n$ were lined up from left to right and indicated by circles in which we indicated their H-eigenvalues $-n, -n+2, \ldots, n-2, n$. The action of X may be depicted

$$y^n \xrightarrow{n} xy^{n-1} \xrightarrow{n-1} x^2y^{n-2} \xrightarrow{n-2} \cdots \xrightarrow{2} x^{n-1}y \xrightarrow{1} x^n \to 0.$$

The action of Y may be depicted

$$0 \leftarrow y^n \xleftarrow{1} xy^{n-1} \xleftarrow{2} x^2y^{n-2} \xleftarrow{3} \cdots \xleftarrow{n-1} x^{n-1}y \xleftarrow{n} x^n.$$

Explicitly, when $n = 3$, we may represent the various actions with respect to the basis x^3, x^2y, xy^2, y^3 by

$$dR(X) = \begin{pmatrix} & 1 & & \\ & & 2 & \\ & & & 3 \\ & & & \end{pmatrix},$$

$$dR(Y) = \begin{pmatrix} & & & \\ 3 & & & \\ & 2 & & \\ & & 1 & \end{pmatrix},$$

$$dR(H) = \begin{pmatrix} 3 & & & \\ & 1 & & \\ & & -1 & \\ & & & -3 \end{pmatrix}.$$

As we saw in Homework 4, the relation

$$[X, Y] = H$$

implies that

$$[dR(X), dR(Y)] = dR(H).$$

It is an instructive exercise to verify this directly from as many perspective as possible (e.g., by direct inspection of the action, by staring at the action on basis vectors using the graph-theoretic depiction described above, by explicitly computing the commutators of the above 4×4 matrices, etc.).

Theorem 96. *W_n is an irreducible representation of $G = \mathrm{SL}_2(\mathbb{C})$.*

By Theorem 95, it is equivalent to show that W_n is irreducible as a representation of $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$. There are a couple ways to show this. Firstly, given any nonzero invariant subspace W of W_n and take a nonzero element $v \in W$, then it follows from the above description of the action that there is a $k \geq 0$ for which $X^{k+1}v = 0$, and moreover, that if k is the smallest integer with this property, then $X^k v$ is a nonzero multiple of x^n; then $Y^m X^k v$ is a nonzero multiple of $x^{n-k}y^k$. Since W is invariant, it contains $Y^m X^k v$, hence contains all the basis elements for W_n, and so $W = W_n$, i.e., W_n is irreducible.

Another way to structure part of the argument is to use the following elementary consequence of the invertibility of the Vandermonde determinant:

Lemma 97. *If V is a representation of $\mathfrak{g}$ and W is an invariant subspaces and $v \in W$ is a vector that may be expressed as a sum $v = v_1 + \cdots + v_n$ where $Hv_i = \lambda_i v_i$ for some $\lambda_i \in \mathbb{C}$ with $\lambda_i \neq \lambda_j$ whenever $i \neq j$, then each v_i also belongs to W.*

This shows that any invariant subspace W of W_n contains the components of each of its vectors wrt the standard basis, and one can then argue as above to get all the basis elements.

16.6. **Classifying finite-dimensional irreducible representations of** $\mathrm{SL}_2(\mathbb{C})$**.** One cares to do this because it shows up all over the place (in studying special functions, in classifying Lie groups and Lie algebras, in studying representations of other Lie groups thanks to the various ways that $\mathrm{SL}_2(\mathbb{C})$ may be embedded in them, in quantum mechanics, Hodge theory, etc.)

Theorem 98. *Let V be any finite-dimensional irreducible representation of $G = \mathrm{SL}_2(\mathbb{C})$. Then V is isomorphic to one of the representations W_n considered in the previous section for some $n \geq 0$.*

By arguing as in the proof of Lemma 95, it suffices to show this for $\mathfrak{g}$-representations instead of G-representations, which makes the problem a bit easier.

Lemma 99. *Let T be a linear transformation on a nonzero finite-dimensional complex vector space V. Then T has an eigenvector, i.e., a nonzero vector $v \in V$ so that $Tv = \lambda v$ for some $\lambda \in \mathbb{C}$.*

Proof. The characteristic polynomial $\det(x - T)$ is monic of degree $\dim(V) \geq 1$, hence has a root λ; then $\det(\lambda - T) = 0$, so $T - \lambda$ is non-invertible, so $\ker(T - \lambda) \neq 0$, i.e., T has an eigenvector. □

Definition 100. Let V be a representation of $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$, and let $\lambda \in \mathbb{C}$. We say that a nonzero vector $v \in V$ has *weight* λ if v is an eigenvector for H with eigenvalue λ, i.e., $Hv = \lambda v$.

Example 101. The vector $x^{n-k}y^k \in W_n$ has weight $n - 2k$.

Remark 102. In what follows, we write (for instance) HX as an abbreviation for $dR(H)dR(X)$; this differs from the matrix product of H and X, which we shall have no occasion to refer to.

Lemma 103. *Suppose $v \in V$ as above has weight λ. Then Xv has weight $\lambda + 2$ and Yv has weight $\lambda - 2$.*

Proof. We will use that

$$[H, X] = 2X, \quad [H, Y] = -2Y.$$

We have

$$\begin{aligned} HXv &= (HX - XH)v + XHv \\ &= [H, X]v + XHv \\ &= 2Xv + X(\lambda v) \\ &= (\lambda + 2)Xv, \end{aligned}$$

and similarly $HYv = (\lambda - 2)Yv$. □

Lemma 104. *Let V be a finite-dimensional representation of $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$. Then there is a nonzero $v \in V$ and $\lambda \in \mathbb{C}$ so that*

$$Hv = \lambda v,$$
$$Xv = 0.$$

Proof. By Lemma 99, there exists some nonzero $u \in V$ with some weight $\mu \in \mathbb{C}$. The vectors $X^k u$ have weight $\mu + 2k$. Since V is finite-dimensional, H has only finitely many eigenvalues, so we have $X^{k+1}u = 0$ for large enough k. Choosing k minimal with this property and taking $v := X^k u$ gives $Hv = (\mu+2k)v$ and $Xv = 0$, as required. □

Remark 105. Although Lemma 103 is basically trivial, it is one of the most frequently applied calculations in Lie theory, and deserves careful study.

To prove Theorem 98, we now take for V any irreducible finite-dimensional representation of $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ and let $v_0 \in V$ be a nonzero element satisfying

$$Hv = \lambda_0 v, \quad Xv = 0.$$

Such a vector exists by the previous lemma. For $k \geq 0$, set

$$v_k := Y^k v_0.$$

Then v_k has weight $\lambda_0 - 2k$, so the various v_k are all linearly independent. Let $W := \oplus \mathbb{C} v_k$ be the span of the v_k. We claim that W is $\mathfrak{g}$-invariant. To that end, it suffices by the linearity of the action to show for each basis element $Z \in \{H, X, Y\}$ of $\mathfrak{g}$ and basis element v_k of W that $Zv_k \in W$. Clearly

$$Hv_k = (\lambda_0 - 2k)v_k \in W$$

and

$$Yv_k = v_{k+1} \in W.$$

We now verify by induction on k that

$$Xv_k = c_k v_{k-1} \in W$$

with $c_k := k(\lambda_0 - k + 1)$ and (by convention) $v_{-1} := 0$. When $k = 0$, this is clear. For $k \geq 0$, it follows by our inductive hypothesis and using the trick $XY = (XY - YX) + YX$ as in the proof of Lemma 103 that

$$\begin{aligned} Xv_{k+1} &= XYv_k \\ &= [X, Y]v_k + YXv_k \\ &= Hv_k + Yc_k v_{k-1} \\ &= (\lambda - 2k + c_k)v_k. \end{aligned}$$

We conclude by checking that $c_{k+1} = \lambda - 2k + c_k$.

Since W is nonzero and $\mathfrak{g}$-invariant and since V is assumed irreducible, we must have $W = V$. Since V is finite-dimensional, we have $v_{n+1} = 0$ for some n. Choosing n minimal with this property implies that

$$v_n \neq 0$$

and

$$v_{n+1} = 0$$

whence

$$0 = Xv_{n+1} = c_{n+1} v_n = (n+1)(\lambda_0 - n)v_n.$$

Since $n+1 \neq 0$ and $v_n \neq 0$, it follows that $\lambda_0 = n$.

In summary, we have shown that V has the basis $v_0, \dots, v_n$ on which the action is given by

$$\begin{aligned} Hv_k &= (n-2k)v_k, \\ Yv_k &= v_{k+1}, \\ Xv_k &= k(n-k+1)v_{k-1}. \end{aligned}$$

But it is easy to check that W_n has the basis $w_0, \dots, w_n$ with $w_k := Y^k x^n$ on which the action is described in the same way. Thus the linear map $V \to W_n$ extending $v_k \mapsto w_k$ is an isomorphism of $\mathfrak{g}$-representations, as required.

16.7. **Complete reducibility.** It is interesting to ask whether one can classify *all* finite-dimensional representations of a group such as $\mathrm{SL}_2(\mathbb{C})$ rather than just the irreducible ones as done in §16.5. For example, can we break arbitrary representations up into sums of irreducible ones? We can also ask the same question for more general Lie groups G: how can we understand general representations $R : G \to \mathrm{GL}(V)$ in terms of the irreducible ones? This was a basic question of late 19th century mathematics known nowadays as classical invariant theory; the representations of interest were as in Example 93.

Definition 106. Let G be any Lie group and $R : G \to \mathrm{GL}(V)$ a finite-dimensional representation. We say that V is *completely reducible* if there are irreducible invariant subspaces $V_1, \dots, V_n$ of V so that $V = V_1 \oplus \cdots \oplus V_n$.

Definition 107. Similarly, let $\mathfrak{g}$ be a Lie algebra and $\rho : \mathfrak{g} \to \mathrm{End}(V)$ a finite-dimensional representation. We say that V is *completely reducible* if there are irreducible invariant subspaces $V_1, \dots, V_n$ of V so that $V = V_1 \oplus \cdots \oplus V_n$.

Lemma 108. *Let G be a connected Lie group, let $R : G \to \mathrm{GL}(V)$ a finite-dimensional representation, and let $dR : \mathfrak{g} \to \mathrm{End}(V)$ be the induced representation of the Lie algebra $\mathfrak{g}$ of G. Then V is completely reducible as a representation of G if and only if V is completely reducible as a representation of $\mathfrak{g}$.*

Proof. Indeed, the invariant subspaces of G and $\mathfrak{g}$ are the same, thanks to Lemma 95. □

Example 109. The zero representation $V := \{0\}$ is completely reducible (take $n := 0$). Any irreducible representation is completely reducible (take $n := 1, V_1 := V$).

Lemma 110. *Let $R : G \to \mathrm{GL}(V)$ be a finite-dimensional representation. The following are equivalent:*

(1) V is completely reducible.
(2) Every invariant subspace W of V has an invariant complement W'.

Recall here that a subspace $W' \leq V$ is said to be a *complement* of a subspace $W \leq V$ if $V = W \oplus W'$, that is to say, if every $v \in V$ may be expressed as $v = w + w'$ for some unique $w \in W, w' \in W'$.

Example 111. Suppose that $R : G \to \mathrm{GL}(V)$ has the property that there is an inner product $\langle , \rangle$ on V that is invariant by G in the sense that

$$\langle R(g)u, R(g)v \rangle = \langle u, v \rangle \tag{70}$$

for all $g \in G$ and all $u, v \in V$. (In other words, after choosing an orthonormal basis of V and using that basis to identify $V \cong \mathbb{C}^n$, we are assuming that $R(G)$ is contained in the unitary group $\mathrm{U}(n)$.) Then every invariant subspace W has an invariant complement W': one can just take for W' the *orthogonal complement*

$$W^{\perp} := \{v \in V : \langle v, w \rangle = 0 \text{ for all } w \in W\}$$

of W. It follows from the definition of an inner product that the orthogonal complement is in fact a complement; what needs to be checked is that it is invariant. Thus, let $v \in W^{\perp}$ and $g \in G$; we want to check that $R(g)v \in W^{\perp}$, i.e., that $\langle R(g)v, w \rangle = 0$ for all $w \in W$. But (70) implise that

$$\langle R(g)v, w \rangle = \langle R(g^{-1})R(g)v, R(g^{-1})w \rangle = \langle v, R(g^{-1})w \rangle,$$

and the invariance of W implies that $R(g^{-1}) \in w$, hence that $\langle v, R(g^{-1})w \rangle = 0$, as required.

We turn to the proof of Lemma 110, which we split into two parts.

Existence of invariant complements implies complete reducibility. Assume first that every invariant subspace of V has an invariant complement; we aim then to show that V is completely reducible. (In following this argument, it may be helpful to pretend that we are in the setting of Example 111.) If $V = \{0\}$, then we are done. Since $\dim(V) < \infty$, there exists a minimal nonzero invariant subspace V_1 of V. If $V_1 = V$, then we are done. Otherwise, let V_1' be an invariant complement of V_1; by assumption, $V_1' \neq 0$ and

$$V = V_1 \oplus V_1'.$$

Let V_2 be a minimal nonzero invariant subspace of V_1'. If $V_2 = V_1'$, then $V = V_1 \oplus V_2$ is a sum of invariant irreducible subspaces, so we are done. If not, let $W_2 \leq V$ be an invariant complement to V_2, and let $V_2' := W_2 \cap V_1 \leq V_1'$ be its intersection with V_1', which is then invariant and satisfies $V_1' = V_2 \oplus V_2'$ (check this; it's easy), hence

$$V = V_1 \oplus V_2 \oplus V_2'.$$

By assumption, W_2 is nonzero, hence V_2' is nonzero. Let V_3 be a minimal nonzero invariant subspace of V_2'. If $V_3 = V_2'$, then we are done as before. If not, let $V_3' := V_2' \cap W_3 \leq V_2'$ be the intersection with V_2' of some invariant complement $W_3' \leq V$ of V_2'; then, as before,

$$V = V_1 \oplus V_2 \oplus V_3 \oplus V_3'.$$

Proceed as above, and invoke that $\dim(V) < \infty$ to know that the process must eventually terminate.

(The argument just presented is "obvious" and fairly natural, but somewhat suboptimal. I'll leave it as an exercise for the interested reader to make it "slicker" by considering in the first step an invariant subspace W of V that is maximal with respect to the property of being a direct sum of irreducible invariant subspaces, and deriving a contradiction if $W \neq V$. This "slicker" formulation of the argument has certain advantages; for instance, it works without any fuss in the infinite-dimensional setting.) $\square$

Complete reducibility implies existence of invariant complements. Okay, now let's show that if V is completely reducible, then every invariant subspace W of V has an invariant complement W'. Thus, suppose we can write $V = V_1 \oplus \cdots \oplus V_n$ as a sum of irreducible invariant subspaces. Let I be a subset of $\{1, \dots, n\}$ that is maximal with respect to the property that

$$W \cap (\oplus_{i \in I} V_i) = \{0\}.$$

(Note that the empty set satisfies this property, and there are only finitely many subsets, so such an I exists.) Set $W' := \oplus_{i \in I} V_i$. We claim that $V = W \oplus W'$. By construction, we have $W \cap W' = \{0\}$, so it suffices to show that $V = W + W'$. To that end, it suffices to show for each $j \in \{1, \dots, n\}$ that

$$V_j \subseteq W + W'. \tag{71}$$

If $j \in I$, then $V_j \subseteq W'$, so 71 holds, so suppose $j \notin I$. If (71) fails, then $V_j \cap (W + W')$ is a proper invariant subspace of V_j, hence

$$V_j \cap (W + W') = 0,$$

or in other words, there is no nontrivial solution to the equation $v = w + w'$ with $v \in V_j$, $w \in W, w' \in W'$, or equivalently (upon replacing w, v by their negatives), there is no nontrivial solution to the equation $w = w' + v$ with $v \in V_j$, $w \in W, w' \in W'$, i.e.,

$$W \cap (W' \oplus V_j) = 0,$$

which says that $W \cap (\oplus_{i \in I \cup \{j\}} V_i) = 0$, contradicting the assumed maximality of I. $\square$

Remark 112. The natural setting for Lemma 110 is the theory of semisimple modules over a ring.

Example 113 (A representation that is not completely reducible)**.** Consider the Lie subgroup P of $\mathrm{SL}_2(\mathbb{C})$ consisting of matricse of the form $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$. Let $R : P \to \mathrm{GL}(V)$ be the standard representation of P on $V := \mathbb{C}^2$, thus

$$R(\begin{pmatrix} a & b \\ & d \end{pmatrix}) \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + b \\ dy \end{pmatrix}.$$

Then V is not completely reducible. Indeed, consider the subspace $W := \mathbb{C}e_1 \leq V$ consisting of column vectors of the form

$$\begin{pmatrix} x \\ 0 \end{pmatrix}.$$

It is easy to check that W is invariant, and that every complement W' of W has the form

$$W' = \left\{ \begin{pmatrix} cy \\ y \end{pmatrix} : y \in \mathbb{C} \right\}$$

for some $c \in \mathbb{C}$. But it is equally clear that W' is not invariant; for instance, one has

$$\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} \in P, \quad \begin{pmatrix} c \\ 1 \end{pmatrix} \in W, \quad \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} \begin{pmatrix} c \\ 1 \end{pmatrix} = \begin{pmatrix} c+1 \\ 1 \end{pmatrix} \notin W.$$

Thus not all representations are completely reducible.

One gets more general examples of this sort by replacing P by the stabilizer P of any nontrivial flag of vector spaces $0 = V_0 \subset V_1 \subset \cdots \subset V_r = \mathbb{C}^n$.

Example 114. The representation

$$\mathbb{R} \to \mathrm{GL}_2(\mathbb{R})$$
$$x \mapsto \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix}$$

is not completely reducible, for the same reason as in the previous example. Similarly for the representation $\mathbb{C} \to \mathrm{GL}_2(\mathbb{C})$ defined by the same formula. Likewise for the representation

$$\mathbb{R}^\times \to \mathrm{GL}_2(\mathbb{R})$$
$$x \mapsto \begin{pmatrix} 1 & \log|x| \\ & 1 \end{pmatrix}.$$

Likewise for the representation

$$\mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}_2(\mathbb{R})$$
$$g \mapsto \begin{pmatrix} 1 & \log|\det(g)| \\ & 1 \end{pmatrix}.$$

Definition 115. Let G be a Lie group. We say that G is *linearly reductive* if every finite-dimensional representation of G is completely reducible.

Example 116. We have seen that the groups P from Example 113 are *not* linearly reductive. Similarly, we see from Example 114 that the real Lie groups $\mathbb{R}$ and $\mathrm{GL}_n(\mathbb{R})$ and the complex Lie group $\mathbb{C}$ are not linearly reductive.

A minor caution regarding terminology: We are working here in the category of Lie groups over the field $\mathbf{k} = \mathbb{R}$ or $\mathbb{C}$. When we speak of any property of a Lie group, it matters which we field we regard it as being defined over. For example, we will show eventually that $\mathbb{C}^\times$, regarded as a complex Lie group is linearly reductive. However, if we instead regard $\mathbb{C}^\times$ as a real Lie group, then it is not linearly reductive: the representation

$$\mathbb{C}^\times \ni z \mapsto \begin{pmatrix} 1 & \log|z| \\ & 1 \end{pmatrix}$$

is not completely reducible. That representation is not smooth in the complex sense (i.e., is not holomorphic), and so does not define a representation of $\mathbb{C}^\times$ when we regard it as a complex Lie group.

Similarly, although we are working in this course in the category of Lie groups, we could instead work in the category of *algebraic groups*, which are obtained by replacing manifolds with solution spaces to polynomial equations (called *varieties*) and replacing smooth maps between manifolds with maps described by polynomials (called *morphisms of varieties*). The group $\mathrm{GL}_n(\mathbb{R})$ can be regarded either as a Lie group or as an algebraic group. In the category of Lie groups, it is not linearly reductive. But the counter-example we gave involved the logarithm function, which is not algebraic. It turns out that when $\mathrm{GL}_n(\mathbb{R})$ is regarded as an algebraic group, it is linearly reductive. This means that one can't construct non-completely-reducible representations of $\mathrm{GL}_n(\mathbb{R})$ using only polynomials; to put it another way, it turns out that any representation $R : \mathrm{GL}_n(\mathbb{R}) \to \mathrm{GL}_N(\mathbb{R})$ whose matrix coefficients $R_{ij}(g)$ are polynomial functions of the coordinates g_{kl} is completely reducible.

16.8. **Linear reductivity of compact groups.**

Theorem 117 (Maschke)**.** *Any finite group G is linearly reductive.*

Proof. Let V be a complex vector space, and let $R : G \to \mathrm{GL}(V)$ be a representation. We wish to show that V is completely reducible. There are a couple ways to phrase the argument; I'll record both for the sake of variety.

First, by Lemma 110 and Example 111, it will suffice to show that there exists a invariant inner product on V. To show this, let $\langle , \rangle_0$ be any inner product on V (just fix a linear isomorphism $V \cong \mathbb{C}^n$ and take the standard one), and then define the averaged inner product $\langle , \rangle$ by the formula

$$\langle u, v\rangle := \frac{1}{|G|} \sum_{g \in G} \langle R(g)u, R(g)v\rangle_0.$$

It is then easy to check that $\langle , \rangle$ is the required an invariant inner product.

We now phrase the argument another way by making more direct use of the criterion of Lemma 110. It will suffice to show that each invariant subspace W of V has an invariant complement. To that end, it will suffice to construct an *equivariant projection* $p : V \to W$. (A *projection* from a vector space V to a subspace W is a linear map $p : V \to W$ whose restriction to W is the identity map. A linear map p between representations is said to be *equivariant* if it is a morphism of representations, i.e., if $p(gv) = gp(v)$ for all $v \in V, g \in G$.) Assuming we have constructed such a projection, we may take $W' := \ker(p)$. Since p is a projection, we then have

$$V = W \oplus W'.$$

On the other hand, since p is equivariant, its kernel W' is invariant. We thereby obtain the required invariant complement of W, assuming the existence of an equivariant projection.

To produce an equivariant projection, start with any projection $\phi_0 : V \to W$ (e.g., by taking a basis $e_1, \dots, e_d$ for W, extending it to a basis $e_1, \dots, e_n$ for V, and defining $\phi_0(e_j)$ to be e_j if $j \leq d$ and 0 otherwise) and define the average $\phi : V \to W$ by

$$\phi(v) := \frac{1}{|G|} \sum_{g \in G} R(g)\phi_0(R(g)^{-1}v).$$

Then it is easy to check that ϕ is still a projection, and also that ϕ is equivariant: for $h \in G$, one obtains using the change of variables $g \mapsto hg$ on G and the homomorphism property of representations that

$$\begin{aligned}
\phi(R(h)v) &= \frac{1}{|G|} \sum_{g \in G} R(g)\phi_0(R(g)^{-1}R(h)v) \\
&= \frac{1}{|G|} \sum_{g \in G} R(g)\phi_0(R(h^{-1}g)^{-1}v) \\
&= \frac{1}{|G|} \sum_{g \in G} R(hg)\phi_0(R(g)^{-1}v) \\
&= R(h)\phi(v),
\end{aligned}$$

so ϕ is equivariant, as required. ∎

Theorem 118. *Any compact Lie group G is linearly reductive. More generally, if G is a compact topological group and $R : G \to \mathrm{GL}(V)$ is any continuous finite-dimensional representation, then R is completely reducible.*

Proof. We will use the following fact, whose proof is sketched in §10, §10.4, §10.6: there is a unique Radon probability measure μ which is left and right invariant under G in the sense that $\mu(Eg) = \mu(gE) = \mu(E)$ for all Borel subsets $E \subseteq G$ and $g \in G$, or equivalently,

$$\int_{g\in G} f(g)\, d\mu(g) = \int_{g\in G} f(hg)\, d\mu(g) = \int_{g\in G} f(gh)\, d\mu(g)$$

for all $h \in G$ and all continuous functions $f : G \to \mathbb{C}$. For example, if G is a finite group, one can take for μ the normalized counting measure. We may then argue exactly as in either of the proofs of Theorem (117) by replacing averaging over the group with averaging with respect to μ, i.e., taking

$$\langle u, v\rangle := \int_{g\in G} \langle R(g)u, R(g)v\rangle_0 \, d\mu(g)$$

or

$$\phi(v) := \int_{g\in G} R(g)\phi_0(R(g)^{-1}v)\, d\mu(g)$$

instead of what we did above. □

The tools of §17 will give us a number of examples of linearly reductive complex Lie groups ($\mathrm{GL}_n(\mathbb{C})$, $\mathrm{SL}_n(\mathbb{C})$, $\mathrm{SO}_n(\mathbb{C})$, etc.) and also linearly reductive real Lie groups ($\mathrm{SL}_n(\mathbb{R})$, $\mathrm{SO}(p,q)^0$ for $(p,q) \neq (1,1)$) in addition to the compact groups ($\mathrm{U}(n), \mathrm{SO}(n)$, etc.) already covered above. The following result was established in the above, and is of independent interest:

Theorem 119. *Let $R : G \to \mathrm{GL}(V)$ be a finite-dimensional representation of a compact group G. Then there exists an invariant inner product on V.*

The proofs given above were self-contained except that we punted the existence of the Haar measure μ to §10. Here we record a self-contained way (learned from Onishchik–Vinberg) to "get around" constructing such a μ. (I put "get around" in quotes because the ideas here are similar to those used in §10.6 to construct such a μ; however, they are somewhat simpler in the present context.)

Theorem 120. *Let S be a finite-dimensional vector space, let G be a compact group, and let $\alpha : G \to \mathrm{GL}(S)$ be a representation. Let $M \subseteq S$ be a nonempty convex G-invariant subset. Then M contains a fixed point of G.*

Proof. We reduce first to the case that M is bounded by replacing M as necessary by the convex hull of some orbit of G on M; such an orbit is compact because of the compactness of G. The idea is then that M has a well-defined "center of mass" which, being canonically defined, will turn out to be G-invariant.

Turning to details, let $W \leq S$ denote the span of all differences of pairs of elements of M. Let P denote the smallest plane containing M, or equivalently, the union of all lines in S containing at least two elements of M. Then P is a coset of W. Since M is G-invariant and the associations $M \mapsto W, P$ are canonical, we know that W and P are G-invariant, too. The vector space W has a Lebesgue measure μ. For $g \in G$, let $J(g)$ denote the Jacobian of the linear transformation of W induced by $R(g)$. Then $J : G \to \mathbb{R}_+^\times$ is a homomorphism. Since G is compact, J is trivial.

Thus the Lebesgue measure μ is G-invariant. By fixing a basepoint on P, we can transfer μ to a measure ν on P, which is again G-invariant. The *center of mass*

$$c(M) := \frac{\int_{x \in M} x \, d\nu(x)}{\nu(M)}$$

of M then makes sense (because M is bounded) and is G-invariant (because ν is G-invariant). Since M is convex, we have moreover that $c(M)$ belongs to M: this is clear if M is a point, while otherwise the interior M^0 of M (as defined using the topology on P) is nonempty, so if $c(M) \notin M$, then (by the separating hyperplane theorem, i.e., Hahn–Banach in finite dimensions) there is an affine-linear function $\ell : S \to \mathbb{R}$ satisfying $\ell(M^0) \subseteq \mathbb{R}_{>0}$ but $\ell(c(M)) = 0$, which leads to a contradiction upon applying ℓ to the definition of $c(M)$. Thus $c(M) \in M$ gives the required G-fixed point. □

Example 121. Let G be a compact group, let $R : G \to \mathrm{GL}(V)$ be a finite-dimensional representation, let $W \leq V$ be an invariant subspace, and let S denote the set of all linear maps $V \to W$ and $M \subseteq S$ the subset consisting of projections $\phi : V \to W$. Then G acts on S by the rule $g \cdot \phi := R(g) \circ \phi \circ R(g)^{-1}$, and M is a nonempty convex G-invariant subset. A projection $\phi \in M$ is equivariant if and only if it is a fixed point for this action, so Theorem 120 tells us that an equivariant projection exists. Using this, we can complete the proof of Theorem 118 without directly establishing the existence of μ.

Example 122. Let G be a compact group, let $R : G \to \mathrm{GL}(V)$ be a finite-dimensional representation, let S be the space of hermitian forms on V, and let $M \subseteq S$ be the subset of positive definite hermitian forms (i.e., inner products). Then G acts on S by $(g \cdot B)(v_1, v_2) := B(R(g)^{-1} v_1, R(g)^{-1} v_2)$, and the subspace M is convex, nonempty and G-invariant. An inner product $B \in M$ is G-invariant in the present sense if and only if it is invariant in the sense defined above, so Theorem 120 tells us that an invariant inner product exists.

16.9. **Constructing new representations from old ones.**

16.9.1. *Direct sum.* Given a pair of representations V_1, V_2 of a Lie group G, we get a representation on their direct sum $V_1 \oplus V_2$ by

$$g(v_1 \oplus v_2) := gv_1 \oplus gv_2.$$

Similarly for representations of a Lie algebra. The two constructions are compatible under differentiation.

16.9.2. *Tensor product.* Given a pair of representations V_1, V_2 of a Lie group G, we get a representation on their tensor product $V_1 \otimes V_2$ by

$$g(v_1 \otimes v_2) := gv_1 \otimes gv_2.$$

If V_1, V_2 are instead representations of a Lie algebra $\mathfrak{g}$, then the natural action to take on their tensor product is

$$X(v_1 \otimes v_2) := Xv_1 \otimes v_2 + v_1 \otimes Xv_2.$$

One of the homework problems for this week is to check that this in fact defines a Lie algebra representation. We checked in class that if we differentiate the first action, we get the second:

$$\frac{d}{dt}(\exp(tX)v_1 \otimes \exp(tX)v_2)|_{t=0} = Xv_1 \otimes v_2 + v_1 \otimes Xv_2.$$

16.9.3. *Symmetric power representations.* Given a complex vector space V and $t \in \mathbb{Z}_{\geq 0}$, the *symmetric power* $\mathrm{Sym}^t(V)$ is the space of homogeneous polynomials of degree t on the dual space V^*. Each element of V may be identified with a linear polynomial on V^*. Any element of $\mathrm{Sym}^t(V)$ may be written as a finite linear combination of monomials $v_1 \cdots v_t$ with each $v_j \in V$. If $e_1, \ldots, e_n$ is a basis for V, then the polynomials $e_{i_1} \cdots e_{i_t}$ taken over all indices satisfying $1 \leq i_1 \leq \cdots \leq i_t \leq n$ form a basis for $\mathrm{Sym}^t(V)$.

Given a representation $R : G \to \mathrm{GL}(V)$ of a Lie group G, one obtains a symmetric power representation $\mathrm{Sym}^n(R) : G \to \mathrm{GL}(\mathrm{Sym}^n(V))$ by setting: for $v_1, \ldots, v_t \in V$,

$$(\mathrm{Sym}^n(R)(g))v_1 \cdots v_t := (R(g)v_1) \cdots (R(g)v_t).$$

(See Wikipedia or Google for more details on this construction.)

16.10. **Characters.** For each representation V of $\mathfrak{g} := \mathfrak{sl}_2(\mathbb{C})$, define the *character* of V to be the Laurent polynomial

$$\mathrm{ch}(V) \in A := \mathbb{Z}[z, z^{-1}]$$

given by

$$\mathrm{ch}(V) := \sum_{m \in \mathbb{Z}} (\dim V[m]) z^m,$$

where $V[m] := \{v \in V : Hv = mv\}$ with $H = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \in \mathfrak{g}$ as before. (By this point in lecture, we saw that the action of H on any finite-dimensional representation is diagonalizable, so $V = \oplus V[m]$.) For example, for the irreducible representations W_m ($m \in \mathbb{Z}_{\geq 0}$) considered in lecture, we have $\mathrm{ch}(W_m) = \sum_{-m \leq j \leq m : j \equiv m(2)} z^m = z^m + z^{m-2} + \cdots + z^{-m}$. Such functions are *symmetric* (i.e., invariant under $z \mapsto z^{-1}$.) The *Weyl denominator* is the element

$$D := z - z^{-1} \in A.$$

It is the simplest example of an *anti-symmetric* element of A. One has

$$D \cdot \mathrm{ch}(W_m) = z^{m+1} - z^{-(m+1)}.$$

In other words, *as V traverses the set of isomorphism classes of irreducible representations, $D \cdot \mathrm{ch}(V)$ traverses the "obvious" basis for the space of anti-symmetric elements of A.*

For a finite-dimensional representation $R : G \to \mathrm{GL}(V)$ of $G := \mathrm{SL}_2(\mathbb{C})$, one has

$$\mathrm{ch}(V)|_{z=e^{i\theta}} = \mathrm{trace}(R(\begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{pmatrix})).$$

The character satisfies

$$\mathrm{ch}(V_1 \oplus V_2) = \mathrm{ch}(V_1) + \mathrm{ch}(V_2),$$
$$\mathrm{ch}(V_1 \otimes V_2) = \mathrm{ch}(V_1)\,\mathrm{ch}(V_2).$$

This is easily seen by taking a basis $e_1, \ldots, e_m$ of H-eigenvectors for V_1 and a basis $f_1, \ldots, f_m$ of H-eigenvectors for V_2 and using that $e_1, \ldots, e_m, f_1, \ldots, f_m$ is then a basis of H-eigenvectors for $V_1 \oplus V_2$ while $e_i \otimes f_j$ give a basis of H-eigenvectors for $V_1 \otimes V_2$.

If we're given a representation V of $\mathfrak{g}$, we know that we can decompose

$$V \cong \oplus_{m \geq 0} W_m^{\mu(m)}$$

for some multiplicities $\mu(m) \geq 0$. We can determine the multiplicities easily if we know the character of V, and particularly easily if we multiply first by the Weyl denominator: we have

$$D \cdot \mathrm{ch}(V) = \sum_{m \geq 0} \mu(m) D \cdot \mathrm{ch}(W_m) = \sum_{m \geq 0} \mu(m)(z^{m+1} - z^{-m-1}),$$

so we can read off the multiplicity $\mu(m)$ of W_m inside V as the coefficient of z^{m+1} in the anti-symmetric Laurent polynomial $D \cdot \mathrm{ch}(V)$. For example, in this way (or in others) we can easily derive the *Clebsh–Gordon decomposition*

$$W_m \otimes W_n \cong W_{m+n} \oplus W_{m+n-2} \oplus \cdots \oplus W_{|m-n|} = \oplus_{\substack{|m-n| \leq j \leq m+n: \\ j \equiv m+n(2)}} W_j$$

from the polynomial identity: for $m \geq n$ (say),

$$(z^{m+1} - z^{-m-1})(z^n + z^{n-2} + \cdots + z^{-n}) = \sum_{\substack{m-n \leq j \leq m+n: \\ j \equiv m+n(2)}} (z^{j+1} - z^{-j-1}).$$

We can make this more explicit, e.g., the isomorphism

$$W_2 \oplus W_0 \cong W_1 \otimes W_1$$

can be given by identifying $W_2 = \mathbb{C}[z,w]_{(2)}$ and $W_0 = \mathbb{C}$ and $W_1 = \mathbb{C}[x,y]_{(1)}$ (where a subscripted (n) denotes homogeneous elements of order n) and the map

$$W_0 \to W_1 \otimes W_1$$

is given by

$$1 \mapsto (x \otimes y - y \otimes x)/2$$

and the map

$$W_2 \to W_1 \otimes W_1$$

by

$$z^2 \mapsto x \otimes x,$$
$$w^2 \mapsto y \otimes y,$$
$$zw \mapsto (x \otimes y + y \otimes x)/2.$$

See any introductory textbook on quantum mechanics for more on the importance of these sorts of decompositions in physics.

17. The unitary trick

Definition 123. Let $\mathfrak{g}$ be a complex Lie algebra. Let $\mathfrak{h}$ be a real Lie algebra. We say that

- $\mathfrak{h}$ is a *real form* of $\mathfrak{g}$, or equivalently, that
- $\mathfrak{g}$ is the *complexification* of $\mathfrak{h}$,

if $\mathfrak{h}$ is (isomorphic to) a real Lie subalgebra of $\mathfrak{g}$ for which

$$\mathfrak{g} = \mathfrak{h} \oplus i\mathfrak{h},$$

or equivalently, for which the natural map $\mathfrak{h} \otimes_{\mathbb{R}} \mathbb{C} \to \mathfrak{g}$ is an isomorphism. In other words, every $z \in \mathfrak{g}$ may be expressed uniquely as $x + iy$ with $x, y \in \mathfrak{h}$.

Example 124. $\mathfrak{sl}_n(\mathbb{R})$ and $\mathfrak{su}(n)$ are real forms of $\mathfrak{sl}_n(\mathbb{C})$; $\mathfrak{so}(n)$ is a real form of $\mathfrak{so}_n(\mathbb{C})$. $\mathfrak{gl}_n(\mathbb{R})$ and $\mathfrak{u}(n)$ are real forms of $\mathfrak{gl}_n(\mathbb{C})$. A bit less obviously, $\mathfrak{so}(p,q)$ is (isomorphic to) a real form of $\mathfrak{so}_n(\mathbb{C})$ (try to check this!).

Lemma 125. *Let L be any complex Lie algebra. Let $\mathfrak{g}$ be a complex Lie algebra and $\mathfrak{h}$ a real form of $\mathfrak{g}$. Then the natural restriction map*

$$\left\{\begin{matrix}\text{morphisms of complex Lie algebras}\\ \Phi : \mathfrak{g} \to L\end{matrix}\right\} \to \left\{\begin{matrix}\text{morphisms of real Lie algebras}\\ \phi : \mathfrak{h} \to L\end{matrix}\right\}$$

is a bijective.

Proof. Given Φ, one defines ϕ by restriction. Given ϕ, one defines $\Phi(x+iy) := \phi(x) + i\phi(y)$. One checks that the associations $\phi \mapsto \Phi$ and $\Phi \mapsto \phi$ are mutually inverse, and that one defines a Lie algebra morphism (over the relevant field) if and only if the other does. (One could alternatively take the conclusion of this lemma as the *definition* of real form/complexification, as in the functorial characterization of tensor product.) □

Example 126. Suppose $L = \mathrm{End}(V)$ for a complex vector space V. Then Lemma 125 says that for a complex Lie algebra $\mathfrak{g}$ with real form $\mathfrak{h}$, the following sets are in natural bijection:

(1) holomorphic representations $\rho : \mathfrak{g} \to \mathrm{End}(V)$.
(2) representations $\rho_0 : \mathfrak{h} \to \mathrm{End}(V)$.

Moreover, the invariant subspaces W for ρ and ρ_0 are the same. In particular, ρ is irreducible if and only if ρ_0 is irreducible, and also ρ is completely reducible if and only if ρ_0 is completely reducible.

Definition 127. Let G be a connected complex Lie group. A *real form* of G is a connected real Lie subgroup $H \leq G$ with the property that the Lie algebra $\mathfrak{h}$ of H is a real form of the Lie algebra $\mathfrak{g}$ of G.

Example 128. $\mathrm{GL}_n(\mathbb{R})$ and $\mathrm{U}(n)$ are real forms of $\mathrm{GL}_n(\mathbb{C})$. $\mathrm{SL}_n(\mathbb{R})$ and $\mathrm{SU}(n)$ are real forms of $\mathrm{SL}_n(\mathbb{C})$. $\mathrm{SO}(n)$ and $\mathrm{SO}(p,q)^0$ $(p+q=n)$ are real forms of $\mathrm{SO}_n(\mathbb{C})$.

Theorem 129. *Let G be a connected complex Lie group. Suppose that G has a compact (connected) real form. Then G is linearly reductive.*

Proof. Let H be a compact (connected) real form of G, and denote Lie algebras in the usual way. Let $R : G \to \mathrm{GL}(V)$ be a finite-dimensional representation of G. We wish to show that V is completely reducible under G. We have seen in Lemma 108 that V is completely reducible under G if and only if it is under $\mathfrak{g}$, and likewise that V is completely reducible under H if and only if it is under $\mathfrak{h}$. By Lemma 125, we know that V is completely reducible under $\mathfrak{g}$ if and only if it is under $\mathfrak{h}$. In summary, V is completely reducible under any one of $G, \mathfrak{g}, \mathfrak{h}, H$ if and only if it is under all of them. Since compact Lie groups as linearly reductive (Theorem 118), we know that V is completely reducible under H, so we are done. □

Example 130. It follows that the groups $\mathrm{GL}_n(\mathbb{C})$ (e.g., $\mathbb{C}^\times$) $\mathrm{SL}_n(\mathbb{C})$, $\mathrm{SO}_n(\mathbb{C})$ are linearly reductive.

