LIE GROUPS

Course **C3.5** HT 2022

Notes by Nigel Hitchin 2015

Lecturer: Frances Kirwan

kirwan@maths.ox.ac.uk

Contents

1 Introduction

A first course on group theory introduces a group as a set with two operations – multiplication and inversion – satisfying appropriate axioms. However the subject only begins to come to life when groups act on other sets: it is easier to think of the symmetric group S_n as the group of permutations of the numbers $\{1, 2, \ldots, n\}$ than as a multiplication table. But these same groups also appear acting on other sets, for example the symmetric group S_4 is isomorphic to the symmetries of the cube, and so acts on $\mathbf{R}^3$ by rotations. Such linear actions of groups are frequent in real life. They are called *representations* and they are studied side-by-side together with the groups considered as objects in their own right.

The symmetric groups are finite groups but there are infinite groups such as the full group of rotations which is what this course is about. Roughly speaking they are infinite but with a finite number of degrees of freedom: a rotation in $\mathbf{R}^3$ is given by an axis (equivalently a unit vector) with two degrees of freedom and an angle of rotation, making three in all. Or the isometries of $\mathbf{R}^2$ – two degrees of freedom for translations and one for rotations making again three. The structure that makes this a consistent mathematical concept is that of a manifold, and that is what a Lie group is: a manifold for which the group operations are smooth maps.

Lie groups are very special manifolds, however. If you think, rightly, of manifolds as higher-dimensional generalizations of surfaces then the only compact, connected surface which is a Lie group is a torus, and the only spheres which are Lie groups are those of dimension one and three. So, although we shall have to review the basic features of manifolds it will be via a viewpoint which is slanted towards this special case.

Manifolds are also topological spaces and the two examples of Lie groups above are quite different: the rotation group is compact and the isometry group of the plane noncompact. The representation theory of compact Lie groups has a lot in common with that of finite groups, and for the most part this course will deal with compact groups and their representations.

This is a rich and well-studied area and a 16-lecture course, which also has to prove the fundamental results, is necessarily restricted in its scope. We introduce roots and the Weyl group but stop short of studying irreducible representations by highest weights. This is a beautiful subject and provides a means of reading off information about irreducibles from combinatorial formulae, but time does not allow us to pursue that here.

We begin in Section 2 with the basics of manifold theory, as applied to Lie groups.

Section 3 defines the Lie algebra of a Lie group and the exponential map which links the two. Section 4 deals with Lie subgroups where we prove in particular the useful result that closed subgroups of Lie groups are Lie groups. Section 5 gives a precise description of the Lie groups which have isomorphic Lie algebras.

The remaining four sections are about representations of compact Lie groups. Section 6 introduces integration on a compact Lie group, orthogonality of characters and the Peter-Weyl theorem which shows that every irreducible representation appears within the Hilbert space of L^2 functions on the group. Section 7 introduces maximal tori and proves the crucial result that all such tori are conjugate. We use a proof based on the degree of a smooth map between manifolds of the same dimension, the higher-dimensional analogue of the winding number. Given that result, many features follow rapidly: the Weyl group, roots and weights. Section 8 introduces the representation ring or character ring and its relation with the action of the Weyl group. The final section states without proof the classification of simply-connected compact simple Lie groups. I have also added a concrete description of the exceptional group G_2 to illustrate some of the features encountered within the course.

2 Manifolds

2.1 Definitions

The concept of a manifold starts with defining the notion of a *coordinate chart*.

Definition 1 *A coordinate chart on a set M is a subset $U \subseteq M$ together with a bijection*

$$\varphi : U \to \varphi(U) \subseteq \mathbf{R}^n$$

onto an open set $\varphi(U)$ in $\mathbf{R}^n$.

Thus we can parametrize points x of U by n coordinates $\varphi(x) = (x_1, \ldots, x_n)$.

Definition 2 *A smooth manifold of dimension n is a set M with a collection of coordinate charts $\{U_\alpha, \varphi_\alpha\}_{\alpha \in I}$ such that*

- *M is covered by the $\{U_\alpha\}_{\alpha \in I}$*
- *for each $\alpha, \beta \in I$, $\varphi_\alpha(U_\alpha \cap U_\beta)$ is open in $\mathbf{R}^n$*
- *the map*

  $$\varphi_\beta \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

  is C^∞ with C^∞ inverse.

Remark: Recall that $F(x_1, \ldots, x_n) \in \mathbf{R}^n$ is C^∞ if it has derivatives of all orders. In manifold theory one can have lower degrees of differentiability with different outcomes, but it is a deep theorem that groups which are manifolds only assuming continuity of the maps above generates the same theory as asserting smoothness.

A manifold is automatically a topological space. Recall what a topological space is: a set M with a distinguished collection of subsets called *open sets* such that

1. $\emptyset$ and M are open
2. an arbitrary union of open sets is open
3. a finite intersection of open sets is open

For a manifold we shall say that a subset $V \subseteq M$ is open if, for each α, $\varphi_\alpha(V \cap U_\alpha)$ is an open set in $\mathbf{R}^n$. It is straightforward to see that this defines a topology, and furthermore the coordinate charts are homeomorphisms. To avoid pathological cases, we make the assumption that the topological space is *Hausdorff*, and has a *countable basis of open sets.*

With this definition one can easily see that the product $M \times N$ of manifolds is another manifold where $\dim(M \times N) = \dim M + \dim N$. All you need to do is look at the product $\varphi_\alpha \times \psi_i$ for charts $\{U_\alpha, \varphi_\alpha\}, \{V_i, \psi_i\}$ on M, N respectively.

Here is the definition of a smooth map between manifolds:

Definition 3 *A map $F : M \to N$ of manifolds is a smooth map if for each point $x \in M$ and chart $(U_\alpha, \varphi_\alpha)$ in M with $x \in U_\alpha$ and chart (V_i, ψ_i) of N with $F(x) \in V_i$, the set $F^{-1}(V_i)$ is open and the composite function*

$$\psi_i F \varphi_\alpha^{-1}$$

on $\varphi_\alpha(F^{-1}(V_i) \cap U_\alpha)$ is a C^∞ function.

A smooth map is continuous in the manifold topology.

The natural notion of equivalence between manifolds is the following:

Definition 4 *A diffeomorphism $F : M \to N$ is a smooth map with smooth inverse.*

The composition of two diffeomorphisms is again a diffeomorphism and in particular $\mathrm{Diff}(M)$, the set of all diffeomorphisms $F : M \to M$ is a group, but far too large to be a Lie group.

2.2 Lie groups

Finally we can define a Lie group:

Definition 5 *A Lie group G is a smooth manifold which is also a group and is such that*

- *the multiplication map $\mu : G \times G \to G$*
- *and the inversion map $i : G \to G$*

are smooth maps of manifolds.

Then we have the natural notion:

Definition 6 *A Lie group homomorphism $\gamma : G \to H$ is a smooth map which is also a group homomorphism.*

Examples:

1. The *general linear group* $GL(n, \mathbf{R})$ of all invertible $n \times n$ matrices is the open subset of the n^2-dimensional vector space of all $n \times n$ matrices given by $\det A \neq 0$. The inclusion $GL(n, \mathbf{R}) \subset \mathbf{R}^{n^2}$ is a single chart: the coordinates are the entries $a_{ij}, 1 \leq i, j \leq n$ of the matrix A. The entries of the product AB are polynomials in $a_{ij}, b_{k\ell}$ and so μ is smooth. Similarly the determinant is a smooth function. We can represent $A^{-1} = (\det A)^{-1} \operatorname{adj} A$ where A, the adjugate matrix, is the matrix of signed cofactors. The entries of $\operatorname{adj} A$ are polynomials in a_{ij} of degree $(n-1)$ and $\det A$ is a non-vanishing polynomial of degree n so the inversion map, a ratio of polynomials, is smooth. So $G = GL(n, \mathbf{R})$ is a Lie group (noncompact since $\det A$ is an unbounded continuous function).

2. A standard way to produce manifolds is via the implicit function theorem. A precise statement is the following:

Theorem 2.1 *Let $F : U \to \mathbf{R}^m$ be a C^∞ function on an open set $U \subseteq \mathbf{R}^{n+m}$ and take $c \in \mathbf{R}^m$. Assume that for each $a \in F^{-1}(c)$, the derivative*

$$DF_a : \mathbf{R}^{n+m} \to \mathbf{R}^m$$

is surjective. Then $F^{-1}(c)$ has the structure of an n-dimensional manifold. Moreover the manifold topology is the induced topology which is therefore Hausdorff and has a countable basis of open sets.