Remark 131. Let G be a connected complex Lie group which has a compact (connected) real form, then we have seen that G, as well as its compact (connected) real form, is linearly reductive. However, it need not be the case that every real form H of G is linearly reductive; consider for instance the case $G = \mathbb{C}^\times, H = \mathbb{R}^\times$. (This is another example where things become nicer by working in the category of algebraic groups; compare with Remark 116.)

In the proof of Theorem 129, the notion of "completely reducible" relevant for the representation V of $\mathfrak{h}$ is that there exist no invariant complex subspaces. One can ask what happens if one works instead with invariant real subspaces:

Exercise 19. Let V be a complex vector space and $\mathrm{End}(V)$ the Lie algebra of $\mathbb{C}$-linear endomorphisms of V. Let $\mathfrak{h}$ be a real Lie algebra and $\rho : \mathfrak{h} \to \mathrm{End}(V)$ a morphism.

Let $W \leq V$ be a real subspace (i.e., a subspace of the real vector space underlying V).

(1) Show that if W is invariant or irreducible (under the action by $\mathfrak{h}$), then so is iW.
(2) Suppose that W is invariant and irreducible. Show that either
   (a) $W = iW$, in which case W is an invariant irreducible complex subspace, or
   (b) $W \cap iW = \{0\}$, in which case $W + iW$ is an invariant irreducible complex subspace.

   Deduce that if V decomposes as a direct sum of invariant irreducible real subspaces, then it also decomposes as a direct sum of invariant irreducible complex subspaces.

Let us work out the complexification of $H := \mathrm{SO}(p,q)^0$. Recall that $\mathrm{SO}(p,q) = \{g \in \mathrm{SL}_{p+q}(\mathbb{R}) : gI_{p,q}g^t = I_{p,q}\}$ where $I_{p,q} := \mathrm{diag}(1,\dots,1,-1,\dots,-1)$, thus

$$\mathfrak{h} = \{X \in \mathrm{SL}_{p+q}(\mathbb{R}) : XI_{p,q}+I_{p,q}X^t = 0\} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in M_{p+q}(\mathbb{R}) : A^t = -A, D^t = -D, B^t = C \right\}.$$

The complexification is given by

$$\mathfrak{h}\otimes_{\mathbb{R}}\mathbb{C} \cong \mathfrak{h}\oplus i\mathfrak{h} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in M_{p+q}(\mathbb{C}) : A^t = -A, D^t = -D, B^t = C \right\} \subseteq M_{p+q}(\mathbb{C}).$$

If $\varepsilon := \mathrm{diag}(i,\dots,i,1,\dots,1)$, then we can check easily that

$$\varepsilon(\mathfrak{h} \oplus i\mathfrak{h})\varepsilon^{-1} = \mathfrak{so}_{p+q}(\mathbb{C}).$$

Thus the complexification of $\mathfrak{so}(p,q)$ is isomorphic to $\mathfrak{so}_{p+q}(\mathbb{C})$. Similarly, $\mathrm{SO}(p,q)^0$ is (isomorphic to) a real form of $\mathrm{SO}_{p+q}(\mathbb{C})$; just consider

$$\varepsilon\, \mathrm{SO}(p,q)^0 \varepsilon^{-1} \leq \mathrm{SO}_{p+q}(\mathbb{C}) \leq \mathrm{SL}_{p+q}(\mathbb{C}).$$

To cook up some more interesting examples, let

$$\mathbb{H} := \left\{ \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} : z, w \in \mathbb{C} \right\} \subseteq M_2(\mathbb{C})$$

denote Hamilton's quaternion algebra over the reals. (It is an associative algebra with center $\mathbb{R}$ and of dimension 4 over its center.) There is a natural involution $x \mapsto x^*$ on $\mathbb{H}$, given by $x^* := \overline{x}^t$, or equivalently,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^* = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Then $\mathbb{H}^\times$ is a Lie group and

$$\mathrm{Lie}(\mathbb{H}^\times) = \mathbb{H}.$$

The space $M_n(\mathbb{H})$ of $n \times n$ matrices with quaternionic entries can be regarded as a subspace of $M_{2n}(\mathbb{C})$. Set

$$\mathrm{SL}_n(\mathbb{H}) := M_n(\mathbb{H}) \cap \mathrm{SL}_{2n}(\mathbb{C})$$

and

$$\mathrm{U}_n(\mathbb{H}) := \{g \in M_n(\mathbb{H}) : g\overline{g^*} = 1\}.$$

Recall that $\mathrm{SU}(p,q) = \{g \in \mathrm{SL}_{p+q}(\mathbb{C}) : gI_{p,q}\overline{g}^t = I_{p,q}\}$.

Exercise 20. Determine the complexifications of the Lie algebras of the following real Lie groups:

(1) $\mathrm{SL}_n(\mathbb{H})$
(2) $\mathrm{SU}(p,q)$
(3) $\mathrm{U}_m(\mathbb{H})$

(Each is isomorphic to a classical complex Lie group we're already familiar with.)

It turns out that with this exercise and the examples given previously, we've found all of the real forms of the complex Lie algebras $\mathfrak{sl}_n(\mathbb{C}), \mathfrak{so}_n(\mathbb{C})$.

Exercise 21. Let $\mathfrak{h} := \mathfrak{sl}_n(\mathbb{C})$, but regarded as a real Lie algebra rather than a complex one. (Concretely, we can think of $\mathfrak{h}$ as a real Lie subalgebra of $\mathfrak{sl}_{2n}(\mathbb{R})$.) Show that one has an isomorphism of complex Lie algebras $\mathfrak{h} \otimes_{\mathbb{R}} \mathbb{C} \cong \mathfrak{sl}_n(\mathbb{C}) \oplus \mathfrak{sl}_n(\mathbb{C})$.

Remark 132. Let $\mathfrak{h}$ be a real form of a complex Lie algebra $\mathfrak{g}$. Then $\mathfrak{h}$ is the fixed point set of the automorphism $\sigma : \mathfrak{g} \to \mathfrak{g}$ given by identifying $\mathfrak{g}$ with $\mathfrak{h} \otimes_{\mathbb{R}} \mathbb{C}$ and requiring that for $X \in \mathfrak{h}$ and $z \in \mathbb{C}$, one has $\sigma(X \otimes z) := X \otimes \overline{z}$; in other words, if we identify $\mathfrak{g}$ with $\mathfrak{h} \oplus i\mathfrak{h}$, then $\sigma(x + iy) := x - iy$ for $x, y \in \mathfrak{h}$. Then σ is an involution on $\mathfrak{g}$ (i.e., σ^2 is the identity transformation); moreover, σ is a Lie algebra automorphism that is anti-linear with respect to scalar multiplication by complex numbers, i.e., for all $c \in \mathbb{C}$, $Z \in \mathfrak{g}$, $\sigma(cZ) = \overline{c}\sigma(Z)$, and satisfies $\sigma^2 = 1$.

Conversely, given an anti-linear involution $\sigma : \mathfrak{g} \to \mathfrak{g}$, we claim that its fixed point subspace $\mathfrak{h} : \{X \in \mathfrak{g} : \sigma(X) = X\}$ is a real form of $\mathfrak{g}$. Well, since σ is a real Lie algebra automorphism, we know at least that $\mathfrak{h}$ is a real Lie subalgebra. Observe that $i\mathfrak{h} = \{X \in \mathfrak{g} : \sigma(X) = -X\}$. Since $\mathfrak{g}$ is the sum of the $+1$ and -1 eigenspaces of σ, we deduce that $\mathfrak{g} = \mathfrak{h} \oplus i\mathfrak{h}$. Hence $\mathfrak{h}$ is a real form.

18. Ad and ad

18.1. **Basic definitions.** When one learns basic group theory (say of finite groups), one studies groups G acting on sets X. A particularly important action is the conjugation action of G on itself, given by $(g, x) \mapsto gxg^{-1}$. The orbits for this action are the conjugacy classes in G. Much nontrivial information about G can be extracted from a careful study of the conjugation action of G on itself. For example, the Sylow theorems are proved in this way.

When G is a Lie group, one can again consider the conjugation action of G on itself, but it turns out to be more useful to differentiate this action a bit, so that tools from linear algebra become at our disposal.

Definition 133. Given a Lie group G with Lie algebra $\mathfrak{g}$, the *adjoint representation* of G is the map

$$\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$$

is defined by

$$\mathrm{Ad}(g)X := gXg^{-1},$$

where the RHS may be interpreted in various ways:

(1) If G is a subgroup of $\mathrm{GL}_n(\mathbf{k})$, then $\mathfrak{g}$ is a subalgebra of $M_n(\mathbf{k})$, and so we can interpret gXg^{-1} as a product of matrices.
(2) We can use the trick of §15 ("pretending that every Lie group is a matrix Lie group") to embed both G and $\mathfrak{g}$ inside $\mathrm{End}(C^\infty(G))$; in that optic, the product gXg^{-1} is given by composition.
(3) (I don't recommend spending too much time studying this interpretation.) In general, for $g \in G$ and $x_0 \in G$ and $X \in T_{x_0}(G)$, we can define $gX \in T_{gx_0}(G)$ to be the image of X under the differential of left translation by g, i.e.,: if $\psi : G \to G$ is the map $\psi(x) := gx$, then $gX := (T_{x_0}\psi)(X)$ where $T_{x_0}\psi : T_{x_0}(G) \to T_{gx_0}(G)$ is the derivative. We can similarly define $Xg \in T_{x_0 g}(G)$ using right translations. This makes sense in particular when $x_0 = e$ is the identity element, so that $T_e G = \mathfrak{g}$; then for $X \in \mathfrak{g}$ we have $gX \in T_g G$ and thus $(gX)g^{-1} \in \mathfrak{g}$. Alternatively, we may first form $Xg^{-1} \in T_{g^{-1}}G$ and then $g(Xg^{-1}) \in \mathfrak{g}$. The two answers are the same because left and right translations commute with one another. (See §11.6 for related discussion.)
(4) For any $g \in G$ and $X \in \mathfrak{g}$, the map $\mathbf{k} \ni t \mapsto g\exp(tX)g^{-1}$ is a one-parameter subgroup, so its initial velocity is an element of the Lie algebra:
$$gXg^{-1} = \partial_{t=0} g\exp(tX)g^{-1}.$$

It is clear that $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ is a morphism of Lie groups.

Note that if G is a real Lie group, then Ad is a real representation, not a complex representation of the sort that we have primarily been considering. If G is a complex Lie group, then Ad is a holomorphic representation.

Exercise 22. Let $g \in G, X \in \mathfrak{g}$. Show that
$$g\exp(X)g^{-1} = \exp(\mathrm{Ad}(g)X).$$

Hint: one can either

(1) appeal to uniqueness of one-parameter subgroups (§13.2), or
(2) apply the result of §13.6 to $f : G \to G$ given by $f(a) := gag^{-1}$.

Since the Lie algebra $\mathfrak{g}$ of a Lie group is a vector space, we may form its linear dual $\mathfrak{g}^* := \mathrm{Hom}_{\mathbf{k}}(\mathfrak{g}, \mathbf{k})$. From some perspectives (which we might discuss eventually), the following action is better behaved than the adjoint action:

Definition 134. The *coadjoint representation* of a Lie group G is the map
$$\mathrm{Ad}^* : G \to \mathrm{GL}(\mathfrak{g}^*)$$
given for $g \in G, \lambda \in \mathfrak{g}^*, X \in \mathfrak{g}$ by
$$(\mathrm{Ad}^*(g)\lambda)(X) := \lambda(\mathrm{Ad}(g)^{-1}X).$$

Exercise 23. Check that Ad^* is a representation, but that it wouldn't be in general had we omitted the inverse in the definition.

Definition 135. Given any Lie algebra $\mathfrak{g}$ with Lie bracket $\mathfrak{g}$, we can define
$$\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$$
by the formula
$$(\mathrm{ad}(X))(Y) := [X, Y].$$

We usually abbreviate the LHS to $\mathrm{ad}(X)Y$. We might sometimes also abbreviate $\mathrm{ad}_X := \mathrm{ad}(X)$, so that $\mathrm{ad}_X Y = [X, Y]$. Recall that the Jacobi identity may be written in the following equivalent ways:

$$[[X, Y], Z] = [[X, Z], Y] + [X, [Y, Z]], \tag{72}$$

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]. \tag{73}$$

The first identity (72) may be interpreted as saying that $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ is a morphism of Lie algebras, that is to say, that

$$\mathrm{ad}([X, Y]) = [\mathrm{ad}(X), \mathrm{ad}(Y)],$$

since indeed $\mathrm{ad}([X, Y])Z = [[X, Y], Z]$ and $[\mathrm{ad}(X), \mathrm{ad}(Y)]Z = \mathrm{ad}(X)\,\mathrm{ad}(Y)Z - \mathrm{ad}(Y)\,\mathrm{ad}(X)Z = \mathrm{ad}(X)[Y, Z] - \mathrm{ad}(Y)[X, Z] = [X, [Y, Z]] - [Y, [X, Z]]$. The second identity may be interpreted as saying that that ad_X is a derivation for each $X \in \mathfrak{g}$, i.e., that

$$\mathrm{ad}_X[Y, Z] = [\mathrm{ad}_X Y, Z] + [Y, \mathrm{ad}_X Z],$$

i.e., that $\mathrm{ad}(\mathfrak{g}) \subseteq \mathrm{Der}(\mathfrak{g})$. So in summary, ad defines a morphism of Lie algebras

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

One can already get *huge* mileage out of this simple statement (see the tricky problem on this week's homework). Note in particular that it contains two different interpretations of the Jacobi identity.

18.2. **Relationship between** Ad **and** ad**.** The above definitions are related as follows:

Lemma 136. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then the differential $d\,\mathrm{Ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ of the morphism of Lie groups $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ is the morphism of Lie algebras $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$, that is to say, $d\,\mathrm{Ad} = \mathrm{ad}$, or more verbosely, for any $X, Y \in \mathfrak{g}$,*

$$\partial_{s=0} \mathrm{Ad}(\exp(sX))Y = \mathrm{ad}(X)Y.$$

Proof. The RHS is $\mathrm{ad}(X)Y = [X, Y]$. We expand out its definition and compute, obtaining

$$\begin{aligned} [X, Y] &:= \partial_{s=0}\partial_{t=0}(e^{sX}, e^{tY}) \\ &= \partial_{s=0}\partial_{t=0} e^{sX} e^{tY} e^{-sX} e^{-tY} \\ &= \partial_{s=0}(e^{sX} Y e^{-sX} - Y) \\ &= \partial_{s=0} \mathrm{Ad}(e^{sX})Y, \end{aligned}$$

as required. □

18.3. **Interpretation of the Jacobi identity.**

Remark 137. Let G be a Lie group. When we defined its Lie algebra $\mathfrak{g}$ and defined the Lie bracket $[,]$, we *promised* that the Jacobi identity followed from the associativity of the group law on G, but didn't prove it. We can now give a rigorous and fairly conceptual proof: Recall that we proved that any morphism $f : G \to H$ of Lie groups induces a morphism $df : \mathfrak{g} \to \mathfrak{h}$ of Lie algebras. Applying this fact with $H := \mathrm{GL}(\mathfrak{g})$ to the adjoint representation $f := \mathrm{Ad}$ implies that $\mathrm{ad} = d\,\mathrm{Ad}$ is a morphism of Lie algebras. But we saw above that this last assertion is equivalent to the Jacobi identity. (Exercise: check carefully that we haven't used circular reasoning here.)

18.4. Ad, ad **are intertwined by the exponential map.** Let G be a Lie group with Lie algebra $\mathfrak{g}$. One has maps

$$\exp : \mathfrak{g} \to G$$

and also a map

$$\exp : \mathrm{End}(\mathfrak{g}) \to \mathrm{GL}(\mathfrak{g})$$

(the matrix exponential). The adjoint maps defined above are intertwined by these exponential maps, that is to say, $\mathrm{Ad} \circ \exp = \exp \circ \mathrm{ad}$, or more verbosely, for each $X \in \mathfrak{g}$,

$$\mathrm{Ad}(\exp(X)) = \exp(\mathrm{ad}(X)) := \sum_{n=0}^{\infty} \frac{\mathrm{ad}(X)^n}{n!},$$

or even more verbosely, for each $X, Y \in \mathfrak{g}$,

$$\exp(X) Y \exp(-X) = \mathrm{Ad}(\exp(X))Y = \sum_{n=0}^{\infty} \frac{\mathrm{ad}(X)^n Y}{n!} = \sum_{n=0}^{\infty} \frac{[X, [X, \dots, [X, Y]]]}{n!}$$

where there are n copies of X in the iterated commutator on the RHS. One can prove this by writing the LHS and RHS as $\Phi_1(1)$ and $\Phi_2(1)$ respectively, where $\Phi_1(t) := \mathrm{Ad}(\exp(tX))$ and $\Phi_2(t) := \exp(\mathrm{ad}(tX)) = \exp(t\,\mathrm{ad}(X))$ are one-parameter subgroups in $\mathrm{GL}(\mathfrak{g})$; by Lemma 136, we have $\Phi_1'(0) = d\,\mathrm{Ad}(X) = \mathrm{ad}(X) = \Phi_2'(0)$, hence by the uniqueness of one-parameter subgroups the two sides coincide. This identity is already not entirely obvious in the matrix case $G = \mathrm{GL}_n(\mathbf{k})$; it is instructive to verify it directly in that case.

18.5. **Some low-rank exceptional isomorphisms induced by the adjoint representation.** The adjoint representation Ad of the groups $\mathrm{SL}_2(\mathbb{C}), \mathrm{SU}(2), \mathrm{SL}_2(\mathbb{R})$ induce the exceptional isomorphisms

$$\mathrm{PSL}_2(\mathbb{C}) := \mathrm{SL}_2(\mathbb{C})/\{\pm 1\} \cong \mathrm{SO}_3(\mathbb{C}),$$
$$\mathrm{SU}(2)/\{\pm 1\} \cong \mathrm{SO}(3),$$
$$\mathrm{PSL}_2(\mathbb{R}) := \mathrm{SL}_2(\mathbb{R})/\{\pm 1\} \cong \mathrm{SO}(1,2)^0.$$

We explained this in detail in class for the first two examples and left the third as an exercise. The adjoint representations ad of the corresponding Lie algebras likewise induce isomorphisms

$$\mathfrak{sl}_2(\mathbb{C}) \cong \mathfrak{so}_3(\mathbb{C}),$$
$$\mathfrak{su}(2) \cong \mathfrak{so}(3),$$
$$\mathfrak{sl}_2(\mathbb{R}) \cong \mathfrak{so}(1,2).$$

(Note that $\mathrm{Lie}(\mathrm{PSL}_2(\mathbb{C})) = \mathrm{Lie}(\mathrm{SL}_2(\mathbb{C}))$, etc., because $\{\pm 1\}$ is discrete.)

We spent some time in class introducing some terminology for interpreting the above isomorphisms in a natural way.

Definition 138. Let $\mathbf{k}$ be a field, perhaps of characteristic $\neq 2$. A (non-degenerate) *quadratic space* V *over* $\mathbf{k}$ is a pair $V = (V, Q)$, where

1. V is a finite-dimensional $\mathbf{k}$-vector space, and
2. Q is a map $Q : V \to \mathbf{k}$ for which the map $B := B_Q : V \times V \to \mathbf{k}$ defined by $B(x,y) := Q(x+y) - Q(x) - Q(y)$ has the properties:
   - (a) B is *bilinear*, i.e., $B(a_1x_1 + a_2x_2, b_1y_1 + b_2y_2) = \sum_{i=1,2} \sum_{j=1,2} a_i b_j B_Q(x_i, y_j)$ for all $a_i, b_j \in \mathbf{k}$ and $x_i, y_j \in V$;

(b) B is *non-degenerate* in the sense that for each nonzero $x \in V$ there exists a nonzero $y \in V$ so that $B(x, y) \neq 0$; equivalently, the map $x \mapsto B(x, \cdot)$ defines a linear isomorphism from V to its linear dual V^*.

A *morphism* of quadratic spaces $f : (V_1, Q_1) \to (V_2, Q_2)$ is a linear map $f : V_1 \to V_2$ so that $Q_2 \circ f = Q_1$. Two such quadratic spaces are thus *isomorphic* if there exists a linear isomorphism $f : V_1 \to V_2$ satisfying $Q_2 \circ f = Q_1$.

Example 139. If $\mathbf{k} = \mathbb{C}$ and $n \in \mathbb{Z}_{\geq 0}$, then $V := \mathbb{C}^n$ with $Q_n(x) := \sum_{i=1}^n x_i^2$ is a quadratic space, called the *standard n-dimensional quadratic space over $\mathbb{C}$*.

Example 140. If $\mathbf{k} = \mathbb{R}$ and $p, q \in \mathbb{Z}_{\geq 0}$, then $V := \mathbb{R}^{p+q}$ with $Q_{p,q}(x) := \sum_{i=1}^p x_i^2 - \sum_{j=1}^q x_{p+j}^2$ is a quadratic space, called the *standard quadratic space over $\mathbb{R}$ of signature (p, q)*.

Theorem 141. *Over $\mathbf{k} = \mathbb{R}$ or $\mathbb{C}$, every quadratic space is isomorphic to one of the above examples.*

Definition 142. Given a quadratic space $V = (V, Q)$ over $\mathbf{k}$ (take $\mathbf{k} = \mathbb{R}$ or $\mathbb{C}$ for the purposes of this course, although the construction applies more generally) we may define its *orthogonal group* $\mathrm{O}(V) = \{g \in \mathrm{GL}(V) : Q(gv) = Q(v) \text{ for all } v \in V\}$ and *special orthogonal group* $\mathrm{SO}(V) = \mathrm{O}(V) \cap \mathrm{SL}(V)$.

If two quadratic spaces are isomorphism, it is clear that their (special) orthogonal groups are likewise isomorphic. The above definition thus introduces no new groups beyond the examples $\mathrm{O}_n(\mathbb{C})$, $\mathrm{O}(n)$, $\mathrm{O}(p, q)$, $\mathrm{SO}_n(\mathbb{C})$, $\mathrm{SO}(n)$, $\mathrm{SO}(p, q)$ that we have already seen, but it is sometimes convenient to be able to refer to them in a coordinate-free manner.

Example 143. Over $\mathbf{k} = \mathbb{C}$, the space $\mathfrak{sl}_2(\mathbb{C})$ with the quadratic form det is a quadratic space. In fact, the linear map $j : \mathbb{C}^3 \to \mathfrak{sl}_2(\mathbb{C})$ given by

$$(x, y, z) \mapsto \begin{pmatrix} ix & y + iz \\ -y + iz & -ix \end{pmatrix} \in \mathfrak{sl}_2(\mathbb{C}) \tag{74}$$

satisfies $\det(j(x, y, z)) = x^2 + y^2 + z^2$, and hence induces an explicit isomorphism of quadratic spaces

$$(\mathbb{C}^3, Q_3) \cong (\mathfrak{sl}_2(\mathbb{C}), \det).$$

Thus, in particular,

$$\mathrm{SO}(\mathfrak{sl}_2(\mathbb{C}), \det) \cong \mathrm{SO}_3(\mathbb{C}).$$

Example 144. Over $\mathbf{k} = \mathbb{R}$, the space $\mathfrak{su}(2)$ with the quadratic form det is a quadratic space. In fact, the linear map $j : \mathbb{R}^3 \to \mathfrak{su}(2)$ given by (74) satisfies $\det(j(x, y, z)) = x^2 + y^2 + z^2$, and hence induces an explicit isomorphism of quadratic spaces

$$\mathrm{SO}(\mathfrak{su}(2), \det) \cong \mathrm{SO}(3).$$

One of the homework problems for this week is to work out something similar for $\mathfrak{sl}_2(\mathbb{R})$.

Consider now the adjoint map

$$\mathrm{Ad} : \mathrm{SL}_2(\mathbb{C}) \to \mathrm{GL}(\mathfrak{sl}_2(\mathbb{C})).$$

Since $\det(\mathrm{Ad}(g)X) = \det(X)$ for all $g \in \mathrm{SL}_2(\mathbb{C})$ and $X \in \mathfrak{sl}_2(\mathbb{C})$, we have in fact

$$\mathrm{Ad}(\mathrm{SL}_2(\mathbb{C})) \subseteq \mathrm{O}(\mathfrak{sl}_2(\mathbb{C}), \det) \cong \mathrm{O}_3(\mathbb{C}).$$

Since (as we have shown) $\mathrm{SL}_2(\mathbb{C})$ is connected, so is its image under Ad, thus in fact

$$\mathrm{Ad}(\mathrm{SL}_2(\mathbb{C})) \subseteq \mathrm{SO}_3(\mathbb{C}).$$

Similarly

$$\mathrm{ad}(\mathfrak{sl}_2(\mathbb{C})) \subseteq \mathfrak{so}(\mathfrak{sl}_2(\mathbb{C}), \det) \cong \mathfrak{so}_3(\mathbb{C}).$$

We may check easily that $\ker(\mathrm{Ad}) = \{\pm 1\}$ and $\ker(\mathrm{ad}) = \{0\}$; from this and a dimensionality check it follows that ad is a linear isomorphism and hence (using that $\mathrm{SO}_3(\mathbb{C})$ is connected and that the exponential map has the properties that it has) that $\mathrm{Ad} : \mathrm{SL}_2(\mathbb{C}) \to \mathrm{SO}_3(\mathbb{C})$ is surjective. We thereby obtain an isomorphism

$$\mathrm{SL}_2(\mathbb{C})/\{\pm 1\} \cong \mathrm{SO}_3(\mathbb{C}).$$

Similar arguments give the other isomorphisms claimed above. The homework problems for this week give some other interpretations.

Remark 145. One can check that the inverse isomorphism $\mathfrak{so}_3(\mathbb{C}) \xrightarrow{\cong} \mathfrak{sl}_2(\mathbb{C})$ is not of the form df for some morphism f of Lie groups $\mathrm{SO}_3(\mathbb{C}) \to \mathfrak{sl}_2(\mathbb{C})$; we shall return to this point later.

Using the above exceptional isomorphisms, together with the fact that $-1 \in \mathrm{SL}_2(\mathbb{C})$ acts on the irreducible representation W_n by the sign $(-1)^n$, we deduce that the irreducible representations of $\mathrm{SO}_3(\mathbb{C})$ are given by the W_{2n} for all $n \geq 0$; the action of $g \in \mathrm{SO}_3(\mathbb{C})$ on W_{2n} is defined to be $R(\tilde{g})$ for some preimage $\tilde{g} \in \mathrm{SL}_2(\mathbb{C})$. We should see next time a bit more explicitly how this goes.

19. Maurer–Cartan equations and applications

19.1. **The equations.** Let G be a Lie group and a smooth map

$$g : \mathbf{k}^2 \dashrightarrow G.$$

We can think of g as a smooth parametrized surface, or as a family of (possibly disconnected) curves $t \mapsto g(t, s)$ indexed by a deformation parameter s. The velocity of the curve with parameter s may be described by the smooth map

$$\xi : \mathbf{k}^2 \dashrightarrow \mathfrak{g}$$

characterized by the equation

$$\frac{\partial g}{\partial t} = g\xi.$$

Similarly, for a fixed time t, the velocity of the deformation is quantified by the map

$$\eta : \mathbf{k}^2 \dashrightarrow \mathfrak{g}$$

characterized by

$$\frac{\partial g}{\partial s} = g\eta.$$

Since g is smooth, we have

$$\frac{\partial^2 g}{\partial s \partial t} = \frac{\partial^2 g}{\partial t \partial s}.$$

Equating the two expressions obtained by expanding out the LHS and RHS and making use of the definition of the Lie bracket give what is known as the Maurer–Cartan equation.

For the remaining discussion, we either assume that G is a matrix Lie group (so that everything is a matrix for us to multiply willy-nilly) or use the trick of §15. Then

$$\frac{\partial^2 g}{\partial s \partial t} = \frac{\partial (g\xi)}{\partial s} = g\eta\xi + g\frac{\partial \xi}{\partial s},$$

while

$$\frac{\partial^2 g}{\partial t \partial s} = \frac{\partial (g\eta)}{\partial t} = g\xi\eta + g\frac{\partial \eta}{\partial t}.$$

Equating the two and using the key relation

$$g\eta\xi - g\xi\eta = g[\eta, \xi]$$

now gives

$$\frac{\partial \eta}{\partial t} - \frac{\partial \xi}{\partial s} = [\eta, \xi], \tag{75}$$

which we may expand a bit more verbosely as

$$\frac{\partial}{\partial t}(g^{-1}\frac{\partial g}{\partial s}) - \frac{\partial}{\partial s}(g^{-1}\frac{\partial g}{\partial t}) = [g^{-1}\frac{\partial g}{\partial s}, g^{-1}\frac{\partial g}{\partial t}]. \tag{76}$$

One can rewrite everything we've said here in the language of differential geometry in terms of derivatives of a certain natural $\mathfrak{g}$-valued 1-form on G; see Google for details.

19.2. **Lifting morphisms of Lie algebras.**

Theorem 146. *Let G, H be Lie groups. Consider the natural map $j : \operatorname{Hom}(G, H) \to \operatorname{Hom}(\mathfrak{g}, \mathfrak{h})$ given by $f \mapsto df$.*

(1) If G is connected, then j is injective.
(2) If G is simply-connected, then j is surjective.

We have already shown the first part. It remains to show when G is simply-connected that for each morphism $\phi : \mathfrak{g} \to \mathfrak{h}$ of Lie algebras that there is a morphism $f : G \to H$ of Lie groups so that $df = \phi$.

To construct f, start with an element $x \in G$. Since G is connected, we can find a smooth curve γ in G with $\gamma(0) = e, \gamma(1) = x$. This curve will satisfy a differential equation

$$\frac{\partial \gamma}{\partial t} = \gamma\xi$$

for some $\xi : \mathbb{R} \dashrightarrow \mathfrak{g}$. We take $\phi(\xi) := \phi \circ \xi : \mathbb{R} \dashrightarrow \mathfrak{h}$ and consider the curve $\delta : \mathbb{R} \dashrightarrow H$ satisfying the initial condition $\delta(0) = e$ and the differential equation

$$\frac{\partial \delta}{\partial t} = \delta\phi(\xi);$$

this can always be solved, by an argument similar to that used to construct the exponential map. We now attempt to define

$$f(x) := \delta(1).$$

The issue is that this definition is not obviously independent of the choice of path γ.

However, since G is *simply-connected*, we can join any two such paths γ_0, γ_1 by a smooth homotopy $g : \mathbb{R}^2 \dashrightarrow G$ satisfying

$$g(t, 0) = \gamma_0(t), \quad g(t, 1) = \gamma_1(t), \quad g(0, s) = e, \quad g(1, s) = x.$$

It will again satisfy a differential equation

$$\frac{\partial g}{\partial t} = g\xi$$

now for some $\xi : \mathbb{R}^2 \dashrightarrow \mathfrak{g}$. We can take the composition $\phi(\xi) : \mathbb{R}^2 \dashrightarrow \mathfrak{h}$ and solve for a function $h : \mathbb{R}^2 \dashrightarrow H$ satisfying the initial condition $h(0, s) = e$ and the differential equation

$$\frac{\partial h}{\partial t} = h\phi(\xi). \tag{77}$$

Then $h(t, 0) = \delta_0(t)$ and $h(t, 1) = \delta_1(t)$ where δ_0, δ_1 are attached to γ_0, γ_1 as above. Our aim is to show that $\delta_0(1) = \delta_1(1)$. To that end, it will suffice to show that

$$h(1, s) \text{ is independent of } s. \tag{78}$$

We can express what we are given and what we want to show more succinctly in terms of the deformation velocities $\eta : \mathbb{R}^2 \dashrightarrow \mathfrak{g}$ of g and $\zeta : \mathbb{R}^2 \dashrightarrow \mathfrak{h}$ characterized by

$$\frac{\partial g}{\partial s} = g\eta \tag{79}$$

and

$$\frac{\partial h}{\partial s} = g\zeta. \tag{80}$$

With this notation, what we are given is that

$$\eta(1, s) = 0 \tag{81}$$

and what we want to show is that

$$\zeta(1, s) = 0. \tag{82}$$

The required implication will follow in a slightly stronger form if we can show that

$$\zeta = \phi(\eta). \tag{83}$$

We are given the compatible inital conditions

$$\zeta(0, s) = 0 = \phi(\eta)(0, s). \tag{84}$$

The Maurer–Cartan equation now gives

$$\frac{\partial \eta}{\partial t} - \frac{\partial \xi}{\partial s} = [\eta, \xi]. \tag{85}$$

and

$$\frac{\partial \zeta}{\partial t} - \frac{\partial \phi(\xi)}{\partial s} = [\zeta, \phi(\xi)].$$

By applying ϕ to (85) and using that ϕ preserves brackets, we obtain

$$\frac{\partial \phi(\eta)}{\partial t} - \frac{\partial \phi(\xi)}{\partial s} = [\phi(\eta), \phi(\xi)].$$

Thus for each s, the functions $t \mapsto \zeta(t, s)$ and $t \mapsto \phi(\eta)(t, s)$ satisfy the same initial conditions at $t = 0$ and the same differential equation. The required identity (83) follows finally from the basic uniqueness theorem for ODEs.

We've shown that the map $f : G \to H$ is well-defined, i.e., independent of the choice of path.

Exercise 24. Show that f is a group homomorphism, i.e., $f(g_1 g_2) = f(g_1)f(g_2)$. [Hint: use the uniqueness of paths involved in the definition of f.]

It's also not hard to verify that f is smooth.

20. The universal covering group

We record the basic definitions and results from class (in the order essentially opposite to that in which they were presented)

Definition 147. Let $p_1 : X_1 \to Y$ and $p_2 : X_2 \to Y$ be a pair of maps. A map $f : X_1 \to X_2$ will be said to *commute with* p_1 *and* p_2 if the only conceivable condition relating the three maps is satisfied: $p_1 = p_2 \circ f$.

Definition 148. Let Z be a manifold and let Y be a connected manifold. The *trivial fiber bundle over* Y *with fiber* Z is the map $\mathrm{pr}_1 : Y \times Z \to Y$ given by taking the first coordinate. More generally, a *trivial fiber bundle over* Y *with fiber* Z is a smooth map of manifolds $p : X \to Y$ such that there exists a diffeomorphism $\iota : X \to Y \times Z$ commuting with p and pr_1.

Definition 149. Let Z be a manifold and let Y be a connected manifold. A *locally trivial fiber bundle over* Y *with fiber* Z is a map $p : X \to Y$, where X is a manifold and p is a smooth map with the property that each element of Y is contained in an open neighborhood U for which the induced map $p : p^{-1}(U) \to U$ is a trivial fiber bundle over U with fiber Z.

Definition 150. Let X, Y be connected manifolds. A *cover* $p : X \to Y$ is a locally trivial fiber bundle with discrete fiber. This means more concretely that every element of Y has an open neighborhood U so that $p^{-1}(U)$ is a disjoint union of open subsets V_α with the property that $p : V_\alpha \to U$ is a diffeomorphism. (A picture involving a "stack of pancakes" is appropriate here.)

Remark 151. One can speak about covers of much more general topological spaces; we won't ened such notions here.

Example 152. The natural map $p : \mathbb{R} \to \mathbb{R}/\mathbb{Z} \cong U(1) \cong S^1$ is a non-trivial locally trivial fiber bundle over the circle S^1 with fiber $\mathbb{Z}$.

Definition 153. A morphism $f : G \to H$ between connected Lie groups is a *covering morphism* if it is a morphism of Lie groups that is also a cover in the above sense.

Exercise 25. The following are equivalent for a morphism $f : G \to H$ between connected Lie groups:

(1) f is a cover in the sense of Definition 150, i.e., a locally trivial fiber bundle.
(2) $\ker(f)$ is discrete and f is onto.
(3) f is a local homeomorphism.
(4) $df : \mathfrak{g} \to \mathfrak{h}$ is an isomorphism.

(In all cases, f is surjective.)

Lemma 154. *If N is a discrete normal subgroup of a connected Lie group G, then N is contained in the center of G.*

In particular, the kernel of any covering morphism is a discrete subgroup of the center

Proof. Let $n \in N$. To show that n belongs to the center of G, we must verify that the set $S := \{gng^{-1} : g \in G\} \subseteq N$ satisfies $S = \{n\}$. Indeed, S is a subset of the *discrete* topological space N that contains n, but S is the continuous image under the map $g \mapsto gng^{-1}$ of the *connected* topological space G, so S is connected and thus $S = \{n\}$. $\square$

The following key result classifies connected Lie groups as quotients of simply-connected Lie groups by discrete central subgroups:

Theorem 155. *Let G be a connected Lie group. Then there exists a simply-connected (connected) Lie group $\tilde{G}$ and a covering morphism $p : \tilde{G} \to G$. The kernel N of p is a discrete subgroup of the center of G. One has $\pi_1(G) \cong N$. The pair $(\tilde{G}, N)$ is unique up to isomorphism.*

By combining with the result from last lecture, together with the (not proved yet and not obvious) fact that every finite-dimensional Lie algebra arises as the Lie algbebra of some Lie group, we obtain:

Corollary 156. *The category of simply-connected Lie groups is equivalent to the category of finite-dimensional Lie algebras, that is to say:*

- *Every simply-connected Lie group has a finite-dimensional Lie algebra, and every finite-dimensional Lie algebra arises (up to isomorphism) in this way.*
- *If G, H are simply-connected Lie groups with Lie algebras $\mathfrak{g}, \mathfrak{h}$, then the map $f \mapsto df$ induces a bijection $\operatorname{Hom}(G, H) = \operatorname{Hom}(\mathfrak{g}, \mathfrak{h})$. In particular, $G \cong H$ if and only if $\mathfrak{g} \cong \mathfrak{h}$.*

In lecture we gave the following examples:

(1) $G = \mathbb{R}/\mathbb{Z} \cong \mathrm{U}(1) \cong S^1$, $\tilde{G} = \mathbb{R}$, $N = \mathbb{Z}$
(2) $G = \mathrm{SO}(3), \tilde{G} = \mathrm{SU}(2), N = \{\pm 1\}$.
(3) $G = \mathbb{C}^\times, \tilde{G} = \mathbb{C}, p = \exp : \tilde{G} \to G$, $N = 2\pi i \mathbb{Z}$

We gave a few other similar examples. See the homework for further examples.

The proof of the theorem relies on the construction of the universal cover $p : \tilde{X} \to X$ of a connected manifold X, which is a cover in the sense of Definition 150 with the property: for any cover $q : \tilde{Y} \to Y$ and any smooth map $f : X \to Y$ and any $x \in X, y \in Y, \tilde{x} \in \tilde{X}, \tilde{y} \in \tilde{Y}$ satisfying the compatibility conditions $x = p(\tilde{x})$ and $y = f(x) = q(\tilde{y})$ there exists a unique smooth map $\tilde{f} : \tilde{X} \to \tilde{Y}$ such that

(1) $\tilde{f}$ lifts f in the sense that $q \circ \tilde{f} = p \circ f$, and
(2) $\tilde{f}(\tilde{x}) = \tilde{y}$.

The construction of $p : \tilde{X} \to X$ (which is not terribly important for our purposes) is as follows: One fixes a basepoint $x_0 \in X$ and defines $\tilde{X}$ as a set to be the set of all pairs $(x, [\gamma])$, where $x \in X$ and $[\gamma]$ is a homotopy class of smooth paths $\gamma : [0, 1] \to X$ with $\gamma(0) = x_0$ and $\gamma(1) = x$. The map $p : \tilde{X} \to X$ is given by $p((x, [\gamma])) := x$. We must now define a smooth structure on $\tilde{X}$ and verify that it is indeed a locally trivial fiber with discrete fiber and that the universal property is satisfied. We will content ourselves here to define the smooth structure; the verifications are then routine. Given a point $x_1 \in X$, there is a small open neighborhood U of x_1 that is diffeomorphic to an open Euclidean ball B. For each $x \in U$, each homotopy class $[\gamma]$ of paths γ from x_0 to x may be factored uniquely as $[\delta_x \circ \gamma_1]$, where δ_x is given by the straight line from x to x_1 (under the identification $U \cong B$) and $[\gamma_1]$ traverses the set A of homotopy classes of paths γ_1 from x_0 to x_1. We obtain in this way an identification of sets $p^{-1}(U) = \bigsqcup_{\alpha \in A} V_\alpha$, where each V_α identifies naturally with B. We use this identification to define the smooth structure on $p^{-1}(U)$.