In this construction, functions of $\mathbf{R}^{n+m}$ which are smooth restrict to smooth functions on $F^{-1}(c)$. The standard example is the sphere where we take $F : \mathbf{R}^{n+1} \to \mathbf{R}$ to be $F(x) = x_1^2 + \cdots + x_{n+1}^2$ and $c = 1$. If $n = 1$ then the sphere is the unit complex numbers, which is a group. Moreover complex multiplication $(x_1 + ix_2)(y_1 + iy_2)$ is smooth, as is inversion $(x_1 + ix_2) \mapsto (x_1 - ix_2)$ so the circle is a Lie group.

The 3-sphere is also a group: it is the group $SU(2)$ of 2×2 unitary matrices with determinant one. To see this take

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

If A is unitary then the first row is a unit complex vector so $a\bar{a}+b\bar{b}=1$. This is the unit sphere $(a,b)\in\mathbf{C}^2$. But by unitarity $a\bar{c}+b\bar{d}=0$ and since $\det A=1$, $ad-bc=1$. Solve these equations for (c,d) and we get $c=-\bar{b}, d=\bar{a}$ so the first row determines everything. Multiplication and inversion $A^{-1}=A^*$ are clearly smooth.

The circle S^1 and $SU(2)$ are the simplest compact Lie groups and we shall frequently use them as examples. The circle is abelian but $SU(2)$ is not.

3. The product of m copies of the circle S^1 is a *torus* which we shall denote by T^m. It is abelian, and is important since we shall see that every compact Lie group contains distinguished tori.

4. The standard use of Theorem 2.1 is to the group $O(n)$ of orthogonal matrices: the space of $n\times n$ real matrices such that $AA^T=1$. Take the vector space $\mathbf{R}^{n^2}$ of all real $n\times n$ matrices and define the function

$$F(A)=AA^T$$

to the vector space of *symmetric* $n\times n$ matrices. This has dimension $n(n+1)/2$. Then $O(n)=F^{-1}(I)$.

We have

$$F(A+H)=AA^T+HA^T+AH^T+R(A,H)$$

where the remainder term $\|R(A,H)\|/\|H\|\to 0$ as $H\to 0$. So the derivative at A is

$$DF_A(H)=HA^T+AH^T$$

and putting $H=KA$ this is

$$KAA^T+AA^TK^T=K+K^T$$

if $AA^T=I$, i.e. if $A\in F^{-1}(I)$. But given any symmetric matrix S, taking $K=S/2$ shows that DF_A is surjective and so, applying Theorem 2.1 we find that $O(n)$ is a manifold. Its dimension is $n^2-n(n+1)/2=n(n-1)/2$. The smoothness of multiplication and inversion follow from that of $GL(n,\mathbf{R})$.

Remark: The role of K in this last example illustrates a particular point about Lie groups as manifolds. If A was the identity then the derivative would be just $H+H^T$. What we have done to get K from H is to right-multiply by A^{-1}. In other words translated back from A to the identity. Here is a point which distinguishes Lie groups from other manifolds – every point looks like every other one by translation. In particular once we know we have a manifold, a chart in a neighbourhood of the identity can be translated to a chart anywhere else on the manifold.

5. Here is another chart for $O(n)$. Let S be a real skew-symmetric $n \times n$ matrix, so $S^T = -S$. Then S has eigenvalues which are either complex or zero so $I + S$ is invertible. Then

$$((I - S)(I + S)^{-1})^T = ((I + S)^{-1})^T (I - S)^T = (I - S)^{-1}(I + S).$$

But $(I + S), (I - S)$ commute, so this is

$$(I + S)(I - S)^{-1} = ((I - S)(I + S)^{-1})^{-1}.$$

So $A = (I - S)(I + S)^{-1}$ is orthogonal. Now rewrite this as $S(I + A) = I - A$. The matrix $I + A$ is invertible if A has no eigenvalue -1, and this is a long way from the identity, so S is uniquely determined by A and we have a chart for a neighbourhood of the identity. By translation we get similar charts everywhere, with local coordinates skew-symmetric matrices.

Each Lie group has some natural diffeomorphisms: for a fixed element $h \in G$ we have *left translation* $L_h x = hx$. From the smoothness of multiplication this is a smooth map and it is invertible with inverse $L_{h^{-1}}$ of the same form so it is a diffeomorphism. Since $L_g L_h x = L_{gh} x$ this is a homomorphism of groups $G \to \mathrm{Diff}(G)$. We also have *right translation* $R_g x = xg$ and *conjugation* $x \mapsto gxg^{-1}$. Note that $R_{g^{-1}} R_{h^{-1}} x = xh^{-1}g^{-1} = x(gh)^{-1}$ so a $R_{g^{-1}}$ defines an action of G.

2.3 Functions and vector fields

We defined above a smooth map between manifolds. The simplest ones are smooth maps to $\mathbf{R}$. For example the trace $\operatorname{tr} A = \sum_i A_{ii}$ for any of the groups of matrices above is a natural smooth function. For a rotation in $\mathbf{R}^3$, $\operatorname{tr} A = 1 + 2\cos\theta$ where θ is the angle of rotation. But if a function is differentiable we have to know what its derivative is.

The most convenient way to do this is to generalize the *directional derivative* of a function f of several variables

$$\sum_i c_i \frac{\partial f}{\partial x_i}$$

– the derivative of f in the direction $(c_1, \ldots, c_n)$. Since manifolds do not sit naturally in $\mathbf{R}^N$ in which the direction is a vector in the ambient space we *define* a tangent to a manifold by a linear map with the same formal properties as the directional derivative.

Definition 7 *A tangent vector at a point $a \in M$ is a linear map $X_a : C^\infty(M) \to \mathbf{R}$ such that*

$$X_a(fg) = f(a)X_a g + g(a)X_a f.$$

This is the formal version of the Leibnitz rule for differentiating a product, and in a chart $\{U, \varphi\}$ where $\varphi = (x_1, \ldots, x_n)$

$$X_a(f) = \sum_i c_i \frac{\partial f}{\partial x_i}(\varphi(a)) \stackrel{\text{def}}{=} \sum_i c_i \left(\frac{\partial}{\partial x_i}\right)_a f$$

The tangent space at $a \in M$ is the vector space $T_a M$ of all tangent vectors at a and in a coordinate system has as a basis the n tangent vectors

$$\left(\frac{\partial}{\partial x_1}\right)_a, \ldots, \left(\frac{\partial}{\partial x_n}\right)_a.$$

Defining the tangent space this way provides an abstract, coordinate-free definition of the derivative of a map of manifolds:

Definition 8 *The derivative at $a \in M$ of the smooth map $F : M \to N$ is the homomorphism of tangent spaces*

$$DF_a : T_a M \to T_{F(a)} N$$

defined by

$$DF_a(X_a)(f) = X_a(f \circ F).$$

Concretely, this becomes

$$\begin{aligned} DF_a \left(\frac{\partial}{\partial x_i}\right)_a (f) &= \frac{\partial}{\partial x_i}(f \circ F)(a) \\ &= \sum_j \frac{\partial F_j}{\partial x_i}(a) \frac{\partial f}{\partial y_j}(F(a)) = \sum_j \frac{\partial F_j}{\partial x_i}(a) \left(\frac{\partial}{\partial y_j}\right)_{F(a)} f \end{aligned}$$

Thus the derivative of F is an invariant way of defining the Jacobian matrix. Moreover the derivative of the composition of two maps is the composition of the derivatives. This in coordinates is the chain rule.

A particular case is when $N = \mathbf{R}$ and then the derivative of a smooth function f, denoted df, is a linear map from $T_a M$ to $\mathbf{R}$: an element of the dual space $T_a^* M$ called the cotangent space.

We need the notion of a smoothly varying family of tangent vectors, like the wind velocity at each point of the earth's surface. This is provided by the following

Definition 9 *A vector field is a linear map $X : C^\infty(M) \to C^\infty(M)$ which satisfies the Leibnitz rule*

$$X(fg) = f(Xg) + g(Xf)$$

so for each $a \in M$ the evaluation of Xf at a gives a tangent vector. In local coordinates

$$X(f) = \sum_i c_i(x) \left(\frac{\partial f}{\partial x_i}\right)_x$$

where each $c_i(x)$ is a smooth function.

Any object like a vector field which is defined intrinsically in a coordinate-free fashion can be transformed by a diffeomorphism. For a vector field X and $F \in \mathrm{Diff}(M)$ we define a vector field F_*X by using the derivative $DF_a : T_aM \to T_{F(a)}M$:

$$(F_*X)_{F(x)} = DF_x(X_x).$$

If $F_*X = X$ then we say X is invariant by F. We then define

Definition 10 *On a Lie group G, a vector field X is left-invariant if $(L_g)_*X = X$ for each $g \in G$, where L_g is the left-translation diffeomorphism.*

The left-invariant vector fields are very important. We can construct one by taking a tangent vector $X_e \in T_eG$, the tangent space at the identity $e \in G$, and defining

$$X_g = (DL_g)X_e$$

for then

$$((L_g)_*X)_{gh} = DL_g(X_h) = DL_g(DL_hX_e) = D(L_g \circ L_h)X_e = DL_{gh}(X_e) = X_{gh}.$$

Conversely, if X is left-invariant then $(L_g)_*X = X$ so $X_g = DL_gX_e$ and this is true for all g. Thus the vector space of left-invariant vector fields is isomorphic to T_eG and is a vector space of dimension $\dim G$.

Remark: If $X_e^1, \ldots, X_e^n$ is a basis for T_eG then the corresponding left-invariant vector fields $X^1, \ldots, X^n$ form a basis for each tangent space T_gG. In particular each X^i is non-vanishing so the Euler characteristic of G vanishes. This is the first evidence that Lie groups are very special manifolds. The topological consequences are much stronger since we have not just one but a basis – *any* vector field can be written uniquely as

$$X = \sum_i f_iX^i$$

for globally defined smooth functions f_i.

Taking the example of a vector field as the wind velocity (assuming it is constant in time) on the surface of the earth, we can go further to say that a particle at position x moves after time t seconds to a position $\varphi_t(x)$. After a further s seconds it is at

$$\varphi_{t+s}(x) = \varphi_s(\varphi_t(x)).$$

What we get this way is a homomorphism of groups from the additive group $\mathbf{R}$ to $\mathrm{Diff}(M)$. Since $\mathrm{Diff}(M)$ is not a Lie group we can't use our definition of a smooth map but the technical definition is the following:

Definition 11 *A one-parameter group of diffeomorphisms of a manifold M is a smooth map*

$$\varphi : M \times \mathbf{R} \to M$$

such that (writing $\varphi_t(x) = \varphi(x,t)$)

- $\varphi_t : M \to M$ *is a diffeomorphism*
- $\varphi_0 = id$
- $\varphi_{s+t} = \varphi_s \circ \varphi_t$.

To a one-parameter group of diffeomorphisms we can associate a vector field: given φ_t and f a smooth function, then for each $a \in M$

$$f(\varphi_t(a))$$

is a smooth function of t and we write

$$\frac{\partial}{\partial t} f(\varphi_t(a))|_{t=0} = X_a(f).$$

It is straightforward to see that, differentiating a product with respect to t, the Leibnitz rule holds and since $\varphi_0(a) = a$ this is a tangent vector at a. So as $a = x$ varies we have a vector field. In local coordinates we have

$$\varphi_t(x_1, \ldots, x_n) = (y_1(x,t), \ldots, y_n(x,t))$$

and

$$\begin{aligned} \frac{\partial}{\partial t} f(y_1, \ldots, y_n) &= \sum_i \frac{\partial f}{\partial y_i}(y) \frac{\partial y_i}{\partial t}(x)|_{t=0} \\ &= \sum_i c_i(x) \frac{\partial f}{\partial x_i}(x) \end{aligned}$$

which yields the vector field

$$X = \sum_i c_i(x) \frac{\partial}{\partial x_i}.$$

We now want to reverse this: go from the vector field to the diffeomorphism. The first point is to track the "trajectory" of a single particle.

Definition 12 *An integral curve of a vector field X is a smooth map of an interval $\gamma : (\alpha, \beta) \subset \mathbf{R} \to M$ such that*

$$D\gamma \left(\frac{d}{dt} \right) = X_{\gamma(t)}.$$

In a coordinate chart (U, ψ) around a then if

$$X = \sum_i c_i(x) \frac{\partial}{\partial x_i}$$

the equation for γ to define an integral curve can be written as the system of ordinary differential equations

$$\frac{dx_i}{dt} = c_i(x_1, \ldots, x_n).$$

The existence and uniqueness theorem for ODE's asserts that there is some interval on which there is a unique solution with initial condition

$$(x_1(0), \ldots, x_n(0)) = \psi(a).$$

A further theorem on ODEs says that in a suitably small neigbourhood the solution has smooth dependence on the initial conditions. The curve γ then depends on the initial point x and we write $\gamma(t) = \varphi_t(x)$ where $(t, x) \mapsto \varphi_t(x)$ is smooth. Now consider $\varphi_t \circ \varphi_s(x)$. If we fix s and vary t, then this is the unique integral curve of X through $\varphi_s(x)$. But $\varphi_{t+s}(x)$ is an integral curve which at $t = 0$ passes through $\varphi_s(x)$. By uniqueness they must agree so that $\varphi_t \circ \varphi_s = \varphi_{t+s}$. (Note that $\varphi_t \circ \varphi_{-t} = id$ shows that we have a diffeomorphism wherever it is defined).

Our conclusion is that we have the group law for these diffeomorphisms but only for s, t, x in a small neighbourhood.

Example: Take M to be the one-dimensional manifold $(0, \infty) \subset \mathbf{R}$ and $X = d/dx$. Then the ODE is

$$\frac{dx}{dt} = 1$$

and so $\varphi_t(x) = t + x$ and this is only defined for $t > -x$. On the other hand if we take $X = xd/dx$ we have

$$\frac{dx}{dt} = x$$

and $\varphi_t(x) = e^t x$ which is positive if $x > 0$ and so is defined for all t.

In fact $\lambda x d/d(\lambda x) = xd/dx$. If we regard $(0, \infty)$ as a Lie group with multiplication of positive reals as the operation, then this says that X is left-invariant. Moreover if we set $x = 1$, the identity element, then the integral curve $\varphi_t(1) = e^t$ is a smooth map from $\mathbf{R}$ to $(0, \infty)$ which is a group homomorphism.

The example above is true more generally:

Theorem 2.2 *Let G be a Lie group and X a left-invariant vector field. Then*

- *the integral curve $\varphi_t(a)$ through $a \in G$ exists for all $t \in \mathbf{R}$*
- *when $x = e$ (the identity element)*

  $$\varphi_t(e) : \mathbf{R} \to G$$

  is a group homomorphism
- *any Lie group homomorphism $\gamma : \mathbf{R} \to G$ arises this way*

Proof: (i) First note that if $\varphi_t(a)$ is an integral curve for X through $a \in G$ then $g\varphi_t(a)$ is an integral curve through ga. This holds because

$$D\varphi_t(x)\left(\frac{d}{dt}\right) = X_{\varphi_t(x)}$$

and so, because X is left-invariant,

$$(DL_g)_{\varphi_t(x)} \circ D\varphi_t\left(\frac{d}{dt}\right) = (DL_g)_{\varphi_t(x)}(X_{\varphi_t(x)}) = X_{g\varphi_t(x)}.$$

But by the chain rule the left hand side is

$$D(L_g\varphi_t)\left(\frac{d}{dt}\right)$$

which shows that $L_g\varphi_t$ is an integral curve, and since $\varphi_0(a) = a$, $L_g\varphi_0(a) = ga$.

This gives us two facts: the chart on which we have local existence and uniqueness can be shifted around by left translation to give the same result in a neighbourhood of any point; and if $t \in (-\epsilon, \epsilon)$ is the interval on which the existence theorems work, the same interval works for any of these neighbourhoods.

Consider now the curve

$$\psi_t = \varphi_{\epsilon/2}(a)a^{-1}\varphi_{t-\epsilon/2}(a).$$

This is well-defined for $t \in (-\epsilon/2, 3\epsilon/2)$ and when $t = \epsilon/2$, $\psi_t = \varphi_{\epsilon/2}(a)a^{-1}a = \varphi_{\epsilon/2}(a)$. It defines an integral curve because it is a left-translate of an integral curve. However, it agrees with φ_t at $t = \epsilon/2$ and so by uniqueness it extends the solution to the larger interval $(-\epsilon, 3\epsilon/2)$. Continuing in this way at both ends of the interval, suppose that (a, b) is a maximal interval on which $\varphi_t(a)$ is defined. Then if b is finite we have a solution on $(a, b-\epsilon/4)$ but then the same argument extends it to $(a, b+\epsilon/4)$ giving a contradiction. So the integral curve is defined for all $t \in \mathbf{R}$.