Given a connected Lie group G, one obtains first a cover of manifolds $p : \tilde{G} \to G$. To define the group structure on $\tilde{G}$, one first selects an arbitrary preimage $\tilde{e} \in \tilde{G}$

of the identity element $e \in G$. Using the universal property, one obtains unique lifts $\tilde{m} : \tilde{G} \times \tilde{G} \to \tilde{G}$ and $\tilde{i} : \tilde{G} \to \tilde{G}$ of the multiplication and inversion maps $m : G \times G \to G$ and $i : G \to G$ satisfying $\tilde{m}(\tilde{e}, \tilde{e}) = \tilde{e}$ and $\tilde{i}(\tilde{e}) = \tilde{e}$. One can now check that the usual axioms (associativity, smoothness, etc.) are satisfied, and that $p : \tilde{G} \to G$ is a covering morphism. The kernel N of p is obviously discrete, so by Lemma 154, it is central.

One can also check, using the universal property, that $\pi_1(G) \cong N$: The isomorphism $j : \pi_1(G) \to N$ is obtained by defined by taking a homotopy class $[\gamma]$ of loops γ in G with basepoint e, using the universal property to lift them uniquely to paths $\tilde{\gamma} : \tilde{G} \to \tilde{G}$ with $\tilde{\gamma}(0) = \tilde{e}$, and setting $j([\gamma]) := \tilde{\gamma}(1)$. This is well-defined. Conversely, given $n \in N$, we can take a path $\tilde{\gamma}$ in $\tilde{G}$ from $\tilde{e}$ to $n \in N$ and project it down via p to obtain a loop γ in G based at the identity element $e \in G$. The maps $\pi_1(G) \to N$ and $N \to \pi_1(G)$ constructed in this way are mutually inverse group homomorphisms.

For more details, I recommend consulting the first reference listed on the course homepage.

21. Quotients of Lie groups

Theorem 157. *Let G be a Lie group and H a Lie subgroup. There is a unique smooth structure on the set G/H so that the natural map $G \to G/H$ is a submersion, or equivalently, a quotient map in the smooth category. If moreover H is a normal subgroup, then G/H is naturally a Lie group and $G \to G/H$ is a surjective morphism of Lie groups. Moreover, the map $G \to G/H$ is a locally trivial fiber bundle with fiber H (and also what is known as a "principal H-bundle"). Moreover, if the Lie group G acts transitively on a smooth manifold X and if H is the stabilizer in G of some point $x \in X$, then the natural map $G/H \to X$ given by $gH \mapsto gx$ is a diffeomorphism.*

The final assertion that $G/H \cong X$ is related to Exercise 11; its proof in the present setting requires the second-countability assumption on our manifolds. We sketched the construction of the smooth structure on G/H in some detail in lecture, leaving the verification of the properties as an exercise; see the first reference on the course webpage for more details.

The argument here is very similar (basically "dual to") that concerning submanifolds given by Theorem 24. In this analogy, "immersion" is to "submanifold" as "submersion" is to "quotient manifold."

By an argument dual to that of Corollary (25), we know that a smooth surjection $p : X \to Y$ is a submersion if and only if it is a *quotient map* in the category of smooth manifolds, that is to say, if and only if smooth maps of manifolds $Y \to Z$ are in natural bijection with smooth maps $X \to Z$ that factor set-theoretically through p. (This is easy to see in local coordinates, using the local description of submersions as surjective linear maps.) The proof of uniqueness of smooth structure on G/H is then dual to that given in the proof of Theorem 24.

For existence, one first takes a small enough submanifold S of G that is transversal to H at the identity element (draw a picture). By the inverse function theorem, the multiplication map $\mu : S \times H \to G$ is then a local diffeomorphism at the identity, and is in particular injective in a neighborhood of the identity. (In a bit more detail: the transversality assumption says that that the derivative $T_{(e,e)}\mu : T_e S \times \mathfrak{h} \to \mathfrak{g}$ is an isomorphism, hence by the inverse function theorem, it is a diffeomorphism in a neighborhood of the origin.)

It follows that if S is small enough, then μ is actually a diffeomorphism onto its image: for else we may (as explained in more detail in lecture) find arbitrarily small distinct $s_1, s_2 \in S$ for which $s_1^{-1} s_2 \in H$, contrary to the local injectivity of μ. In particular, after shrinking S as necessary, the map of sets $\pi : S \to G/H$ is injective. Denote by $U \subseteq G/H$ the image of S and by gU the image of the translate gS by an element $g \in G$. We equip gU with the smooth structure transferred from gS. Then π is a submersion over gU. The sets gU cover G/H, and the smooth structures on their overlaps are compatible thanks to the uniqueness established before (compare with the proof of Theorem 24). One can check using the definition of quotient maps that the natural map $G \times G/H \to G/H$ is smooth, and that if H is normal, then the induced multiplication map on G/H is smooth, so G/H is naturally a Lie group.

22. Homotopy exact sequence

If G, H are Lie groups with G connected, one has an exact sequence

$$\pi_2(G/H) \to \pi_1(H) \to \pi_1(G) \to \pi_1(G/H) \xrightarrow{\delta} H/H^0 \to 0$$

where δ sends the homotopy class $[\gamma]$ of a loop $\gamma : [0,1] \to H$ based at the identity element to $\delta([\gamma]) := \tilde{\gamma}(1)^{-1}$, where $\tilde{\gamma} : [0,1] \to G$ is the unique lift of γ to a path in G satisfying $\tilde{\gamma}(0) = e$.

Corollary 158. *If $\pi_1(G/H) = \pi_2(G/H) = 0$, then $\pi_1(G) \cong \pi_1(H)$.*

Corollary 159. *If $\pi_1(G) = 0$, then $\pi_1(G/H) \cong H/H^0$.*

We explained how this may be used to compute inductively the fundamental groups of the classical groups; see the first reference on the course webpage for more details.

23. Cartan decomposition

Let G be a real Lie subgroup of $\mathrm{GL}_N(\mathbb{C})$ with the property that $\Theta(G) = G$, where $\Theta : \mathrm{GL}_N(\mathbb{C}) \to \mathrm{GL}_N(\mathbb{C})$ is the involution given by inverse conjugate transpose:

$$\Theta(g) := {}^t\overline{g}^{-1}.$$

Then the set

$$K := \{g \in G : \Theta(g) = g\} = \mathrm{U}(N) \cap G$$

of elements in G fixed by Θ, or equivalently, belonging to the unitary rgoup, may be shown to be a real Lie subgroup with Lie algebra

$$\mathfrak{k} = \{X \in \mathfrak{g} : \theta(X) = X\}$$

where $\theta := d\Theta$ is given by

$$\theta(X) := -\overline{X}^t.$$

Set

$$\mathfrak{p} := \{X \in \mathfrak{g} : \theta(X) = -X\}.$$

Then $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. In words, $\mathfrak{k}$ consists of the skew-hermitian elements of $\mathfrak{g}$, while $\mathfrak{p}$ consists of the hermitian elements of $\mathfrak{g}$; the assumption $\Theta(G) = G$ implies also that $\theta(\mathfrak{g}) = \mathfrak{g}$, hence (easily) that such a decomposition exists.

Example 160. If $G = \mathrm{GL}_n(\mathbb{C})$, then $K = \mathrm{U}(n)$ and $\mathfrak{p}$ consists of all hermitian matrices and $\exp(\mathfrak{p})$ consists of all positive-definite hermitian matrices.

In particular, if $n = 1$, then $G = \mathbb{C}^\times, K = U(1)$, and $\mathfrak{p}$ consists of all 1×1 real matrices.

Example 161. If $G = \mathrm{GL}_n(\mathbb{R})$, then $K = \mathrm{O}(n)$ and $\mathfrak{p}$ consists of all symmetric matrices and $\exp(\mathfrak{p})$ consists of all positive-definite symmetric matrices.

Example 162. One of the homework problems this week is to verify that if $G = \mathrm{O}(p, q)$, then $K = \mathrm{O}(p) \times \mathrm{O}(q)$.

Definition 163. We say that G is (real) *algebraic* if it may be defined inside $\mathrm{GL}_N(\mathbb{C})$ by a system of polynomial equations in the real and imaginary parts of group elements and their inverses. (Every example we've seen has this property.)

Theorem 164 (Cartan decomposition)**.** *Let G, K be as above. Assume that G is algebraic. Then the natural map*

$$K \times \mathfrak{p} \to G$$
$$(k, Y) \mapsto k \exp(Y)$$

is a diffeomorphism.

Remark 165. If G has finitely many connected components, then the assumption that G is algebraic turns out to hold automatically in this setup (TODO: double-check this), but we probably won't have time to prove this. In any event, it's true in all of the examples we've seen.

Example 166. In the context of Example 160, this amounts to the "polar decomposition" of an invertible complex matrix g as a product of a unitary matrix k and a positive-definite hermitian matrix $\exp(Y)$. In the special case $n = 1$, this is just the polar decomposition of a nonzero complex numbers as $e^{i\theta} r$ (writing $r = e^y, y \in \mathbb{R}$).

Corollary 167. *Let G, K be as above. Then K is a deformation retract of G. In particular, $\pi_i(G) \cong \pi_i(K)$ for all $i \geq 0$.*

(One of the homework problems involves applying this last corollary in the special case $i = 0$ to relate the connected components of G and K.)

Remark 168. This corollary "explains" some of the "coincidences" such as

$$\pi_1(\mathrm{SL}_2(\mathbb{R})) \cong \mathbb{Z} \cong \pi_1(\mathrm{SO}(2)),$$
$$\pi_1(\mathrm{SL}_n(\mathbb{R})) \cong \mathbb{Z}/2 \cong \pi_1(\mathrm{SO}(n)) \text{ for } n \geq 3,$$
$$\pi_1(\mathrm{SU}(n)) \cong \{0\} \cong \pi_1(\mathrm{SL}_n(\mathbb{C})),$$
$$\pi_1(\mathrm{U}(n)) \cong \mathbb{Z} \cong \pi_1(\mathrm{GL}_n(\mathbb{C})),$$
$$\pi_1(\mathrm{SO}_n(\mathbb{C})) \cong \pi_1(\mathrm{SO}(n))$$

that we observed empirically at the beginning of the lecture.

Remark 169. Let G be a connected complex Lie group, and suppose it has a compact real form K, so that $\mathfrak{g} \cong \mathfrak{k} \oplus i\mathfrak{k}$. The point of this remark is to indicate briefly (as is evident in all examples) why the Cartan decomposition should always apply to G and K. It turns out (as we'll show later in the course) that we may always realize K as a subgroup of $\mathrm{U}(N)$ for some N. We can then realize G as a subgroup of $\mathrm{GL}_N(\mathbb{C})$ in such a way that $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$; given what we've seen in the course, we can verify this already in the special case G is simply-connected (by lifting the inclusion map $\mathfrak{g} \hookrightarrow \mathfrak{gl}_N(\mathbb{C})$), and by the end of the course we should also be able to reduce the general case to that special one. Then $K = G \cap \mathrm{U}(N)$.

We briefly indicate the proof of Theorem 164. One basically takes the proof of the special case (see Example 166) concerning polar decomposition on $\mathrm{GL}_N(\mathbb{C})$ (perhaps seen in a linear algebra course?) and checks that it descends to G. So, let's see. There are a few things to check.

(1) The map $K \times \mathfrak{p} \ni (k, Y) \mapsto k \exp(Y) \in G$ is bijective.
(2) The map $Y \mapsto \exp(Y)$ is a diffeomorphism onto its image.

Using the first two assertions and the inverse function theorem, one gets that the map $(k, Y) \mapsto k \exp(Y)$ is itself a diffeomorphism onto its image. To verify the second assertion, we should first compute the differential of the exponential map (a result of independent interest). The answer is that for any $X, Y \in \mathfrak{g}$,

$$\exp(-X)\frac{d}{dt}\exp(X + tY)|_{t=0} = \Psi(\mathrm{ad}_X)Y \tag{86}$$

where

$$\Psi(z) := \sum_{n=1}^{\infty} \frac{(-z)^{n-1}}{n!} = \frac{1 - \exp(-z)}{z}.$$

This is obtained by applying Homework 9 to the map $f : \mathbb{R} \to \mathfrak{g}$ given by $f(t) := X + tY$; one then has $f(0) = X$ and $f'(0) = Y$, and so (86) follows from Homework 9. It follows from the second description of Ψ that $(d\exp)_X$ is injective provided that ad_X has no eigenvalues of the form $2\pi i k$ with k a nonzero integer. Since hermitian matrices have *real* eigenvalues (under the standard representation as well as the adjoint representation), it follows in particular that $\exp : \mathfrak{p} \to \mathrm{GL}_N(\mathbb{C})$ is everywhere regular. Finally, we observe that every positive hermitian matrix g may be written uniquely as the exponential of a symmetric matrix: any such matrix is diagonal with respect to some basis and has positive real entries on the diagonal, etc. In summary, $\exp : \mathfrak{p} \to \mathrm{GL}_N(\mathbb{C})$ is a diffeomorphism onto its image, as required.

All that remains now is the bijectivity. We verify first the injectivity, which will serve as useful motivation. Suppose that $g = k \exp(Y)$. Since $\Theta(k) = k$ and $\Theta(\exp(Y)) = \exp(\theta(Y)) = \exp(-Y) = \exp(Y)^{-1}$ and Θ is a homomorphism, it follows that $g^{-1}\Theta(g) = \exp(2Y)$. Since $\exp : \mathfrak{p} \to \exp(\mathfrak{p})$ is bijective, it follows that Y is uniquely determined by g, hence so is $k = g\exp(-Y)$.

We turn finally to surjectivity. Given $g \in G \leq \mathrm{GL}_N(\mathbb{C})$, we can verify directly that $\Theta(g)^{-1}\Theta(g)$ is a positive definite hermitian matrix, and so we can define $Y \in \mathfrak{gl}_N(\mathbb{C})$ to be the unique hermitian matrix for which $\exp(2Y) = \Theta(g)^{-1}\Theta(g)$. It is then not hard to verify that $k := g\exp(-Y)$ is unitary, i.e., $\Theta(k) = k$. This gives the required decomposition in the Lie group $\mathrm{GL}_N(\mathbb{C})$; the problem is to show that in fact $k \in K$ and $Y \in \mathfrak{p}$, or equivalently, that $k \in G$ and $Y \in \mathfrak{g}$. Since $\exp(\mathfrak{g}) \subseteq G$ and because of the way k was defined, it will suffice to verify that $Y \in \mathfrak{g}$. What we know (from our assumptions $g \in G$ and $\Theta(G) = G$) is that $\exp(2Y) \in G$. Since G is a group, we can raise the last assertion to any integer power $t \in \mathbb{Z}$ to see that $\exp(2tY) \in G$. Since Y is hermitian, we may choose a basis with respect to which it is diagonal and suppose that $Y = \mathrm{diag}(y_1, \ldots, y_N)$ for some real numbers $y_1, \ldots, y_N \in \mathbb{R}$. Since G was assumed to be algebraic (defined by polynomial equations), the condition $\exp(2tY) \in G$ is then a system of polynomial equations involving the positive real numbers $\exp(2ty_1), \dots, \exp(2ty_N)$. Since this polynomial system is satisfied for all integers t, one can show (see below) that it is satisfied also for all real numbers t. Thus $\exp(2tY) \in G$ for all $t \in \mathbb{R}$. By differentiating this last fact we deduce as required that $Y \in \mathfrak{g}$.

For completeness, we record the algebraic fact that we used in the proof:

Exercise 26. Let $a_1, \dots, a_n \in \mathbb{R}$ and let $F \in \mathbb{R}[x_1, \dots, x_n]$ be a polynomial satisfying

$$F(e^{a_1 t}, \dots, e^{a_n t}) = 0 \tag{87}$$

for all $t \in \mathbb{Z}$. Show that (87) holds also for all $t \in \mathbb{R}$.

(We apply this with $n := 2N$ and $(a_1, \dots, a_n) := (y_1, \dots, y_N, -y_1, \dots, -y_N)$.)

24. BCHD

Let G_1, G_2 be Lie groups.

Definition 170. We say that G_1, G_2 are *locally isomorphic* if there are open neighborhoods $U_i \subseteq G_i$ of the identity elements and a diffeomorphism $f : U_1 \xrightarrow{\cong} U_2$ so that whenever $x, y, xy \in U_1$, one has $f(xy) = f(x)f(y)$. In that case, f is said to be (guess!) a *local isomorphism.*

Example 171.

(1) Any covering homomorphism $f : G_1 \to G_2$ induces a local isomorphism.
(2) If G_1 is the connected component of G_2, then the inclusion $G_1 \hookrightarrow G_2$ defines a local isomorphism.
(3) The relation of being locally isomorphic is obviously an equivalence relation, i.e., is reflexive and transitive. In verifying this it is convenient to note that one can always shrink the subsets U_1, U_2 suitably.

Theorem 172. *G_1, G_2 are locally isomorphic if and only if $\mathfrak{g}_1, \mathfrak{g}_2$ are isomorphic.*

Proof. The forward direction is easy: given $f : U_1 \to U_2$ with inverse $f^{-1} : U_2 \to U_1$ as in the definition of "locally isomorphic," the differentials at the identity $df : \mathfrak{g}_1 \to \mathfrak{g}_2$ and $d(f^{-1}) : \mathfrak{g}_2 \to \mathfrak{g}_1$ define morphisms of Lie algebras (by the same proof as in §14.4) which are mutually inverse.

The converse direction is more subtle, but follows from what we have seen already. Namely, let G_1^0 denote the connected component and $\widetilde{G_1^0}$ its simply-connected covering group. Then $\mathfrak{g}_1 = \mathrm{Lie}(G_1) = \mathrm{Lie}(G_1^0) = \mathrm{Lie}(\widetilde{G_1^0})$. Let $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ be an an isomorphism. Then ϕ is of the form df for some $f : \widetilde{G_1^0} \to G_2$. Since ϕ is an isomorphism, we know by (for instance) Exercise 25 that f is a covering morphism. By Example 171, it follows that G_1 is locally isomorphic to G_1^0 which is in turn locally isomorphic to $\widetilde{G_1^0}$ which is finally (via f) locally isomorphic to G_2, as required. □

It is natural to ask for a more "local" proof of Theorem 172 that does not require topological considerations or global constructions involving universal covers, etc. The BCHD formula gives such a proof. To motivate that, recall that $\exp : \mathfrak{g} \to G$ is a local diffeomorphism near the origin; for this reason, it makes sense to define for small enough $x, y \in \mathfrak{g}$ the quantity

$$x * y := \log(\exp(x)\exp(y)).$$

(We explained this in somewhat more detail in class.) Thus $x * y$ is the group law on G expressed in local coordinates defined via the exponential map. One has identities like

$$x * (y * z) = (x * y) * z$$

whenever all involved quantities make sense (e.g., whenever x, y, z are all small enough). We also have

$$x * (-x) = 0,$$

etc. If G is abelian, then $x * y = x + y$. In general, it is somewhat more complicated. (See the homework problems this week for some examples where it isn't too complicated.)

Exercise 27. If $\phi : \mathfrak{g} \to \mathfrak{h}$ is the differential of some morphism of Lie groups $G \to H$ and $x, y \in \mathfrak{g}$ are small enough, then

$$\phi(x) * \phi(y) = \phi(x * y).$$

Suppose now temporarily that $G = \mathrm{GL}_n(\mathbb{R})$, so that $\mathfrak{g} = M_n(\mathbb{R})$. Using the series expansions $\log(z) = \sum_{m \geq 1} (-1)^{m-1} (z-1)^m / m$ and $\exp(x) = \sum_{p \geq 0} x^p / p!$, we obtain

$$\exp(x)\exp(y) - 1 = \sum_{\substack{p,q \geq 0: \\ (p,q) \neq (0,0)}} \frac{x^p y^q}{p! q!}$$

and thus

$$x * y = \sum_{m \geq 1} \frac{(-1)^{m-1}}{m} \sum_{\substack{p_1, q_1, \dots, p_m, q_m \geq 0 \\ (p_j, q_j) \neq (0,0)}} \frac{x^{p_1} y^{q_1} \cdots x^{p_m} y^{q_m}}{p_1! q_1! \cdots p_m! q_m!}. \tag{88}$$

For $n \geq 1$, let z_n denote the nth homogeneous component of the sum on the RHS, so that

$$x * y = \sum_{n \geq 1} z_n. \tag{89}$$

If we play around for a bit (as done in class), we find quickly that

$$z_1 = x + y,$$

$$z_2 = \frac{1}{2}[xy]$$

where $[xy] := [x, y]$. The verification of this involved a "miraculous" coincidence of the shape

$$\left(\frac{x^2}{2} + xy + \frac{y^2}{2} \right) - \frac{1}{2}(x^2 + xy + yx + y^2) = \frac{xy - yx}{2}.$$

We indicated that with more work one can show that

$$z_3 = \frac{1}{12}([x[xy]] + [y[yx]])$$

and

$$z_4 = \frac{1}{24}[y[x[yx]]].$$

Theorem 173 (BCH "formula")**.** *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then for small enough $x, y \in \mathfrak{g}$, the identity* (89) *holds for some degree n Lie polynomial z_n in x, y (i.e., z_n is a linear combination of iterated n-fold commutators as above).*

There is a more explicit version of this:

Theorem 174 (BCHD formula)**.** *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then for small enough $x, y \in \mathfrak{g}$, the identity* (89) *holds with*

$$z_n = \frac{1}{n} \sum_{m \geq 1} \frac{(-1)^{m-1}}{m} \sum_{\substack{p_1, q_1, \ldots, p_m, q_m \geq 0 \\ (p_j, q_j) \neq (0,0) \\ p_1 + q_1 + \cdots + p_m + q_m = n}} \frac{[x^{p_1} y^{q_1} \cdots x^{p_m} y^{q_m}]}{p_1! q_1! \cdots p_m! q_m!} \tag{90}$$

where, for instance,

$$[x^3 y^2 x^4 y^1] := [x[x[x[y[y[x[x[x[xy]]]]]]]].$$

The similarity between (88) and (90) is no coincidence:

Exercise 28. Convince yourself that the problem involving commutators on Homework 6 (known as something like *Dynkin's lemma*) allows one to deduce Theorem 174 from Theorem 173.

To prove Theorem 173, define

$$f(t) := x * ty$$

for small $t \in \mathbb{R}$. Then f is smooth, and

$$x * y = f(1) = f(0) + \int_{t=0}^{1} f'(t)\, dt.$$

Since $\exp f(t) = \exp(x) \exp(ty)$ we have $\partial_t \exp f(t) = f(t) y$ and thus by Homework 9,

$$y = \exp(-f(t)) \partial_t \exp(f(t)) = \Psi(\mathrm{ad}_{f(t)}) f'(t),$$

where

$$\Psi(z) = \sum_{m=1}^{\infty} \frac{(-z)^{m-1}}{m!} = \frac{1 - \exp(-z)}{z}.$$

We form the inverse power series

$$\Psi(z)^{-1} = \frac{z}{1 - \exp(-z)} = 1 + \frac{z}{2} + \cdots.$$

We then have

$$f'(t) = \Psi(\mathrm{ad}_{f(t)})^{-1} y.$$

Using Exercise 27, we have

$$\mathrm{ad}_{f(t)} = \mathrm{ad}_x * \mathrm{ad}_{ty} = \log(e^{\mathrm{ad}(x)} e^{t\,\mathrm{ad}(y)})),$$

hence

$$\Psi(\mathrm{ad}_{f(t)})^{-1} = \psi(e^{\mathrm{ad}(x)} e^{t\,\mathrm{ad}(y)})$$

where

$$\psi(w) := \Psi(\log(w))^{-1} = \frac{w \log w}{w - 1} = 1 + (w-1)/2 + \cdots$$

(whose coefficients are what are called Bernoulli numbers). In summary,

$$x * y = x + \int_{t=0}^{1} \psi(e^{\mathrm{ad}(x)} e^{t\,\mathrm{ad}(y)}) y\, dt.$$

We can now expand the integrand out into a power series and integrate term-by-term; we then obviously get an analytic expression of the required form. (In fact, it's not too hard to push this analysis a bit further to derive (90) directly, without using Dynkin's trick; just expand everything out.)

Remark 175. The most "conceptual" perspective on the BCH theorem in its qualitative form may be found in Serre's book on Lie algebras and Lie groups in one of the final chapters on Lie algebras; see also around p72 of the book by Onischik–Vinberg–Gourbatsevich.

Note that Theorem 172 follows directly from the BCHD formula: if $x, y \in \mathfrak{g}_i$ are small enough then the product $\exp(x)\exp(y)$ is determined entirely by the Lie bracket on $\mathfrak{g}_i$, so an isomorphism $\mathfrak{g}_1 \cong \mathfrak{g}_2$ obviously lifts to a local isomorphism between neighborhoods of the identity elements in G_1, G_2.

There is a lot more to say about the BCHD formula; a taste is given on the homework set for this lecture. For some problems it may help to note that for any fixed norm $|.|$ on $\mathfrak{g}$ and x, y small enough, one has

$$x * y = O(|x|\,|y|).$$

Next lecture we should state some further consequences of BCHD.

25. Some more ways to produce and detect Lie groups

25.1. **Summary.** Recall that we have called a subgroup H of a Lie group G a *Lie subgroup* if it is a submanifold, and that for this to hold, it suffices to verify that H is locally a submanifold near the identity element of G. Checking this condition over and over again eventually becomes tedious, so we ask for some more systematic ways to detect it. Here are a few, to be developed in detail throughout this section:

(1) It was observed in §11.5 that Lie subgroups are automatically closed. Much more interestingly and perhaps surpririnsgly, the converse is true over $k = \mathbb{R}$: any closed subgroup H of a *real* Lie group G is automatically a Lie subgroup (Theorem 176). This is very powerful, and implies most of the criteria discussed below.

   This criterion does not apply directly to complex Lie groups: for instance a real Lie subgroup of a complex Lie group is seldom a complex Lie subgroup (think $\mathbb{R} \hookrightarrow \mathbb{C}$ or $\mathrm{GL}_n(\mathbb{R}) \hookrightarrow \mathrm{GL}_n(\mathbb{C})$). But it's not hard to verify (e.g., by inspecting the proof of what we are talking about) that if H is a real Lie subgroup of a complex Lie group G with the property that $\mathfrak{h}$ is a complex vector space of $\mathfrak{g}$, then H is a complex Lie subgroup of G. A fairly good rule of thumb is that if a subgroup H of a complex Lie group G has the properties
   (a) H is closed, and
   (b) the definition of H does not make reference to the real numbers, complex conjugation or similar "non-holomorphic" notions,

   then H is probably a complex Lie subgroup.

   Some of the methods used to prove the criteria to be given below are of independent interest, even in the real case, because they give convenient ways to compute Lie algebras in many common situations.

(2) "Stabilizers" of any sort (of points in a manifold, of vectors in a representation, etc.) are, in practice, obviously closed, hence are Lie subgroups by the previous item. Moreover,their Lie algebras tend to be "the obvious thing." Many subgroups can be somehow interpreted as stabilizers:
   (a) kernels of morphisms of Lie groups,
   (b) stabilizers of subspaces in representations,
   (c) intersections of Lie subgroups,

(d) etc.

(3) An interesting result that does not follow from the above criteria is that in a *simply-connected* Lie group G, the commutator subgroup $G' := [G, G]$ is a Lie subgroup. This conclusion fails in general, although it remains true that the commutator subgroup is an *immersed* Lie subgroup.

(4) Given a Lie subgroup H of a Lie group G, one can naturally construct the quotient manifold G/H; if H is normal, then G/H is also a Lie group.

25.2. **Closed subgroups of real Lie groups.**

25.2.1. *Statement of the key result.*

Theorem 176. *Let G be a real Lie group, and let $H \subseteq G$ be a subset. The following are equivalent:*

(1) H is a closed subgroup of G, that is to say,
   (a) H is a closed subset of G, and
   (b) $e \in H$, and $h^{-1}, h_1h_2 \in H$ whenever $h, h_1, h_2 \in H$.

(2) H is a Lie subgroup of G, that is to say,
   (a) H is a submanifold of G, and
   (b) $e \in H$, and $h^{-1}, h_1h_2 \in H$ whenever $h, h_1, h_2 \in H$.

25.2.2. *A toy example.* One can already illustrate the basic idea behind Theorem 176 in the case $G = \mathbb{R}$, where it amounts to the following:

Theorem 177. *Let H be a closed subgroup of the real line $\mathbb{R}$. Then exactly one of the following possibilities occur:*

- *$H = \mathbb{R}$, or*
- *H is discrete, or equivalently, one has $(-\varepsilon, \varepsilon) \cap H = \{0\}$ for some $\varepsilon > 0$.*

Proof. If the second possibility does not occur, then we can find a sequence $x_n \in H$ with $x_n \to 0$ and $x_n \neq 0$. Let $x \in \mathbb{R}$ be given. We can find a sequence of reals c_n so that $c_n x_n \to x$; for instance, one can take $c_n = x/x_n$. But since $x_n \to 0$, the conclusion $c_n x_n \to x$ is unaffected by rounding c_n to the nearest integer. We can thus find a sequence of *integers* c_n so that $c_n x_n \to x$. Since H is a subgroup, we then have $c_n x_n \in H$ for all n. Since H is closed, it follows that $x \in H$. Since x was arbitrary, we conclude as required that $H = \mathbb{R}$. □

25.2.3. *Proof of the key result.*

Proof of Theorem 176. Thanks to ??? and ???, all we need to show is that if H is a closed subgroup of G, then H is locally a submanifold of G at the identity. Let $\mathfrak{g}$ denote the Lie algebra of $\mathfrak{g}$. Set

$$\mathfrak{h} := \left\{ x \in \mathfrak{g} : \exists c_n \in \mathbb{R}, x_n \in \mathfrak{g} \cap \exp^{-1}(H) \text{ so that } x_n \to 0, c_n x_n \to x \right\}.$$

Two quick remarks before continuing with the proof:

- If we somehow knew already that H were a Lie subgroup, then $\mathfrak{h}$ would of course be its Lie algebra, as any x_n in the definition would belong to $\mathfrak{h}$ for n large enough.
- A good "enemy scenario" to keep in mind is when $G = (\mathbb{R}/\mathbb{Z})^2$ and H is the image of the map $x \mapsto (x, \alpha x)$, where $\alpha \in \mathbb{R} - \mathbb{Q}$. Then H is a subgroup (indeed, an immersed Lie subgroup), but fails to be closed. The set $\mathfrak{h}$ defined as above is all of $\mathfrak{g}$, and so has nothing to do with the Lie algebra of H. The argument to follow will need to rule out this scenario.

We show now that

$$\mathfrak{h} \text{ is a vector subspace of } \mathfrak{g} \text{ for which } \exp(\mathfrak{h}) \subseteq H.$$

(1) It is clear that $\mathfrak{h}$ is stable under scalar multiplication.
(2) Let $x \in \mathfrak{h}$, so that $x = \lim c_n x_n$ for some c_n, x_n as in the definition of $\mathfrak{h}$. The condition $x_n \to 0$ implies that rounding c_n to the nearest integer does not affect the condition $c_n x_n \to x$, so we may assume without loss of generality that $c_n \in \mathbb{Z}$. Since H is a closed subgroup, we then have $\exp(x) = \lim \exp(x_n)^{c_n} \in H$.
(3) Let $x, y \in \mathfrak{h}$. For n large enough, set $z_n := \log(\exp(x/n)\exp(y/n))$. By the previous item, $\exp(z_n) = \exp(x/n)\exp(y/n) \in H$. As we've seen earlier in the course (during the discussion of the exponential map; what we need here also follows easily from BCH), we have $z_n = x/n + y/n + O(1/n^2)$, hence $z_n \to 0$ and $nz_n \to x + y$. Therefore $x + y \in \mathfrak{h}$.

Let $\mathfrak{h}' \leq \mathfrak{g}$ be any vector space complement to $\mathfrak{h}$, so that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}'$. The map $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}' \ni (v, w) \mapsto \exp(v)\exp(w)$ has derivative 1 at the origin, hence is a local diffeomorphism. There is thus a small open neighborhood U of the identity element in G and a smooth chart

$$(\alpha, \alpha') : U \to \mathfrak{h} \oplus \mathfrak{h}'$$

characterized by the identity

$$g = \exp(\alpha(g))\exp(\alpha'(g)) \text{ for all } g \in U.$$

We claim that if U is small enough, then

$$U \cap H = \{g \in U : \alpha'(g) = 0\}.$$

This shows that H is locally a submanifold of G at the identity, as required.

To prove the claim, note first that if $\alpha'(g) = 0$, then $g = \exp(\alpha(g)) \in \exp(\mathfrak{h}) \subseteq H$. This establish one inclusion.

Conversely, if the reverse inclusion fails for arbitrarily small U, then we can find a sequence $h_n \in U \cap H$ with $h_n \to 1$ so that $\alpha'(h_n) \neq 0$. Set

$$x_n := \alpha'(h_n).$$

Since $\exp(x_n)$ belongs to the group generated by $h \in H$ and $\exp(\alpha(h_n)) \in \exp(\mathfrak{h}) \subseteq H$, it belongs to H, and so

$$x_n \in \mathfrak{h}' \cap \exp^{-1}(H), \quad x_n \to 0, x_n \neq 0.$$

By passing to a subsequence, we have $|x_n|^{-1}x_n \to x$ for some nonzero $x \in \mathfrak{h}'$, but then also $x \in \mathfrak{h}$; since $\mathfrak{h} \cap \mathfrak{h}' = 0$, we obtain the required contradiction. □

A simple corollary that already illustrates the basic idea is the following:

Corollary 178. *Let H be a closed subgroup of $\mathbb{R}^n$. Then there is a vector space $V \leq \mathbb{R}^n$ and an open neighborhood $0 \in U \subseteq \mathbb{R}^n$ so that $H \cap U = V \cap U$.*

Proof. We take $G := \mathbb{R}^n$ and note that $\mathfrak{g} = \mathbb{R}^n$ and that the exponential map $\mathfrak{g} \to G$ is the identity. The conclusion then follows from Theorem 176. □

Exercise 29. Write down a direct proof of Corollary 178. (The proof of Theorem 176 simplifies a bit in this special case while retaining its basic flavor; it is instructive to work out exactly how it simplifies.)

25.3. **Stabilizers.**

25.3.1. *The basic result.* Let M be a manifold and G a Lie group acting on M, i.e., equipped with a smooth map $G \times M \to M$ satisfying the usual requirements of an action (see Definition 59). Let $x \in M$. Consider the orbit map $\alpha : G \to M$ given by $\alpha(g) := gx$. A crucial property of this map is that it has *constant rank*. Indeed, its rank at g_2 is the same as its rank at $g_1 g_2$ thanks to the identity $\alpha(g_1 g_2) = \alpha(g_1)\alpha(g_2)$ and the fact that $\alpha(g_1)$ is a diffeomorphism. Using the constant rank theorem from multivariable calculus, it follows that α is linearizable in a neighborhood of any point of G, and in particular, near the identity element. It follows that

$$\mathrm{Stab}_G(x) = \{g \in G : \alpha(g) = x\}$$

is a submanifold of G, hence a Lie subgroup, with Lie algebra

$$\mathrm{stab}_{\mathfrak{g}}(x) := \{X \in \mathfrak{g} : d\alpha(X) = 0\}.$$

25.3.2. *Application to kernels.* If $f : G \to H$ is a morphism of Lie groups, then we may regard G as acting on H via $g \cdot x := f(g)x$. The stabilizer of the identity element of H under this action is then the kernel of f, so by the result of the previous section,

$$\ker(f) = \{g \in G : f(g) = e\}$$

is a Lie subgroup of G with Lie algebra

$$\ker(df) : \{X \in \mathfrak{g} : df(X) = 0\}.$$

25.3.3. *Application to preimages.* Given a morphism $f : G \to H$ of Lie groups and a Lie subgroup H_1 of H, the preimage $f^{-1}(H_1)$ may be interpreted as the stabilizer in G of the identity element in the quotient manifold H/H_1 under the action afforded by f. Thus $f^{-1}(H_1)$ is a Lie subgroup of G with Lie algebra $(df)^{-1}(\mathfrak{h}_1) \leq \mathfrak{g}$.

25.3.4. *Application to intersections.* If H_1, H_2 are Lie subgroups of a Lie group G, then $H_1 \cap H_2$ is the preimage of H_2 under the inclusion map $H_1 \hookrightarrow G$, and is thus itself a Lie subgroup.

25.3.5. *Application to stabilizers of vectors or subspaces in representations.* Let G be a Lie group and $R : G \to \mathrm{GL}(V)$ a representation. For $v \in V$, we know by §25.3.1 that $\mathrm{Stab}_G(v)$ is a Lie subgroup with Lie algebra $\mathrm{stab}_G(v)$. For a subspace U of V, we can apply the considerations of §25.3.1 to a suitable Grassmannian manifold (consisting of subspaces of V of given dimension) to see that

$$\{g \in G : R(g)U \subseteq U\} \tag{91}$$

is a Lie subgroup of G with Lie algebra

$$\{X \in \mathfrak{g} : dR(X)U \subseteq U\}.$$

Alternatively, we can note that $\{g \in \mathrm{GL}(V) : gU \subseteq U\}$ is a Lie subgroup of $\mathrm{GL}(V)$ (consisting of block upper-triangular matrices); it follows then from from the discussion of §25.3.3 that its preimage (91) is a Lie subgroup of G with Lie algebra as indicated.

25.4. **Commutator subgroups.** For a general Lie group G, the subgroup $G' := [G, G]$ of commutators need not be closed, hence need not be a Lie subgroup. But if G is *simply-connected*, then G' is indeed a Lie subgroup. To see this, denote by $\mathfrak{g}$ the Lie algebra of G and by $\mathfrak{g}' := [\mathfrak{g}, \mathfrak{g}]$ the subalgebra generated by the commutators. Then $\mathfrak{g}/\mathfrak{g}'$ is an abelian Lie algebra, hence is the Lie algebra of a vector space V. Since G is simply-connected, the natural Lie algebra morphism $df : \mathfrak{g} \to \mathfrak{g}/\mathfrak{g}'$ lifts to a Lie group morphism $f : G \to V$. Then $\ker(f)$ is a Lie subgroup with Lie algebra $\mathfrak{g}'$; moreover, it is clear that $\ker(f) \supseteq G'$. In the opposite direction, we can play around with commutators of paths near the identity in G and the inverse function theorem to see that G' contains a neighborhood of the identity in $\ker(f)$. It follows that G' and $\ker(f)$ coincide near the identity. In particular, G' is a Lie subgroup.

In fact, since $G/\ker(f) \cong V$ is a vector space, it follows from the short exact sequence $\cdots \to \pi_1(G/\ker(f)) \to \pi_0(\ker(f)) \to 0$ that $\pi_0(\ker(f)) = 0$, i.e., that $\ker(f)$ is connected. So we actually have $G' = \ker(f)$.

26. Immersed Lie subgroups

We've described thus far a fair bit of the basic Lie-theoretic dictionary: simply-connected Lie groups correspond to Lie algebras, etc. We've also seen that for a Lie group G, the connected Lie subgroups H of G are determined by their Lie algebras $\mathfrak{h} \leq \mathfrak{g}$. It's natural to ask which $\mathfrak{h}$ arise in this way. The subtlety of the problem can be seen by considering simple examples such as $G = \mathbb{R}^2$ and $G = (\mathbb{R}/\mathbb{Z})^2$, as in lecture.

An easier question is to ask whether arbitrary Lie subalgebras $\mathfrak{h} \leq \mathfrak{g}$ of the Lie algebra $\mathfrak{g}$ of a Lie group G correspond to "something" involving G. The answer is that they are in natural bijection with *connected immersed Lie subgroups* H of G (recall Definition 63, which I've gone back in the notes and modified since we started the course in order to make things work here). By definition, the latter are abstract subgroups H of G with the property that there exists a manifold structure on H with respect to which H is a Lie group and so that the inclusion $H \to G$ is an injective (immersive) morphism of Lie groups.

To put it a bit more verbosely: a subset H of a Lie group G is an immersed Lie subgroup if there exists a Lie group $\hat{H}$ and an injective immersion $\iota : \hat{H} \to G$ with image H. In that case $d\iota : \mathrm{Lie}(\hat{H}) \to \mathfrak{g}$ is an injective morphism of Lie algebras whose image $\mathfrak{h}$ we define to be the Lie algebra of H. This gives one direction of the above correspondence.