(ii) If $\varphi_t(e)$ is the integral curve through e, then consider $\varphi_s(e)\varphi_t(e)$ (group multiplication) for fixed s. Since $\varphi_0(e) = e$, $\varphi_s(e)\varphi_t(e)$ and $\varphi_{s+t}(e)$ agree at $t = 0$. Moreover $\varphi_s(e)\varphi_t(e)$ is left translation by $g = \varphi_s(e)$ and is therefore an integral curve of X. From uniqueness we must then have $\varphi_s(e)\varphi_t(e) = \varphi_{s+t}(e)$ and a group homomorphism.

(iii) If $\gamma : \mathbf{R} \to G$ is a Lie group homomorphism then $R_{(\gamma_t)^{-1}}$ (the *right action* $\varphi_t(x) = x\gamma^{-1}(t)$) is a one-parameter group of diffeomorphisms. Since $(gx)h = g(xh)$, left and right actions commute. So the one-parameter group φ_t commutes with all left translations. The vector field obtained by differentiation with respect to t at $t = 0$ is therefore left-invariant. From part (ii) we see that the integral curve through $\varphi_s(e)$ is $\varphi_s(e)\varphi_t(e)$ i.e. right multiplication by the group so this is obtained by the same process.

□

3 Lie groups and Lie algebras

3.1 The Lie bracket

We saw in the previous section how a one-parameter group of diffeomorphisms on a manifold gives rise to a vector field by

$$\frac{\partial}{\partial t} f(\varphi_t(x))|_{t=0} = Xf(x).$$

We can do the same with vector fields, namely use the diffeomorphism φ_t to transform the vector field Y into $(\varphi_t)_* Y$ and differentiate at $t = 0$. If this sounds worrying – differentiating a map from $\mathbf{R}$ into the infinite-dimensional space of vector fields – just note that these are vector fields with a parameter t so we are just differentiating a tangent vector at each point:

$$\sum_i \frac{\partial c_i(x,t)}{\partial t} \frac{\partial f}{\partial x_i}.$$

The result is called the *Lie derivative* of Y in the direction X, denoted by $\mathcal{L}_X Y$. From the action of diffeomorphisms on vector fields we have

$$(\varphi_t)_* Y(f(\varphi_t(x))) = (Yf)(\varphi_t(x))$$

and differentiating at $t = 0$ we obtain

$$(\mathcal{L}_X Y)f + YXf = XYf.$$

Thus, as operators on the space of smooth functions we have $\mathcal{L}_X Y = XY - YX$. This is also denoted by $[X, Y]$ and called the *Lie bracket* of two vector fields. (One can also check directly that $[X, Y]$ satisfies the Leibnitz formula which characterizes vector fields, but we have given here a more general context for this.)

Example: In $\mathbf{R}^n$ we have

$$\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right] f = \frac{\partial^2 f}{\partial x_i \partial x_j} - \frac{\partial^2 f}{\partial x_j \partial x_i} = 0$$

for all f so the Lie bracket of these standard vector fields vanishes.

Remark: If $[X, Y] = 0$ more generally, then $\partial(\varphi_{s+t})_* Y/\partial t = 0$ at $t = 0$ so $((\varphi_s)_* Y)_a$ is the value Y at $\varphi_s(a)$. In other words the one parameter group φ_t defined by X preserves the vector field Y. This means the integral curve $\psi_t(\varphi_s(a))$ of Y through $\varphi_s(a)$ is the transform $\varphi_s(\psi_t(a))$ of the integral curve through a. In other words the two one parameter groups commute, at least where they are both defined.

Expanding all terms $[X, Y] = XY - YX$ and cancelling pairs one can easily see that

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Now suppose we consider a Lie group and X, Y are left-invariant vector fields. In Theorem 2.2 we saw that the one-parameter group for X consists of right multiplication so $(\varphi_t)_* Y$ is left-invariant, and differentiating at $t = 0$ we see that the Lie bracket $[X, Y]$ is also left-invariant.

Definition 13 *A Lie algebra is a vector space V with a map $[\ ,\ ] : V \times V \to V$ such that*

- *the bracket $[\ ,\ ]$ is bilinear*
- $[X, Y] = -[Y, X]$
- *Jacobi's identity holds:* $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$

So we see that the left-invariant vector fields on a Lie group G form a finite-dimensional Lie algebra $\mathfrak{g}$.

Example: Take $G = GL(n, \mathbf{R})$. This is an open set in $\mathbf{R}^{n^2}$ and so the tangent space at any point can be considered as an element in this vector space. In particular, the Lie algebra of G is isomorphic to the tangent space at the identity T_I so it may be considered as the space of all $n \times n$ matrices. We need the Lie bracket, so take the matrix C to lie in the tangent space at the identity. Then AC, the left translate by A, is the value of the corresponding left-invariant vector field Y at $A \in GL(n, \mathbf{R})$. To compute the Lie bracket we need to right multiply by γ^{-1} for a homomorphism $\gamma : \mathbf{R} \to G$.

For an $n \times n$ matrix B define

$$\exp tB = I + tB + \frac{t^2}{2!}B^2 + \frac{t^3}{3!}B^3 + \ldots$$

Using the usual estimates $\|AB\| \leq \|A\|\|B\|$ one can show that this is a smooth map from $\mathbf{R}$ to the space of matrices and it is invertible with inverse $\exp(-tB)$. It is a Lie group homomorphism $\gamma : \mathbf{R} \to GL(n, \mathbf{R})$ and

$$\frac{\partial}{\partial t} \exp tB|_{t=0} = B.$$

Right multiplying by γ^{-1} gives $AC\exp(-tB)$ so this is the value of the vector field $(\varphi_t)_*Y$ at $A\exp -tB$ where the diffeomorphism φ_t is $\varphi_t(g) = g\gamma^{-1}(t)$. Left-translating back to the origin gives

$$(A\exp(-tB))^{-1}AC\exp(-tB) = (\exp tB)C(\exp(-tB)) = I + t(BC - CB) + \dots$$

Differentiating at $t = 0$ we see that $\mathcal{L}_X Y = [X, Y]$ is the left translate of the commutator $BC - CB$.

So for $GL(n, \mathbf{R})$ the Lie algebra is the space of $n \times n$ matrices with $[A, B] = AB - BA$. Virtually all our examples are subgroups of $GL(n, \mathbf{R})$ where the same picture holds.

3.2 Examples of Lie groups

1. $\mathbf{SL(n, R)}$

The *special linear group* is the group of $n \times n$ matrices with determinant 1. It is non-compact since $\operatorname{tr}(\operatorname{diag}(\lambda, \lambda^{-1}, 1 \dots, 1)) = \lambda + \lambda^{-1}$ so $\operatorname{tr} : G \to \mathbf{R}$ is an unbounded continuous function. To prove this is a manifold we view it as $\det^{-1}(1)$ for the real-valued smooth function det on the space of $n \times n$ matrices. To use the implicit function theorem, as in the case of $O(n)$, it suffices to work at the identity and

$$\det(I + H) = 1 + \operatorname{tr} H + R(H)$$

for a remainder term where $|R(H)|/\|H\| \to 0$. The tangent space at the identity is the kernel of the derivative map, the space of matrices of trace 0 and this is the Lie algebra. The Lie bracket is again the commutator of the matrices. We could also use the complex numbers to obtain $SL(n, \mathbf{C})$.

2. $\mathbf{SO(n)}$

The determinant of an orthogonal matrix is ± 1 and det is continuous, so $O(n)$ is not connected as a topological space. The group $SO(n)$ the *special orthogonal group* is the subgroup with determinant $+1$. In fact $SO(n)$ is connected and is the connected component of $O(n)$ which contains the identity. This can be seen be using the eigenvalues of $A \in SO(n)$, which are either ± 1 or complex conjugates $e^{\pm i\theta_j}$.

If A preserves a subspace $V \subset \mathbf{R}^n$ then it preserves the orthogonal complement. Suppose it has a complex eigenvalue $e^{i\theta_j}$ with eigenvector v. Then there is a real invariant 2-dimensional subspace spanned by the real and imaginary parts of v and A acts as

$$\begin{pmatrix} \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j \end{pmatrix}$$

Then

$$\begin{pmatrix} \cos t\theta_j & \sin t\theta_j \\ -\sin t\theta_j & \cos t\theta_j \end{pmatrix}$$

with $t \in [0,1]$ connects A in $SO(n)$ to a matrix which is the identity on this subspace. Continuing this way, we end up with $A' \in SO(n)$ with eigenvalues ± 1. But $\det A' = 1$ so there is an even number of -1s. But then note that $\theta_j = \pi$ in the above works the same way for every 2-dimensional eigenspace with eigenvalue -1. So we can connect to the identity.