The reverse direction is more subtle: given a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, one takes for H the subgroup generated by the image of $\mathfrak{h}$ under the exponential map. One then attempts to define a manifold structure on H to be that generated for small open $0 \in U \subseteq \mathfrak{h}$ and $h \in H$ by the charts $h\exp(U) \ni h\exp(X) \mapsto X \in U$.

Another part of the correspondence is that the Lie group structure on any immersed Lie subgroup $H \leq G$ is uniquely determined by the subset H.

See Chapter 1, Sections 2.4 and 5.3 of the first reference on the course page for more details.

27. Simple Lie groups

Recall that a subset H of an abstract group G is called

- a *subgroup* if $e \in H$ and $h_1, h_2 \in H \implies h_1 h_2 \in H$ and $h \in H \implies h^{-1} \in H$, and is in that case called

- *normal* if $gHg^{-1} \subseteq H$ for all $g \in G$,
- *trivial* if $H = \{1\}$,
- *proper* if $H \neq G$,

and that G is called *simple* if it has no nontrivial proper normal subgroups. For Lie theory, a slightly modified definition turns out to be convenient.

Definition 179. Let $\mathfrak{g}$ be a Lie algebra, thus $\mathfrak{g}$ is a vector space (over $\mathbf{k} = \mathbb{R}$ or $\mathbb{C}$, say) equipped with a bracket operation $[,]$ that is bilinear, antisymmetric, and satisfies the Jacobi identity. A vector subspace $\mathfrak{h}$ of $\mathfrak{g}$ is called

- a *subalgebra* if $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$, and is
- an *ideal* if $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$.

Here $[\mathfrak{g}, \mathfrak{h}]$ denotes the span of the commutators $[x, y]$ with $x \in \mathfrak{g}, y \in \mathfrak{h}$. It is clear that an ideal is a subalgebra.

We denote the relationship that $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ symbolically by $\mathfrak{h} \triangleleft \mathfrak{g}$.

Henceforth denote by G a *connected* Lie group and by $\mathfrak{g}$ its Lie algebra. Recall that there is a natural bijection

$$\{ \text{ subalgebras } \mathfrak{h} \text{ of } \mathfrak{g}\} \cong \{ \text{ connected virtual Lie subgroups } H \text{ of } G \ \}$$

given in the forward direction by $\mathfrak{h} \mapsto H := \langle \exp_G(\mathfrak{h}) \rangle$.

Exercise 30. Let H be a connected virtual Lie subgroup of G with Lie algebra $\mathfrak{h}$. The following are equivalent:

(1) H is a normal subgroup of G.
(2) $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

[Use the standard differentiation/exponentiation technique.]

Definition 180. $\mathfrak{g}$ is *abelian* if $[\mathfrak{g}, \mathfrak{g}] = 0$.

Exercise 31. G is abelian if and only if $\mathfrak{g}$ is abelian. [Use the standard differentiation/exponentiation technique.]

Definition 181. $\mathfrak{g}$ is *simple* if it is non-abelian and has no nontrivial proper ideals.

Definition 182. G is *simple* if it is non-abelian it has no nontrivial proper normal connected virtual Lie subgroups.

Lemma 183. *The following are equivalent:*

(i) G is simple.
(ii) $\mathfrak{g}$ is simple.
(iii) Every proper normal subgroup of G is discrete.

Proof. The equivalence of (i) and (ii) is immediate from Exercise 30. It is clear that (iii) implies (i): if every proper normal subgroup of G is discrete, and if H is a proper normal connected virtual Lie subgroup, then H is in particular a proper normal subgroup, hence is discrete; since H is then discrete and connected, it is trivial, and since H was arbitrary we conclude that G is simple.

The interesting implication is thus that (i) and (ii) imply (iii). To see that, let K be a normal subgroup of G that is not discrete; we wish to show that $K = G$. The closure $\overline{K}$ of K is (by continuity) a closed normal subgroup of G. Since K is not discrete, neither is $\overline{K}$, hence neither is the connected component $\overline{K}^0$. Since $\overline{K}^0$ is a characteristic subgroup of $\overline{K}$, it is a closed normal subgroup of G, hence a non-discrete normal Lie subgroup of G; since G is simple, the only possibility is that $\overline{K}^0 = G$, hence $\overline{K} = G$.

In summary, K is dense in G. We claim that there is $k \in K$ and $X \in \mathfrak{g}$ so that $\mathrm{Ad}(k)X \neq X$. If not, then it would follow by continuity that $\mathrm{Ad}(G)$ is trivial, hence that $\mathfrak{g}$ is abelian, contrary to our assumption that $\mathfrak{g}$ is simple, hence non-abelian.

Consider the curve $\gamma(t) := (k, \exp_G(tX))$, where $(,)$ denotes the commutator. We then have $\gamma(0) = e$, $\gamma(t) \in K$ for all t, and $Y := \gamma'(0) = \mathrm{Ad}(k)X - X \neq 0$. Since $\mathfrak{g}$ is simple, its center $\mathfrak{z}(\mathfrak{g}) = \{Z \in \mathfrak{g} : [Z, \mathfrak{g}] = 0\}$ is trivial (for else its center would be a nontrivial ideal, hence $\mathfrak{g}$ would coincide with its center, i.e., $\mathfrak{g}$ would be abelian; contradiction). In particular, $[Y, \mathfrak{g}] \neq 0$. It follows by the standard differentiate/exponentiate technique that the subspace $\mathfrak{a}$ of $\mathfrak{g}$ spanned by $\mathrm{Ad}(G)Y$ is a nonzero ideal (check this). Since $\mathfrak{g}$ is simple, it follows that $\mathfrak{a} = \mathfrak{g}$. We can thus find $g_1, \ldots, g_n \in G$, where $n = \dim(G)$, so that the elements $\mathrm{Ad}(g_1)\gamma'(0), \ldots, \mathrm{Ad}(g_n)\gamma'(0)$ span $\mathfrak{g}$. Also, the curves $t \mapsto \mathrm{Ad}(g_j)\exp(tX)$ all lie in K, since $\exp(tX) \in K$ and K is normal. The map

$$(t_1, \ldots, t_n) \mapsto (\mathrm{Ad}(g_1)\exp(t_1 X)) \cdots (\mathrm{Ad}(g_n)\exp(t_n X))$$

then has differntial at $(0, \ldots, 0)$ given by an invertible linear map, hence (by the inverse function theorem) defines a local diffeomorphism; since its image lies in K, we deduce that K contains a neighborhood of the identity in G, and since G is connected, it follows that $K = G$, as required. □

Thus apart from excluding abelian examples and possible discrete normal subgroups, the notions of a connected Lie group being simple as an abstract group or simple as a Lie group are the same.

In the rest of the lecture, we described which classical Lie groups/algebras are simple and what the isomorphisms between them are. This will be discussed in subsequent lectures.

28. Simplicity of the special linear group

28.1. **Some linear algebra.** Let V be a complex vector space (not necessarily finite-dimensional, for now). Given an operator $x \in \mathrm{End}(V)$ and $\lambda \in \mathbb{C}$, we may define the *eigenspace*

$$V^\lambda := \{v \in V : xv = \lambda v\}.$$

Lemma 184. *The spaces V^λ are linearly independent, that is to say, if $n \geq 1$ and $\lambda_1, \ldots, \lambda_n$ are distinct complex numbers and $v_1 \in V^{\lambda_1}, \ldots, v_n \in V^{\lambda_n}$ satisfy $v_1 + \cdots + v_n = 0$, then $v_1 = \cdots = v_n = 0$.*

Proof. We induct on n. When $n = 1$, the required conclusion is clear: if $v_1 \in V^{\lambda_1}$ satisfies $v_1 = 0$, then certainly $v_1 = 0$. Suppose now that $n \geq 2$, and let $v_1, \in V^{\lambda_1}, \ldots, v_n \in V^{\lambda_n}$ with $v_1 + \cdots + v_n = 0$. Then certainly

$$0 = \lambda_n(v_1 + \cdots + v_n)$$

and also

$$0 = x(v_1 + \cdots + v_n) = \lambda_1 v_1 + \cdots + \lambda_n v_n,$$

hence upon taking differences,

$$(\lambda_1 - \lambda_n)v_1 + (\lambda_2 - \lambda_n)v_2 + \cdots + (\lambda_{n-1} - \lambda_n)v_{n-1} = 0.$$

By our inductive hypothesis, $(\lambda_j - \lambda_n)v_j = 0$ for $j = 1, \ldots, n-1$. Since $\lambda_j \neq \lambda_n$, it follows that $v_1 = \cdots = v_{n-1} = 0$ and hence also that $v_n = 0$, as required. $\square$

Thus the sum $\sum_{\lambda \in \mathbb{C}} V^\lambda$ is in fact a direct sum $\oplus V^\lambda$.

Definition 185. An operator $x \in \mathrm{End}(V)$ is *semisimple* (or *diagonalizable*, or *completely reducible*; depending upon my mood I alternate between the various terminologies) if

$$V = \oplus_{\lambda \in \mathbb{C}} V^\lambda.$$

Assume henceforth that V is *finite-dimensional*.

Definition 186. For $x \in \mathrm{End}(V)$, we say that a subspace $W \leq V$ is x-invariant if $xW \subseteq W$.

Exercise 32. The following are equivalent:

(1) x is semisimple.
(2) There is a basis of V with respect to which x is represented by a diagonal matrix.
(3) In the Jordan decomposition of x as a sum of Jordan blocks Jordan blocks

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \cdots & \cdots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix},$$

only 1×1 blocks appear.
(4) The characteristic polynomial and minimal polynomial of x are the same.
(5) Every x-invariant subspace $W \leq V$ has an x-invariant complement W', i.e., a subspace $W' \leq V$ for which $V = W \oplus W'$.
(6) The minimal polynomial of x is squarefree, i.e., of the form $(X - a_1) \cdots (X - a_r)$ for some *distinct* complex numbers $a_1, \ldots, a_r$.
(7) There exists a squarefree polynomial that annihilates x.

(1) iff (2): take as the basis for V a union of arbitrary bases of V^λ. (2) iff (3): clear. (3) iff (4): clear. (1) iff (5): argue as in the proof of Lemma 110. It is clear that (4), (6) and (7) are equivalent.

Exercise 33. Suppose $x \in \mathrm{End}(V)$ is semisimple.

(1) For each eigenvalue λ of x, let W^λ be a subspace of the λ-eigenspace V^λ of x. Verify that $\oplus_\lambda W^\lambda$ is an x-invariant subspace of V.
(2) Let $W \leq V$ be an x-invariant subspace. Show that W is of the form $\oplus_\lambda W^\lambda$, where $W^\lambda = W \cap V^\lambda$ is a subspace of V^λ. [Hint: one can either appeal to the previous lemma or argue as in the proof of Lemma 184; other arguments are probably also possible.]
(3) Let $W \leq V$ be an x-invariant subspace.
(4) Show that $x|_W$ is semisimple.

Theorem 187. *Let $x_1, \ldots, x_n \in \mathrm{End}(V)$ be semisimple commuting elements. Then there is a basis of V with respect to which all of $x_1, \ldots, x_n$ are diagonalizable.*

Proof. We induct on n. For $n = 1$, everything is clear, so suppose $n \geq 2$. Let $V = \oplus V^\lambda$ be the decomposition of V into eigenspaces for x_1. For $j \in \{2..n\}$, we claim that $x_j V^\lambda \subseteq V^\lambda$. Indeed, for $v \in V^\lambda$, we have $x_1 x_j v = x_j x_1 v = x_j \lambda v = \lambda x_j v$, hence $x_j v \in V^\lambda$, as required. By Exercise 33, the restrictions to V^λ of the operators $x_2, \ldots, x_n$ are semisimple (and commuting), hence by our inductive hypothesis may be simultaneously diagonalized; we now conclude by taking a union over λ of bases of V^λ with respect to which the $x_1, \ldots, x_n$ are diagonalized. □

28.2. **Recap on $\mathfrak{sl}_2(\mathbb{C})$.** Recall that $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ has the basis elements

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

which satisfy the relations

$$[X, Y] = H, \quad \mathrm{ad}(H)X = 2X, \quad \mathrm{ad}(H)Y = -2Y.$$

We can view

$$\mathfrak{h} = \mathbb{C}X \oplus \mathbb{C}H \oplus \mathbb{C}Y$$

as the eigenspace decomposition of $\mathrm{ad}(H)$ with eigenvalues $2, 0, -2$.

Theorem 188. *Let $\rho : \mathfrak{g} \to \mathrm{End}(V)$ be a finite-dimensional representation. Then $\rho(H)$ is semisimple.*

Proof. Since V decomposes as a direct sum of the irreducible representations W_m, we may assume $V = W_m$. Then the elements $x^m, x^{m-1}y, \ldots, y^m$ give a basis of V with respect to which $\rho(H)$ is diagonal. □

We record for future reference that the element

$$w := e^{-X} e^{Y} e^{-X} = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C}) \tag{92}$$

has the property that

$$Ad(w)H = -H. \tag{93}$$

28.3. **Reformulation in terms of representations of abelian Lie algebras.** In this section all Lie algebras are finite-dimensional over the complex numbers and all vector spaces are complex and finite-dimensional. Recall that a Lie algebra $\mathfrak{h}$ is *abelian* if $[\mathfrak{h}, \mathfrak{h}] = 0$.

Let $\mathfrak{h}$ be an abelian Lie algebra and $\rho : \mathfrak{h} \to \mathrm{End}(V)$ a representation. Let $\mathfrak{h}^* := \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$ denote the dual vector space. For $\lambda \in \mathfrak{h}^*$, define the *eigenspace*

$$V^\lambda := \{v \in V : Hv = \lambda v \text{ for all } H \in \mathfrak{h}\}$$

where we abbreviate $Hv := \rho(H)v$.

Lemma 189. *The spaces V^λ are linearly independent, i.e., $\sum_{\lambda \in \mathbb{C}} V^\lambda = \oplus_{\lambda \in \mathbb{C}} V^\lambda$.*

Proof. Same proof as Lemma 184. □

Lemma 190. *Let $\mathfrak{h}$ be an abelian Lie algebra and $\rho : \mathfrak{h} \to \mathrm{End}(V)$ a representation. The following are equivalent:*

1. *$\rho(H)$ is semisimple for all $H \in \mathfrak{h}$.*
2. *There is a basis of V with respect to which every element of $\rho(\mathfrak{h})$ is diagonal.*
3. *$V = \oplus V^\lambda$.*

Proof. Take a basis $H_1, \dots, H_n$ and consider $x_1 := \rho(H_1), \dots, x_n := \rho(H_n) \in \mathrm{End}(V)$. If each x_j is semisimple, then Theorem 187 tells us that there is a basis with respect to which they are (and hence every element of $\rho(\mathfrak{h})$ is) diagonal. The converse and the equivalence with $V = \oplus V^\lambda$ are left to the reader. □

Definition 191. We say that a representation $\rho : \mathfrak{h} \to \mathrm{End}(V)$ of an abelian Lie algebra $\mathfrak{h}$ is *semisimple* if teh equivalent conditions of the previous lemma are satisfied.

Definition 192. Let $\rho : \mathfrak{h} \to \mathrm{End}(V)$ be a representation of an abelian Lie algebra $\mathfrak{h}$. A *weight* of ρ is an element $\lambda \in \mathfrak{h}^*$ for which $V^\lambda \neq 0$. The space V^λ is then called the *weight space of weight* λ. (The definition is most interesting when ρ is semisimple, so that $V = \oplus V^\lambda$.)

Example 193. Let $V = \mathbb{C}^n$, let $\mathfrak{h}$ be the space of diagonal matrices in $M_n(\mathbb{C}) = \mathrm{End}(V)$, and let $\rho : \mathfrak{h} \to \mathrm{End}(V)$ be the identity map. Then the weights of ρ are the functionals $\lambda_1, \dots, \lambda_n : \mathfrak{h} \to \mathbb{C}$ giving diagonal coordinates on $\mathfrak{h}$, i.e.,

$$H = \begin{pmatrix} \lambda_1(H) & & \\ & \cdots & \\ & & \lambda_n(H) \end{pmatrix}.$$

The corresponding weight spaces V^{λ_j} are the one-dimensional subspaces $\mathbb{C}e_j$ spanned by the standard basis elements $e_1, \dots, e_n$ of $\mathbb{C}^n$.

28.4. **Roots of an abelian subalgebra.** Let $\mathfrak{g}$ be a Lie algebra (always finite-dimensional and over the complex numbers, for the present purposes) and $\mathfrak{h} \leq \mathfrak{g}$ an abelian subalgebra.

Recall that adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ given by $\mathrm{ad}(X)Y := [X, Y]$. We restrict this to $\mathfrak{h}$ to obtain a representation

$$\mathrm{ad} : \mathfrak{h} \to \mathrm{End}(\mathfrak{g}).$$

The set R of *roots* of the pair $(\mathfrak{g}, \mathfrak{h})$ is defined to be the set of nonzero weights of this representation. (We might more verbosely write $R = R(\mathfrak{g}, \mathfrak{h})$ when we wish to make explicit the dependence.) In other words, for $\lambda \in \mathfrak{h}^*$, set

$$\mathfrak{g}^\lambda := \{X \in \mathfrak{g} : [H, X] = \lambda(H)X \text{ for all } H \in \mathfrak{h}\}.$$

Then

$$R = \{\alpha \in \mathfrak{h}^* - \{0\} : \mathfrak{g}^\alpha \neq 0\}. \tag{94}$$

For $\alpha \in R$, we call $\mathfrak{g}^\alpha$ the corresponding *root space.*

Exercise 34. Write $0 \in \mathfrak{h}^*$. Then

$$\mathfrak{h} \subseteq \mathfrak{g}^0.$$

Here is a very useful general observation:

Lemma 194. *For any $\alpha, \beta \in \mathfrak{h}^*$, one has*

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

Proof. **It is probably easier to work this out as an exercise than to read the proof.** It suffices to show for all $X \in \mathfrak{g}^\alpha$ and $Y \in \mathfrak{g}^\beta$ that $[X, Y] \in \mathfrak{g}^{\alpha+\beta}$, i.e., that for all $H \in \mathfrak{h}$, one has $\operatorname{ad}(H)[X,Y] = (\alpha+\beta)(H)[X,Y]$. This follows from the Jacobi identity in the form

$$\operatorname{ad}(H)[X,Y] = [\operatorname{ad}(H)X, Y] + [X, \operatorname{ad}(H)Y], \tag{95}$$

like so:

$$\operatorname{ad}(H)[X,Y] = [\alpha(H)X, Y] = [X, \beta(H)Y] = (\alpha(H)+\beta(H))[X,Y] = (\alpha+\beta)(H)[X,Y]. \tag{96}$$

□

More generally:

Lemma 195. *Let* $\rho : \mathfrak{g} \to \operatorname{End}(V)$ *be any linear representation. For* $\lambda \in \mathfrak{h}^*$*, set* $V^\lambda := \{v \in V : \rho(H)v = \lambda(H)v \text{ for all } v \in V\}$*. Let* $\alpha, \lambda \in \mathfrak{h}^*$*. Then*

$$\rho(\mathfrak{g}^\alpha)V^\lambda \subseteq V^{\alpha+\lambda}.$$

Proof. This is essentially the same proof as the previous lemma, but with the adjoint representation replaced by ρ; it should also remind the reader of stuff we did awhile ago involving raising/lowering operators acting on representations of $\mathfrak{sl}_2(\mathbb{C})$:

It is probably easier to work this out as an exercise than to read the proof. Let $X \in \mathfrak{g}^\alpha$ and $v \in V^\lambda$. To establish the required membership $\rho(X)v \in V^{\alpha+\lambda}$, we must verify for each $H \in \mathfrak{h}$ that $\rho(H)\rho(X)v = (\alpha+\lambda)(H)\rho(X)v$. Indeed,

$$\rho(H)\rho(X) = \rho([H,X]) + \rho(X)\rho(H).$$

Since $[H,X] = \alpha(H)X$ and ρ is linear, we have

$$\rho([H,X])v = \rho(\alpha(H)X)v = \alpha(H)\rho(X)v.$$

Since $\rho(H)v = \lambda(H)v$ by assumption and $\rho(X)$ is linear, we have

$$\rho(X)\rho(H)v = \rho(X)\lambda(H)v = \lambda(H)\rho(X)v.$$

Summing up, we obtain

$$\rho(H)\rho(X) = \alpha(H)\rho(X)v + \lambda(H)\rho(X)v,$$

which simplifies to give the required identity. □

In other words, weight spaces V^λ are permuted by root spaces $\mathfrak{g}^\alpha$ and root spaces $\mathfrak{g}^\alpha, \mathfrak{g}^\beta$ interact nicely with respect to the commutator. Note that Lemma 194 is the specialization of 195 to the adjoint representation.

28.5. **The roots of the special linear Lie algebra.** Let $\mathfrak{g} := \mathfrak{sl}_n(\mathbb{C})$. Let $\mathfrak{h} \leq \mathfrak{g}$ be the subspace of diagonal matrices. One has $\dim(\mathfrak{h}) = n - 1$. Every element $H \in \mathfrak{h}$ is of the form

$$H = \begin{pmatrix} \lambda_1(H) & & \\ & \cdots & \\ & & \lambda_n(H) \end{pmatrix} \tag{97}$$

with $\lambda_1(H) + \cdots + \lambda_n(H) = 0$. The functionals $\lambda_1, \ldots, \lambda_n \in \mathfrak{h}^*$ defined by (97) span $\mathfrak{h}^*$ and satisfy the relation $\lambda_1 + \cdots + \lambda_n = 0$.

For $j, k \in \{1..n\}$, let $E_{jk} \in M_n(\mathbb{C})$ denote the elementary matrix with entries $(E_{jk})_{mn} := \delta_{jm}\delta_{kn}$. If $j \neq k$, then $E_{jk} \in \mathfrak{g}$. One has in general $E_{jj} - E_{kk} \in \mathfrak{g}$, but $E_{jj} \neq \mathfrak{g}$ (by the trace condition). In general,

$$[E_{ij}, E_{kl}] = \delta_{jk}E_{il} - \delta_{il}E_{kj}. \tag{98}$$

For $H \in \mathfrak{h}$ and any j, k, we compute directly that $HE_{jk} = \lambda_j(H)E_{jk}$ and $E_{jk}H = \lambda_k(H)E_{jk}$, hence

$$\mathrm{ad}(H)E_{jk} = (\lambda_j(H) - \lambda_k(H))E_{jk}. \tag{99}$$

Thus for all $j \neq k$, we see that $\lambda_j - \lambda_k$ is a root. Observe also that

$$\mathfrak{g}^0 = \mathfrak{h}; \tag{100}$$

said another way, $\mathfrak{h}$ is a *maximal* abelian subalgebra of $\mathfrak{g}$. Since

$$\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{j \neq k} \mathbb{C}E_{jk}) \tag{101}$$

we see that

$$R = \{\lambda_j - \lambda_k : j \neq k\}. \tag{102}$$

We may rewrite (101) in either of the forms

$$\mathfrak{g} = \oplus_{\lambda \in \mathbb{C}} \mathfrak{g}^\lambda \tag{103}$$

or

$$\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{\alpha \in R} \mathfrak{g}^\alpha). \tag{104}$$

In particular:

Lemma 196. $\mathrm{ad} : \mathfrak{h} \to \mathrm{End}(\mathfrak{g})$ *is semisimple.*

Lemma 197. *The set R of roots spans $\mathfrak{h}^*$.*

Proof. We must verify that if $H \in \mathfrak{h}$ satisfies $\alpha(H) = 0$ for all $\alpha \in R$, then $H = 0$. Indeed, given such an H, we have (by taking $\alpha = \lambda_j - \lambda_k$) that $\lambda_j(H) = \lambda_k(H)$ for all $i \neq j$, hence all the entries of H are the same; since H has trace zero, it follows as required that $H = 0$. □

For $\alpha = \lambda_j - \lambda_k \in R$, the corresponding root space is

$$\mathfrak{g}^\alpha = \mathbb{C}X_\alpha \text{ where } X_\alpha := E_{jk}.$$

Note that

$$\alpha \in R \implies -\alpha \in R \tag{105}$$

since indeed $-\alpha = \lambda_k - \lambda_j$ for $\alpha = \lambda_j - \lambda_k$, but that

$$n\alpha \notin R \text{ for any } n \in \mathbb{Z} - \{\pm 1\}.$$

Define $Y_\alpha := E_{kj} \in \mathfrak{g}^{-\alpha}$; by (98), the element

$$H_\alpha := [X_\alpha, Y_\alpha]$$

is given explicitly by $H_\alpha = E_{jj} - E_{kk} \in \mathfrak{h}$ and satisfies

$$\alpha(H_\alpha) = 2$$

because $\alpha(H_\alpha) = (\lambda_j - \lambda_k)(E_{jj} - E_{kk}) = 1 - (-1) = 2$. (One has here that $Y_\alpha = X_{-\alpha}$, but what matters most is that they both span the same one-dimensional space.) One has

$$H_{-\alpha} = -H_\alpha \text{ for all } \alpha \in R.$$

It is easy to see that

$$\sum_{\alpha \in R} \mathbb{C}H_\alpha = \mathfrak{h}, \tag{106}$$

but that the sum is not direct for $n > 2$.

Lemma 198. *Each $\mathfrak{g}^\alpha$ is one-dimensional. For all $\alpha, \beta \in R$,*

$$[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \begin{cases} \mathfrak{g}^{\alpha+\beta} & \text{if } \alpha+\beta \in R \\ \mathbb{C}H_\alpha & \text{if } \alpha+\beta = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Proof. This reduces to an explicit computation using (98). It is instructive to note the consistency with the conclusion $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] \subseteq \mathfrak{g}^{\alpha+\beta}$ of Lemma 194 and the identity (100). $\square$

Here are some basic consequences:

Lemma 199.

(1) *If $\alpha, \beta, \alpha+\beta$ are all roots, then the map*

$$\mathrm{ad}(X_\beta) : \mathfrak{g}^\alpha \to \mathfrak{g}^{\alpha+\beta}$$

is an isomorphism (of one-dimensional vector spaces).

(2) *If $\alpha, \beta, \alpha-\beta$ are all roots, then the map*

$$\mathrm{ad}(Y_\beta) : \mathfrak{g}^\alpha \to \mathfrak{g}^{\alpha-\beta}$$

is an isomorphism (of one-dimensional vector spaces).

(3) *For any roots α, β, the composition*

$$\mathrm{ad}(X_\beta) \circ \mathrm{ad}(X_\alpha) : \mathfrak{g}^{-\alpha} \to \mathfrak{g}^\beta$$

is an isomorphism (of one-dimensional vector spaces) if and only if $\beta(H_\alpha) \neq 0$.

(4) *For any roots α, β, the composition*

$$\mathrm{ad}(Y_\beta) \circ \mathrm{ad}(Y_\alpha) : \mathfrak{g}^\alpha \to \mathfrak{g}^{-\beta}$$

is an isomorphism (of one-dimensional vector spaces) if and only if $\beta(H_\alpha) \neq 0$.

Proof. The first two assertions are immediate from Lemma 198. For the third assertion, we factor the map of interest as

$$\mathfrak{g}^{-\alpha} \xrightarrow{\mathrm{ad}(X_\alpha)} \mathbb{C}H_\alpha \xrightarrow{\mathrm{ad}(X_\beta)} \mathfrak{g}^\beta$$

and observe (using Lemma 198) that $\mathrm{ad}(X_\alpha) : \mathfrak{g}^{-\alpha} \to \mathbb{C}H_\alpha$ is always an isomorphism of one-dimensional vector spaces and noting that $\mathrm{ad}(X_\beta)H_\alpha = -[H_\alpha, X_\beta] = -\beta(H_\alpha)X_\beta$, which vanishes if and only if $\beta(H_\alpha) = 0$.

The proof of the fourth assertion is similar: we factor the map $\mathfrak{g}^\alpha \to \mathfrak{g}^{-\beta}$ in question as

$$\mathfrak{g}^{\alpha} \xrightarrow{\mathrm{ad}(Y_\alpha)} \mathbb{C}H_\alpha \xrightarrow{\mathrm{ad}(Y_\beta)} \mathfrak{g}^{-\beta},$$

observe that the first map in this composition is always an automorphism, and then observe that the second map sends H_α to $[Y_\beta, H_\alpha] = -[H_\alpha, Y_\beta] = \beta(H_\alpha)Y_\beta$. $\square$

One can check directly using (102) that for all $\alpha, \beta \in R$,

$$\beta(H_\alpha) = 0 \iff \alpha(H_\beta) = 0. \tag{107}$$

Definition 200. Two roots α, β are called *orthogonal* if either of the equivalent conditions (107) hold. They are called *non-orthogonal* if $\beta(H_\alpha) \neq 0$ (equivalently, $\alpha(H_\beta) \neq 0$).

Explicitly, if $\alpha = \lambda_i - \lambda_j$ and $\beta = \lambda_k - \lambda_l$, then α, β are orthogonal if and only if $\{i, j\} \cap \{k, l\} = \emptyset$.

Lemma 201. *Suppose given roots $\beta, \alpha_1, \dots, \alpha_r \in R$ with the property that $\beta + \alpha_1 + \dots + \alpha_r$ is a root and that for each $s = 1, 2, \dots, r$, the partial sum $\beta + \alpha_1 + \dots + \alpha_s$ is either a root, or is zero, and if it is zero that the roots α_s, α_{s+1} are non-orthogonal. Then the compositions*

$$\mathfrak{g}^{\beta} \xrightarrow{X_{\alpha_1}} \mathfrak{g}^{\beta+\alpha_1} \xrightarrow{X_{\alpha_2}} \cdots \xrightarrow{X_{\alpha_r}} \mathfrak{g}^{\beta+\alpha_r}$$

$$\mathfrak{g}^{\beta+\alpha_r} \xrightarrow{Y_{\alpha_r}} \cdots \xrightarrow{Y_{\alpha_2}} \mathfrak{g}^{\beta+\alpha_1} \xrightarrow{Y_{\alpha_1}} \mathfrak{g}^{\beta}$$

(where we abbreviate $\xrightarrow{X}$ for $\xrightarrow{\mathrm{ad}(X)}$, etc) are isomorphisms of one-dimensional vector spaces.

Proof. Follows immediately by repeated application of Lemma 199. □

28.6. **The simplicity of $\mathfrak{sl}_n(\mathbb{C})$.** Set $\mathfrak{g} := \mathfrak{sl}_n(\mathbb{C})$. Let $\mathfrak{h}$ denote its diagonal subalgebra, and R the set of roots for $(\mathfrak{g}, \mathfrak{h})$.

Lemma 202. *Set*

$$\lambda_{\max} := \lambda_1 - \lambda_n \in R.$$

Let $\beta \in R$. Then there is a nonnegative integer $r \geq 0$ and roots $\alpha_1, \dots, \alpha_r$ so that

$$\beta + \alpha_1 + \cdots + \alpha_r = \lambda_{\max}$$

and so that for $s \in \{1..r-1\}$, the partial sum $\beta + \alpha_1 + \dots + \alpha_s$ is either a root, or zero, and if it is zero, then the roots

$$\alpha_{s+1}(H_{\alpha_s}) \neq 0, \quad \alpha_s(H_{\alpha_{s+1}}) \neq 0, \tag{108}$$

i.e., the roots α_s, α_{s+1} are non-orthogonal in the sense of Definition 200. (We can always take $r \leq 2$, but that doesn't matter so much.)

Proof. Write $\beta = \lambda_j - \lambda_k$. The proof is basically to stare at the diagrams

$$\lambda_j - \lambda_k \xrightarrow{\lambda_k - \lambda_n} \lambda_j - \lambda_n \xrightarrow{\lambda_1 - \lambda_j} \lambda_1 - \lambda_n$$

and

$$\lambda_j - \lambda_k \xrightarrow{\lambda_1 - \lambda_j} \lambda_1 - \lambda_k \xrightarrow{\lambda_k - \lambda_n} \lambda_1 - \lambda_n.$$

We omit $\xrightarrow{\mu}$ from the above diagrams if $\mu = 0$. What needs to be checked is that for each j, k, at least one of the above diagrams has the property that whenever it looks like

$$\lambda_j - \lambda_k \xrightarrow{\mu} 0 \xrightarrow{\nu}, \lambda_1 - \lambda_n,$$

the roots μ and ν are non-orthogonal, i.e., satisfy $\mu(H_\nu) \neq 0$.

We turn to the details:

- If $k = n$ and $j = 1$, we take $r = 0$.
- If $k = n$ and $j > 1$, we take $r := 1$ and $\alpha_1 := \lambda_1 - \lambda_j \in R$; then $\beta + \alpha_1 = \lambda_{\max}$.
- If $k < n$ and $j = 1$, we take $r := 1$ and $\alpha_1 := \lambda_k - \lambda_n \in R$; then $\beta + \alpha_1 = \lambda_{\max}$.
- If $k < n$ and $1 < j < n$, we take $r := 2$ and $\alpha_1 := \lambda_k - \lambda_n \in R$ and $\alpha_2 := \lambda_1 - \lambda_j \in R$. Then $\beta + \alpha_1 + \alpha_2 = \lambda_{\max}$. The partial sum $\beta + \alpha_1 = \lambda_j - \lambda_n$ is always a root, since $j < n$.

- If $k < n$ and $j = n$, we take $r := 2$ and $\alpha_1 := \lambda_1 - \lambda_j \in R$ and $\alpha_2 := \lambda_k - \lambda_n \in R$. Then $\beta + \alpha_1 + \alpha_2 = \lambda_{\max}$. The partial sum $\beta + \alpha_1 = \lambda_1 - \lambda_k$ is either a root or zero; it is zero iff $k = 1$, in which case
$$\alpha_2(H_{\alpha_1}) = (\lambda_k - \lambda_n)(E_{11} - E_{jj}) = 2 \neq 0,$$
$$\alpha_1(H_{\alpha_2}) = (\lambda_1 - \lambda_j)(E_{kk} - E_{nn}) = 2 \neq 0,$$
so the roots α_1, α_2 are non-orthogonal in that case, as required.

□

Theorem 203. $\mathfrak{g}$ *is simple.*

Proof. In lecture we gave a more "brute force" proof; here we will clean it up a bit by appeal to Lemmas 199 and 202.

Let $\mathfrak{a} \leq \mathfrak{g}$ be a nonzero ideal, or equivalently, an $\mathrm{ad}(\mathfrak{g})$-invariant subspace. We must show that $\mathfrak{a} = \mathfrak{g}$.

Suppose first that $\mathfrak{a} \subseteq \mathfrak{h}$, and let $0 \neq H \in \mathfrak{a}$ be given. Since the set R of roots spans $\mathfrak{h}^*$ (Lemma 197), we can find $\alpha \in R$ so that $\alpha(H) \neq 0$. But then $X_\alpha \in \mathfrak{g}^\alpha$ (see §28.5) has the property that $[H, X_\alpha] = \alpha(H) X_\alpha$ is a nonzero element of $\mathfrak{g}^\alpha$. On the other hand, we have $[H, X_\alpha] \in \mathfrak{a} \subseteq \mathfrak{h}$ because $\mathfrak{a}$ is an ideal. Since $\mathfrak{h} \cap \mathfrak{g}^\alpha = 0$, we obtain the required contradiction.

Suppose next that $\mathfrak{a}$ is not contained in $\mathfrak{h}$. By the semisimplicity of $\mathrm{ad} : \mathfrak{h} \to \mathrm{End}(\mathfrak{g})$ (see Lemma 196) and the fact that $\mathrm{ad}(\mathfrak{h})$-semisimplicity is preserved upon passage to $\mathrm{ad}(\mathfrak{h})$-invariant subspaces (see Exercise 33), we have
$$\mathfrak{a} = (\mathfrak{a} \cap \mathfrak{h}) \oplus (\oplus_{\alpha \in R} \mathfrak{a} \cap \mathfrak{g}^\alpha). \tag{109}$$
Since $\mathfrak{a}$ is not contained in $\mathfrak{h}$, there is some $\beta \in R$ for which have $\mathfrak{a}$ intersects and hence contains the one-dimensional space $\mathfrak{g}^\beta$.

We claim now with notation as in Lemma 202 that $\mathfrak{g}^{\lambda_{\max}} \subseteq \mathfrak{a}$. To see this, let us write
$$\beta + \alpha_1 + \cdots + \alpha_r = \lambda_{\max} \tag{110}$$
as in Lemma 202. We know that $\mathfrak{a}$ is an ideal, i.e., is $\mathrm{ad}(\mathfrak{g})$-invariant, and that $\mathfrak{a}$ contains $\mathfrak{g}^\beta$, so it will suffice to show that the composition
$$\mathfrak{g}^\beta \xrightarrow{X_{\alpha_1}} \mathfrak{g}^{\beta+\alpha_1} \xrightarrow{X_{\alpha_2}} \cdots \xrightarrow{X_{\alpha_{r-1}}} \mathfrak{g}^{\beta+\alpha_1+\cdots+\alpha_{r-1}} \xrightarrow{X_{\alpha_r}} \mathfrak{g}^{\lambda_{\max}} \tag{111}$$
is an isomorphism (because then $\mathfrak{a} \supseteq \mathfrak{g}^{\lambda_{\max}}$), which follows from Lemma 201.

Now let $\beta \in R$ be arbitrary; we will show that $\mathfrak{a} \supseteq \mathfrak{g}^\beta$. We again write (110) as in Lemma 202. Since we already know that $\mathfrak{a} \supseteq \mathfrak{g}^{\lambda_{\max}}$ and since $\mathfrak{a}$ is $\mathrm{ad}(\mathfrak{g})$-stable, it will suffice to show that the composition
$$\mathfrak{g}^{\lambda_{\max}} \xrightarrow{Y_{\alpha_r}} \mathfrak{g}^{\beta+\alpha_1+\cdots+\alpha_{r-1}} \xrightarrow{Y_{\alpha_{r-1}}} \cdots \xrightarrow{Y_{\alpha_2}} \mathfrak{g}^{\beta+\alpha_1} \xrightarrow{Y_{\alpha_1}} \mathfrak{g}^\beta \tag{112}$$
is an isomorphism, which follows again from (108) and Lemma 199.

In summary, we have seen that $\mathfrak{a}$ contains $\mathfrak{g}^\alpha$ and hence X_α for all $\alpha \in R$; since $\mathfrak{a}$ is an ideal, it contains also $[X_\alpha, Y_\alpha] = H_\alpha$. Since the H_α span $\mathfrak{h}$ (see (106) and surrounding), we have $\mathfrak{a} \supseteq \mathfrak{h}$. Thus $\mathfrak{a} \supseteq \mathfrak{h} \oplus (\oplus_{\alpha \in R} \mathfrak{g}^\alpha)$, i.e., $\mathfrak{a} = \mathfrak{g}$, as required. □

28.7. **The proof given in lecture.** Here are some notes indicating the more "brute force" approach to Theorem 203 presented in lecture. It may be instructive to compare this approach to that above (they differ primarily in notation).

With notation as in the proof of Theorem 187, set

$$\Gamma := \{\alpha \in R : \mathfrak{a} \cap \mathfrak{g}^\alpha \neq 0\}.$$

Then $\mathfrak{a} \supseteq \oplus_{\alpha \in \Gamma} \mathfrak{g}^\alpha$. The key step in the proof was to show that if Γ is nonempty, then

$$\Gamma = R. \tag{113}$$

To see this, let $\beta \in \Gamma$ be given. Suppose that $\beta = \lambda_i - \lambda_j$ with $i > j$; a similar but slightly simpler argument applies if instead $i < j$. We have

$$[E_{ij}, E_{jn}] = \begin{cases} E_{in} & \text{if } i \neq n, \\ E_{in} - E_{jj} & \text{if } i = n, \end{cases} \tag{114}$$

and in either case it follows that

$$[E_{1i}, [E_{ij}, E_{jn}]] = E_{1n}. \tag{115}$$

Since $\mathfrak{a}$ is an ideal, we deduce that $E_{1n} \in \mathfrak{a}$, i.e.,

$$\lambda_1 - \lambda_n \in \Gamma. \tag{116}$$

Now let $\alpha \in R$ be arbitrary, say $\alpha = \lambda_j - \lambda_k$. If $j < k$, then we have

$$[E_{1n}, E_{nk}] = E_{1k},$$

hence $\lambda_1 - \lambda_k \in \Gamma$, and

$$[E_{j1}, E_{1k}] = E_{jk}$$

hence $\lambda_j - \lambda_k \in \Gamma$, as required. Suppose instead that $j > k$. If $k = 1$, then we have

$$[E_{n1}, E_{1n}] = E_{nn} - E_{11}$$

and hence

$$[E_{n1}, [E_{n1}, E_{1n}]] = -2E_{n1},$$

hence $\lambda_n - \lambda_1 \in \Gamma$, as required. In the remaining case that $j > k > 1$, we have

$$[E_{1n}, E_{nk}] = E_{1k}$$

and hence

$$[E_{j1}, [E_{1n}, E_{nk}]] = E_{jk},$$

hence $\lambda_j - \lambda_k \in \Gamma$, as required.