The Lie algebra is the space of skew symmetric matrices with commutator as bracket.

This group is compact: in fact each row of an orthogonal matrix has unit length and so $O(n)$ is a closed subspace of the product of spheres $S^{n-1} \times \cdots \times S^{n-1} \subset \mathbf{R}^{n^2}$ which is closed and bounded hence compact.

Instead of the Euclidean positive definite bilinear form $(\ ,\ )$ we can take one with p positive terms and q negative ones, and then the group of matrices such that $(Ax, Ay) = (x, y)$ is denoted $O(p, q)$. This group is noncompact and there may be more than two components in this case.

3. $\mathbf{Sp(2m, R)}$

This, the *real symplectic group*, is the group of $2m \times 2m$ matrices preserving a nondegenerate skew-symmetric bilinear form $(\ ,\ )$, i.e. $(Ax, Ay) = (x, y)$. It is noncompact. If $(x, y) = \sum B_{ij} x_i y_j$ for some basis and skew-symmetric matrix B, then C is in the Lie algebra if $(Cx, y) + (x, Cy) = 0$ or $C^T B + BC = 0$. But

$$(BC)^T = C^T B^T = -C^T B = BC$$

and hence the Lie algebra is isomorphic to the space of symmetric $2m \times 2m$ matrices S. If $S = BC, T = BD$ then the Lie bracket is $B(CD - DC)$.

3. $\mathbf{U(m)}$

The *unitary group* is the group of unitary $m \times m$ matrices, i.e. $A^* = \bar{A}^T = A^{-1}$. This is compact and connected by a similar argument to $SO(n)$. Its Lie algebra is the space of skew-Hermitian matrices $C^* = -C$. The determinant is now a Lie group homomorphism to the unit complex numbers, another Lie group.

3. $\mathbf{SU(m)}$

The *special unitary group* is the subgroup of $U(m)$ for which $\det A = 1$. Its Lie algebra is the space of skew-Hermitian matrices of trace zero.

3.3 The adjoint representation

If V is a vector space (real or complex) then $\mathrm{Aut}(V)$ is the group of invertible linear transformations from V to V. By choosing a basis it is isomorphic to either $GL(n,\mathbf{R})$ or $GL(n,\mathbf{C})$ where $n = \dim V$.

Definition 14 *A representation of a Lie group G on V is a Lie group homomorphism $G \to \mathrm{Aut}(V)$.*

All our examples above come with a particular representation – they are defined as subgroups of $\mathrm{Aut}(V)$ for $V = \mathbf{R}^n$ or $\mathbf{C}^n$ for some n. But for any Lie group we can take $V = \mathfrak{g}$ to be the space of left-invariant vector fields and the action of right translation makes this a representation of G. It is called the *adjoint representation*. For $GL(n,\mathbf{R})$ it has dimension n^2 – much bigger than the defining representation – but for some groups (especially one called E_8 of dimension 248) it is the smallest non-trivial representation. The action of g on $X \in \mathfrak{g}$ is denoted by

$$\mathrm{Ad}\, g\,(X).$$

Since X is left-invariant we can also consider the right action as the conjugation action:

$$\mathrm{Ad}\, g = (L_g)_*(R_{g^{-1}})_* X$$

and conjugation $C_g(x) = gxg^{-1}$ fixes the identity so identifying $\mathfrak{g}$ with T_e the adjoint action can also be written

$$DC_g : T_e \to T_e.$$

In particular, if $\gamma : \mathbf{R} \to G$ is a homomorphism with tangent vector X at e, then $g\gamma g^{-1}$ has tangent vector $Adg(X)$.

In this case the representation space has the extra structure of the Lie bracket and we need to know whether $\mathrm{Ad}(g)$ preserves this. So given left-invariant vector fields X, Y let γ be the homomorphism determined by X then

$$[\mathrm{Ad}\, g(X), \mathrm{Ad}\, g(Y)] = \frac{\partial}{\partial t}(R_{(g\gamma g^{-1})^{-1}})_*(R_{g^{-1}})_* Y|_{t=0} = \frac{\partial}{\partial t}(R_{g^{-1}})_*(R_{\gamma^{-1}})_* Y|_{t=0}$$

and the right hand term is

$$(R_{g^{-1}})_* \frac{\partial}{\partial t}(R_{\gamma^{-1}})_* Y|_{t=0} = \mathrm{Ad}\, g([X,Y]).$$

Consider now the homomorphism of Lie groups $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$. The group $\mathrm{Aut}(\mathfrak{g})$ is just the invertible linear transformations from $\mathfrak{g}$ to itself. The tangent space at I is all such linear transformations $\mathrm{End}(\mathfrak{g})$. The derivative at the identity of Ad is then a linear map

$$D\,\mathrm{Ad}_e : T_e(\cong \mathfrak{g}) \to \mathrm{End}(\mathfrak{g}).$$

The Lie bracket $[X, Y]$ of left-invariant vector fields is the derivative at $t = 0$ of $(R_{\gamma^{-1}})_* Y$ where γ is the one-parameter subgroup which is an integral curve of X. Since the adjoint action is right translation on left-invariant vector fields $(D\,\mathrm{Ad}_e)X_e$ is the directional derivative of $(R_{\gamma^{-1}})$ at the identity, which is $[X, Y]_e$. This endomorphism of the Lie algebra is denoted by $\mathrm{ad}\, X$ so that

$$\mathrm{ad}\, X(Y) = [X, Y].$$

Examples:

1. If $G = GL(n, \mathbf{R})$ then the Lie algebra is the space of all $n \times n$ matrices X and $AdA\,(X) = AXA^{-1}$. And then $\mathrm{ad}\, X(Y) = XY - YX$.

2. Take $G = SO(3)$, the Lie algebra is the 3-dimensional space of all 3×3 skew symmetric matrices. Recall the vector cross-product

$$\mathbf{a} \times \mathbf{b} = (\|\mathbf{a}\| \|\mathbf{b}\| \sin\theta)\mathbf{n}$$

where $\mathbf{n}$ is a unit vector orthogonal to $\mathbf{a}$ and $\mathbf{b}$ and such that $\mathbf{a}, \mathbf{b}, \mathbf{n}$ has a right-handed orientation and θ is the angle between $\mathbf{a}$ and $\mathbf{b}$. If $A \in SO(3)$, it preserves lengths, angles and orientation and so

$$A(\mathbf{a} \times \mathbf{b}) = (A\mathbf{a} \times A\mathbf{b}).$$

Now $\mathbf{x} \mapsto \mathbf{a} \times \mathbf{x}$ is skew-symmetric with respect to the symmetric bilinear form $\mathbf{x} \cdot \mathbf{y}$ since

$$(\mathbf{a} \times \mathbf{x}) \cdot \mathbf{y} = [\mathbf{a}, \mathbf{x}, \mathbf{y}]$$

which is skew in all its entries, so we can describe the Lie algebra as vectors $\mathbf{a} \in \mathbf{R}^3$. Then for $A \in SO(3)$

$$AdA(\mathbf{a})(\mathbf{x}) = A(\mathbf{a} \times A^{-1}\mathbf{x}) = A\mathbf{a} \times AA^{-1}\mathbf{x} = A\mathbf{a} \times \mathbf{x}$$

and the adjoint action is just the usual action on vectors. So

$$\mathrm{ad}\, \mathbf{a}(\mathbf{x}) = \mathbf{a} \times \mathbf{x}.$$

We saw above that $\mathrm{Ad}\, g$ preserves the Lie bracket on $\mathfrak{g}$. This is part of a more general fact:

Proposition 3.1 *Let $\varphi : G \to H$ be a Lie group homomorphism, then the linear map $D\varphi_e : T_eG \to T_eH$ preserves the Lie bracket.*

In other words $D\varphi$ defines a homomorphism of Lie algebras $\mathfrak{g} \to \mathfrak{h}$. The case of Ad is the conjugation homomorphism $\varphi : G \to G$ for $\varphi(x) = gxg^{-1}$.

Proof: Take $Y_e \in TG_e$. Then $(L_g)_* Y_e$ is the left-invariant vector field Y defined by Y_e. Given $X_e \in T_eG$ let γ be the one-parameter subgroup tangent to X_e at e. Then at $g \in G$ by definition

$$[X, Y]_g = \frac{\partial}{\partial t}(R_{\gamma^{-1}})_*(L_g)_* Y_e.$$

Now apply $D\varphi$. The chain rule for derivatives shows that

$$D\varphi([X, Y]_g) = \frac{\partial}{\partial t}(R_{\varphi(\gamma)^{-1}})_*(L_{\varphi(g)})_* D\varphi_e(Y_e)|_{t=0}. \tag{1}$$

Then $(L_{\varphi(g)})_* D\varphi_e(Y_e)$ is the left-invariant vector field $\tilde{Y}$ on H defined by $D\varphi_e(Y_e)$ at the point $\varphi(g)$.