29. Classification of the classical simple complex Lie algebras

29.1. **Recap.** We've seen in lecture that the Lie algebra $\mathfrak{sl}_n(\mathbb{C})$ is simple ($n \geq 2$), and on the homework that $\mathfrak{sp}_{2n}(\mathbb{C})$ is simple ($n \geq 1$). Similar arguments imply that $\mathfrak{so}_{2n+1}(\mathbb{C})$ is simple for $n \geq 1$ and that $\mathfrak{so}_{2n}(\mathbb{C})$ is simple for $n \geq 3$; the handout (§29.4) from lecture (available now also on the course homepage) describes the root systems, and I'll leave it as an exercise to adapt the techniques used to prove the simplicity of $\mathfrak{sl}_n(\mathbb{C})$ and $\mathfrak{sp}_{2n}(\mathbb{C})$ to the orthogonal case. There is one trick in that case which is very handy. Recall from §18.5 the notion of a *quadratic space* over $\mathbb{C}$, and that any n-dimensional quadratic space is isomorphic to the standard one. It is convenient to equip $\mathbb{C}^{2n}$ with the structure of a quadratic space for which the associated non-degenerate symmetric bilinear form $\langle , \rangle$ (denoted B in §18.5) is given by

$$\langle x, y \rangle = \sum_{i=1}^{n} (x_i y_{n+i} + x_{n+i} y_i) \tag{117}$$

and $\mathbb{C}^{2n+1}$ that for which

$$\langle x, y \rangle = \sum_{i=1}^{n} (x_i y_{n+i} + x_{n+i} y_i) + x_{2n+1} y_{2n+1}. \tag{118}$$

It is then easy to see that $\mathrm{SO}_m(\mathbb{C}) := \{g \in \mathrm{SL}_m(\mathbb{C}) : \langle gx, gy \rangle \text{ for all } x, y \in \mathbb{C}^m\}$, when defined with respect to the above inner products, contains in the case $m = 2n$ the diagonal subgroup $H = \{\mathrm{diag}(z_1, \ldots, z_n, z_1^{-1}, \ldots, z_n^{-1}) : z_1, \ldots, z_n \in \mathbb{C}^\times\}$ and in the case $m = 2n+1$ the diagonal subgroup $H = \{\mathrm{diag}(z_1, \ldots, z_n, z_1^{-1}, \ldots, z_n^{-1}, 1) : z_1, \ldots, z_n \in \mathbb{C}^\times\}$ whose Lie algebras $\mathfrak{h}$ are given respectively by $\mathfrak{h} = \{\mathrm{diag}(Z_1, \ldots, Z_n, -Z_1, \ldots, -Z_n) : Z_1, \ldots, Z_n \in \mathbb{C}\}$ and $\mathfrak{h} = \{\mathrm{diag}(Z_1, \ldots, Z_n, -Z_1, \ldots, -Z_n, 0) : Z_1, \ldots, Z_n \in \mathbb{C}\}$. The Lie algebra of $\mathfrak{g}$ is described on the handout (§29.4) and will not be repeated here.

29.2. **Classical simple complex Lie algebras.**

Definition 204. By a *classical simple complex Lie algebra* we shall mean a complex Lie algebra of one of the following forms:

- $A_n := \mathfrak{sl}_{n+1}(\mathbb{C})$ for $n \geq 1$,
- $B_n := \mathfrak{so}_{2n+1}(\mathbb{C})$ for $n \geq 1$,
- $C_n := \mathfrak{sp}_{2n}(\mathbb{C})$ for $n \geq 1$,
- $D_n := \mathfrak{so}_{2n}(\mathbb{C})$ for $n \geq 3$.

Remark 205. $D_2 \cong A_1 \times A_1$ is not simple (it is a direct sum of two simple Lie algebras). $D_1 = \mathfrak{so}_2(\mathbb{C})$ is not simple (it is abelian).

Recall that our motivating goal for the past few lectures has been to prove the following theorem:

Theorem 206. *There are no isomorphisms between the classical simple complex Lie algebras except possibly those of the form*

$$A_1 \cong B_1 \cong C_1, \tag{119}$$

$$B_2 \cong C_2, \tag{120}$$

$$A_3 \cong D_3. \tag{121}$$

Remark 207. In fact, the isomorphisms (119), (120) and (121) all hold. We have proven that $A_1 \cong B_1$. It is immediate from the definition that $A_1 \cong C_1$. We have not yet proven the exceptional isomorphisms (120) and (121), but they exist, and are not inordinately complicated to establish.

Remark 208. We may reformulate Theorem 206 in terms of the simply-connected complex Lie groups having the indicated Lie algebras.

Remark 209. We record some motivation for caring about Theorem 206.

(1) It's interesting in its own right; it's natural to ask for a complete list of isomorphisms between some naturally occurring groups.

(2) The techniques involved in the proof (roots, weights, reflections, Dynkin diagrams, ...) are very important in all of Lie theory and its applications in other fields of mathematics. The specific groups involved are also universally important. Our primary goal is really to introduce those techniques by application to a motivating problem.
(3) Theorem 206 is weaker than the full classification theorem for *all* complex simple Lie algebras (not just the classical ones), which says that the above list is complete with five exceptions, denoted G_2, F_4, E_6, E_7, E_8. That full classification is not inordinately difficult, but would probably take most of a semester to present properly, and it's very easy to get lost in the middle of it and lose the big picture. On the other hand, we should be able to complete the proof of Theorem 206 in a couple lectures.
(4) It takes a lot of experience to gain intuition for working with roots, weights, reflections, etc. It seems best to introduce them as explicitly as possible. That way, many properties that would normally require laborious and unenlightening proofs can be discovered by inspection; one can then later learn proofs of such properties that apply more generally.

29.3. **How to classify them (without worrying about why it works).** We now outline the structure of the proof of Theorem 206. Let $\mathfrak{g}$ be a classical simple complex Lie algebra. We will attach to $\mathfrak{g}$ a certain oriented multigraph, called a *Dynkin diagram*. (See the handout (§29.4) for what these look like in all cases. We explained in class how coincidences between "small" Dynkin diagrams explain all of the exceptional isomorphisms (119), (120) and (121).)

The procedure by which we will attach the Dynkin diagram will involve several choices. To make the proof of Theorem 206 rigorous, we will later have to go back and check that these choices did not affect the final result.

Let us note right away that because $\mathfrak{g}$ is simple, its center (equivalently, the kernel of $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$) is trivial. Indeed, $\mathfrak{g}$ is non-abelian (by the definition of "simple"), so the center $\mathfrak{z}$ of $\mathfrak{g}$ satisfies $\mathfrak{z} \neq \mathfrak{g}$. On the other hand, the center is an ideal; since $\mathfrak{g}$ is simple, we must have $\mathfrak{z} = 0$. In other words,

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}) \text{ is injective.} \tag{122}$$

(1) First, introduce the following definition:

Definition 210. Let $\mathfrak{g}$ be a simple complex Lie algebra (the case that $\mathfrak{g}$ is "classical" is all we will use for now). A *Cartan subalgebra*[5] of $\mathfrak{g}$ is a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ for which
(a) $\mathfrak{h}$ is abelian,
(b) $\mathfrak{h}$ consists entirely of ad-semisimple elements (that is to say, for each $X \in \mathfrak{h}$, the linear endomorphism $\mathrm{ad}_X \in \mathrm{End}(\mathfrak{h})$ is diagonalizable, i.e., admits a basis of eigenvectors), and
(c) $\mathfrak{h}$ is its own centralizer: if $X \in \mathfrak{g}$ satisfies $[X, H] = 0$ for all $H \in \mathfrak{h}$, then $X \in \mathfrak{h}$.

[5] What I've recorded here is not the standard definition of "Cartan subalgebra," but is equivalent to a specialization of that definition, and is convenient for our immediate purposes; we may return to the more general notion later.

For example, the subalgebra $\mathfrak{h}$ defined on the handout (§29.4) is a Cartan subalgebra, and it turns out that all other Cartan subalgebras are "conjugate" to it; we will explain this more in a bit. The following definition thus depends only upon $\mathfrak{g}$, not upon $\mathfrak{h}$:

Definition 211. The *rank* of $\mathfrak{g}$ is defined to the dimension of $\mathfrak{h}$.

Let R denote the set of roots for $\mathrm{ad} : \mathfrak{h} \to \mathrm{End}(\mathfrak{g})$, thus R is the set of all nonzero $\alpha \in \mathfrak{h}^*$ for which the subspace $\mathfrak{g}^\alpha := \{X \in \mathfrak{g} : [H, X] = \alpha(H)X \text{ for all } H \in \mathfrak{h}\}$ of $\mathfrak{g}$ is nonzero. We describe R explicitly on the handout (§29.4).

For example, for $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$, we can take for $\mathfrak{h}$ the standard diagonal subalgebra. We saw in the lecture on the simplicity of $\mathfrak{g}$ that $\mathfrak{h}$ consists entirely of ad-semisimple elements; indeed,

$$\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{\alpha\in R}\mathfrak{g}^\alpha) \tag{123}$$

where each space on the RHS is an eigenspace for $\mathrm{ad}(\mathfrak{h})$. It is clear that $\mathfrak{h}$ is its own centralizer. Indeed, suppose $Z \in \mathfrak{g}$ commutes with every element of $\mathfrak{h}$. We can decompose Z using (123) as a sum $Z_0 + \sum_{\alpha\in R} Z_\alpha$, where $Z_0 \in \mathfrak{h}$ and $Z_\alpha \in \mathfrak{g}^\alpha$. For each $H \in \mathfrak{h}$, we have $[H, Z] = 0$, by assumption; on the other hand,

$$[H, Z] = [H, Z_0] + \sum_{\alpha\in R} [H, Z_\alpha] = \sum_{\alpha\in R} \alpha(H)Z_\alpha. \tag{124}$$

Since the Z_α are linearly independent of one another (as they belong to distinct root spaces), we must have $\alpha(H)Z_\alpha = 0$ for all $H \in \mathfrak{h}$. Since each root $\alpha \in R$ is nonzero, we can find $H \in \mathfrak{h}$ so that $\alpha(H) \neq 0$, thus $Z_\alpha = 0$ for all $\alpha \in R$ and thus $Z = Z_0$ belongs to $\mathfrak{h}$. Since Z was arbitrary, we conclude that $\mathfrak{h} \subseteq \mathfrak{h}'$, as required.

More generally, one can verify that the subaglebras $\mathfrak{h}$ defined on the handout (§29.4) in the cases A_n, B_n, C_n, D_n are in fact Cartan subalgebras. One sees also that A_n, B_n, C_n, D_n have rank n. This explains the indexing.

We observe (by inspecting each family) that the set R of roots for $(\mathfrak{h}, \mathfrak{g})$ has the following properties (noted earlier for $\mathfrak{sl}_n(\mathbb{C})$ and $\mathfrak{sp}_{2n}(\mathbb{C})$):

(a) For $\alpha \in R$, one has $\{n \in \mathbb{Z} : n\alpha \in R\} = \{\pm 1\}$.
(b) $\dim \mathfrak{g}^\alpha = 1$ for all $\alpha \in R$.
(c) Let $X_\alpha \in \mathfrak{g}^\alpha$ be nonzero, so that $\mathfrak{g}^\alpha = \mathbb{C}X_\alpha$. There exists a unique $Y_\alpha \in \mathfrak{g}^{-\alpha}$ so that the element $H_\alpha \in \mathfrak{h}$ defined by $H_\alpha := [X_\alpha, Y_\alpha]$ satisfies $\alpha(H_\alpha) = 2$.
(d) For all $\alpha, \beta \in R$,

$$[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \begin{cases} \mathfrak{g}^{\alpha+\beta} & \text{if } \alpha + \beta \in R \\ \mathbb{C}H_\alpha & \text{if } \alpha + \beta = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Explicit choices for the $X_\alpha, Y_\alpha, H_\alpha$ in all cases are given on the handout (§29.4).

(2) Next, we introduce the following definition:

Definition 212. A *base* (or *simple system* or *system of simple roots*) $S \subseteq R$ is a subset with the following properties:

(a) S is a basis of $\mathfrak{h}^*$.
(b) For each $\beta \in R$, if one writes

$$\beta = \sum_{\alpha \in S} c_\alpha \alpha \tag{125}$$

with $c_\alpha \in \mathbb{C}$ (as one can, because S is a basis), then the c_α are integers and all have the same sign (i.e., either $c_\alpha \geq 0$ for all α or $c_\alpha \leq 0$ for all α).

On the handout (§29.4), an explicit choice of a simple system S is given for each of the classical complex simple Lie algebras. There are in fact many possible choices, and we will have to argue later that they are all "sufficiently equivalent" for the purposes of the construction to follow.

The set of *positive roots* (with respect to the given simple system S) is the set R^+ consisting of all $\beta \in R$ for which in the decomposition (125), one has $c_\alpha \geq 0$ for all $\alpha \in S$. The set R^- of *negative roots* is defined analogously. One has $R = R^+ \sqcup R^-$ and $R^- = (-R^+) := \{-\alpha : \alpha \in R^+\}$. It's worth going through the examples and seeing what R^+ looks like. For example, in the case $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ and for the standard choice of S recorded on the handout (§29.4), R^+ consists of those α for which the corresponding root space $\mathfrak{g}^\alpha$ belongs to the space of strictly upper-triangular matrices; by contrast, the negative roots correspond to strictly lower-triangular root spaces.

(3) One now writes down the *Cartan matrix* $N = (\alpha(H_\beta))_{\alpha,\beta \in S}$. This is straightforward, but it's worth working through all of the examples to make sure you understand it. Two Cartan matrices are *equivalent* if one is obtained from the other by relabeling the indices. (There is no preferred ordering among elements of the finite set S, so this is a natural notion of equivalence.)

(4) Finally, for convenience, one converts the Cartan matrix into a *Dynkin diagram.* This is a finite graph whose vertices are given by the elements of S. (It is convenient in practice because most entries of the Cartan matrix turn out to be zero.) Given distinct elements $\alpha, \beta \in S$, one sees by inspection that the ordered pair of integers $(\alpha(H_\beta), \beta(H_\alpha))$ is of the following form:
(a) $(0, 0)$: in this case we draw *no* edges connecting α, β.
(b) $(-1, -1)$: in this case we draw one undirected edge between α, β.
(c) $(-2, -1)$: in this case we draw a double edge directed from α to β; see the handout
(d) $(-1, -2)$: in this case we draw a double edge directed from β to α; see the handout
(e) $(-3, -1)$: in this case we draw a triple edge directed from α to β; this case doesn't occur for classical Lie algebras (but does for the exceptional Lie algebra G_2)
(f) $(-1, -3)$: in this case we draw a triple edge directed from β to α; same comments apply.

It is obvious that the Dynkin diagram determines the Cartan matrix (and vice-versa, of course). Dynkin diagrams are nicer to work with because their equivalences are easier to spot.

To complete the proof of Theorem 206, we need to check that the Dynkin diagram (up to equivalence, i.e., relabeling of the vertices) is independent of the choice of Cartan subalgebra $\mathfrak{h}$ and simple system S made in the above construction. This will occupy the next several sections.

Exercise 35. Show that if $\mathfrak{g} := \mathfrak{so}_m(\mathbb{C})$ is defined using the *standard* scalar product $\langle x, y\rangle := \sum_{i=1}^m x_i y_i$ on $\mathbb{C}^m$ (as opposed to that in §29.1), so that $\mathfrak{g} = \{X \in \mathfrak{sl}_m(\mathbb{C}) : X^t + X = 0\}$, then $\mathfrak{g}$ contains no nonzero diagonal elements, but that the following subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is a Cartan subalgebra (see Definition 210): if $m = 2n$, then

$$\mathfrak{h} = \left\{ \begin{pmatrix} 0 & i\lambda_1 & & & & & & \\ -i\lambda_1 & 0 & & & & & & \\ & & 0 & i\lambda_2 & & & & \\ & & -i\lambda_2 & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & i\lambda_n \\ & & & & & & -i\lambda_n & 0 \end{pmatrix}, \right\}$$

while if $m = 2n+1$, then

$$\mathfrak{h} = \left\{ \begin{pmatrix} 0 & i\lambda_1 & & & & & & & \\ -i\lambda_1 & 0 & & & & & & & \\ & & 0 & i\lambda_2 & & & & & \\ & & -i\lambda_2 & 0 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \ddots & & & \\ & & & & & & 0 & i\lambda_n & \\ & & & & & & -i\lambda_n & 0 & \\ & & & & & & & & 0 \end{pmatrix}. \right\}$$

29.4. **Dynkin diagrams of classical simple Lie algebras.** As an exercise, reproduce the following in private.

$A_{n-1} : \mathrm{SL}_n(\mathbb{C}), \mathrm{SU}(n), \mathfrak{g} = \mathfrak{sl}_n(\mathbb{C}) = \{a \in M_n(\mathbb{C}) : \operatorname{trace}(a) = 0\}$

$\varepsilon_i := E_{ii}$,

$$\mathfrak{h} = \left\{ \sum_{i=1}^n a_i \varepsilon_i : \sum a_i = 0 \right\} = \left\{ H = \begin{pmatrix} \lambda_1(H) & & \\ & \cdots & \\ & & \lambda_n(H) \end{pmatrix} : \lambda_1(H) + \cdots + \lambda_n(H) = 0 \right\}$$

$$\mathfrak{h}^* = \frac{\mathbb{C}\lambda_1 \oplus \cdots \oplus \mathbb{C}\lambda_n}{\mathbb{C}(\lambda_1 + \cdots + \lambda_n)}, \quad R = \{\pm(\lambda_j - \lambda_k) : j < k\}$$

$$X_{\lambda_j - \lambda_k} = E_{jk}, \quad Y_{\lambda_j - \lambda_k} = E_{kj}, \quad H_{\lambda_j - \lambda_k} = \varepsilon_j - \varepsilon_k \quad (j \neq k)$$

$$S = \{\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, \ldots, \lambda_{n-1} - \lambda_n\}, \quad R^+ = \{(\lambda_j - \lambda_k) : j < k\}$$

$$N = (\alpha(H_\beta))_{\alpha,\beta \in S} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}$$

 (A_5)

e.g., $(\lambda_{j+1} - \lambda_j)(H_{\lambda_j - \lambda_{j-1}}) = -1$ and $(\lambda_j - \lambda_{j-1})(H_{\lambda_{j+1} - \lambda_j}) = -1$

$B_n : \mathrm{Spin}_{2n+1}(\mathbb{C}), \mathrm{Spin}(2n+1),$[6]

$$\mathfrak{g} = \mathfrak{so}_{2n+1}(\mathbb{C}) = \left\{ \begin{pmatrix} a & b & x \\ c & -a^t & y \\ -y^t & -x^t & 0 \end{pmatrix} \in M_{2n+1}(\mathbb{C}) : \begin{matrix} a, b, c \in M_{n\times n}(\mathbb{C}), \\ x, y \in M_{n\times 1}(\mathbb{C}), b^t = -b, c^t = -c \end{matrix} \right\}$$

$\varepsilon_i := E_{ii} - E_{n+i,n+i},$

$$\mathfrak{h} = \mathbb{C}\varepsilon_1 \oplus \cdots \oplus \mathbb{C}\varepsilon_n = \left\{ H = \begin{pmatrix} \lambda_1(H) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_n(H) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_1(H) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda_n(H) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \right\}$$

$$\mathfrak{h}^* = \mathbb{C}\lambda_1 \oplus \cdots \oplus \mathbb{C}\lambda_n, \quad R = \{\pm(\lambda_j \pm \lambda_k) : j < k\} \sqcup \{\pm\lambda_j\}$$

$$X_{\lambda_j-\lambda_k} = E_{jk} - E_{n+k,n+j}, \quad Y_{\lambda_j-\lambda_k} = E_{kj} - E_{n+j,n+k}, \quad H_{\lambda_j-\lambda_k} = \varepsilon_j - \varepsilon_k \quad (j \neq k)$$

$$X_{\lambda_j+\lambda_k} = E_{j,n+k} - E_{k,n+j}, \quad Y_{\lambda_j+\lambda_k} = -E_{n+j,k} + E_{n+k,j}, \quad H_{\lambda_j+\lambda_k} = \varepsilon_j + \varepsilon_k \quad (j < k)$$

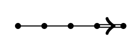
$$X_{-\lambda_j-\lambda_k} = E_{n+j,k} - E_{n+k,j}, \quad Y_{-\lambda_j-\lambda_k} = -E_{j,n+k} + E_{k,n+j}, \quad H_{-\lambda_j-\lambda_k} = -\varepsilon_j - \varepsilon_k \quad (j < k)$$

$$X_{\lambda_j} = E_{j,2n+1} - E_{2n+1,n+j} \quad Y_{\lambda_j} = 2(E_{2n+1,j} - E_{n+j,2n+1}), \quad H_{\lambda_j} = 2\varepsilon_j$$

$$X_{-\lambda_j} = E_{2n+1,j} - E_{n+j,2n+1} \quad Y_{-\lambda_j} = 2(E_{j,2n+1} - E_{2n+1,n+j}), \quad H_{-\lambda_j} = -2\varepsilon_j$$

$$S = \{\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, \ldots, \lambda_{n-1} - \lambda_n, \lambda_n\}, \quad R^+ = \{(\lambda_j \pm \lambda_k) : j < k\} \sqcup \{\lambda_j\}$$

$$N = (\alpha(H_\beta))_{\alpha,\beta\in S} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -2 \\ & & & -1 & 2 \end{pmatrix}$$



(B_5)

e.g., $(\lambda_n)(H_{\lambda_{n-1}-\lambda_n}) = -1$ and $(\lambda_{n-1} - \lambda_n)(H_{\lambda_n}) = -2$.

$C_n : \mathrm{Sp}_{2n}(\mathbb{C}), \mathrm{Sp}(2n), \quad \mathfrak{g} = \mathfrak{sp}_{2n}(\mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ c & -a^t \end{pmatrix} \in M_{2n}(\mathbb{C}) : b^t = b, c^t = c \right\},$

$\varepsilon_i := E_{ii} - E_{n+i,n+i},$

$$\mathfrak{h} = \mathbb{C}\varepsilon_1 \oplus \cdots \oplus \mathbb{C}\varepsilon_n = \left\{ H = \begin{pmatrix} \lambda_1(H) & 0 & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_n(H) & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_1(H) & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda_n(H) \end{pmatrix} \right\}$$

$$\mathfrak{h}^* = \mathbb{C}\lambda_1 \oplus \cdots \oplus \mathbb{C}\lambda_n, \quad R = \{\pm(\lambda_j \pm \lambda_k) : j < k\} \sqcup \{\pm 2\lambda_j\}$$

$$X_{\lambda_j-\lambda_k} = E_{jk} - E_{n+k,n+j}, \quad Y_{\lambda_j-\lambda_k} = E_{kj} - E_{n+j,n+k}, \quad H_{\lambda_j-\lambda_k} = \varepsilon_j - \varepsilon_k \quad (j \neq k)$$

$$X_{\lambda_j+\lambda_k} = E_{j,n+k} + E_{k,n+j}, \quad Y_{\lambda_j+\lambda_k} = E_{n+j,k} + E_{n+k,j}, \quad H_{\lambda_j+\lambda_k} = \varepsilon_j + \varepsilon_k \quad (j < k)$$

$$X_{-\lambda_j-\lambda_k} = E_{n+j,k} + E_{n+k,j}, \quad Y_{-\lambda_j-\lambda_k} = E_{j,n+k} + E_{k,n+j}, \quad H_{-\lambda_j-\lambda_k} = -\varepsilon_j - \varepsilon_k \quad (j < k)$$

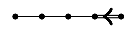
$$X_{2\lambda_j} = E_{j,n+j}, \quad Y_{2\lambda_j} = E_{n+j,j}, \quad H_{2\lambda_j} = \varepsilon_j$$

$$X_{-2\lambda_j} = E_{n+j,j}, \quad Y_{-2\lambda_j} = E_{j,n+j}, \quad H_{-2\lambda_j} = -\varepsilon_j$$

$$S = \{\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, \ldots, \lambda_{n-1} - \lambda_n, 2\lambda_n\}, \quad R^+ = \{(\lambda_j \pm \lambda_k) : j < k\} \sqcup \{2\lambda_j\}$$

---

[6]Defined via the scalar product $\langle x, y\rangle := \sum_{j=1}^n (x_j y_{n+j} + x_{n+j} y_j) + x_{2n+1} y_{2n+1}$

$$N = (\alpha(H_\beta))_{\alpha,\beta\in S} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -2 & 2 \end{pmatrix}$$  (C_5)

e.g., $(2\lambda_n)(H_{\lambda_{n-1}-\lambda_n}) = -2$ and $(\lambda_{n-1}-\lambda_n)(H_{2\lambda_n}) = -1$

$D_n : \mathrm{Spin}_{2n}(\mathbb{C}), \mathrm{Spin}(2n),$[7] $\quad \mathfrak{g} = \mathfrak{so}_{2n}(\mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ c & -a^t \end{pmatrix} \in M_{2n}(\mathbb{C}) : b^t = -b, c^t = -c \right\},$

$\varepsilon_i := E_{ii} - E_{n+i,n+i},$

$$\mathfrak{h} = \mathbb{C}\varepsilon_1 \oplus \cdots \oplus \mathbb{C}\varepsilon_n = \left\{ H = \begin{pmatrix} \lambda_1(H) & 0 & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_n(H) & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_1(H) & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda_n(H) \end{pmatrix} \right\}$$

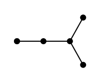
$$\mathfrak{h}^* = \mathbb{C}\lambda_1 \oplus \cdots \oplus \mathbb{C}\lambda_n, \quad R = \{\pm(\lambda_j \pm \lambda_k) : j < k\}$$

$X_{\lambda_j-\lambda_k} = E_{jk} - E_{n+k,n+j}, \quad Y_{\lambda_j-\lambda_k} = E_{kj} - E_{n+j,n+k}, \quad H_{\lambda_j-\lambda_k} = \varepsilon_j - \varepsilon_k \quad (j \neq k)$

$X_{\lambda_j+\lambda_k} = E_{j,n+k} - E_{k,n+j}, \quad Y_{\lambda_j+\lambda_k} = -E_{n+j,k} + E_{n+k,j}, \quad H_{\lambda_j+\lambda_k} = \varepsilon_j + \varepsilon_k \quad (j < k)$

$X_{-\lambda_j-\lambda_k} = E_{n+j,k} - E_{n+k,j}, \quad Y_{-\lambda_j-\lambda_k} = -E_{j,n+k} + E_{k,n+j}, \quad H_{-\lambda_j-\lambda_k} = -\varepsilon_j - \varepsilon_k \quad (j < k)$

$S = \{\lambda_1 - \lambda_2, \lambda_2 - \lambda_3, \ldots, \lambda_{n-2} - \lambda_{n-1}, \lambda_{n-1} - \lambda_n, \lambda_{n-1} + \lambda_n\}, \quad R^+ = \{(\lambda_j \pm \lambda_k) : j < k\}$

$$N = (\alpha(H_\beta))_{\alpha,\beta\in S} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & -1 \\ & & -1 & 2 & 0 \\ & & -1 & 0 & 2 \end{pmatrix}$$  (D_5)

e.g., $(\lambda_{n-2}-\lambda_{n-1})(H_{\lambda_{n-1}-\lambda_n}) = (\lambda_{n-2}-\lambda_{n-1})(H_{\lambda_{n-1}+\lambda_n}) = (\lambda_{n-1}-\lambda_n)(H_{\lambda_{n-2}-\lambda_{n-1}}) = (\lambda_{n-1}+\lambda_n)(H_{\lambda_{n-2}-\lambda_{n-1}}) = -1$ and $(\lambda_{n-1}-\lambda_n)(H_{\lambda_{n-1}+\lambda_n}) = (\lambda_{n-1}+\lambda_n)(H_{\lambda_{n-1}-\lambda_n}) = 0$

29.5. **Classical algebras come with faithful representations and are cut out by anti-involutions.** Let $\mathfrak{g}$ be a classical simple complex Lie algebra. Then $\mathfrak{g}$ comes equipped with a defining faithful representation $\mathfrak{g} \hookrightarrow \mathrm{End}(V)$, where $V = \mathbb{C}^{n+1}, \mathbb{C}^{2n+1}, \mathbb{C}^{2n}, \mathbb{C}^{2n}$ according as $\mathfrak{g} = A_n, B_n, C_n, D_n$.

We now record a property of $\mathfrak{g}$ that will allow us to give "ad hoc" proofs of some assertions in the sections to follow. Set $\mathfrak{sl}(V) := \{x \in \mathrm{End}(V) : \mathrm{trace}(x) = 0\}$. There is a linear anti-involution $\sigma : \mathrm{End}(V) \to \mathrm{End}(V)$ for which

$$\mathrm{trace}(\sigma(x)) = \mathrm{trace}(x) \text{ for all } x \in \mathrm{End}(V) \tag{126}$$

and

$$\mathfrak{g} = \{x \in \mathfrak{sl}(V) : \sigma(x) = -x\}. \tag{127}$$

Namely:

(1) If $\mathfrak{g} = A_n$, then we take for σ the identity map.

---

[7]Defined via the scalar product $\langle x, y\rangle := \sum_{j=1}^n (x_j y_{n+j} + x_{n+j} y_j)$.

(2) If $\mathfrak{g} = B_n, C_n, D_n$, we take

$$\sigma(x) := J^{-1} x^t J$$

where

$$J = \begin{pmatrix} 0_{n\times n} & 1_{n\times n} & 0_{n\times 1} \\ 1_{n\times n} & 0_{n\times n} & 0_{n\times 1} \\ 0_{1\times n} & 0_{1\times n} & 1_{1\times 1} \end{pmatrix} \text{ if } \mathfrak{g} = B_n,$$

$$J = \begin{pmatrix} 0_{n\times n} & 1_{n\times n} \\ -1_{n\times n} & 0_{n\times n} \end{pmatrix} \text{ if } \mathfrak{g} = C_n,$$

$$J = \begin{pmatrix} 0_{n\times n} & 1_{n\times n} \\ 1_{n\times n} & 0_{n\times n} \end{pmatrix} \text{ if } \mathfrak{g} = D_n,$$

By "linear anti-involution" we mean that

$$\sigma(ax + by) = a\sigma(x) + b\sigma(y), \quad \sigma(xy) = \sigma(y)\sigma(x), \quad \sigma(\sigma(x)) = x.$$

29.6. **Diagonalization in classical Lie algebras.** Let $\mathfrak{g}$ be a classical simple complex Lie algebra and let $g \hookrightarrow \mathrm{End}(V)$ be as in §29.5. Let G denote a simply-connected complex Lie group having Lie algebra $\mathfrak{g}$. (Thus G is one of $\mathrm{SL}_{n+1}(\mathbb{C}), \mathrm{Spin}_{2n+1}(\mathbb{C}), \mathrm{Sp}_{2n}(\mathbb{C}), \mathrm{Spin}_{2n}(\mathbb{C})$ according as we are in case A_n, B_n, C_n, D_n.) Let $x \in \mathfrak{g}$. We can think of it as a linear transformation $x : V \to V$.

Lemma 213. *If $x : V \to V$ is semisimple, then there exists $g \in G$ so that $\mathrm{Ad}(g)x \in \mathfrak{g} \subseteq \mathrm{End}(V)$ is diagonal, i.e., represented by a diagonal matrix with respect to the standard basis of V.*

Proof. Suppose first that $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$. Let $\mathcal{B} = (v_1, \dots, v_n)$ be a basis of eigenvectors for x. Let $g \in \mathrm{GL}_n(\mathbb{C})$ be the "change of basis matrix" from $\mathcal{B}$ to the standard basis. Then gxg^{-1} is diagonal. This conclusion is unaffected by replacing g with a scalar multiple; by doing so, we may arrange that g belongs to $\mathrm{SL}_n(\mathbb{C})$.

TODO (or Exercise): discussion of other cases.

□

29.7. **Semisimplicity of elements of classical Lie algebras.** Let $\mathfrak{g} \hookrightarrow \mathrm{End}(V)$ be as in §29.5. We have seen (see (122)) that $\mathrm{ad} : \mathfrak{g} \hookrightarrow \mathrm{End}(\mathfrak{g})$ is another faithful representation. Given an element $x \in \mathfrak{g}$, it thus makes sense to compare properties of the linear transformation $x : V \to V$ with those of $\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$. The following comparison will be of particular use (see §28.1 to refresh the terminology):

Lemma 214. *$x : V \to V$ is semisimple if and only if $\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$ is semisimple.*

For the proof, it will be convenient to recall a standard fact from linear algebra:

Theorem 215. *Let $x \in \mathrm{End}(V)$ be a linear endomorphism of a finite-dimensional complex vector space V. Then there exist unique $s, n \in \mathrm{End}(V)$ so that*

(1) s is semisimple,
(2) n is nilpotent, and
(3) $[s, n] = 0$.

Moreover, there exist polynomials S and N (depending upon x) so that $S(0) = 0 = N(0)$ and $s = S(x)$ and $n = N(x)$ and so that S, T are both odd (i.e., they are sums of monomials of odd degree).

Proof. See for instance p40 of the book by Serre on the course reference. The final condition is not stated there, but may be obtained by inspection of the proof. (We note that the "existence" may be obtained by taking for s the "diagonal part" and n the "off-diagonal part" of the Jordan normal form of x.) $\square$

We turn to the proof of Lemma 214 (which wasn't presented correctly in lecture). Before embarking, we note that the same conclusion holds for any simple complex Lie algebra but with a slightly more complicated proof; since we are focusing on the classical case for now, we will freely make use of the assumption that $\mathfrak{g}$ is classical. *This proof is not particularly important*; I am including it here only because it is short and suffices for our present focused goal of classifying *classical* simple complex Lie algebras. (I'll also remark that it may be possible to check it by hand more simply than how I have argued here.)

We observe first that it is easy to see that any $x \in \mathfrak{g}$ for which $x : V \to V$ is semisimple has the property that $\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$ is semisimple: we may assume (by §29.6) that x is diagonal, in which case the required conclusion is clear by the root space decomposition as computed explicitly in the handout (§29.4). It remains to establish the converse.

Let σ be the anti-involution discussed in §29.5 that defines $\mathfrak{g}$. Let $x \in \mathfrak{g}$ be any element. Consider its Jordan decomposition $x = s + n$ in $\mathrm{End}(V)$. Write $s = S(x), n = N(x)$ as above. Since S, N are odd, we have $\sigma(S(x)) = S(\sigma(x)) = -S(x)$ and $\sigma(N(x)) = N(\sigma(x)) = -N(x)$, whence s, n also belong to $\mathfrak{g}$.

We've seen already that as s is semisimple, so is ad_s. Moreover, since n is nilpotent, it's not hard to see that also ad_n is nilpotent: if $n^m = 0$ for some $m \in \mathbb{Z}_{\geq 1}$, then for any $y \in \mathfrak{g}$, we have $\mathrm{ad}_n^{2m}(y) = [n, [n, \dots, [n, y]]] = 0$, since when we expand out in terms of monomials, we always have at least m copies of n occurring consecutievly.

Finally, since ad is a Lie algebra morphism, we have $[\mathrm{ad}_s, \mathrm{ad}_n] = \mathrm{ad}_{[s,n]} = \mathrm{ad}_0 = 0$.

In summary, $\mathrm{ad}_s, \mathrm{ad}_n$ satisfy we see that the decomposition $\mathrm{ad}_x = \mathrm{ad}_s + \mathrm{ad}_n$ satisfies the assumptions of Theorem 215.

Assume finally that ad_x is semisimple. Then (by the uniqueness assertion of Theorem 215) we must have $\mathrm{ad}_x = \mathrm{ad}_s$ and $\mathrm{ad}_n = 0$; since ad is injective, we must have $n = 0$; therefore $x = s$ is semisimple, as required.

29.8. **Conjugacy of Cartan subalgebras.** Let $\mathfrak{g}$ be a classical simple complex Lie algebra. Here's the key to showing the independence of the constructions given above with respect to the choice of $\mathfrak{h}$:

Lemma 216. *For any two Cartan subalgebras $\mathfrak{h}, \mathfrak{h}'$ of $\mathfrak{g}$, there exists $g \in G$ so that $\mathrm{Ad}(g)\mathfrak{h} = \mathfrak{h}'$. (Here we can take for G any Lie group having $\mathfrak{g}$ as its Lie algebra; the simply-connected one will do. One could alternatively and more naturally take for G the* inner automorphism group $\mathrm{Int}(\mathfrak{g}) := \langle \exp(\mathrm{ad}_X) : X \in \mathfrak{g} \rangle \leq \mathrm{Aut}(\mathfrak{g}) \leq \mathrm{GL}(\mathfrak{g})$ *as defined in lecture.)*

Proof. Recall from §29.5 the embedding $\mathfrak{g} \hookrightarrow \mathrm{End}(V)$. We have seen in §214 that an element $x \in \mathfrak{g}$, regarded as a linear transformation $x : V \to V$, is semisimple if and only if $\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$ is semisimple.

We turn to the proof. It will suffice to consider the case that $\mathfrak{h}$ is the standard diagonal Cartan subalgebra and $\mathfrak{h}'$ is arbitrary. Since the elements of $\mathfrak{h}'$ are all commuting and ad-semisimple, we may simultaneously diagonalize their adjoint action as a sum of eigenspaces, i.e., we may write down a root space decomposition of $\mathfrak{h}'$; see §28.4, and recall that elements of $\mathfrak{h}'$ are ad-semisimple). Let $Z \in \mathfrak{h}'$ be any element with the property that $\alpha'(Z) \neq 0$ for all roots α' of $\mathfrak{h}'$. (Such an element exists because each α' is nonzero, and so its kernel has codimension one, and a finite union of codimension one subspaces of a vector space over an infinite field is properly contained in that vector space.) Then the only elements of $\mathfrak{g}$ that commute with Z are those in $\mathfrak{h}'$ (compare with discussion surrounding (124)). By §29.6, there is an element $g \in G$ (the simply-connected complex Lie group having Lie algebra $\mathfrak{g}$) for which $gZg^{-1} := \mathrm{Ad}(g)Z \in \mathfrak{h}$. Thus every $H \in \mathfrak{h}$ commutes with gZg^{-1}. It follows that Z commutes with $g^{-1}Hg$ for all $H \in \mathfrak{h}$. By the property of Z just mentioned, it follows that $\mathrm{Ad}(g^{-1})\mathfrak{h} \subseteq \mathfrak{h}'$, hence $\mathfrak{h} \subseteq \mathrm{Ad}(g)\mathfrak{h}'$. By the maximality condition in the definitino of "Cartan subalgebra," it follows that $\mathfrak{h} = \mathrm{Ad}(g)\mathfrak{h}'$, as required. $\square$

29.9. **Interpretation of Cartan matrix in terms of inner products.** Let $\mathfrak{g}$ be one of A_n, B_n, C_n, D_n. We can think of $\mathfrak{h}^*$ as a subspace of $\mathbb{C}^n$ by writing each $\lambda \in \mathfrak{h}^*$ in the form $\lambda = l_1\lambda_1 + \cdots + l_n\lambda_n$ and associating to λ the element $(l_1, \ldots, l_n) \in \mathbb{C}^n$. In the case $\mathfrak{g} = A_n$, we have $\lambda_1 + \cdots + \lambda_n = 0$, so there is some ambiguity in this assignment; we pin it down by requiring that $l_1 + \cdots + l_n = 0$. Using this assignment, we can define an inner product by the formula

$$(\lambda, \mu) := l_1 m_1 + \cdots + l_n m_n \text{ if } \lambda = \sum l_i \lambda_i, \mu = \sum m_j \lambda_j.$$

By inspection of the formulas on the handout (§29.4), we have

$$\alpha(H_\beta) = 2\frac{(\alpha, \beta)}{(\beta, \beta)}.$$

(We will explain this properly later; for now, we stick to the narrow goal of classifying the classical simple complex Lie algebras.) Thus the Cartan matrix can be described in terms of inner products involving the simple roots.

29.10. **Independence with respect to the choice of simple system.** In the context of §29.3, suppose $S, S' \subseteq R$ are two simple systems. We want to know that the Cartan matrices N, N' that they define are equivalent (i.e., that they coincide after relabeling the indices).