Evaluating (1) at $t = 0$ we get on the right hand side $[\tilde{X}, \tilde{Y}]_{\varphi(g)}$ where $\tilde{X}$ is the vector field with tangent $D\varphi_e(X_e) \in T_eH$. Take $g = e$ and we obtain

$$D\varphi([X, Y]_e) = [\tilde{X}, \tilde{Y}]_e$$

as required. □

Remark: In the case of the adjoint representation $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ this derivative was ad. The Lie bracket on $\mathrm{End}(\mathfrak{g})$ is just the commutator of linear transformations so from Proposition 3.1 we must have

$$\operatorname{ad} X \operatorname{ad} Y - \operatorname{ad} Y \operatorname{ad} X = \operatorname{ad}[X, Y].$$

Apply this to Z and it is equivalent to the Jacobi identity.

3.4 The exponential map

We defined the exponential of a matrix above:

$$\exp B = I + B + \frac{1}{2!}B^2 + \frac{1}{3!}B^3 + \dots.$$

This is a smooth function from the space of $n \times n$ matrices to the invertible ones: from the Lie algebra of $GL(n, \mathbf{R})$ to the Lie group. There is a generalization of the to an arbitrary Lie group:

Definition 15 *The exponential map for a Lie group G is the map* $\exp : T_eG \to G$ *defined by*

$$\exp(X_e) = \gamma(1)$$

where $\gamma(t)$ is the one-parameter subgroup with tangent vector X_e at the identity.

The map exp is smooth. More explicitly it is the composition of $X_e \mapsto (1, e, X_e) \in \mathbf{R} \times G \times T_e$ with $(t, g, X_e) \mapsto (L_g)_* \gamma_X(t)$ and the latter is smooth by smooth dependence of solutions to ODEs on initial conditions.

It is also a local diffeomorphism by the inverse function theorem, for the derivative at $0 \in T_e$ in the direction X_e is by definition X_e, so $D\exp_e = id$. It gives therefore a natural local coordinate system near the identity and by translation we get similar coordinate neighbourhoods at all points.

Examples:

1. For the unit complex numbers, S^1, the exponential map is $x \mapsto e^{2\pi i x}$. This is a diffeomorphism for $x \in (0, 1)$ but of course $\exp^{-1}(1) = \mathbf{Z} \subset \mathbf{R}$, so it is not a global diffeomorphism. It is surjective however.

2. If $G = SU(2)$, the exponential map is surjective (in fact for any compact Lie group) but note that skew-hermitian matrices of the form

$$B = \begin{pmatrix} ia & b \\ -\bar{b} & -ia \end{pmatrix}$$

satisfy $B^2 = -1$ if $a^2 + b\bar{b} = 1$. So, following the usual algebra of $e^{i\pi} = -1$, we obtain $\exp \pi B = -I$ and the inverse image of $-1 \in SU(2)$ is a 2-dimensional sphere inside a 3-dimensional sphere.

3. If $A = \exp B$ then $(\exp(B/2))^2 = A$ so A has a square root. If $G = SL(2, \mathbf{R})$ and $A = \mathrm{diag}(\lambda, \lambda^{-1})$ where λ is negative and not -1, then A has no square root and hence is not in the image of exp. This can be seen by looking at

$$C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with $ad - bc = 1$. If $C^2 = A$ then $b(a + d) = 0$ and $a^2 + bc = \lambda$. So $b \neq 0$ since $\lambda < 0$ hence $a + d = 0$ and $1 = -a^2 - bc = -\lambda$.

The exponential map intertwines the associated homomorphisms of Lie groups and Lie algebras:

Proposition 3.2 *Let $\varphi : G \to H$ be a Lie group homomorphism, and let $D\varphi_e : T_eG \to T_eH$ be its derivative at the identity. Then*

$$\exp(D\varphi_e(X_e)) = \varphi(\exp X_e).$$

Proof: A one-parameter subgroup with tangent X_e maps under φ to a one-parameter subgroup with tangent $D\varphi_e(X_e)$. □

It also tells us that a homomorphism $\varphi : G \to H$ of Lie groups is *determined* by $D\varphi_e$. This needs another result:

Proposition 3.3 *Let $V \subset G$ be an open set containing the identity, and suppose G is connected. Then every element of G is the product of a finite number of elements in V, together with their inverses.*

Proof: Since left translation is a homeomorphism each set gV is open. Let $F \subseteq G$ be the subset obtained from products and inverses in V, and $g \in F$. Then $gV \subseteq F$ since we are multiplying on the right by a further element of V. This shows F is open. Now if g lies in the closure of F, the open set gV intersects F in some point h, but then $g = hv^{-1}$ for some $v \in V$ and $g \in F$. Since F is open and closed and G is connected $F = G$. □

It follows that a homomorphism $\varphi : G \to H$ is determined uniquely by its restriction to V. But there exists V such that the exponential map is a diffeomorphism from an open set in T_e onto V and further $\exp(D\varphi_e(X_e)) = \varphi(\exp X_e)$. This means that φ on V is uniquely determined by $D\varphi_e$.

Remark: The above result does not say that a Lie algebra homomorphism $\mathfrak{g} \to \mathfrak{h}$ defines a Lie group homomorphism $G \to H$. For example, take the following basis for the Lie algebra $\mathfrak{su}(2)$:

$$X = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad Y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

These satisfy $[X, Y] = 2Z, [Y, Z] = 2X, [Z, X] = 2Y$. Recall that the Lie algebra $\mathfrak{so}(3)$ could be identified with $\mathbf{R}^3$ and the action $\mathbf{x} \mapsto \mathbf{a} \times \mathbf{x}$ and then $[\mathbf{a}, \mathbf{b}] = \mathbf{a} \times \mathbf{b}$. So putting $X = 2\mathbf{i}, Y = 2\mathbf{j}, Z = 2\mathbf{k}$ using the standard basis for $\mathbf{R}^3$, we get an isomorphism of Lie algebras.

But $SO(3)$ is not isomorphic to $SU(2)$: the matrix $-I \in SU(2)$ commutes with everything but there is no such rotation in $SO(3)$.

The exponential map is however useful in describing abelian Lie groups:

Theorem 3.4 *Let G be a connected Lie group, then*

- $\exp : \mathfrak{g} \to G$ *is a group homomorphism if and only if G is abelian*
- *G is abelian if and only if it is isomorphic to $T^r \times \mathbf{R}^k$ where T^r is the r-dimensional torus.*

Proof:

(i) Since the additive group $\mathfrak{g}$ is abelian, if exp is a homomorphism then

$$\exp(X)\exp(Y) = \exp(X+Y) = \exp(Y+X) = \exp(Y)\exp(X).$$

So since exp is a diffeomorphism near 0, there is a neighbourhood V of e all of whose elements commute. Now apply Proposition 3.3 using the fact that G is connected. Finite products of elements in V must commute, so G is abelian.

Conversely suppose G is abelian then the multiplication map $\mu : G \times G \to G$ is a Lie group homomorphism, since $\mu(g_1h_1, g_2h_2) = g_1h_1g_2h_2 = g_1g_2h_1h_2$. But $D\mu(X,Y) = X + Y$ so from Proposition 3.2 $\exp(X+Y) = \exp(X)\exp(Y)$.

(ii) From (i) any element is a product of terms $\exp(\pm X_i)$ where each $\exp(\pm X_i) \in V$, but since exp is a homomorphism these are of the form $\exp(\pm X_1 \pm X_2 \cdots \pm X_k)$ and hence exp is a surjective homomorphism. We need to identify the quotient group $G \cong \mathfrak{g}/K$ where K is the kernel of exp.

Since $\exp(A+X) = \exp(A)\exp(X)$ and exp is a local diffeomorphism at 0, it is a local diffeomorphism at every point, in particular at $A \in K$, the kernel of exp. This means there is a neighbourhood of $A \in \mathfrak{g}$ which only intersects K in the point A. We need now to identify the structure of K, an additive subgroup of $\mathbf{R}^n$.