We will do this as follows:

Proposition 217. *Let $S, S' \subseteq R$ be simple systems. There is a linear transformation $w : \mathfrak{h}^* \to \mathfrak{h}^*$ that is orthogonal with respect to the pairing $(,)$ on $\mathfrak{h}^*$ defined in §29.9 and for which $wS = S'$.*

Assuming Proposition 217, we may complete the proof of Theorem 206 as follows: We need to check that N, N' are equivalent. Let $\alpha, \beta \in S$. Then

$$\begin{aligned}\alpha(H_\beta) &= 2\frac{(\alpha, \beta)}{(\beta, \beta)} \\ &= 2\frac{(w\alpha, w\beta)}{(w\beta, w\beta)} \\ &= (w\alpha)(H_{w\beta}).\end{aligned}$$

Since $w\alpha$ traverses S' as α traverses S, we deduce that N and N' are equivalent, as required.

To complete the proof of Theorem 206, it remains only to prove Proposition 217, i.e., to produce w. We do so as follows. For each $\alpha \in R$, let $s_\alpha : \mathfrak{h}^* \to \mathfrak{h}^*$ denote the linear transformation given by

$$s_\alpha \lambda := \lambda - \langle \lambda | \alpha \rangle \alpha,$$

where

$$\langle \lambda | \alpha \rangle := \lambda(H_\alpha) = 2\frac{(\lambda, \alpha)}{(\alpha, \alpha)}.$$

Geometrically, s_α is a reflection with respect to the hyperplane orthogonal to α. It follows by inspection of the formulas on the handout (§29.4) that

$$s_\alpha(R) = R \text{ for all } \alpha \in R. \tag{128}$$

We will see that this is an astonishingly powerful condition; we will explain it properly in due course.

The s_α may be described as follows, as detailed in lecture:

- If $\alpha = \lambda_j - \lambda_k$, then $s_\alpha : \lambda_j \mapsto \lambda_k, \lambda_k \mapsto \lambda_j$ (with all other λ_i left fixed) Here and henceforth we assume that $j \neq k$.
- If $\alpha = \lambda_j + \lambda_k$, then $s_\alpha : \lambda_j \mapsto -\lambda_k, \lambda_k \mapsto -\lambda_j$ (with all other λ_i left fixed).
- If $\alpha = \lambda_j$ or $2\lambda_j$, then $s_\alpha : \lambda_j \mapsto -\lambda_j$ (with all other λ_i left fixed).

Definition 218. Let W be the group generated by the root reflections s_α for $\alpha \in R$; it is called the *Weyl group* and is finite, as it is a subgroup of the permutation group of the spanning set R for $\mathfrak{h}^*$. Since it is generated by reflections, it consists of orthogonal transformations.

As we explained in lecture, it is described as follows:

- A_n: one has $|W| = n!$; for each permutation $j \mapsto j'$ of $(1, 2, \dots, n)$, one has the element $w \in W$ given bye $\lambda_j \mapsto \lambda_{j'}$.
- B_n: one has $|W| = 2^n n!$; for each permutation $j \mapsto j'$ of $(1, 2, \dots, n)$ and collection of signs $\pm$ (indexed by j), one has the element $w \in W$ given by $\lambda_j \mapsto \pm\lambda_{j'}$.
- C_n: same description as for B_n.
- D_n: same description as for B_n, C_n, except require that the product of all signs be $+1$ (i.e., that the number of minus signs be even).

Now let $S \subseteq R$ be a simple system. Let $\mathfrak{h}^*_{\mathbb{R}}$ denote the $\mathbb{R}$-span of R, or equivalently, the $\mathbb{R}$-span of the elements $\lambda_1, \dots, \lambda_n$; it is a real vector space of dimension $\dim_{\mathbb{C}}(\mathfrak{h}^*)$.

Definition 219. An element $\lambda \in h^*_{\mathbb{R}}$ is said to be S-*nonnegative* (or simply *nonnegative* when the simple system S is clear by context), denoted $\lambda \geq 0$, if when we write $\lambda = \sum_{\alpha \in S} c_\alpha \alpha$ with each $c_\alpha \in \mathbb{R}$, then we actually have $c_\alpha \geq 0$ fro all $\alpha \in S$.

Given $\lambda, \mu \in h^*_{\mathbb{R}}$, we say that λ is S-*higher than* μ (or simply *higher than* μ when S is clear), denoted $\lambda \geq \mu$, if $\lambda - \mu$ is nonnegative.

We write $\lambda > \mu$ if $\lambda \geq \mu$ and $\lambda \neq \mu$, etc.

Remark 220. Note that $\lambda \geq \mu$ defines a *partial order* on $\mathfrak{h}^*_{\mathbb{R}}$: there are plenty of pairs of elements that are incomparable. On the other hand, for any $\lambda \in \mathfrak{h}^*_{\mathbb{R}}$ and $\alpha \in R$, the elements λ and $s_\alpha\lambda$ are always comparable: one has $\lambda \geq s_\alpha\lambda$ if and only if $\langle \lambda | \alpha \rangle \alpha \geq 0$.

Example 221. R^+ is precisely the set of S-nonnegative elements of R.

Definition 222. An element $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$ is said to be *S-dominant* (or simply *dominant*, when the simple system S is clear by context) provided that any of the following evidently equivalent conditions are satisfied:

(1) $\lambda(H_\alpha) \geq 0$ for all $\alpha \in S$.
(2) $(\lambda, \alpha) \geq 0$ for all $\alpha \in S$.
(3) $(\lambda, \alpha) \geq 0$ for all $\alpha \in R^+$.
(4) $\lambda(H_\alpha) \geq 0$ for all $\alpha \in R^+$.
(5) $(\lambda, \alpha) \leq 0$ for all $\alpha \in R^-$.
(6) $\lambda(H_\alpha) \leq 0$ for all $\alpha \in R^-$.
(7) $\lambda \geq s_\alpha \lambda$ for all $\alpha \in S$.
(8) $\lambda \geq s_\alpha \lambda$ for all $\alpha \in R$.

[The following equivalences are clear: $(1) \iff (2), (3) \iff (4), (5) \iff (6)$. We have $(2) \iff (3) \iff (5)$ by linearity of the inner product. We have $(1) \iff (7)$ and $(4), (6) \iff (7)$ by definition of s_α and the partial relation "$\geq$."]

Definition 223. An element $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$ is *regular* if $(\lambda, \alpha) \neq 0$ for all $\alpha \in R$.

It is easy to see that regular dominant elements exist: they are just those elements belonging to a suitable "upper-right quadrant" (TODO: explain better).

Example 224. Supopse S is the "standard" simple system described on the handout (§29.4). Then it's easy to see that the elements $\lambda = \sum l_i \lambda_i \in \mathfrak{h}_{\mathbb{R}}^*$ that are S-dominant are precisely those satisfying the following conditions in the respective cases:

(A_n) $l_1 \geq l_2 \geq l_3 \geq \cdots \geq l_{n-1} \geq l_n$
(B_n) $l_1 \geq l_2 \geq l_3 \geq \cdots \geq l_{n-1} \geq l_n \geq 0$
(C_n) $l_1 \geq l_2 \geq l_3 \geq \cdots \geq l_{n-1} \geq l_n \geq 0$
(D_n) $l_1 \geq l_2 \geq l_3 \geq \cdots \geq l_{n-1} \geq |l_n|$

The regular dominant elements are those for which every "$\geq$" is actually a strict inequality "$>$."

Lemma 225. *Let S be a simple system with associated set R^+ of positive roots. Let λ be regular and S-dominant. Then $R^+ = \{\alpha \in R : (\alpha, \lambda) > 0\}$. Moreover, S is the set of elements $\alpha \in R^+$ that are* indecomposable *in the sense that one cannot write $\alpha = \beta_1 + \cdots + \beta_k$ for some $\beta_1, \ldots, \beta_k \in R^+$ with $k \geq 2$.*

In particular, S is determined by any regular S-dominant element λ.

Proof. The first assertion is clear: if $\alpha \in R$ satisfies $(\alpha, \lambda) > 0$, then in the decomposition $\alpha = \sum_{\beta \in S} c_\beta \beta$ (where the c_β are integers all of the same sign, and $(\beta, \lambda) > 0$) we deduce that each $c_\beta \geq 0$, etc. The second assertion follows immediately from the definition of "simple system." □

We can now prove Proposition 217. Let S_0, S_1 be two simple systems; we want to show that there is $w \in W$ so that $wS_1 = S_0$. Let λ be regular and S_1-dominant. Choose $w \in W$ so that $w\lambda$ is *maximal* with respect to the partial order given by Definition 219 *with respect to the simple system* S_0. Thus, in paricular, $s_\alpha w \lambda \leq w\lambda$ for all $\alpha \in R$ (cf. Remark 220). Actually, we must have $s_\alpha w\lambda < w\lambda$: for if instead we had $s_\alpha w \lambda = w\lambda$, then we'd have $\lambda = w^{-1} s_\alpha w \lambda = s_{w^{-1}\alpha} \lambda$ and so $(w^{-1}\alpha, \lambda) = 0$, contrary to the assumption that λ is regular. It follows from the equivalence of the various conditions in Definition 222 that $w\lambda$ is S_0-dominant.

In summary, λ is S_1-dominant and $w\lambda$ is S_0-dominant. Using Lemma 225, it follows easily that $wS_1 = S_0$, as required. TODO: explain more.

We record a few other facts of independent interest. (Most of these assertions can be deduced by inspection for the "standard" simple system and then deduced for general simple systems from the fact that the Weyl group acts transitively on them; there are also more natural but lengthier proofs that apply more generally.)

(1) Let S be a simple system. Then the Weyl group is generated by the root reflections s_α, $\alpha \in S$.
(2) A *Weyl chamber* is a connected component C of the set $\mathfrak{h}_{\mathbb{R}}^{\mathrm{reg}} := \{\lambda \in \mathfrak{h}_{\mathbb{R}} : (\alpha, \lambda) \neq 0 \text{ for all } \alpha \in R\}$ of regular elements. The set of simple systems S is in natural (W-equivariant) bijection with the set of Weyl chambers:
   (a) Given S, one takes for C the set $C := \{\lambda \in \mathfrak{h}_{\mathbb{R}} : (\alpha, \lambda) > 0 \text{ for all } \alpha \in S\}$ of regular S-dominant elements; that set is called (naturally) the *S-dominant Weyl chamber*. (The S-dominant Weyl chamber is, of course, a Weyl chamber: it is connected, or even path-connected by straight line segments; it is also maximal among connceted subsets, by (say) the intermediate value theorem.)
   (b) Given C, one takes for S the set of indecomposable elements in $R^+ := \{\alpha \in R : (\alpha, \lambda) > 0\}$.
(3) The Weyl group acts *simply transitively* on the set of Weyl chambers, for each pair S, S' of simple systems there exists a unique $w \in W$ for which $wS = S'$. In particular, if $w \in W$ satisfies $wS = S$ for some simple system S, then $w = 1$. (This follows from the argument given above, together with the empirical observation that the only $w \in W$ which stabilizes the "standard" Weyl chamber is $w = 1$.)
(4) The Weyl group acts *simply transitively* on the set of simple systems, for each pair S, S' of simple systems there exists a unique $w \in W$ for which $wS = S'$. In particular, if $w \in W$ satisfies $wS = S$ for some simple system S, then $w = 1$. This follows from the previous few points.
(5) Let S be a simple system. For each $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$ there exists a *unique* $w \in W$ so that $w\lambda$ is S-dominant.

In lecture, we presented (some of) the above material in a slightly different order; namely, we first stated the bijection between simple systems and Weyl chambers (after working out enough examples to make it seem obvious).

30. Why simple Lie algebras give rise to root systems

30.1. **Overview.** In the previous section, we explained how Dynkin diagrams may be used to classify the classical complex simple Lie algebras A_n, B_n, C_n, D_n. That explanation involved a fair number of "empirical observations:"

(1) We observed (without "explanation") that Cartan subalgebras $\mathfrak{h}$ of $\mathfrak{g}$ exist and are unique.
(2) We observed that the root spaces $\mathfrak{g}^\alpha$ of $\mathfrak{h}$ are one-dimensional and satisfy

$$\alpha \in R \implies \{n \in \mathbb{Z} : n\alpha \in R\} = \{\pm 1\}.$$

   We observed also that there exist $X_\alpha \in \mathfrak{g}^\alpha, Y_\alpha \in \mathfrak{g}^{-\alpha}$ and $H_\alpha \in \mathfrak{h}$ so that

$$[X_\alpha, Y_\alpha] = H_\alpha$$

and

$$\alpha(H_\alpha) = 2$$

and

$$[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \begin{cases} \mathfrak{g}^{\alpha+\beta} & \alpha + \beta \in R \\ \mathbb{C}H_\alpha & \alpha + \beta = 0 \\ 0 & \text{otherwise.} \end{cases}$$

(3) We observed the relation

$$\langle \beta | \alpha \rangle := \beta(H_\alpha) = 2\frac{(\beta, \alpha)}{(\alpha, \alpha)}$$

for any $\alpha, \beta \in R$.

(4) We observed that the root reflections $s_\alpha : \mathfrak{h}^* \to \mathfrak{h}^*$ defined for $\alpha \in R$ by

$$s_\alpha(\beta) := \beta - \langle \beta | \alpha \rangle \alpha,$$

satisfy $s_\alpha(R) = R$.

We would like now to go back and "explain" the above observations a bit more "conceptually." This will involve an application of some properties of representations of $\mathfrak{sl}_2(\mathbb{C})$ that we established long ago.

For the remainder of this section, "Lie algebra" always means "over the complex numbers."

30.2. **The basic theorem on Cartan subalgebras.** Recall Definition 210; it applies to any simple Lie algebra $\mathfrak{g}$.

Theorem 226. *(1) There exists a Cartan subalgebra $\mathfrak{h} \leq \mathfrak{g}$.*

(2) Any two Cartan subalgebras $\mathfrak{h}, \mathfrak{h}'$ are conjugate in the sense that for any Lie group G with $\mathrm{Lie}(G) = \mathfrak{g}$ *(such as the inner automorphism group $G = \mathrm{Int}(\mathfrak{g})$ as defined in Lemma 216), there exists $g \in G$ so that* $\mathrm{Ad}(g)\mathfrak{h}' = \mathfrak{h}$.

(3) There is a scalar product $(,)$ on $\mathfrak{g}$ (i.e., a non-degenerate symmetric bilinear form $\mathfrak{g} \otimes \mathfrak{g} \to \mathbb{C}$) and a real form $\mathfrak{h}_\mathbb{R} \leq \mathfrak{h}$ with the following properties:

(a) The roots α of $\mathrm{ad} : \mathfrak{h} \to \mathrm{End}(\mathfrak{g})$ *satisfy $\alpha(\mathfrak{h}_\mathbb{R}) \subseteq \mathbb{R}$.*

(b) The restriction of $(,)$ to $\mathfrak{h}_\mathbb{R}$ is real-valued and positive-definite.

(c) $(,)$ is $\mathfrak{g}$-invariant, i.e., for all $x, y, z \in \mathfrak{g}$,

$$([z, x], y) + (x, [z, y]) = 0.$$

(Think of this condition as the "$t = 0$ derivative" of a condition like $(e^{tz}x, e^{tz}y) = (x, y)$.)

Example 227. (1) One can take $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$, $\mathfrak{h} \leq \mathfrak{g}$ the diagonal subalgebra, $\mathfrak{h}_\mathbb{R} \leq \mathfrak{h}$ the real Lie subalgebra consisting of elements with real entries, and $(x, y) := \mathrm{trace}(xy)$ for $x, y \in \mathfrak{g}$. Similar choices apply for all of the classical simple algebras.

(2) For any simple $\mathfrak{g}$, it turns out that one can take $(x, y) := \mathrm{trace}(\mathrm{ad}_x \, \mathrm{ad}_y)$ (which is called the *Killing form*).

As noted earlier, Theorem 226 is easy to establish for the classical simple algebras. We will not prove it in general for the following reasons:

(1) The proof might take a couple weeks, and I think that it is not as interesting or useful as the other topics that I plan to cover in the remaining time.

(2) The conclusion is not difficult in the primary examples of interest (the classical families). We may thus interpret it as telling us that we might as well have *defined* a simple Lie algebra to be a simple Lie algebra in the ordinary sense with the additional property that it possesses a Cartan subalgebra satisfying the above properties; such a definition would apply to the primary examples of interest, and the above theorem may be interpreted as giving a weaker condition under which it holds.

A good reference for the proof of Theorem 226 is Chapter 3 of Serre's *Complex semisimple Lie algebras*.

In the following sections, we will include in some hypotheses phrases like "Let $\mathfrak{g}$ be a simple Lie algebra (that satisfies the Cartan subalgebra theorem)." Theorem 226 says that the parenthetical hypothesis is unnecessary; we include it only to keep track of what we have actually proven in the course.

30.3. Abstract root systems.

Definition 228. Let V be a finite-dimensional real inner product space. A *root system* is a subset R of V such that

(1) R is finite;
(2) R does not contain 0;
(3) for any $\alpha, \beta \in R$, the quantity
$$\langle \beta | \alpha \rangle := 2\frac{(\beta, \alpha)}{(\alpha, \alpha)}$$
is an integer;
(4) for any $\alpha \in R$, the map
$$s_\alpha : V \to V$$
$$s_\alpha(\lambda) := \lambda - \langle \lambda | \alpha \rangle \alpha$$
satisfies $s_\alpha(R) = R$.

We say that R is *reduced* if for all $\alpha \in R$,
$$\{n \in \mathbb{R} : n\alpha \in R\} = \{\pm 1\}. \tag{129}$$
The notion of an *isomorphism of root systems* is clear.[8]

In §29, we saw several examples of root systems (without referring to them by that name).

Example 229. Let $R_1 \subseteq V_1, R_2 \subseteq V_2$ be two root systems. We may regard V_1, V_2 as being embedded in the direct sum inner product space $V := V_1 \oplus V_2$ by the maps $v_1 \mapsto (v_1, 0)$, $v_2 \mapsto (0, v_2)$. We may then define their *disjoint union* $R_1 \cup R_2 \subset V$ to be the set of images of R_1, R_2 under such maps. It is easily seen to define a root system.

Definition 230. A root system R is *irreducible* if it is not isomorphic to a disjoint union of nonempty root systems.

[8]A morphism is a map that preserves all relevant structure. An isomorphism is a morphism with a two-sided inverse morphism.

30.4. **Illustration of the root system axioms.** The root system axioms have a number of consequences; we illustrate a few of them here, referring to the second reference by Serre on the course homepage for details and further discussion.

As illustration, let us first verify that the axiom (129) can be weakened (assuming the other axioms) to

$$\alpha \in R \implies 2\alpha \notin R. \tag{130}$$

To that end, suppose $\alpha, c\alpha \in R$ for some nonzero scalar $c \in \mathbb{R}$. Then the quantities

$$\langle c\alpha|\alpha\rangle = 2c, \quad \langle\alpha|c\alpha\rangle = 2c^{-1}$$

are both integers, hence

$$c \in \{\pm 1/2, \pm 1, \pm 2\}. \tag{131}$$

It is clear that (130) and (131) imply (129).

As our next illustration:

Lemma 231. *Let α, β be non-proportional roots (i.e., elements of the given root system R that are not multiples of one another). Then the unordered pair of integers $\{\langle\alpha|\beta\rangle, \langle\beta|\alpha\rangle\}$ is of the form $\{0,0\}$ or $\{\varepsilon, \varepsilon n\}$ for some $\varepsilon \in \{\pm 1\}$ and $n \in \{1,2,3\}$; in other words, it belongs to the following list:*

- $\{0,0\}$
- $\{1,1\}$
- $\{1,2\}$
- $\{1,3\}$
- $\{-1,-1\}$
- $\{-1,-2\}$
- $\{-1,-3\}$

Proof. By elementary geometry, we have $\langle\alpha|\beta\rangle\langle\beta|\alpha\rangle = 4\cos^2(\phi)$, where $\pm\phi$ denotes the angle between the vectors α, β. Since α, β are non-proportional, we have $\cos^2(\phi) < 1$. If $\cos(\phi) = 0$, then α, β are orthogonal and so both quantities are zero. Otherwise $\langle\alpha|\beta\rangle$ and $\langle\beta|\alpha\rangle$ are integers whose product belongs to $\{1,2,3\}$, for which the only possibilities are those listed. □

Lemma 232. *Suppose that α, β are non-proportional roots.*

- *If $(\alpha, \beta) < 0$, then $\alpha + \beta$ is a root.*
- *If $(\alpha, \beta) > 0$, then $\alpha - \beta$ is a root.*

Proof. In the first case, we see from Lemma 231 that after possibly swapping α and β, we have $\langle\beta|\alpha\rangle = -1$ (and $\langle\alpha|\beta\rangle = -n$ for some $n \in \{1,2,3\}$). Then $s_\alpha(\beta) = \beta - \langle\beta|\alpha\rangle\,\alpha = \beta + \alpha$ is a root thanks to the axiom $s_\alpha(R) = R$. A similar argument applies in the second case. □

We may define a *base* (or *simple system*) exactly as before to be a subset S of R that is a basis for the underlying inner product space V with the property that for each $\alpha \in R$, the coefficients c_β in the expansion $\alpha = \sum_{\beta \in S} c_\beta \beta$ either all belong to $\mathbb{Z}_{\geq 0}$ or all belong to $\mathbb{Z}_{\leq 0}$. One can show directly from the root system axioms that bases exist and have the properties established previously for the classical families and on the homeworks; see the second reference by Serre on the course webpage for more details. We note for now just that the observation from the homework that $(\alpha, \beta) \leq 0$ for $\alpha, \beta \in S$ follows from Lemma 232: if otherwise $(\alpha, \beta) > 0$, then $\alpha - \beta$ would be a root, contrary to the defining property of the simple system S.

One can likewise define the Weyl group W of a root system to be the subgroup of the orthogonal group of V generated by the root reflections s_α; the properties we established previously for the root systems arising from classical families can also be established directly from the root system axioms (see Serre for details).

Finally, one can attach to each reduced root system a Cartan matrix and a Dynkin diagram; the diagram turns out to be connected if and only if the root system is irreducible, and one can show (by elaborate application of the root system axioms) that all irreducible reduced root systems belong either to the classical families A_n, B_n, C_n, D_n or belong to an exceptional set $\{G_2, F_2, E_6, E_7, E_8\}$.

30.5. **Simple Lie algebras give rise to root systems.** The "unexplained observations" recorded in §30.1 are all contained in the following result, which is our next target:

Theorem 233. *Let $\mathfrak{g}$ be a simple Lie algebra (that satisfies the Cartan subalgebra theorem). Let $\mathfrak{h}$ be a Cartan subalgebra. Let $R \subseteq \mathfrak{h}_{\mathbb{R}}^*$ be the set of roots for* $\mathrm{ad} : \mathfrak{h} \to \mathrm{End}(\mathfrak{g})$. *Then R is a reduced root system. Moreover:*

(1) For each $\alpha \in R$, one has $\dim \mathfrak{g}^\alpha = 1$.
(2) For each $\alpha \in R$ there is a unique $H_\alpha \in \mathfrak{h}_{\mathbb{R}}$ with $\alpha(H_\alpha) = 2$ so that for each nonzero $X_\alpha \in \mathfrak{g}^\alpha$ there is a unique $Y_\alpha \in \mathfrak{g}^{-\alpha}$ so that $H_\alpha = [X_\alpha, Y_\alpha]$.
(3) One has

$$[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \begin{cases} \mathfrak{g}^{\alpha+\beta} & \alpha + \beta \in R \\ \mathbb{C}H_\alpha & \alpha + \beta = 0 \\ 0 & \textit{otherwise.} \end{cases}$$

Remark 234. We will discuss the proof of Theorem 233 in detail for the following reasons:

(1) Although the proof gives us nothing "new" for the classical families (we have already checked all of the conclusions by hand), it tells us *why* they are true, gives a less computational explanation, etc.
(2) The techniques involved in the proof of Theorem 233 are of general use.
(3) The proof of Theorem 233 will give us the opportunity to apply some properties of representations of $\mathfrak{sl}_2(\mathbb{C})$ that we established earlier in the course.

Remark 235. One can also establish the following complements to Theorem 233

(1) The root systems arising from simple Lie algebras are irreducible.
(2) Two simple Lie algebras are isomorphic if and only if their associated root systems are isomorphic.
(3) Every irreducible reduced root system arises from a (unique) simple Lie algebra.

We might discuss the first couple of these if we have time; see the second reference by Serre on the course homepage for further discussion of the final point.

30.6. **Some stuff about scalar products and inner products.** Let $\mathfrak{g}$ be a simple Lie algebra (that satisfies the Cartan subalgebra theorem). Let $\mathfrak{h}$ be a Cartan subalgebra, and let $(,)$ be an invariant scalar product on $\mathfrak{g}$ that is real-valued and positive-definite on $\mathfrak{h}_{\mathbb{R}}$. Since $(,)$ is nondegenerate, it induces a linear isomorphism $\mathfrak{g} \to \mathfrak{g}^*$ given by $x \mapsto (x, \cdot)$. We can thus transfer $(,)$ to an scalar product (also denoted $(,)$) on $\mathfrak{g}^*$ by requiring that $((x,\cdot),(y,\cdot)) = (x,y)$ for all $x, y \in \mathfrak{g}$. We may restrict the scalar product $(,)$ on $\mathfrak{g}^*$ to

$$\mathfrak{h}_{\mathbb{R}}^* := \mathrm{Hom}_{\mathbb{R}}(\mathfrak{h}_{\mathbb{R}}, \mathbb{R}) \cong \{\lambda \in \mathfrak{h}^* : \lambda(h_{\mathbb{R}}) \subseteq \mathbb{R}\}.$$

Since the scalar product that we started with on $\mathfrak{g}$ has positive-definite restriction to $\mathfrak{h}_{\mathbb{R}}$, we know also that the scalar product on $\mathfrak{g}^*$ that we just defined has positive-definite restriction to $\mathfrak{h}_{\mathbb{R}}^*$, hence defines an inner product on that space.

In what follows, we shall always regard $\mathfrak{h}_{\mathbb{R}}^*$ as an inner product space with respect to an inner product as constructed above.

30.7. **Some recap on** SL(2)**.** Let's recall a few facts we learned long ago. Recall the standard basis elements

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

of $\mathfrak{g} := \mathfrak{sl}_2(\mathbb{C})$, and that any finite-dimensional representation V of $\mathfrak{g}$ breaks up into weight spaces $V = \oplus_{m \in \mathbb{C}} V[m]$ where $V[m] := \{v \in V : Hv = mv\}$. (In other words, H acts semisimply in any finite-dimensional representation. We proved this awhile ago. Another quick proof: we can first decompose with respect to the action of the diagonal subgroup of SU(2), since the latter is compact; the Lie algebra of that subgroup is generated by H, so we get a decomposition with respect to H.)

Recall that the m for which $V[m] \neq 0$ are called the *weights* of V; the spaces $V[m]$ are then called *weight spaces*.

The set of weights of the irreducible representations W_m of dimension $m+1$ is

$$\{-m, -m+2, \ldots, m-4, m-2, m\}.$$

The set of weights of a direct sum of several copies of W_m is the union of the sets of weights of the W_m that occur. Since any such V is isomorphic to such a finite direct sum, we know a lot about the set of weights.

Lemma 236. *The weights are integers. An integer m is a weight if and only if $-m$ is a weight.*

Proof. This follows (among other ways) from the classification: V is a finite direct sum of copies of the W_m, and each of those irreducible representations has the above property. $\square$

Lemma 237. *If the weights of V all have the same parity (i.e., are all even or all odd), then the set of weights is of the form*

$$\{-m, -m+2, \ldots, m-4, m-2, m\}$$

for some $m \in \mathbb{Z}_{\geq 0}$.

In that case, let ℓ be any weight of V. Let $p, q \geq 0$ be the largest positive integers for which $\ell - 2p$ and $\ell + 2q$ are weights. Then

$$\ell = p - q. \tag{132}$$

Proof. The first assertion follows from the classification. For the second assertion, we must have $\ell - 2p = -m$ and $\ell + 2q = m$, whence $2\ell = 2(p-q)$, as required. $\square$

Lemma 238. *If $m, m+2$ are weights of V, then the maps $X : V[m] \to V[m+2]$ and $Y : V[m+2] \to V[m]$ are not identically zero.*

Proof. Again follows by reducing to the irreducible case and then inspecting. $\square$

Lemma 239. *Suppose V has the following properties:*

1. *The weights of V are all even.*
2. *$V[0]$ is one-dimensional.*
3. *$V[2]$ is nonzero.*
4. *There exists a nonzero element $v \in V[2]$ such that $Xv = 0$.*

Then the set of weights is $\{-2, 0, 2\}$ and each weight space is one-dimensional.

Proof. The first three conditions tell us that we must have $V \cong W_m$ for some even integer $m \geq 2$. The fourth condition implies that $m = 2$. □

30.8. **Proof of Theorem 233.** Let notation and assumptions be as in the statement of that theorem. Let $(,)$ denote a $\mathfrak{g}$-invariant scalar product on $\mathfrak{g}$ whose restriction to $\mathfrak{h}_{\mathbb{R}}$ is real-valued and positive-definite. (We will only use this final property of the inner product at one point in the proof to follow, and it could be avoided at the cost of a bit more work.) As in §30.6, let $(,)$ denote also the inner product induced on $\mathfrak{g}^*$ by duality, whose restriction to $\mathfrak{h}_{\mathbb{R}}^*$ is then real-valued and positive-definite. (Revisit the examples of classical families.) Since $\mathfrak{h}$ is a Cartan subalgebra, we have a root space decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{\alpha \in R} \mathfrak{g}^{\alpha})$$

for some finite set $R \subseteq \mathfrak{h}_{\mathbb{R}}^* - \{0\}$ of roots. Here $[H, X] = \alpha(H)X$ for all $H \in \mathfrak{h}, X \in \mathfrak{g}^{\alpha}$.

We verify first that

$$[\mathfrak{g}^{\alpha}, \mathfrak{g}^{\beta}] \begin{cases} \subseteq \mathfrak{g}^{\alpha+\beta} & \text{if } \alpha + \beta \in R \\ \subseteq \mathfrak{h} & \text{if } \alpha + \beta = 0 \\ = \{0\} & \text{otherwise.} \end{cases} \tag{133}$$

This is the same verification we've done by now many times: for $x \in \mathfrak{g}^{\alpha}, y \in \mathfrak{g}^{\beta}, H \in \mathfrak{h}$, we have by the Jacobi identity

$$[H, [x, y]] = [[H, x], y] + [x, [H, y]] = \alpha(H)[x, y] + \beta(H)[x, y] = (\alpha + \beta)(H)[x, y],$$

giving what we want.

We show next that

$$(\mathfrak{g}^{\alpha}, \mathfrak{g}^{\beta}) = 0 \text{ unless } \alpha + \beta = 0. \tag{134}$$

Indeed, assume $\alpha + \beta \neq 0$. We must show for $x \in \mathfrak{g}^{\alpha}, y \in \mathfrak{g}^{\beta}$ that $(x, y) = 0$. Choose $H \in \mathfrak{h}$ so that $(\alpha + \beta)(z) \neq 0$. By the $\mathfrak{g}$-invariance of $(,)$, we then have

$$0 = ([H, x], y) + (x, [H, y]) = (\alpha + \beta)(H) \cdot (x, y),$$

hence $(x, y) = 0$, as required.

As a consequence, we see that the decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{\pm\alpha \in R} \mathfrak{g}^{\alpha} \oplus \mathfrak{g}^{-\alpha}) \tag{135}$$

is orthogonal with respect to $(,)$. In particular:

1. $(,)$ has non-degenerate restriction to $\mathfrak{h} \times \mathfrak{h}$, although this follows already from our assumptions.
2. $(,)$ induces a duality between $\mathfrak{g}^{\alpha}$ and $\mathfrak{g}^{-\alpha}$. In particular, $\alpha \in R$ if and only if $-\alpha \in R$.

Let $\mathfrak{h}^* \ni \lambda \mapsto u_\lambda \in \mathfrak{h}$ denote the isomorphism induced by $(,)$, so that $\mu(u_\lambda) = \langle \lambda, \mu \rangle$ for all $\mu \in \mathfrak{h}^*$. We claim next that

$$[x, y] = (x, y) u_\alpha. \tag{136}$$

For the proof, let $H \in \mathfrak{h}$; since $(,)$ has nondegenerate restriction to $\mathfrak{h}$, it will suffice to verify that

$$(H, [x, y]) = (H, (x, y) u_\alpha).$$

Since the $\mathfrak{g}$-invariance gives $(H, [x, y]) = ([H, x], y) = \alpha(H)(x, y)$ and the linearity gives $(H, (x, y) u_\alpha) = (H, u_\alpha)(x, y) = \alpha(H)(x, y)$, we are done.

Note in particular that for $\alpha \in \mathfrak{h}_{\mathbb{R}}^* - \{0\}$, one has

$$0 < (\alpha, \alpha) = \alpha(u_\alpha).$$

Hence for each $\alpha \in R \subseteq \mathfrak{h}_{\mathbb{R}}^* - \{0\}$, it makes sense to define

$$H_\alpha := 2 \frac{u_\alpha}{(\alpha, \alpha)} \in \mathfrak{h}_{\mathbb{R}}.$$

With this definition, we then have

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

Moreover, let us fix a nonzero element $X_\alpha \in \mathfrak{g}^\alpha$. Since $\mathfrak{g}^{-\alpha}$ is in duality with $\mathfrak{g}^\alpha$, there then exists $Y_\alpha \in \mathfrak{g}^{-\alpha}$ so that $(X_\alpha, Y_\alpha) = 2/(\alpha, \alpha)$. By (136), it follows that

$$[X_\alpha, Y_\alpha] = H_\alpha,$$

which is consistent with what we stipulated in the classical examples.

Now let $\mathfrak{s}_\alpha \leq \mathfrak{g}$ denote the three-dimensional vector subspace

$$\mathfrak{s}_\alpha := \mathbb{C} X_\alpha \oplus \mathbb{C} H_\alpha \oplus \mathbb{C} Y_\alpha.$$

Recall that $[X_\alpha, X_\alpha] = [Y_\alpha, Y_\alpha] = [H_\alpha, H_\alpha] = 0$ and $[H_\alpha, X_\alpha] = 2X_\alpha$, $[H_\alpha, Y_\alpha] = -2Y_\alpha$, $[X_\alpha, Y_\alpha] = H_\alpha$. It follows that the map

$$\phi_\alpha : \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{s}_\alpha \subseteq \mathfrak{g}$$

given by sending the standard basis elements

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

in the evident way $(H, X, Y \mapsto H_\alpha, X_\alpha, Y_\alpha)$ is an isomorphism of Lie algebras.

Lemma 240. *Let* $\alpha \in R$. *Then* $\dim \mathfrak{g}^\alpha = 1$ *and* $\{n \in \mathbb{Z} : n\alpha \in R\} = \{\pm 1\}$. *In particular,* $2\alpha \notin R$.

Proof. Set

$$V := \mathbb{C} H_\alpha \oplus (\oplus_{n \in \mathbb{Z}_{\neq 0}} \mathfrak{g}^{n\alpha}),$$

where by convention $\mathfrak{g}^{n\alpha} := 0$ if $n\alpha \notin R$. Observe that V is stable under $\mathfrak{s}_\alpha$; this follows from what was shown above, together with the observation that $H_{n\alpha} \in \mathbb{C} H_\alpha$ for all n. We may thus regard V as a representation of $\mathfrak{sl}_2(\mathbb{C})$ via the map ϕ_α. Equivalently, the map $\rho : \mathfrak{sl}_2(\mathbb{C}) \to \operatorname{End}(V)$ is given for $x \in \mathfrak{sl}_2(\mathbb{C})$ and $v \in V$ by

$$\rho(x) v := [\phi_\alpha(x), v].$$

The possibly nonzero weight spaces are $V[0] = \mathbb{C} H_\alpha$, which is one-dimensional, and $V[2n] = \mathfrak{g}^{n\alpha}$ for $n \neq 0$. We observe that $V[2]$ is nonzero (since it contains X_α) and that there exists $v \in V[2]$ for which $\rho(X) v = 0$ (take $v := X_\alpha$ and use that $[X_\alpha, X_\alpha] = 0$). Lemma 239 applies, telling us that $V \cong W_2$. The various conclusions follow from the description of the weight spaces of W_2. $\square$

For $\alpha \in R$ and $\lambda \in \mathfrak{h}^*$, set
$$\langle \lambda | \alpha \rangle := \lambda(H_\alpha)$$
and define the root reflection
$$s_\alpha : \mathfrak{h}^* \to \mathfrak{h}^*$$
by
$$s_\alpha(\lambda) := \lambda - \langle \lambda | \alpha \rangle \alpha.$$
Observe that $s_\alpha(s_\alpha(\lambda)) = \lambda$.

Lemma 241.
$$\langle \lambda | \alpha \rangle = 2 \frac{(\lambda, \alpha)}{(\alpha, \alpha)}.$$

Proof. Well, by definition, we have
$$u_\alpha = \frac{2}{(\alpha, \alpha)} H_\alpha.$$
If we apply λ to both sides, we get
$$(\lambda, \alpha) = \lambda(u_\alpha) = 2 \frac{\lambda(H_\alpha)}{(\alpha, \alpha)},$$
which rearranges to
$$\lambda(H_\alpha) = 2 \frac{(\lambda, \alpha)}{(\alpha, \alpha)},$$
as required. $\square$

Lemma 242. *Let $\alpha, \beta \in R$. Then $\beta(H_\alpha) \in \mathbb{Z}$ and $s_\alpha(\beta) \in R$. Thus $s_\alpha(R) = R$.*

Moreover, let $p, q \geq 0$ be the largest nonnegative integers for which $\beta - q\alpha$ and $\beta + p\alpha$ are roots. Then $\beta + k\alpha$ is a root for all integers $k \in \{-q..p\}$, and we have $\beta(H_\alpha) = p - q$.

Proof. We argue as in the proof of Lemma 240, but now with
$$V := \oplus_{k \in \mathbb{Z}} \mathfrak{g}^{\beta + k\alpha},$$
regarded as an $\mathfrak{sl}_2(\mathbb{C})$-module via ϕ_α as before. The element $H \in \mathfrak{sl}_2(\mathbb{C})$ acts on $\mathfrak{g}^{\beta+k\alpha}$ by the eigenvalues $(\beta + k\alpha)(H_\alpha) = \beta(H_\alpha) + 2k$; in particular, $\beta(H_\alpha)$ is an H-weight of V. By Lemma 237 (applied to $\ell := \beta(H_\alpha)$), we have $\beta(H_\alpha) = p - q \in \mathbb{Z}$. The H-weight of $\mathfrak{g}^{\beta - \beta(H_\alpha)\alpha}$ is $\beta(H_\alpha) - 2\beta(H_\alpha) = -\beta(H_\alpha)$, which shows that $\mathfrak{g}^{\beta - \beta(H_\alpha)\alpha} \neq 0$, or equivalently, that $s_\alpha(\beta) \in R$. $\square$

Lemma 243. *Let $\alpha, \beta \in R$ such that $\alpha + \beta \in R$. Then $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$.*

Proof. Since the root spaces are all one-dimensional, it suffices to show that the map $\mathrm{ad}_{X_\alpha} : \mathfrak{g}^\beta \to \mathfrak{g}^{\alpha+\beta}$ is nonzero. This follows from Lemma 238 upon taking $V = \oplus \mathfrak{g}^{\beta + k\alpha}$ as above. $\square$

All assertions in Theorem 233 have now been established. (The uniqueness of Y_α follows from the one-dimensionality of $\mathfrak{g}^{-\alpha}$.)

Remark 244. Lemma 232 was proved using only the root system axioms. It may be alternatively deduced "directly" from Lemma 242.

31. Serre relations and applications

31.1. **Generators and relations for simple complex Lie algebras.** Recall from §30.5 and following that one can[9] attach root systems to simple Lie algebras over $\mathbb{C}$. We also mentioned briefly that root systems can be classified in terms of their Cartan matrices, or equivalently, their Dynkin diagrams.