Let r be the dimension of the vector subspace spanned by K, and choose r linearly independent elements $w_1, \ldots, w_r \in K$. Consider the set

$$F = \{x \in K : x = x_1w_1 + \cdots + x_rw_r, 0 \leq x_i \leq 1\}.$$

This is closed, bounded and hence compact, but each point has a neighbourhood which contains one point of K. By compactness this open covering has a finite subcovering and F therefore must be finite.

For each $1 \leq i \leq r$ choose $v_i = x_iw_i + \cdots + x_rw_r$ with $x_i > 0$ and minimal. These are clearly linearly independent and so any $v \in K$ can be written as $v = \lambda_1v_1 + \cdots + \lambda_rv_r$.

An integer linear combination of v_i lies in K so taking the integer part $[\lambda_i]$ we see that $v' = (\lambda_1 - [\lambda_i])v_1 + \cdots + (\lambda_r - [\lambda_r])v_r \in K$. Suppose for a contradiction that some $(\lambda_r - [\lambda_r])$ (which is non-negative of course) is not zero, and choose a minimal index j. Then

$$v' = x_j(\lambda_j - [\lambda_j])w_j + \cdots \in F$$

which contradicts the minimality in the definition of v_j. Hence each λ_i is an integer and K consists of all integer linear combinations of $v_1, \ldots, v_r$.

Extend this to a basis $v_1, \ldots, v_n$ of $\mathfrak{g}$ and then the map

$$\sum_1^n x_i v_i \mapsto (e^{2\pi i x_1}, \ldots, e^{2\pi i x_r}, x_{r+1}, \ldots, x_n)$$

gives an isomorphism of $G \cong \mathfrak{g}/K$ to $T^r \times \mathbf{R}^{n-r}$. □

The exponential map, even for matrices, does not behave as well as the exponential of real or complex numbers. In particular if A and B don't commute then $\exp(A+B) \neq \exp A \exp B$. Nevertheless, it is a local diffeomorphism so there is a local inverse which we could call log. So for any Lie group we could take $A, B \in \mathfrak{g}$ sufficiently small, and ask for

$$\log(\exp A \exp B) \in \mathfrak{g}.$$

This is a passage from two elements in the Lie algebra to a third, and it is a natural question to ask whether there is a formula for this entirely in terms of the algebra of $\mathfrak{g}$ together with its bracket. There is, and it is called the *Campbell-Baker-Hausdorff* formula and involves just the operation $\operatorname{ad} X(Y) = [X, Y]$. We shall not need it later but here is one explicit (at least in a sense) formula. Define

$$\psi(x) = \frac{x \log x}{x - 1} = 1 - \sum_{n=1}^{\infty} \frac{(1-x)^n}{n(n+1)}$$

The coefficients are *Bernoulli numbers*, important in number theory and algebraic topology. Then

$$\log(\exp X \exp Y) = X + \left(\int_0^1 \psi(\exp(\operatorname{ad} X)\exp(t \operatorname{ad} Y))dt\right) Y.$$

4 Submanifolds, subgroups and subalgebras

4.1 Lie subgroups

In group theory, a subgroup of G is simply defined as a subset closed under the operations of multiplication and inversion. Since Lie groups have the extra structure of being manifolds we need to address the definition of submanifolds. However, we shall prove theorems that show us how fairly minimal assumptions allow us to recognize Lie subgroups.

As far as manifolds are concerned there are two notions of submanifold related to an injective smooth map of manifolds $F : M \to N$. If the derivative DF is injective at each point, it is called an *immersed submanifold*. If, further, the induced topology from N on $F(M)$ is the manifold topology of M then it is called an *embedded* submanifold. There are various features that can cause an immersed submanifold to fail to be embedded but here is one that is particularly relevant for us.

Example: Define $F : \mathbf{R} \to S^1 \times S^1$ by

$$F(x) = (e^{2\pi i x}, e^{2\pi\alpha i x})$$

where α is an irrational number. This is a homomorphism of Lie groups with injective derivative. The map is injective because the kernel requires $x \in \mathbf{Z}$ and $\alpha x \in \mathbf{Z}$. However, α can be rationally approximated by $|\alpha - m/n| < 1/n^2$ for arbitrarily large n which means that $e^{2\pi i n\alpha}$ can be made arbitrarily close to 1. So in the induced topology, any neighbourhood of the identity contains arbitrarily large values of x, which is definitely not the topology of $\mathbf{R}$.

An *embedded* submanifold of dimension m can be described in a coordinate neighbourhood as $F^{-1}(c)$ where F is a smooth function $U \subset \mathbf{R}^n \to \mathbf{R}^{n-m}$ with surjective derivative as in Theorem 2.1. So $O(n)$ for example is an embedded submanifold of $\mathbf{R}^{n^2}$. The example above shows that we cannot discuss Lie groups without encountering immersed submanifolds.

Definition 16 *A Lie subgroup is a Lie group H which is a subgroup of G and is such that the inclusion $j : H \to G$ is a Lie group homomorphism.*

The point here is that the induced topology on $H \subseteq G$ may not be the manifold topology of H.

Examples:

1. Given a 1-parameter subgroup $\gamma : \mathbf{R} \to G$ its image $\mathbf{R}/\operatorname{Ker}\gamma$ (which is either $\mathbf{R}$ or S^1 by Theorem 3.4) is a Lie subgroup.

2. The kernel of a Lie group homomorphism $\varphi : G \to G'$ is a Lie subgroup. In this case it does have the induced topology. Since it is closed this will follow from Theorem 4.1.

3. If a Lie group is not connected, then the component containing the identity is an embedded Lie subgroup.

For Lie algebras we have

Definition 17 *A Lie subalgebra is a vector subspace $\mathfrak{h} \subseteq \mathfrak{g}$ which is closed under the bracket operation.*

and clearly given a Lie subgroup $j : H \to G$, $Dj : \mathfrak{h} \to \mathfrak{g}$ is a Lie subalgebra.

4.2 Continuity and smoothness

These are the definitions, but life is made easier by the following:

Theorem 4.1 *Let G be a Lie group. A subgroup H is an embedded Lie subgroup if and only if H is closed.*

The point to note here is that H is not assumed to be a Lie group, simply a (topologically) closed subset which is (algebraically) closed under multiplication and inversion. In particular, it tells us that all those examples of subgroups of $GL(n, \mathbf{R})$ that we gave are Lie groups, without using arguments like that given for $O(n)$.

Proof:

(i) If H is embedded it is given in a coordinate neighbourhood of the identity as $F^{-1}(c)$ then let B be a closed ball containing the identity so that, F being continuous, $H \cap B$ is closed. There is another such neighbourhood B^{-1} consisting of the inverses of elements of B. We want to show that H itself is closed.

Take $y \in \bar{H}$ then there exists $x \in yB^{-1} \cap H$ which implies $y \in xB \cap \bar{H}$ and hence $x^{-1}y \in B \cap \bar{H}$ since $x \in H$. But $B \cap H$ is closed so $x^{-1}y \in B \cap H$ and $x^{-1}y \in H$ hence $y \in H$ and H is closed.

(ii) Now assume that H is closed. We need to construct first a tangent space for H and then a Lie group. Let log denote the local inverse of exp at the identity, defined on $U \subset G$. Then if $U' = U \cap H$ we need $V' = \log(U')$ to be a neighbourhood of 0 in a vector subspace of T_eG.

Consider a sequence $v_n \in V'$ such that $v_n \to 0$. Normalize $v_n/\|v_n\|$ and we have a sequence on the unit sphere, which is closed and bounded hence there is a subsequence converging to X. Consider such sequences and the tangent vectors X which they generate.

Since $\|v_n\| \to 0$, given $t \neq 0 \in \mathbf{R}$, $|t|/\|v_n\| \to \infty$. Let m_n be the integer part $[t/\|v_n\|]$, then $m_n\|v_n\| \to t$. This means

$$\exp(m_n v_n) = \exp(m_n\|v_n\|(v_n/\|v_n\|)) \to \exp(tX). \tag{2}$$

On the other hand each term $\exp(v_n)$ lies in H. Considering the one-parameter subgroup in G with tangent vector v_n we know that $\exp(m_n v_n) = \exp(v_n)^{m_n}$ and since H is a group this implies $\exp(m_n v_n) \in H$. Equation (2) then says that $\exp(tX)$ lies in the closure of H. But H is assumed closed so $\exp(tX) \in H$.