Conversely, it turns out that one go the other direction: if the root systems of a pair $\mathfrak{g}_1, \mathfrak{g}_2$ of simple Lie algebras over $\mathbb{C}$ are isomorphic, then so are $\mathfrak{g}_1$ and $\mathfrak{g}_2$. One can prove this fairly directly (see p184 of Onishchik–Vinberg), but a particularly convincing way to see it is via the following theorem of Serre:

Theorem 245. *Let $\mathfrak{h}$ be a Cartan subalgebra with associated root system R, let S be any simple subsystem, and let $N = (N_{\alpha\beta})_{\alpha,\beta \in S}$ be the Cartan matrix (thus $N_{\alpha\beta} := \alpha(H_\beta)$, say). Then $\mathfrak{g}$ is generated as a Lie algebra by the symbols $H_\alpha, X_\alpha, Y_\alpha$ $(\alpha \in S)$ subject only to the relations: for all $\alpha, \beta \in S$,*

$$[H_\alpha, X_\beta] = N_{\alpha\beta} X_\beta, \quad [H_\alpha, Y_\beta] = N_{\alpha\beta} Y_\beta, \quad [X_\alpha, Y_\alpha] = H_\alpha;$$

for distinct $\alpha, \beta \in S$,

$$[X_\alpha, Y_\beta] = 0, \quad \mathrm{ad}_{X_\alpha}^{-N_{\alpha\beta}+1}(X_\beta) = 0, \quad \mathrm{ad}_{Y_\alpha}^{-N_{\alpha\beta}+1}(Y_\beta) = 0.$$

The meaning of this is hopefully clear by analogy to presentations and relations for groups; for a more precise statement, see either Onishchik–Vinberg or Serre.

Theorem 245 immediately implies that the isomorphism class of a simple Lie algebra over $\mathbb{C}$ depends only upon its Cartan matrix. It has many other applications to be discussed shortly.

31.2. **Semisimple complex Lie algebras.** These play a central role in the theory. They admit several equivalent definitions. The most convenient one for our immediate purposes is the following:

Definition 246. A Lie algebra $\mathfrak{g}$ over $\mathbb{C}$ is *semisimple* if it is isomorphic to a finite direct sum of simple Lie algebras, or equivalently, if $\mathfrak{g}$ is the direct sum of some finite collection of simple ideals.

We can define Cartan subalgebras $\mathfrak{h}$ of semisimple Lie algebras $\mathfrak{g}$ over $\mathbb{C}$ just as we did in the simple case. Moreover, if $\mathfrak{g} = \oplus \mathfrak{g}_i$ with $\mathfrak{g}_i$ simple and containing a Cartan subalgebra $\mathfrak{h}_i$, then we can take $\mathfrak{h} = \oplus \mathfrak{h}_i$. We can likewise associate root systems in the semisimple case just as in the simple case. The only difference is that now the root systems we obtain are not irreducible; instead, they decompose as finite disjoint unions (in the sense of Example 229) of irreducible root systems, corresponding to the decomposition of the semisimple Lie algebra as a finite direct sum of simple Lie algebras.

The bijection between:

- simple Lie algebras over $\mathbb{C}$
- irreducible reduced root systems
- connected Dynkin diagrams

induces one between:

- semisimple Lie algebras over $\mathbb{C}$
- reduced root systems

[9] We have only proved this for those that satisfy the Cartan subalgebra theorem.

- Dynkin diagrams

The Serre relations apply just as well in the semisimple case.

There is a nontrivial equivalence which seems worth mentioning up front:

Theorem 247. $\mathfrak{g}$ *is semisimple if and only if* $\mathfrak{g}$ *contains no abelian ideals.*

Thus being semisimple is in some sense the "opposite" of being abelian. There are other nice criteria for checking semisimplicity of $\mathfrak{g}$ that one can read about in any of the course references (e.g., either by Serre). One says that there should exist a bilinear form on $\mathfrak{g}$ that is $\mathrm{ad}_{\mathfrak{g}}$-invariant and non-degenerate. For the classical examples, the trace form $(x, y) := trace(xy)$ is easily seen to have such properties. Using this criterion, one can give "another" proof that (say) $\mathfrak{sl}_n(\mathbb{C})$ is simple (see §28) by computing the Cartan matrix and checking that the Dynkin diagram is *connected*.

31.3. **Reductive complex Lie algebras.** Here is another definition which admits many equivalent characterizations; we again give that which is more convenient for our immediate purposes.

Definition 248. Let $(\mathfrak{g}_i)_{i\in I}$ be a family of Lie algebras. The *direct sum Lie algebra* is the direct sum vector space $\mathfrak{g} := \oplus_{i\in I}\mathfrak{g}_i$ equipped with the Lie bracket characterized by:

- for $i \in I$ and $x, y \in \mathfrak{g}_i \hookrightarrow \mathfrak{g}$, one has $[x, y]_{\mathfrak{g}} := [x, y]_{\mathfrak{g}_i}$;
- for $i \neq j \in I$ and $x \in \mathfrak{g}_i, y \in \mathfrak{g}_j$, one has $[x, y]_{\mathfrak{g}} := 0$.

It has the universal property: $\mathrm{Hom}(\oplus\mathfrak{g}_i, \mathfrak{h}) = \prod \mathrm{Hom}(\mathfrak{g}_i, \mathfrak{h})$ for all Lie algebras $\mathfrak{h}$.

Definition 249. A finite-dimensional Lie algebra $\mathfrak{g}$ over $\mathbb{C}$ is *reductive* if it is a direct sum of an abelian Lie algebra and a semisimple Lie algebra.

For example, $\mathfrak{gl}_n(\mathbb{C})$ is reductive (it is the direct sum of ${}_n(\mathbb{C})$ and the central subalgebra $\mathfrak{z}$ consisting of scalar matrices), but not semisimple (it contains the abelian ideal $\mathfrak{z}$).

Abelian Lie algebras over any field are classified by their dimension, so the classification of semisimple Lie algebras readily induces a classification of reductive Lie algebras.

Here's a handy and clarifying lemma, proved in lecture and left here as an exercise (ask me if it's unclear):

Lemma 250. *A complex Lie algebra* $\mathfrak{g}$ *is reductive if and only if its adjoint representation* $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ *is completely reducible.*

We can define Cartan subalgebras $\mathfrak{h}$ of reductive Lie algebras $\mathfrak{g}$ just as we did in the semisimple case. (They always contain the center $\mathfrak{z}$.) We can also define the set R of roots. The only difference with the semisimple case is now that the roots R need not span $\mathfrak{h}^*$. For example, if $\mathfrak{g}$ is abelian, then $R = \emptyset$.

31.4. **Compact complex Lie groups.** We don't talk about these much. There's a good reason:

Theorem 251. *Any compact connected complex Lie group G is abelian.*

Proof. Consider $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$. It is a holomorphic matrix-valued function. Since G is compact, it is bounded. By Liouville's theorem, it must be constant. But it preserves the identity, and so must be trivial. Since G is connected, we conclude that it must be abelian. □

Such G typically go instead by the name "abelian variety" and have an interesting theory orthogonal to the primary aims of this course.

31.5. **Compact real Lie algebras.** From now on, when I write "compact Lie group," I mean "compact real Lie group."

Think of your favorite compact Lie group K (e.g., $K = \mathrm{U}(n)$). Consider its Lie algebra $\mathfrak{k}$. How would you go about telling just from $\mathfrak{k}$ that K was compact?

Definition 252. Let $\mathfrak{k}$ be a real Lie algebra. (Every Lie algebra here and for the rest of the course should be assumed finite-dimensional.) We call $\mathfrak{k}$ *compact* if it admits an $\mathrm{ad}(\mathfrak{k})$-invariant inner product, that is to say, a positive definite symmetric bilinear form $(,) : \mathfrak{k} \otimes \mathfrak{k} \to \mathbb{R}$ with the property that

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

for all $x, y, z \in \mathfrak{k}$.

Lemma 253. *Let K be a compact Lie group. Then $\mathfrak{k}$ is compact.*

Proof. Start with any inner product $(,)_0$ on $\mathfrak{k}$. Average it under $\mathrm{Ad}(K)$ with respect to an invariant measure, as in the discussion of the unitary trick earlier in the course. Call $(,)$ the averaged inner product so obtained (it is still an inner product). Then $(,)$ is $\mathrm{Ad}(K)$-invariant, by construction. By differentiating, we see that it is $\mathrm{ad}(\mathfrak{k})$-invariant, as required. $\square$

Example 254. If $\mathfrak{k}$ is an abelian real Lie algebra, then it is compact. This is easy to see directly. Alternatively, we can write $\mathfrak{k} = \mathbb{R}^n$ and apply the Lemma to $K = (\mathbb{R}/\mathbb{Z})^n$.

31.6. **Complex reductive vs. compact real Lie algebras.** Recall our discussion of complexifications and real forms from §17.

We include the following mainly as a bridge from our discussion of complex simple Lie algebras to our next target (compact Lie groups).

Theorem 255.
1. *Let $\mathfrak{k}$ be a compact real Lie algebra. Then its complexification $\mathfrak{k}_{\mathbb{C}} := \mathfrak{k} \otimes \mathbb{C}$ is a reductive complex Lie algebra.*
2. *Let $\mathfrak{g}$ be a complex reductive Lie algebra. Then it has a compact real form $\mathfrak{k}$.*

Proof.
1. By Lemma 250, we have to show that the adjoint representation $\mathrm{ad} : \mathfrak{k}_{\mathbb{C}} \to \mathrm{End}(\mathfrak{k}_{\mathbb{C}})$ is completely reducible. By linear algebra (cf. Example 126), it suffices to show that $\mathrm{ad} : \mathfrak{k} \to \mathrm{End}(\mathfrak{k})$ is completely reducible. To that end, we make use of the existence of a $\mathfrak{k}$-invariant inner product and argue using orthogonal complements as in Example 111.
2. Note first that if result holds for $\mathfrak{g}_1$ and $\mathfrak{g}_2$, then it also holds for their direct sum (take $\mathfrak{k}_1 \oplus \mathfrak{k}_2 \subseteq \mathfrak{g}_1 \oplus \mathfrak{g}_2$). Since $\mathfrak{g}$ is reductive, it suffices to consider separately the case that $\mathfrak{g}$ is abelian and the case that $\mathfrak{g}$ is semisimple (or indeed, simple). The abelian case is easy (see Example 254), so we focus henceforth on the semisimple case.

   Think of the prototypical example $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$. How would one go about "discovering" the compact real form $\mathfrak{k} = \mathfrak{su}(2)$? Well, we have

   $$\mathfrak{k} = \{Z \in \mathfrak{g} : \sigma(Z) = Z\},$$

where $\sigma(Z) := -\overline{Z}^t$. On the standard basis elements X, Y, H this involution is given by

$$\sigma(X) = -Y, \quad \sigma(Y) = -X, \quad \sigma(H) = -H.$$

It is anti-linear in the sense that

$$\sigma(tZ) = \bar{t}\sigma(Z) \text{ for all } t \in \mathbb{C}, Z \in \mathfrak{g}. \tag{137}$$

Also, let's note in this case for the modified Killing form

$$(x, y) := -\operatorname{trace}(\operatorname{ad}(x)\operatorname{ad}(\sigma(y))) \tag{138}$$

on $\mathfrak{g}$, the basis X, Y, H is orthogormal (i.e., $(X, Y) = (X, H) = (Y, H) = 0$), and also $(X, X) = (Y, Y) = (H, H) = 2$. Thus $(,)$ is positive-definite; it is also clearly $\mathfrak{k}$-invariant. (Note: I defined $(,)$ incorrectly in lecture.)

This suggests the general strategy. Let $\mathfrak{g}$ be a semisimple Lie algebra. We use the Serre relations. It thus has generators $X_\alpha, Y_\alpha, H_\alpha$ $(\alpha \in S)$ satisfying some explicit relations. We try to define an anti-linear involution σ on $\mathfrak{g}$ by requiring that (137) hold and that on the generators, one has

$$\sigma(X_\alpha) = -Y_\alpha, \quad \sigma(Y_\alpha) = -X_\alpha, \quad \sigma(H_\alpha) = -H_\alpha.$$

To check that this definition makes sense (i.e., extends from the generators to a real Lie algebra automorphism), we just need to check that it preserves the Serre relations, which is clear. By Remark 132, we know that $\mathfrak{k} := \{X \in \mathfrak{g} : \sigma(X) = X\}$ is a real form. We now define $(,)$ on $\mathfrak{g}$ by (138) and check that it is positive definite on $\mathfrak{k}$ to see that $\mathfrak{k}$ is compact.

□

32. The center and fundamental group of a compact Lie group

We spent most of the lecture stating and motivating the truth of one theorem.

We record the key definitions here for now; we will have more to say about them later.

Let K be a compact Lie group. Let $\mathfrak{k} := \operatorname{Lie}(K)$; it is a compact real Lie algebra. Let $\mathfrak{g}$ denote its complexification; it is a reductive complex Lie algebra. Let $\mathfrak{h} \leq \mathfrak{g}$ be a Cartan subalgebra, and suppose that $\mathfrak{t} = \mathfrak{k} \cap \mathfrak{t}$ is a real form of $\mathfrak{h}$, thus $\dim_{\mathbb{R}}(\mathfrak{t}) = \dim_{\mathbb{C}}(\mathfrak{h})$; we can arrange this using the presentation given by the Serre generators, for instance. Define $\mathfrak{h}_{\mathbb{R}} := i\mathfrak{t}$ and $\mathfrak{h}_{\mathbb{Z}} := \ker(e)$, where $e : \mathfrak{h}_{\mathbb{R}} \to K$ is the map $e(x) := \exp(2\pi i x)$. Let R be the root system of $\mathfrak{h}$. Then $R \subseteq \mathfrak{h}_{\mathbb{R}}^*$. Set $R^\wedge := \{\alpha^\wedge : \alpha \in R\}$, where $\alpha^\wedge = H_\alpha$. Set

$$\mathfrak{h}_{\mathbb{R}}^* := \operatorname{Hom}_{\mathbb{R}}(\mathfrak{h}_{\mathbb{R}}, \mathbb{R}) \cong \{\lambda \in \mathfrak{h}^* : \lambda(\mathfrak{h}_{\mathbb{R}}) \subseteq \mathbb{R}\}$$

and

$$\mathfrak{h}_{\mathbb{Z}}^* := \operatorname{Hom}_{\mathbb{Z}}(\mathfrak{h}_{\mathbb{Z}}, \mathbb{Z}) \cong \{\lambda \in \mathfrak{h}_{\mathbb{R}}^* : \lambda(\mathfrak{h}_{\mathbb{Z}}) \subseteq \mathbb{Z}\} \cong \{\lambda \in \mathfrak{h}^* : \lambda(\mathfrak{h}_{\mathbb{Z}}) \subseteq \mathbb{Z}\}.$$

We then have

$$\mathbb{Z}R \subseteq \mathfrak{h}_{\mathbb{Z}}^* \subseteq (\mathbb{Z}R^\wedge)^* \tag{139}$$

(called respectively the *root lattice*, the *integers*, and the *weight lattice*) and

$$\mathbb{Z}R^\wedge \subseteq \mathfrak{h}_{\mathbb{Z}} \subseteq (\mathbb{Z}R)^*, \tag{140}$$

(called respectively the *coroot lattice*, the *integers*, and the *coweight lattice*), where $\mathbb{Z}R$ denotes the $\mathbb{Z}$-span of R, $\mathbb{Z}R^\wedge$ denotes the $\mathbb{Z}$-span of $R^\wedge$,

$$(\mathbb{Z}R^\wedge)^* := \{\lambda \in \mathfrak{h}_{\mathbb{R}}^* : \lambda(R^\wedge) \subseteq \mathbb{Z}\}$$

and

$$(\mathbb{Z}R)^* := \{H \in \mathfrak{h}_\mathbb{R} : R(H) \subseteq \mathbb{Z}\}.$$

Pontryagin duality for finite abelian groups give us non-canonical isomorphisms

$$\mathfrak{h}_\mathbb{Z}/\mathbb{Z}R^\wedge \cong (\mathbb{Z}R)^*/\mathfrak{h}_\mathbb{Z}^*, \quad (\mathbb{Z}R)^*/\mathfrak{h}_\mathbb{Z} \cong \mathfrak{h}_\mathbb{Z}^*/\mathbb{Z}R. \tag{141}$$

Theorem 256. *The induced map $e : (\mathbb{Z}R)^*/\mathfrak{h}_\mathbb{Z} \to \mathrm{Center}(K)$ given by $e(x) := \exp(2\pi i x)$ is a well-defined isomorphism.*

The map $f : \mathfrak{h}_\mathbb{Z}/\mathbb{Z}R^\wedge \to \pi_1(K)$, sending H to the homotopy class $[\gamma]$ of the path γ given by $\gamma(t) := e(tH)$, is a well-defined isomorphism.

We then explained in detail how this "recovers" the fact that $\pi_1(\mathrm{SU}(n)) = \{1\}$, $\mathrm{Center}(\mathrm{SU}(n)) \cong \mathbb{Z}/n$.

The theorem will take a bit of preparation to prove; we start in the next section.

33. Tori in compact Lie groups

33.1. Basic definitions.

Definition 257. A *torus* is a Lie group isomorphic to $T^k := (\mathbb{R}/\mathbb{Z})^k$ for some $k \in \mathbb{Z}_{\geq 0}$.

Lemma 258. *Let G be a Lie group. The following are equivalent.*

1. *G is a torus.*
2. *G is connected, compact and abelian.*

Proof. The forward direction is clear. Conversely, suppose G is connected, compact, and abelian.

Since G is connected and abelian, we know (by part of Homework 3) that $\exp : \mathfrak{g} \to G$ is a surjective homomorphism with discrete kernel Γ, thus $G \cong \mathfrak{g}/\Gamma$. Since G is compact, the subgroup Γ is discrete and cocompact. Fix an isomorphism $\mathfrak{g} \cong \mathbb{R}^k$. One can show easily that every discrete cocompact subgroup of $\mathbb{R}^k$ is given by $\mathbb{Z}^k$ after a change of coordinates. Thus $G \cong T^k$. □

Definition 259. Let G be a Lie group. A *torus in G* is a closed subgroup $T \leq G$ (hence a Lie subgroup) that is a torus.

Remark 260. Let T be a torus, let G be a Lie group, and let $j : T \to G$ be a morphism of Lie groups. Since T is compact, connected, and abelian, so is its image under T. Thus $j(T)$ is a torus. In particular, immersed Lie subgroups that are isomorphic to tori are in fact closed subgroups. This explains why we restrict to closed subgroups in the previous definition.

Hence let K be a compact connected Lie group with Lie algebra $\mathfrak{k}$.

Definition 261. A *maximal torus* in K is a torus $T \leq K$ that is not properly contained in any torus in K.

Exercise 36. The following are equivalent for a closed connected subgroup T of K:

1. T is a torus.
2. $\mathfrak{t} := \mathrm{Lie}(T)$ is an abelian subalgebra of $\mathfrak{k}$.

Recall from a long time ago that if H_1, H_2 are two connected Lie subgroups of the same Lie group, then

- $\mathfrak{h}_1 = \mathfrak{h}_2$ if and only if $H_1 = H_2$,
- $\mathfrak{h}_1 \subseteq \mathfrak{h}_2$ if and only if $H_1 \subseteq H_2$,

etc.

Lemma 262. *The following are equivalent for a closed connected subgroup T of K:*

(1) T is a maximal torus.
(2) $\mathfrak{t} := \mathrm{Lie}(T)$ is a maximal abelian subalgebra of $\mathfrak{k}$ (i.e., an abelian subalgebra that is not properly contained in any abelian subalgebra.)

Proof. Suppose T is a maximal torus. Let $\mathfrak{t}' \supseteq \mathfrak{t}$ be an abelian subalgebra that contains $\mathfrak{t}$. Suppose there exists $X \in \mathfrak{t}' - \mathfrak{t}$. Since $\mathfrak{t}'$ is abelian, we have $[X, \mathfrak{t}] = 0$. Since T is connected, it follows (by the usual differentiation/exponentiation technique) that the group

$$H := \{e^{yX} t : y \in \mathbb{R}, t \in T\}$$

is abelian. It is also connected. Hence its closure $\overline{H}$ is abelian, connected, and closed inside the compact Lie group K, hence compact, hence a torus (Lemma 258). By Theorem 176, H is a Lie subgroup, so we may consider its Lie algebra $\mathfrak{h}$; clearly $\mathfrak{h}$ contains X and $\mathfrak{t}$. Therefore the torus H has Lie algebra $\mathfrak{h}$ properly containing $\mathfrak{t}$. By the old result recalled above characterizing containments between closed Lie subgroups in terms of containments of their Lie algebras, we deduce that $\overline{H}$ is a torus properly containing T, which contradicts the assumed maximality of T.

Conversely, if $\mathfrak{t}$ is a maximal abelian subalgebra and T' is a torus properly containing T, then (by the same old fact recalled above) its Lie algebra $\mathfrak{t}'$ properly contains $\mathfrak{t}$ and is abelian, contradicting the assumed maximality of $\mathfrak{t}$. □

Corollary 263. *Let $T \leq K$ be a maximal torus in a compact Lie group. Let $\mathfrak{t} \leq \mathfrak{k}$ be the induced inclusion of LIe algebras. Then $\mathfrak{t}$ is self-centralizing: $\{X \in \mathfrak{g} : [X, \mathfrak{t}] = 0\} = \mathfrak{t}$.*

Proof. Otherwise there is $X \in \mathfrak{g}$ so that $[X, \mathfrak{t}] = 0$, hence $\mathfrak{t}' := \mathbb{R}X + \mathfrak{t}$ is an abelian subalgebra of $\mathfrak{k}$ that properly contains $\mathfrak{t}$. By Lemma 262, this does not happen. □

33.2. **Characters of tori.**

Definition 264. Let T be a torus. A *character* of T is a continuous homomorphism $\chi : T \to \mathbb{C}^{(1)} := \{z \in \mathbb{C}^\times : |z| = 1\}$. The character group of T is the group $\mathfrak{X}(T)$ consisting of all characters; the group law is given by multiplication.

Lemma 265. *Let T be a torus and let $R : T \to \mathrm{GL}(V)$ be a representation on a finite-dimensional complex vector space. Then V decomposes as a direct sum of invariant one-dimensional subspaces on which T acts by characters of T.*

More precisely, one has $V = \oplus_\chi V^\chi$, where χ traverses the set of characters of T and $V^\chi := \{v \in V : R(t) = \chi(t) v\}$. Any subspace of V^χ is invariant; by choosing a basis for each V^χ, we obtain a decomposition of V as a sum of one-dimensional invariant (irreducible) subspaces.

Proof. By §16.8, the representation is completely reducible. To complete the proof, we just need to show that any irreducible representation V of T is one-dimensional. To that end, it suffices to show that each $t \in T$ acts on V by some scalar λ.

Indeed, let $t_0 \in T$ be given. Since $R(t_0) \in \mathrm{GL}(V)$ is a cmoplex matrix, it has some (nonzero) eigenvector $v_0 \in V$ and some eigenvalue $\lambda \in \mathbb{C}$. Consider the eigenspace $W := \{v \in V : R(t_0)v = \lambda v\}$. Our goal is to show that $W = V$. Since V is irreducible and W is nonzero (after all, it contains v_0), it sufficse to show that W is invariant. Here we use the commutativity of T: if $t \in T$, then $R(t_0)R(t) = R(t_0 t) = R(t t_0) = R(t)R(t_0)$, hence for $v \in V$, we have $R(t_0)R(t)v = R(t)\lambda v = \lambda R(t)v$, hence $R(t)v \in W$, and so W is $R(t)$-invariant, as required. □

Lemma 266. *Any continuous homomorphism $\chi : \mathbb{R} \to \mathbb{C}^{(1)}$ is of the form $\chi(x) = e(\xi x) := e^{2\pi i \xi x}$ for some unique $\xi \in \mathbb{R}$.*

Proof. One can certainly do this directly, but we might as well deduce it from stuff we've seen in class:

It is clear that $x \mapsto e(\xi x)$ is a character for each $\xi \in \mathbb{R}$, and that for $\xi_1 \neq \xi_2$, the characters obtained in this way are distinct.

Conversely, recall that shortly after we proved the "closed subgroups are Lie subgroups" theorem, we indicated in class and assigned on the homework that continuous homomorphisms between Lie groups are automatically smooth, hence determined by their differentials. In particular, $\chi : \mathbb{R} \to \mathbb{C}^{(1)}$ is determined by

$$d\chi : \mathbb{R} = \mathrm{Lie}(\mathbb{R}) \to i\mathbb{R} = \mathrm{Lie}(\mathbb{C}^{(1)}),$$

which is then of the form $2\pi i \xi$ for some $\xi \in \mathbb{R}$, etc. □

One obtains an analogous classifciation of the characters of $\mathbb{R}^k$ by taking products. From this we deduce:

Lemma 267. *The character group of T^k is isomorphic to $\mathbb{Z}^k$: to each $\xi = (\xi_1, \dots, \xi_k) \in \mathbb{Z}^k$ one associates the character $T^k = (\mathbb{R}/\mathbb{Z})^k \ni x = (x_1, \dots, x_k) \mapsto e(\sum \xi_i x_i) \in \mathbb{C}^{(1)}$.*

Proof. A character of T^k pulls back under the surjective homomorphism $\mathbb{R}^k \to T^k$ to a character of $\mathbb{R}^k$, which is in turn classified by real numbers $(\xi_1, \dots, \xi_k)$; conversely, such real numbers induce a character of T^k precisely when they are all integers. □

Let T be a torus. Let $\mathfrak{t}$ denote its Lie algebra and $\mathfrak{h} := \mathfrak{t} \otimes_{\mathbb{R}} \mathbb{C}$ the complexification thereof. Set $\mathfrak{h}_{\mathbb{R}} := i\mathfrak{t} \leq \mathfrak{h}$. The map $e : \mathfrak{h}_{\mathbb{R}} \to T$ given by $e(X) := \exp(2\pi i X)$ is a surjective homomorphism with discrete cocompact kernel.

Let $\chi : T \to \mathbb{C}^{(1)}$ be a character of T; as discussed above, it is smooth, so we can consider its differential $d\chi : \mathfrak{t} \to i\mathbb{R}$, which identifies with a linear map $d\chi : \mathfrak{h}_{\mathbb{R}} \to \mathbb{R}$. $\chi(e(H)) = e(\lambda(H))$ for some $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$. Conversely, such a λ defines a character χ if and only if it vanishes on $\mathfrak{h}_{\mathbb{Z}} := \ker(e : \mathfrak{h}_{\mathbb{R}} \to T)$, i.e., if and only if it belnogs to $\mathfrak{h}_{\mathbb{Z}}^*$ as defined earlier. In summary:

Lemma 268. *Let T be a torus. Let $\mathfrak{h}_{\mathbb{R}} := i\mathfrak{t}$, as above, so that $e : \mathfrak{h}_{\mathbb{R}}/\mathfrak{h}_{\mathbb{Z}} \to T$ is an isomorphism.*

Then $\mathfrak{X}(T) \cong \mathfrak{h}_{\mathbb{Z}}^*$ via the bijection $\chi \hookrightarrow \lambda$ characterized by $2\pi\lambda = d\chi$ and $\chi(e(H)) = e(\lambda(H))$ for all $H \in \mathfrak{h}_{\mathbb{R}}^*$.

Definition 269. For $\lambda \in \mathfrak{h}_{\mathbb{Z}}^*$, denote by e^{λ} the character of T associated to it by the above bijection, so that for all $H \in \mathfrak{h}_{\mathbb{R}}$,

$$e^{\lambda}(e(H)) = e(\lambda(H)).$$

This definition applies in particular to each $\alpha \in R \subseteq \mathfrak{h}_{\mathbb{Z}}^*$.

33.3. **Topologies on character groups.** Let's talk briefly about topology. Let G be a topological group. Let $\mathfrak{X}(G)$ denote the set of continuous homomorphisms $\chi : G \to \mathbb{C}^{(1)}$. (The notation is consistent with that used above when G is a torus.) We equip $\mathfrak{X}(G)$ with the "compact-open" topology. This means that a subbasis for the open sets in $\mathfrak{X}(G)$ is given by cosets of sets of the form $V(C, U) := \{\chi \in \mathfrak{X}(G) : \chi(C) \subseteq U\}$, where $C \subseteq T$ is compact and $U \subseteq \mathbb{C}^{(1)}$ is open. Equivalently, a net $\chi^{(\alpha)} \in \mathfrak{X}(G)$ converges to some $\chi \in \mathfrak{X}(G)$ precisely when it converges uniformly on compact sets in the ordinary sense. We may define on $\mathfrak{X}(G)$ the binary operation $\cdot$ given by $(\chi_1 \cdot \chi_2)(g) := \chi_1(g)\chi_2(g)$.

Exercise 37. Show that $\mathfrak{X}(G)$ is a topological group with respect to this operation.

Lemma 270. *Suppose that G is compact. Then $\mathfrak{X}(G)$ is discrete.*

Proof. Take $C := G$ and let $U \subseteq \mathbb{C}^{(1)}$ be an interval of length $1/10$ with center $1 \in \mathbb{C}^{(1)}$. Let $\chi_0 \in \mathfrak{X}(G)$ denote the trivial character $\chi_0(g) := 1$. We claim that $V(C, U) = \{\chi_0\}$. Clearly $\chi_0 \in V(C, U)$. Conversely, let $\chi \in \mathfrak{X}(G) - \{\chi_0\}$ be a nontrivial character, so that there exists $g \in G$ for which $\chi(g) \neq 1$. Since $\chi(g) \in \mathbb{C}^{(1)} - \{1\}$, we can find some power of it, say $\chi(g)^n = \chi(g^n)$, which has negative real part. But then $\chi(g^n) \notin U$, hence $\chi \notin V(C, U)$. We now use that a topological group is discrete if and only if the set consisting of its identity element is open (if this wasn't an exercise before, it could be now). $\square$

The lemma applies notably to the case that G is a compact torus $T = (\mathbb{R}/\mathbb{Z})^k$. We saw above that $\mathfrak{X}(T) \cong \mathbb{Z}^k$ as groups. Lemma 270 tells us moreover that $\mathfrak{X}(T)$ and $\mathbb{Z}^k$ are isomorphic as topological groups, each equipped with the discrete topology.

33.4. **Maximal tori give rise to Cartan subalgebras.**

Theorem 271. *Let T be a maximal torus in the compact connected Lie group K. Let $\mathfrak{t} \leq \mathfrak{k}$ denote their Lie algebras and*

$$\mathfrak{h} := \mathfrak{t} \otimes \mathbb{C} \leq \mathfrak{g} := \mathfrak{k} \otimes \mathbb{C}$$

the complexifications. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. The roots are purely imaginary on $\mathfrak{t}$.

Proof. According to our definition, we must check that $\mathfrak{h}$ is abelian, ad-diagonalizable, and self-centralizing.

(1) Since T is abelian, so is $\mathfrak{t}$, hence $\mathfrak{h}$ is abelian.
(2) Consider the adjoint action $\mathrm{Ad} : K \to \mathrm{End}(\mathfrak{k})$. Restrict it to obtain $\mathrm{Ad} : T \to \mathrm{End}(\mathfrak{k})$. Extend it complex-linearly to obtain $\mathrm{Ad} : T \to \mathrm{End}(\mathfrak{g})$. Since T is compact, this complex linear representation of it is completely reducible. By the previous lemma, it decomposes as a direct sum of one-dimensional invariant subspaces. Differentiating this fact, we see that $\mathrm{ad}(\mathfrak{t})$ and hence (by linearity) $\mathrm{ad}(\mathfrak{h})$ is diagonalizable.

    The functionals $\lambda \in \mathfrak{h}^*$ for $\mathfrak{h}$ acting on $\mathfrak{g}$ by the adjoint map correspond to the characters χ of T occurring in the decomposition described above. It follows from our earlier discussion that each such λ is real-valued on $\mathfrak{h}_{\mathbb{R}}$.

(3) Let $V_0 := \{X \in \mathfrak{g} : \mathrm{Ad}(t)X = X \text{ for all } t \in T\}$ be the subspace on which $\mathrm{Ad}(T)$ acts trivially. By the usual differentiation/exponentiation trick, we have
$$V_0 = \{X \in \mathfrak{g} : [X, \mathfrak{t}] = 0\}$$
and
$$V_0 = \{X \in \mathfrak{g} : [X, \mathfrak{h}] = 0\}$$
. From the first of these last two equations and linear algebra, we have
$$V_0 = \{X \in \mathfrak{k} : [X, \mathfrak{t}] = 0\} \otimes \mathbb{C}.$$
By Corollary 263, we see that $V_0 = \mathfrak{h}$. This gives the required self-centralizing property of $\mathfrak{h}$.

□

In general Lie groups, nontrivial tori (let alone maximal ones) need not exist. But in compact Lie groups, things are better:

Lemma 272. *Let K be a compact connected Lie group. For any torus S in K, there is a maximal torus T in K that contains S. (Note that the trivial torus $S = \{1\}$ always exists.)*

Proof. If S is not maximal, then it is contained in some strictly larger torus S', which then (by consideration of LIe algebras) has strictly larger dimension. Iterating the procedure $S \mapsto S'$ finitely many times, we wind up with a maximal torus. (We can't iterate forever, because $\mathfrak{k}$ is finite-dimensional. □

33.5. **Some notation involving roots.** So now we have the full theory of roots at our disposal. Let's set up some notation. Let K be compact connected, and let $T \leq K$ be a maximal torus. Let $\mathfrak{h}, \mathfrak{g}$ be as above. We can then decompose
$$\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{\alpha \in R} \mathfrak{g}^{\alpha}) \tag{142}$$
where R is a finite subset of $\mathfrak{h}_{\mathbb{R}}^* - \{0\}$. In fact, the discussion above implies that $R \subseteq \mathfrak{h}_{\mathbb{Z}}^* - \{0\}$. For each $H \in \mathfrak{h}_{\mathbb{R}}$ and $X \in \mathfrak{g}^{\alpha}$, we have
$$[H, X] = \alpha(H)X,$$
For $\lambda \in \mathfrak{h}_{\mathbb{Z}}^*$, let $e^{\lambda} \in \mathfrak{X}(T)$ be as in Definitnio 269. This applies in particular to $\alpha \in R \subseteq \mathfrak{h}_{\mathbb{Z}}^*$, and the above identity translates to: for $t \in T$ and $X \in \mathfrak{g}^{\alpha}$,
$$\mathrm{Ad}(t)X = e^{\alpha}(t)X.$$
Note also that if $t = e(H)$ with $H \in \mathfrak{h}_{\mathbb{R}}$, then $e^{\alpha}(t) = e(\alpha(H))$, hence
$$\mathrm{Ad}(e(H))X = e(\alpha(H))X. \tag{143}$$

33.6. **The automorphism group of a compact torus.** Let T be a compact torus. Fix an identification $T = (\mathbb{R}/\mathbb{Z})^k$ for some $k \in \mathbb{Z}_{\geq 0}$. Then T is a compact Lie group. We may speak of its automorphism group $\mathrm{Aut}(T)$. By definition, this consists of continuous homomorphisms $\sigma : T \to T$ that admit continuous inverse homomorphisms. We may identify $\mathfrak{t} := \mathrm{Lie}(T)$ with $\mathbb{R}^k$. Since T is connected, any such σ is determined by its differential $d\sigma : \mathbb{R}^k \to \mathbb{R}^k$, which is a linear map, call it A. Since the exponential map $\mathfrak{t} \to T$ is given with respect to our identifications by the natural projection $\mathbb{R}^k \to \mathbb{R}^k/\mathbb{Z}^k$, we see that σ is the map induced by a linear map $A : \mathbb{R}^k \to \mathbb{R}^k$. For σ to be well-defined, we must have $A(\mathbb{Z}^k) \subseteq \mathbb{Z}^k$. For σ to be an isomorphism, its inverse σ^{-1} should exist and be well-defined, and so we should have $A(\mathbb{Z}^k) = \mathbb{Z}^k$. But $\{A \in \mathrm{GL}_k(\mathbb{R}) : A\mathbb{Z}^k = \mathbb{Z}^k\} = \mathrm{GL}_k(\mathbb{Z})$. We may thereby identify

$$\mathrm{Aut}(T) \cong \mathrm{GL}_k(\mathbb{Z}).$$

We define the topology on $\mathrm{Aut}(T)$ in this case to be the discrete topology.

There's another "transposed" way to make the above identification. Given an automorphism σ of T, we can attach the induced automorphism σ^t of $\mathfrak{X}(T)$, which sends a character χ of T to the new character $\sigma^t\chi \in \mathfrak{X}(T)$ given by $\sigma^t\chi(t) := \chi(\sigma t)$. The map

$$\mathrm{Aut}(T) \ni \sigma \mapsto \sigma^{-t} := (\sigma^t)^{-1} = (\sigma^{-1})^t \in \mathrm{Aut}(\mathfrak{X}(T))$$

is an isomorphism. Since $\mathfrak{X}(T) \cong \mathbb{Z}^k$, we have $\mathrm{Aut}(\mathfrak{X}(T)) = \mathrm{GL}_k(\mathbb{Z})$.

33.7. **Generators.** Recall that an abstract group G is said to be *cyclic* if it admits a *generator*, i.e., an element $g \in G$ for which $G = \{g^n : n \in \mathbb{Z}\}$. There aren't so many cyclic groups; they are all isomorphic either to $\mathbb{Z}$ or $\mathbb{Z}/n$ for some $n \in \mathbb{Z}_{\geq 1}$.

Given a topological group G, one says that G is *topologically cyclic* if it admits a *topological generator*, i.e., an element $g \in G$ for which $G = \overline{\{g^n : n \in \mathbb{Z}\}}$ where $\overline{\cdot}$ denotes closure. For example, any abstract cyclic group (equipped with the discrete topology or any other topology, for that matter) is topologically cyclic, and any generator in the group-theoretic sense is a topological generator, but there are more interesting examples of topologically cyclic groups than just those that are cyclic in the ordinary sense. (For example: the profinite integers $\hat{\mathbb{Z}}$, the p-adic integers $\mathbb{Z}_p$, etc.)

For our purposes, it will be useful to know that compact tori are topologically cyclic:

Lemma 273. *Let $T \cong \mathbb{R}^k/\mathbb{Z}^k$ be a compact torus. Then the set of topological generators of T is dense; in particular, it is nonempty.*

Proof. We presented the simple pigeonholing argument in lecture. Fix a countable basis $B_1, B_2, \ldots$ for T. Take any open subset U of T. We want to show that we can find $g \in U$ so that for each $i \in \mathbb{Z}_{\geq 1}$ there exists an $n \in \mathbb{Z}$ so that $g^n \in B_i$. (Then we're done.)

We now aim to construct such a g. For convenience of notation, let us realize T as the additive group $\mathbb{R}^k/\mathbb{Z}^k$. We aim to find for each $i \in \mathbb{Z}_{\geq 1}$

- a nonempty open set U_i, and
- an integer N_i

so that $U \supseteq U_1 \supseteq U_2 \supseteq U_3 \supseteq \cdots$ and so that the set $N_i U_i := \{n_i c_i : n_i \in N_i, c_i \in U_i\}$ is contained in B_i^0, where B_i^0 denotes an open subset of B_i for which $\overline{B_i^0} \subseteq B_i$. To do this, set $U_0 := U$. For each $i = 1, 2, 3, \ldots$, choose N_i large enough that $N_i U_{i-1} = T$. This is possible because U_{i-1} is open. Then set $U_i := \{u \in U_{i-1} : N_i u \in B_i^0\}$.

The set $\cap_i \overline{U_i}$ is nonempty by the finite intersection property. Any element of it is easily seen to be a generator. $\square$

Since we are primarily interested in topological groups here (or indeed, in Lie groups), we henceforth abuse terminology slightly by saying *generator* when we really mean "topological generator."

Generators are nice. For example, suppose $g \in K$ satisfies $gtg^{-1} = t$ for some generator t of T. Then also $gt^n g^{-1} = t^n$ for all $n \in \mathbb{Z}$. Since the set of all $x \in K$

for which $gxg^{-1} = x$ is closed, we deduce that $gxg^{-1} = x$ holds for all $x \in T$, i.e., that g centralizes T. We shall use arguments along these lines repeatedly.

33.8. **A maximal torus is the connected component of its normalizer.** Let T be a torus in a compact connected Lie group K. Let $N(T) := \{g \in K : gTg^{-1} = T\}$ denote its normalizer. The condition defining $N(T)$ is closed, so $N(T)$ is a closed subgroup of K, hence is a Lie subgroup of K. (One can also show this more directly along the lines of §25.3.) We can thus speak of the connected component $N(T)_0$. In general, $N(T)_0$ can be quite large. For example, if $T = \{1\}$ is the trivial torus, then $N(T) = N(T)_0 = 0$. However:

Lemma 274. *Suppose T is maximal. Then $N(T)_0 = T$.*

Proof. Consider the map $f : N(T)_0 \to \mathrm{Aut}(T)$ given by $f(n) := [t \mapsto ntn^{-1}]$. The domain $N(T)_0$ is connected, and the target $\mathrm{Aut}(T)$ is discrete. Assuming for now that f is continuous, it follows that its image must consist of a point, and so $N(T)_0$ actually centralizes T. Suppose that $N(T)_0$ strictly contains T. Since they are both connected Lie groups, this implies that we can find $X \in \mathfrak{k}, X \notin \mathfrak{t}$ that commutes with all of $\mathfrak{t}$. But this contradicts Corollary 263.

It remains to verify that f is continuous. Let $\chi_1, \dots, \chi_k$ be a $\mathbb{Z}$-basis for the character group $\mathfrak{X}(T) \cong \mathbb{Z}^k$. As in the discussion at the end of §33.6, we can think of $f(n)^{-t} \in \mathrm{Aut}(\mathfrak{X}(T))$. Suppose that $(n_i)_{i \in \mathbb{Z}_{\geq 1}}$ is a sequence of elements in $N(T)$ tending to some limit $n \in N(T)$. Then we have to check that $f(n_i)^{-t} \to f(n)^{-t} \in \mathfrak{X}(T)$ with respect to the discrete topology. This means that we have to show that for i large enough, one has $f(n_i)^{-t} = f(n)^{-t}$. Equivalently, we have to show for each $j \in \{1..k\}$, one has $f(n_i)^{-t}\chi_j = f(n)^{-t}\chi_j$ for i large enough. Since the character group $\mathfrak{X}(T)$ is discrete (see §33.3), it suffices to show that $f(n_i)^{-t}\chi_j$ converges to $f(n)^{-t}\chi_j$ as functions, uniformly on compact sets. This follows immediately from the continuity of the conjugation action of $N(T)$ on T and the compactness of T.

(There are probably simpler or slicker ways to write this proof; I hope in any event that it's clear.) $\square$

Corollary 275. *Let T be a maximal torus. Then the quotient $N(T)/T$ is finite.*

Proof. Indeed, that quotient identifies with the set $N(T)/N(T)_0$ of connected components of $N(T)$. Since K is compact and $N(T)$ is closed, we see also that $N(T)$ is compact, hence has only finitely many connected components. $\square$

33.9. **Conjugacy of maximal tori.** Let K be a compact connected Lie group K. Note that if T is a maximal torus in K, then so is its conjugate gTg^{-1} for any $g \in K$. Here's the big theorem on maximal tori in compact Lie groups:

Theorem 276. *Let T be a maximal torus in a compact connected Lie group K. Then $K = \cup_{g \in K} gTg^{-1}$.*

The (standard) proof we'll record uses the Lefschetz fixed point theorem. That theorem (or one variant of it) says that for a compact manifold M, one can attach to each continuous map $f : M \to M$ an integer $\Lambda(f)$ with the following properties:

(1) $\Lambda(f)$ only depends upon the homotopy class of f.
(2) If f has *isolated simple fixed points*, that is to say, if the set $\mathrm{Fix}(f) := \{x \in M : f(x) = x\}$ is finite and if for each $x \in \mathrm{Fix}(f)$, the linear map

$T_x f : T_x M \to T_{f(x)} M = T_x M$ satisfies $\det(1 - T_x f) \neq 0$, then

$$\Lambda(f) = \sum_{x \in \mathrm{Fix}(f)} \varepsilon_x(f),$$

where $\varepsilon_x(f) \in \{\pm 1\}$ denotes the sign of the nonzero real number $\det(1 - T_x f) \in \mathbb{R}^\times$.

In particular, if $\mathrm{Fix}(f) = \emptyset$, then $\Lambda(f) = 0$. This is a theorem from algebraic topology that we won't prove. We record the definition anyway:

$$\Lambda(f) = \sum_{i \in \mathbb{Z}_{\geq 0}} (-1)^i \operatorname{trace}(f^* | H^i(M, \mathbb{Q}))$$

where the RHS involves singular cohomology groups with rational coefficients. For the identity map $1 : M \to M$, one writes $\chi(M) := \Lambda(1)$. The quantity

$$\chi(M) = \sum_{i \in \mathbb{Z}_{\geq 0}} (-1)^i \operatorname{trace}(f^* | H^i(M, \mathbb{Q}))$$

is called the *Euler characteristic* of M.

Anyway, back to our goal. We want to show that for each $x \in K$, there exists $g \in K$ so that $x \in gTg^{-1}$, or equivalently, so that $xg \in gT$, or equivalently, so that $xgT = gT$. In other words, we want to show that the map

$$f_x : K/T \to K/T$$
$$f_x(gT) := xgT$$

has a fixed point. The manifold K/T is compact (since K is), so we can apply the Lefschetz theorem. Assuming for the sake of contradiction that f_x had no fixed point, we'd deduce from the Lefschetz that $\Lambda(f_x) = 0$. Let's note that since K is a connected manifold, it is path-connected. For any $x, y \in K$, we can find a path connected them; that path induces a homotopy between the maps f_x and f_y, and in particular between them and the identity map f_1, for which $\Lambda(f_1) = \chi(M)$. So we're done if we can show that $\chi(M) \neq 0$. We'll actually show more precisely that

$$\chi(M) = \# N(T)/T. \tag{144}$$

As noted above, we can compute $\chi(M)$ as $\Lambda(f_x)$ for *any* $x \in K$. It is convenient to take for x a generator (see §33.7) of the torus T. What, then, are the fixed points of f_x? Well, $gT \in \mathrm{Fix}(f_x)$ if and only if $xgT = gT$, i.e., $g^{-1}xg \in T$; but since x generates T, it follows then that $g^{-1}Tg \subseteq T$. One can then see in many ways that $g^{-1}Tg = T$. (For example, they are both maximal tori.) Hence $g \in N(T)$. Thus $\mathrm{Fix}(f_x) = N(T)/T$.

Henceforth abbreviate $f := f_x$. For each $g \in N(T)/T$, we have a commutative diagram as drawn in class which shows that $\det(1 - T_{gT} f | T_{gT}(G/T))$ is independent of g, so we henceforth focus on the case that $gT = eT$ is the identity coset.

We can then identify

$$T_{eT}(G/T) = \mathfrak{k}/\mathfrak{t}$$

and hence

$$T_{eT}(G/T)_{\mathbb{C}} = \mathfrak{g}/\mathfrak{h} \cong \oplus_{\alpha \in R} \mathfrak{g}^\alpha.$$

with the usual notation. Let's write $x = e(H)$ for some $H \in \mathfrak{h}_{\mathbb{R}}$. Our assumption that x is a generator entails in particular that $e(\alpha(H)) \neq 1$ for all $\alpha \in R$, as otherwise x would belong to the codimension 1 submanifold of T consisting of elements

$e(H)$ for which $e(\alpha(H)) = 1$. By linear algebra, we can compute determinants after complexifying. We can also write

$$T_x f : T_{eT}(G/T)_{\mathbb{C}} \to T_{eT}(G/T)_{\mathbb{C}}$$

as

$$\mathrm{Ad}(x) : \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$$

because, since $x \in T$, one has

$$xgT = xgx^{-1}T.$$

Thus

$$\det(1 - T_{eT}f|T_{eT}(G/T)) = \prod_{\alpha \in R} \det(1 - \mathrm{Ad}(x)|\mathfrak{g}^\alpha).$$

On the other hand

$$\det(1 - \mathrm{Ad}(x)|\mathfrak{g}^\alpha) = 1 - e(\alpha(H)).$$

We can split the product into a product over pairs $\pm\alpha \in R/\{\pm 1\}$ taken up to sign, giving

$$\det(1 - T_{eT}f|T_{eT}(G/T)) = \prod_{\pm\alpha \in R} (1 - e(\alpha(H)))(1 - e(-\alpha(H))).$$

For each $\alpha \in R$, write $\theta := 2\pi i \alpha(H)$. Then

$$0 \neq (1 - e(\alpha(H)))(1 - e(-\alpha(H))) = (1 - e^{i\theta})(1 - e^{-i\theta}) = 2 - 2\cos(\theta) \geq 0.$$

We conclude that

$$\det(1 - T_{eT}f|T_{eT}(G/T)) > 0$$

hence that (as explained more carefully in class via a commutative diagram)

$$\det(1 - T_{gT}f|T_{eT}(G/T)) > 0 \text{ for each } g \in N(T)/T$$

hence that f has isolated fixed points gT with signs $\varepsilon_{gT}(f) = 1$. Therefore

$$\Lambda(f) = \#N(T)/T,$$

as required.

33.10. **Basic consequences of the conjugacy theorem.** Let K be a connected compact Lie group, and let all other notation be as usual. For now, we indicate some consequecnes relevant for answering the questions raised last time.

Corollary 277. *The center Z of K is the intersection $\cap T$ of all (maximal) tori.*

Proof. Let $z \in Z$, and let T be a maximal torus. By the theorem, we may write $z = gtg^{-1}$ for some $t \in T$. But then $t = g^{-1}zg = z$, because z is in the center. Hence z belongs to T.

Conversely, suppose $z \in K$ belongs to $\cap T$. Let $x \in K$; we must show that x and z commute. To that end, we apply Theorem 276 to find a maximal torus T that contains x. Then T contains z and x; since T is commutative, the elements x and z commute, as required. □

Theorem 278. *Let T be a maximal torus in the compact connecteed Lie group K. Let $\mathfrak{h}_{\mathbb{R}} := i\mathfrak{t}$ and $e : \mathfrak{h}_{\mathbb{R}}/h_{\mathbb{Z}} \xrightarrow{\cong} T$ be as usual. Let R denote the set of roots for $\mathfrak{h} := \mathfrak{t} \otimes \mathbb{C}$ acting on $\mathfrak{g} := \mathfrak{k} \otimes \mathbb{C}$ by the adjiont representation. Let Z denote the center of K. Then e induces an isomorphism*

$$(\mathbb{Z}R)^*/\mathfrak{h}_{\mathbb{Z}} \cong Z.$$

Proof. By Corollary 277, Z is contained in T, hence the image of e contains Z. To complete the proof, all that remains to be showed is the following: for $H \in \mathfrak{h}_{\mathbb{R}}$, we have $e(H) \in Z$ if and only if $H \in (\mathbb{Z}R)^{\wedge}$. This equivalence is demonstrated by noting that each of the following assertions is evidently equivalent to the next:

(1) $e(H) \in Z$.
(2) $e(H) \in \ker(\mathrm{Ad})$ (using here that K is connected)
(3) $e(\alpha(H)) = 1$ for all $\alpha \in R$ (use (142) and (143), and note that $\mathrm{Ad}(e(H))$ acts trivially on $\mathfrak{h}$ because $\mathfrak{t}$ is abelian)
(4) $\alpha(H) \in \mathbb{Z}$ for all $\alpha \in R$ (because $\mathbb{Z} = \{x \in \mathbb{R} : e^{2\pi i x} = 1\}$)
(5) $H \in (\mathbb{Z}R)^*$ (by definition of the latter).

□

Lemma 279. *Let T_1, T_2 be maximal tori in K. Then T_1, T_2 are conjugate.*

Proof. Choose a generator t_1 for T_1. By the cnojugacy theorem, we may find $g \in G$ so that $gt_1g^{-1} \in T_2$. Since t_1 is a generator, it follows that $gT_1g^{-1} \subseteq T_2$. Since T_1, T_2 are both maximal, we conclude that $gT_1g^{-1} = T_2$, as required. □

Definition 280. Given an element $u \in K$, the *centralizer* $Z(u) := Z_K(u)$ is defined to be $Z(u) := \{g \in K : gug^{-1} = u\}$.

Similarly, for any subgroup $U \leq K$, the *centralizer* $Z(U) := Z_K(U)$ is defined to be $Z(U) := \cap_{u \in U} Z(u) = \{g \in K : gug^{-1} = u \text{ for all } u \in U\}$.

Lemma 281. *Let T be a maximal torus. Then $Z(T) = T$.*

Proof. This is a tricky argument, so I've spelled the proof out a bit more verbosely. (This one is worth studying and rewriting on your own, I think.)

Since T is abelian, we have $T \subseteq Z(T)$. Conversely, let $g \in Z(T)$. Let H be the closure of the subgroup of K generated by T and g. Since g commutes with T and T is abelian, we know that H is abelian.

If we were lucky enough that H happened to be connected, then we'd be done: H would then be connected, compact, and abelian, hence a torus, but since T is a maximal torus, the only possiblity is $H = T$, and thus $g \in T$.

Unfortunately, there is no obvious reason for H to be connected. We are led to consider its connected component H_0. Since T is connected, we know that H contains T. Since H_0 is connected, abelian and compact (being closed inside K), it is a torus; since T is maximal, we must have $H_0 = T$.

If we were lucky enough that g happened to belong to H_0, then again, we'd be done. But there is no obvious reason for that be the case. (Think about it.) Fortunately, as we now explain, H/H_0 is not too complicated, so we can make the argument work anyway.

Let's see. Since $H_0 \subseteq H$ is open, the quotient H/H_0 is discrete; it is also compact, hence finite. Moreover, by construction of H, that quotient is generated by the image $\overline{g} \in H/H_0$ of g. Let $m \in \mathbb{Z}_{\geq 1}$ be the smallest natural number for which $\overline{g}^m = 1$, or equivalently, for which $g^m \in H_0 = T$. Let $t \in T$ be a generator of the torus. The torus is a divisible group, so we can find $s \in T$ for which $s^m = tg^{-m}$.

Set $u := sg$. We claim that u is a generator of H. To see that, we must check that the powers of u are dense. We have $u^m = t$, and t is a generator of T, so the closure of the set of powers of u contains T. The full group H is the union over $j \in \mathbb{Z}/m$ of the cosets $g^j H_0 = g^j T$. The set of powers $u^{mn+j} = u^j t^n \in g^j T$ $(n \in \mathbb{Z})$ are dense in that coset. Hence u generates H.

Now we use the big theorem on conjugacy of maximal tori to deduce that u is contained in *some* maximal torus S of G (e.g., a conjugate of T). Since u generates H, we must also have $H \subseteq S$. So now we have the following containments:

$$T \subseteq H \subseteq S.$$

Since T is a *maximal* torus, the only possiblity is that $T = S$, hence that $H = T$, hence that $g \in T$.

Since $g \in Z(T)$ was arbitrary, we conclude finally that $Z(T) = T$, as required. $\square$

34. Regular and singular elements

34.1. **Definitions and basic properties.** Recall that every element of K is contained in *some* maximal torus.

Definition 282. We say that an element $g \in K$ is *regular* if it belongs to exactly one maximal torus; if otherwise it belongs to at least two maximal tori, then we call it *singular*.

Introduce the superscript reg, as in K^{reg} or T^{reg} (for a maximal torus T), to denote "subset of regular elements."

Being regular is a property of conjugacy classes: if $x \in K$ is regular, then so is gxg^{-1} for any $g \in K$, and vice-versa. Since every element is conjugate to some element of any given maximal torus T, we can understand the regular elements pretty well if we understand which elements of T are regular.

To that end, let R denote the set of roots of T. For each $\alpha \in R$, set

$$T_\alpha := \ker(e^\alpha) \leq T$$

where $e^\alpha : T \to \mathbb{C}^{(1)}$ is the character $T \ni e(H) \mapsto e(\alpha(H))$ (here $H \in \mathfrak{h}_\mathbb{R}$). We refer to §33.5 for any unexplained notation.

More verbosely, since $\alpha(H) \in \mathbb{Z}$ iff $e(\alpha(H)) = 1$, one has

$$T_\alpha = \{e(H) : H \in \mathfrak{h}_\mathbb{R}, \alpha(H) \in \mathbb{Z}\}.$$

T_α need not be connected, but its connected component $(T_\alpha)_0$ is easily seen to be codimension one subtorus of T with Lie algebra $\mathfrak{t}_\alpha = \ker(\alpha : \mathfrak{t} \to i\mathbb{R})$.

Proposition 283. *An element of T is regular if and only if it doesn't belong to any the T_α, i.e.,*

$$T^{\text{reg}} = T - \cup_{\alpha \in R} T_\alpha.$$

Proof. For $t \in T$, let $Z(t) := \{g \in K : gtg^{-1} = t\}$ denote its centralizer. It is a Lie subgroup of K with Lie algebra $\mathfrak{z}(t) = \{X \in \mathfrak{k} : \text{Ad}(t)X = X\}$ whose complexification is in turn

$$V := \mathfrak{z}(t)_\mathbb{C} = \{Z \in \mathfrak{g} : \text{Ad}(t)Z = Z\},$$

where $\mathfrak{g} := \mathfrak{k}_\mathbb{C}$ as usual. It is clear that $Z(t) \supseteq T$, hence that $\mathfrak{z}(t) \supseteq \mathfrak{t}$, hence that $V \supseteq \mathfrak{h} := \mathfrak{t}_\mathbb{C}$. Consider the root space decomposition $\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{\alpha \in R} \mathfrak{g}^\alpha)$. If $Z \in \mathfrak{g}$ has the form $Z = Z_0 + \sum Z_\alpha$ with $Z_0 \in \mathfrak{h}$ and $Z_\alpha \in \mathfrak{g}^\alpha$, then

$$\text{Ad}(t)Z_0 = Z_0, \quad \text{Ad}(t)Z_\alpha = e^\alpha(t)Z_\alpha,$$

hence

$$\text{Ad}(t)Z = Z \iff e^\alpha(t) = 1 \text{ whenever } Z_\alpha \neq 0.$$

We deduce that the following are equivalent:

(1) V properly contains $\mathfrak{h}$.
(2) There exists $\alpha \in R$ so that $e^{\alpha}(t) = 1$.
(3) $t \in \cup_{\alpha \in R} T_{\alpha}$.

We now prove the required equivalence. Suppose first that t is not regular. Then it is contained in some maximal torus T' other than T, hence $Z(t) \supseteq T'$, and thus $\mathfrak{z}(t) \supseteq \mathfrak{t}' := \mathrm{Lie}(T')$; consequently $\mathfrak{z}(t)$ properly contains $\mathfrak{t}$ and so V properly contains $\mathfrak{h}$; by the above, this is equivalent to t belonging to $\cup_{\alpha \in R} T_{\alpha}$.

Conversely, suppose t is regular. We claim then that $Z(t)_0 = T$. Clearly $T \subseteq Z(t)_0$. Conversely, let $g \in Z(t)_0$. By the main theorem on maximal tori applied to $Z(t)_0$, we can find a maximal torus S of $Z(t)_0$ containing g. Since t belongs to the center of $Z(t)_0$, we know also that S contains t. Let S' be a maximal torus of G that contains S. Then S' contains t; since t is regular, this implies that $S' = T$. Consequently $g \in S \subseteq S' = T$. Since g was arbitrary, the claim that $Z(t)_0 = T$ is proven. Consequently $V = \mathfrak{h}$ and thus $t \notin \cup_{\alpha \in R} T_{\alpha}$, as required. $\square$

34.2. **Singular elements have codimension at least three.** We know that $K = \cup_{g \in K} gTg^{-1}$. In other words, the well-defined map

$$f : K/T \times T \to K$$

$$f(g, t) := gtg^{-1}$$

is surjective. By the discussion above, the subset K^{sing} of singular elements is the union over $\alpha \in R$ of the images of the well-defined maps

$$f_{\alpha} : K/Z(T_{\alpha}) \times T_{\alpha} \to K.$$

Set $n := \dim(K)$ and $k := \dim(T)$. Clearly $\dim(T_{\alpha}) = k - 1$. On the other hand, as discussed more leisurely in class, $Z(T_{\alpha})_{\mathbb{C}}$ contains $\mathfrak{h} \oplus \mathfrak{g}^{\alpha} \oplus \mathfrak{g}^{-\alpha}$. Thus $\dim(Z(T_{\alpha})) \geq k + 2$, and so

$$\mathrm{image}(f_{\alpha}) \leq (n - (k+2)) + (k - 1) \leq n - 3.$$

Therefore the subset of singular elements in K has codimension ≤ 3.

It is a general fact that given a manifold M with submanifold M_0 for which $M - M_0$ has codimension ≥ 3, the natural map $\pi_1(M_0) \to \pi_1(M)$ is an isomorphism.

A simpler example: if $M - M_0$ has codimension ≥ 2, then the connected components of M and M_0 are in natural bijection (i.e., $\pi_0(M_0) \to \pi_1(M)$ is a bijection).

34.3. **The key covering morphism.** We have a well-defined surjective map

$$f : K/T \times T^{\mathrm{reg}} \to K^{\mathrm{reg}}.$$

The following was explained in lecture, and is not so difficult:

Lemma 284. *This map is a covering map, i.e., a locally trivial fiber bundle with discrete fibers. The fibers have cardinality $|N/T|$, where $N := N(T) := \{g \in K : gTg^{-1} = T\}$.*

34.4. **The affine Weyl group and the components of the set of regular elements.** Let

$$\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}} := \{H \in \mathfrak{h}_{\mathbb{R}} : \alpha(H) \notin \mathbb{Z} \text{ for all } \alpha \in R\}.$$

Equivalently, $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$ is the preimage under $e : \mathfrak{h}_{\mathbb{R}} \to T$ of T^{reg}.

The open subset $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$ of $\mathfrak{h}_{\mathbb{R}}$ is a union of complements of hyperplane. Each connected component P of $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$ is convex, and admits a definition of the shape

$$P = \{H \in \mathfrak{h}_{\mathbb{R}} : n_\alpha < \alpha(H) < n_\alpha + 1 \text{ for all } \alpha \in R\}$$

for some system of integral parameters $n_\alpha \in \mathbb{Z}$ attached to the roots $\alpha \in R$. It is worth trying to draw some pictures of $\mathfrak{h}_{\mathbb{R}}^{\mathrm{reg}}$ in all the rank 2 classical Lie algebras (as attempted in lecture for B_2).

The group $(\mathbb{Z}R)^*$ acts on $\{H \in \mathfrak{h}_{\mathbb{R}} : \alpha(H) \in \mathbb{Z}\}$ by translation: for $Z \in (\mathbb{Z}R)^*$, one has $\alpha(Z) \in \mathbb{Z}$, hence $\alpha(H + Z) \in \mathbb{Z}$ precisely when $\alpha(H) \in \mathbb{Z}$. Consequently $(\mathbb{Z}R)^*$ acts also on $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$, by translation.

Recall that the Weyl group W is generated by the root reflections $s_\alpha : \mathfrak{h}_{\mathbb{R}}^* \to \mathfrak{h}_{\mathbb{R}}^*$ defined for roots $\alpha \in R$ by $s_\alpha \lambda := \lambda - \lambda(H_\alpha)\alpha$. Recall also that $s_\alpha(R) = R$. For each $s \in W$, we may define the transpose element ${}^t s$, which acts now on the space $\mathfrak{h}_{\mathbb{R}}$ dual to the domain $\mathfrak{h}_{\mathbb{R}}^*$ of s; the action of ${}^t s$ is characterized by requiring that for $H \in \mathfrak{h}_{\mathbb{R}}$, one has for each $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$ that $\lambda({}^t sH) = (s\lambda)(H)$. This relation might be more pleasantly written

$$\langle \lambda, {}^t sH \rangle = \langle s\lambda, H \rangle.$$

Exercise 38. Show that for $\alpha \in R$ and $H \in \mathfrak{h}_{\mathbb{R}}$, one has

$${}^t s_\alpha H = H - \alpha(H) H_\alpha.$$

From now on we might abuse notation slightly by writing simply $s_\alpha := {}^t s_\alpha$ and identifying W with the subgroup $\{s := {}^t s : s \in W\}$ of $\mathrm{GL}(\mathfrak{h}_{\mathbb{R}})$. In this way, we regard W as acting on $\mathfrak{h}_{\mathbb{R}}$.

This action of W preserves $\mathbb{Z}R^\wedge$. Indeed, for each $s \in W$, one has $s(R^\wedge) = R^\wedge$, so the generators get permuted.

For this reason, we can form the semidirect product $\mathbb{Z}R^\wedge \rtimes W$. It is the group consisting of all pairs (H, s), where $s \in W, Z \in \mathbb{Z}R^\wedge$. The multiplication law is

$$(Z_1, s_1)(Z_2, s_2) = (Z_1 + s_1 Z_2, s_1 s_2).$$

This group acts naturally on $\mathfrak{h}_{\mathbb{R}}$ by the formula

$$(Z, s) \cdot H := Z + sH.$$

For each $\alpha \in R$ and $n \in \mathbb{Z}$, consider the linear map $s_{\alpha n} : \mathfrak{h}_{\mathbb{R}} \to \mathfrak{h}_{\mathbb{R}}$ given by reflection in the hyperplane $\{H : \alpha(H) = n\}$; explicitly,

$$s_{\alpha n}(H) = H - (\alpha(H) - n) H_\alpha.$$

Clearly $s_{\alpha 0} = s_\alpha \in W$. On the other hand, it is easy to check (either by straightforward algebra or by drawing a picture) that

$$s_{\alpha 1} \circ s_{\alpha 0}(H) = H + H_\alpha$$

and more generally for $n \in \mathbb{Z}$ that

$$s_{\alpha, n+1} \circ s_{\alpha, n}(H) = H + H_\alpha.$$

From these and the identities $s_{\alpha,n}^2 = 1$, we see that the following subgroups of $\mathrm{GL}(\mathfrak{h}_{\mathbb{R}})$ coincide:

(1) The image of $\mathbb{Z}R^\wedge \rtimes W$.
(2) The group generated by the $s_{\alpha,n}$, for $\alpha \in R$ and $n \in \mathbb{Z}$.

Either group is called the *affine Weyl group*. I'll denote that group W_a.

By the above discussion, W_a acts on the hyperplanes $\{H : \alpha(H) = n\}$ and hence on their complement $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$.

Lemma 285. *W_a acts transitively on the set of connected components of $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$.*

Lemma 286. *Given two such connected components P_0, P_1, take some basepoints $H_i \in P_i$ and draw a path H_t from H_0 to H_1 that crosses at most one of the hyperplanes $\{H : \alpha(H) = n\}$ at a time. One obtains in this way a sequence of pairs $(\alpha_1, n_1), \dots, (\alpha_k, n_k)$ so that as t goes from 0 to 1, the point H_t crosses the planes $\{H : \alpha_i(H) = n_i\}$ in order from $i = 1$ to $i = k$. Then the composition $s_{\alpha_k, n_k} \circ \cdots \circ s_{\alpha_1, n_1}$ maps P_0 to P_1.*

35. The distinguished SU(2)'s

Let notation be as in previous sections: K is a compact connected Lie group, T is a maximal torus in K, plus all the other usual notation.

Let $\mathfrak{k} := \mathrm{Lie}(K)$ and $\mathfrak{g} := \mathfrak{k}_{\mathbb{C}} := \mathfrak{k} \otimes_{\mathbb{R}} \mathbb{C}$, as usual. Let $\theta : \mathfrak{g} \to \mathfrak{g}$ denote the involution given by complex conjugation on the second factor of $\mathfrak{k} \otimes_{\mathbb{R}} \mathbb{C}$, so that when we identify $\mathfrak{k}$ with a real Lie subalgebra of $\mathfrak{g}$, we have $\mathfrak{k} = \{X \in \mathfrak{g} : \theta(X) = X\}$. (In typical examples such as $\mathfrak{k} = \mathfrak{u}(n), \mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$, the operator θ is given by $X \mapsto -{}^t\overline{X}$.) Then $\theta(H) = -H$ for all $H \in \mathfrak{h}_{\mathbb{R}} = i\mathfrak{t}$, $\mathfrak{t} := \mathrm{Lie}(T)$.

Let α belong to the set R of roots for T. Recall that $\mathfrak{g}$ contains the subalgebra $\mathfrak{s}_\alpha = \mathbb{C}H_\alpha \oplus \mathbb{C}X_\alpha \oplus \mathbb{C}Y_\alpha$, which is isomorphism in the evident way to $\mathfrak{sl}_2(\mathbb{C})$. The subalgebra $\mathfrak{k}^{(\alpha)} := \{X \in \mathfrak{s}_\alpha : \theta(X) = X\}$ is a real Lie subalgebra of $\mathfrak{s}_\alpha$. We can easily work out a basis for it. Since $H_\alpha \in \mathfrak{h}_{\mathbb{R}}$, one has $\theta(H_\alpha) = -H_\alpha$. One has $\theta(\mathfrak{g}^\alpha) \subseteq \mathfrak{g}^{-\alpha}$ and $\theta(\mathfrak{g}^{-\alpha}) \subseteq \mathfrak{g}^\alpha$, and $\theta^2 = 1$. It follows that the trace of θ acting on $\mathfrak{s}_\alpha = 0$, so it has the same number of $+1$ and -1 eigenvalues, hence $\dim_{\mathbb{R}}(\mathfrak{k}^{(\alpha)}) = 3$. If we let x, y be any $\mathbb{R}$-basis of the 1-dimensional $\mathbb{C}$-vector space $\mathfrak{g}^\alpha$, then it follows easily that $\{iH_\alpha, x + \theta(x), y + \theta(y)\}$ given an $\mathbb{R}$-basis of $\mathfrak{k}^{(\alpha)}$. Using the existence of a positive-definite K-invariant inner product on $\mathfrak{k}$, we can show that $\mathfrak{k}^{(\alpha)} \cong \mathfrak{su}(2)$. (TODO: explain more.) Since SU(2) is simply-connected, we get from this a morphism

$$F_\alpha : \mathrm{SU}(2) \to K$$

so that $(dF_\alpha)_{\mathbb{C}}$ is the natural map $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{s}_\alpha$. (Compare with homework 17.) The image of F_α is contained in $Z(T_\alpha)^0$. We get an element

$$w_\alpha := F_\alpha(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix})$$

for which $\mathrm{Ad}(w_\alpha)$ gives the root reflection s_α.

We (mostly) explained in lecture how we can use this to identify N/T with the Weyl group W (as defined using root reflections). One of the key steps was to show that if an element w of N/T stabilizes a Weyl chamber C, then it is the identity (i.e., $w \in T$). For this we reduced by an averaging trick to the case that w actually fixes some element $H \in C$. (TODO: explain more.)

36. Proofs regarding the basic homomorphism describnig fundamental groups of compact Lie groups

36.1. **Definition.** Let notation be as usual. Recall that we initially define

$$f : \mathfrak{h}_{\mathbb{Z}} \to \pi_1(K)$$

by taking for $f(H)$ the homotopy class $[\gamma]$ of the path $\gamma : [0,1] \to K$ given by

$$\gamma(t) := e(tH) := \exp(2\pi i t H).$$

36.2. **The basic homomorphism: checking that it's a homomorphism.** Given $H_1, H_2 \in \mathfrak{h}_{\mathbb{Z}}$, we obtain paths $\gamma_1, \gamma_2 : [0,1] \to K$ as above. Their composition in the fundamental group is the path (with domain $[0,1]$)

$$t \mapsto \begin{cases} \gamma_1(2t) & t \leq 1/2 \\ \gamma_2(2t-1) & t \geq 1/2 \end{cases}$$

which we can rewrite as

$$t \mapsto \left(\begin{cases} \gamma_1(2t) & t \leq 1/2 \\ 1 & t \geq 1/2 \end{cases} \right) \cdot \left(\begin{cases} 1 & t \leq 1/2 \\ \gamma_2(2t-1) & t \geq 1/2 \end{cases} \right).$$

Introduce a deformation parameter $s \in [0,1]$. Choose a continuous monotonically decreasing function $c : [0,1] \to [1,2]$ for which $c_0 = 2$ and $c_1 = 1$. The above path is then homotopic to

$$t \mapsto \left(\begin{cases} \gamma_1(c_s t) & c_s t \leq 1 \\ 1 & c_s t \geq 1 \end{cases} \right) \cdot \left(\begin{cases} 1 & 1 + c_s(t-1) \leq 0 \\ \gamma_2(1 + c_s(t-1)) & 1 + c_s(t-1) \geq 0. \end{cases} \right).$$

At deformation parameter $s = 1$, the above path is given by

$$t \mapsto \gamma_1(t)\gamma_2(t) = e(t(H_1 + H_2)).$$

Therefore $f : \mathfrak{h}_{\mathbb{Z}} \to \pi_1(K)$ is a homomorphism.

Exercise 39. Use an argument similar to that above to show that the fundamental group of any topological group is abelian.

36.3. **Checking that some stuff is in its kernel.** We now show that for $H \in \mathbb{Z}R^\wedge$ one has $f(H) = 0$. Since f is a homomorphism and $R^\wedge$ gives a basis for $\mathbb{Z}R^\wedge$, our task reduces to verifying for each $\alpha \in R$ that the path

$$[0,1] \ni t \mapsto e(tH_\alpha) \in K \tag{145}$$

is null-homotopic.

To that end, observe first that $e(\frac{1}{2}H_\alpha) \in T_\alpha$; indeed, since $\alpha(H_\alpha) = 2$,

$$e^\alpha(\tfrac{1}{2}H_\alpha) = e(\alpha(\tfrac{1}{2}H_\alpha)) = e(1) = 1.$$

Recall from §35 that there exists $w_\alpha \in Z(T_\alpha)^0$ for which $\mathrm{Ad}(w_\alpha)H_\alpha = -H_\alpha$. For $s \in [0,1]$, let $c_s \in Z(T_\alpha)^0$ be such that $c_0 = 1$ (the identity element) and $c_1 = w_\alpha$. We then have

$$c_s e(\tfrac{1}{2}H_\alpha) c_s^{-1} = e(\tfrac{1}{2}H_\alpha) \text{ for all } s.$$

For this reason, the path (145) may be continuously deformed to

$$t \mapsto \begin{cases} e(tH_\alpha) & t \leq 1/2 \\ c_s e(tH_\alpha) c_s^{-1} & t \geq 1/2. \end{cases}$$

After we deform to $s = 1$, we get

$$w_\alpha e(tH_\alpha) w_\alpha^{-1} = e(t\,\mathrm{Ad}(w_\alpha)H_\alpha) = e(-tH_\alpha) = e((1-t)H_\alpha),$$

so we deduce that the path (145) is homotopic to

$$t \mapsto \begin{cases} e(tH_\alpha) & t \leq 1/2 \\ e((1-t)H_\alpha) & t \geq 1/2. \end{cases} \tag{146}$$

Now introduce another deformation parameter r, starting at $r = 1/2$ and deforming to $r = 0$. The path (146) is then homotopic to

$$t \mapsto \begin{cases} e(tH_\alpha) & t \leq r \\ e((2r-t)H_\alpha) & r \leq t \leq 2r \\ 1 & 2r \leq t. \end{cases} \tag{147}$$

When $r = 0$, we get the trivial path $t \mapsto 1$. Thus the path (145) is nullhomotopic.

36.4. **The basic homomorphism: checking that it's surjective.** We now argue that f is surjective. Recall that $\pi^1(K) = \pi^1(K^{\mathrm{reg}})$. Let H_0 be a small element of $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$, so that $e(H_0)$ is a small element of T^{reg}. Any element of $\pi_1(K)$ can be deformed a bit to start and end at $e(H_0)$, and then deformed a bit more so that it lies entirely in K^{reg}. Using the covering morphism of §34.3, we can then uniquely lift our path to $K/T \times T^{\mathrm{reg}}$; this means concretely that we may express our path uniquely in the form

$$[0, 1] \ni t \mapsto c_t \exp(H_t) c_t^{-1},$$

where $c_t \in K/T$ satisfies $c_0 = eT$, where H_t belongs to the same connected component P of $\mathfrak{h}_{\mathbb{R}}^{\mathrm{sreg}}$ as H_0 and satisfies $H_t|_{t=0} = H_0$, and finally

$$c_1 \exp(H_1) c_1^{-1} = \exp(H_0). \tag{148}$$

This last condition says in particular that $s := c_1$ satisfies $sts^{-1} \in T^{\mathrm{reg}}$ for some $t := e(H_1) \in T^{\mathrm{reg}}$; since t is regular, it belongs to exactly one maximal torus, and so since $t \in T$ and $t \in s^{-1}Ts$ we deduce that $T = s^{-1}Ts$, i.e., that s belongs to the normalizer $N := N_K(T) := \{g \in K : gTg^{-1} = T\}$ of T. We may also rewrite the condition (148) in the form

$$s \cdot H_1 + H_0 \in \mathfrak{h}_{\mathbb{Z}} := \ker(e : \mathfrak{h}_{\mathbb{R}} \to T), \tag{149}$$

where we abbreviate $s \cdot H_1 := \mathrm{Ad}(s)H_1$.

We now want to deform our path so that it is obviously in the image of f. To that end, let us first translate it by $e(-H_0)$ so that its basepoint is at the origin again. We are then left to stare at the path

$$t \mapsto e(-H_0)c_t e(H_t) c_t^{-1}.$$

Since P is convex and $H_0, H_1 \in P$, there is no harm in assuming that H_t is the straight line path from H_0 to H_1, given by $H_t = H_0 + t(H_1 - H_0)$. Using (149) and the fact that s and s^{-1} preserve $\mathfrak{h}_{\mathbb{Z}}$, we may write

$$H_1 = s^{-1} \cdot H_0 + Z$$

for some $Z \in \mathfrak{h}_{\mathbb{Z}}$. We are thus looking at the path

$$t \mapsto e(-H_0)c_t e(H_0 + t(s^{-1} \cdot H_0 + Z - H_0))c_t^{-1}.$$

In the above path, we may continuously deform H_0 to 0. This gives a family of loops based at the origin. When we reach $H_0 = 0$, we end up with the path

$$t \mapsto c_t e(tZ) c_t^{-1}. \tag{150}$$

Since $Z \in \mathfrak{h}_{\mathbb{Z}}$, we have $e(0Z) = e(1Z) = 1$. So we can now deform every element of c_t to the identity element and get a homotopic path. In other words, we can replace the above path by $t \mapsto c_{\varepsilon t} e(tZ) c_{\varepsilon t}^{-1}$ for $0 \leq \varepsilon \leq 1$; we start with $\varepsilon = 1$, giving (150), and then deform to $\varepsilon = 0$, giving the path

$$t \mapsto e(tZ),$$

which is obviously in the image of f.

36.5. **The basic homomorphism: pinning down the kernel.** We've seen that we have a well-defined surjective map

$$f : \mathfrak{h}_{\mathbb{Z}} / \mathbb{Z} R^{\wedge} \to \pi_1(K).$$

We want to show that it's actually injective. Let's observe first also that for any $Z \in \mathfrak{h}_{\mathbb{Z}}$ and $s \in W = N/T$, we have

$$f(sZ) = f(Z).$$

Indeed, $sZ - Z$ belongs to $\mathbb{Z} R^{\wedge}$ (check this on the generators $s = s_\alpha$ of W, using that $H_\alpha \in R^{\wedge}$), and f is a homomorphism. So this tells us that $f(Z)$ doesn't change if we replace Z with anything in the same orbit under the affine Weyl group $W_a := \mathbb{Z} R^{\wedge} \rtimes W$ (see §34.4). Fix some $H_0 \in \mathfrak{h}_{\mathbb{R}}^{\mathrm{reg}}$ and let P denote its connected component. Since W_a acts transitively on the connected components and since the union of their closures is all of $\mathfrak{h}_{\mathbb{R}}$, any $Z \in \mathfrak{h}_{\mathbb{Z}}$ is in the W_a-orbit of some element of the closure of $P - H_0$; since $Z \in \mathfrak{h}_{\mathbb{Z}}$, it can't lie on the boundary of $P - H_0$ (check this; it's easy), and so must lie in $P - H_0$ itself.

What we want to show now is that the above map is an isomorphism. This means we should show that if $Z \in \mathfrak{h}_{\mathbb{Z}}$ has the property that the path $t \mapsto e(tZ)$ is nullhomotopic, then Z belongs to $\mathbb{Z} R^{\wedge}$. By the above discussion, we may assume that $Z \in P - H_0$. Since P is convex, there is then no loss in shifting basepoints a bit to suppose that we are considering the path in K^{reg} given by

$$\gamma(t) := e(H_0 + tZ).$$

Under the covering map $K/T \times T^{\mathrm{reg}} \to K^{\mathrm{reg}}$, the above path lifts uniquely to

$$\tilde{\gamma}(t) \mapsto (eT, H_0 + tZ).$$

The endpoint $\tilde{\gamma}(1) = tZ$ of this lifting is moreover invariant under base-and-end-point-preserving homotopies of γ. So if γ is nullhomotopic, then we must have $Z = 0$. The proof is now complete.

References