We want to show that the Xs produced this way form a vector space. Multiplying by a scalar is just rescaling t above, addition is the problem. Take X and Y and consider

$$\ell(t) = \log(\exp(tX)\exp(tY)).$$

Since $\exp(tX)$ and $\exp(tY)$ lie in H, for small t we have $\ell(t) \in V'$. Take $v_n = \ell(1/n)$ then $v_n \to 0$ and

$$nv_n \to \ell'(0) = X + Y$$

since $D\mu_e(X,Y) = X + Y$ where $\mu : G \times G \to G$ is the multiplication map. Finally

$$v_n/\|v_n\| = (nv_n)/(n\|v_n\|) \to (X+Y)/\|X+Y\|.$$

We therefore have a vector space $W \subset T_e$ defined by sequences as above and such that $\exp(X) \in H$ if $X \in W$. We need to show that this is in some sense the largest possible. So using an inner product decompose $T_e = W \oplus W^\perp$ and consider the map $\psi(w,v) = \exp(w)\exp(v)$ for $w \in W$ and $v \in W^\perp$. Suppose $\psi(w,v)$ lies in H then since we have just shown $\exp(w)$ does, we also have $\exp(v) \in H$. Now repeat the process above with a sequence of such vs. Then $v_n/\|v_n\|$ will converge to a unit vector $u \in W^\perp$. But also by definition $u \in W$ which is a contradiction. Hence $v = 0$.

The map $\log\psi$ gives a local coordinate system in which H is defined as the kernel of projection onto $W^\perp$ and so is an embedded Lie subgroup. $\square$

Example:

1. The *centre* of a group G is the (normal, abelian) subgroup of elements that commute with every element of G. For a fixed element h in a Lie group G the map $g \mapsto ghg^{-1}h^{-1}$ is continuous and so the inverse image of e is closed. This consists of the elements that commute with h. The intersection for all h is the centre which is closed and hence an embedded Lie subgroup.

2. If H is compact, and $\varphi : H \to G$ injective then H is an embedded subgroup.

This theorem has a useful consequence:

Theorem 4.2 *A continuous group homomorphism between two Lie groups is smooth and hence a Lie group homomorphism.*

Proof: Let $\varphi : G \to H$ be a continuous group homomorphism and consider the graph

$$\Gamma = \{(g, \varphi(g)) \in G \times H\}.$$

Since manifolds are Hausdorff, the graph is closed and so by Theorem 4.1 it is a Lie group with a smooth homomorphism to $G \times H$. The projection π from Γ onto the first factor G is smooth because it is the restriction of a smooth map $G \times H \to G$ to Γ. It is moreover a homeomorphism with inverse $g \mapsto (g, \varphi(g))$.

Now if $D\pi_e$ had a non-trivial kernel, the exponential map would define a one-parameter subgroup which mapped to the identity in G. Since π is a bijection this does not hold, so $D\pi_e$ is an isomorphism. By the inverse function theorem this means that π^{-1} is smooth and so defines a Lie group isomorphism between Γ and G.

Now the projection from Γ to the second factor H is smooth and composing with π^{-1} this is the homomorphism φ. $\square$

4.3 Subgroups versus subalgebras

If H is a Lie subgroup of G then its Lie algebra $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$. What about the converse? Is every Lie subalgebra tangential to a Lie subgroup? The answer is:

Theorem 4.3 *There is a one-to-one correspondence between Lie subalgebras $\mathfrak{h} \subseteq \mathfrak{g}$ and connected Lie subgroups $H \subseteq G$.*

The case of the irrational homomorphism $F(x) = (e^{2\pi ix}, e^{2\pi\alpha ix})$ above shows that H may not be closed in G. The subgroup is clearly going to be generated by $\exp(\mathfrak{h}) \subset G$ but this needs to be given the structure of a manifold and of a group.

We start with the vector space $\mathfrak{h}$ of left-invariant vector fields on G. If $\dim \mathfrak{h} = r$ this defines a smoothly varying r-dimensional subspace of each tangent space of G. This makes sense on a more general manifold and is called a *distribution* (not to be confused with the analytical meaning of this term). Since $\mathfrak{h}$ is a subalgebra, it is closed under the Lie bracket, so if $Y_1, \ldots, Y_r$ forms a basis then $[Y_i, Y_j] = c_{ijk}Y_k$ for constants c_{ijk}. This is a Lie bracket of vector fields. If f, g are smooth functions on a manifold and X, Y vector fields then

$$[fX, gY] = f(Xg)Y - g(Yf)X + fg[X, Y] \tag{3}$$

So in our case, introducing coefficients f_i, g_i which are smooth functions on G we see that

$$\sum_{i,j}[f_iY_i, g_jY_j] = \sum_{i,j} f_i(Y_ig_j)Y_j - f_j(Y_jg_i)Y_i + f_ig_jc_{ijk}Y_k.$$

Thus the Lie bracket as vector fields of any linear combinations of the Y_i with functions as coefficients is again a combination of those basis vectors. A distribution on a manifold with this property is called an *integrable distribution.* The basic theorem which relates to this is:

Theorem 4.4 (Frobenius) *If E is an integrable distribution of rank r on an open set in $\mathbf{R}^n$, then through each point in some open subset there is an embedded submanifold of dimension r which is tangential to E. Furthermore we can choose coordinates $x_1, \ldots, x_n$ such that these submanifolds are defined by $x_i = c_i$, for $r + 1 \leq i \leq n$.*

Remark: If $r = 1$, then we can choose a single vector field Y and the integral curve of Y is the submanifold. There is no integrability condition here since $[Y, Y] = 0$. Globally, as with the irrational homomorphism, the full integral curve may not be an embedded submanifold.

Proof:

(i) Let $Y_1, \ldots, Y_r$ be a basis for the distribution: local vector fields which are linearly independent and span E at each point

$$Y_i = \sum_{j=1}^{n} A_{ij}\frac{\partial}{\partial x_j}.$$

The matrix of functions A_{ij} has rank r so by reordering the basis we may assume that on a possibly smaller neighbourhood $A_{ij}, 1 \leq i, j \leq r$ is nonsingular. Call the inverse of this matrix B_{ij} and define

$$X_i = \sum_{j=1}^{r} B_{ij} Y_j.$$

Then observe that, by choosing B as the inverse of A when we expand $[X_i, X_j]$ using $[fX, gY] = f(Xg)Y - g(Yf)X + fg[X, Y]$ the last term only involves $[\partial/\partial x_i, \partial/\partial x_j]$ and so vanishes. Thus

$$[X_i, X_j] = \sum_{k=r+1}^{n} a_{ijk} \frac{\partial}{\partial x_k}.$$

But the integrability condition says that $[X_i, X_j]$ is a linear combination of $X_1, \dots, X_r$ which only involve terms $\partial/\partial x_i$ for $i \leq r$ so $[X_i, X_j] = 0$.

(ii) Let φ_t^i be the (local) one-parameter group of diffeomorphisms for X_i and define the smooth map

$$F(t_1, \dots, t_r) = \varphi_{t_1}^1 \circ \varphi_{t_2}^2 \cdots \circ \varphi_{t_r}^r(a).$$

Since the vector fields commute, so do the one-parameter groups of diffeomorphisms, so F is independent of the ordering. By moving $\varphi_{t_i}^i$ to the beginning and differentiating with respect to t_i we obtain X_i. So we see that F has injective derivative whose image is spanned by $X_1, \dots, X_r$ as required.

Now restrict a to depend on $x_{r+1}, \dots, x_n$: $a = (a_1, \dots, a_r, x_{r+1}, \dots, x_n)$ then by the inverse function theorem the local coordinate system $(t_1, \dots, t_r, x_{r+1}, \dots, x_n)$ satisfies the conditions. $\square$

This theorem is a local one but to make it global we define an *integral manifold* for the distribution to be an immersed submanifold $F : N \to M$ such that DF_a maps the tangent space $T_a N$ isomorphically to the distribution $E_{F(a)} \subset T_{F(a)} M$. This is just a generalization of the integral curve of a vector field. Through each point we can define a maximal integral submanifold and in our case, taking the point as the identity, this will turn out to be the Lie group H. But first a technical point.

Proposition 4.5 *If $F : N \to M$ is an integral submanifold and L is a manifold with map $f : L \to N$, then f is smooth if and only if $F \circ f$ is.*

Proof: One direction is obvious, so suppose that $F \circ f$ is smooth. For each $a \in L$ consider $F(f(a)) \in M$ and a small neighbourhood U of this point for which the

Frobenius theorem holds. Then $F^{-1}(U) \subset N$ is open. Since a manifold has by definition a countable basis of open sets, this is a countable union of open connected manifolds (each given by $x_i = c_i, r+1 \leq i \leq n$ for countably many c_i). The map $F \circ f$, being continuous, maps a connected open neighbourhood of a to just one of these components given by a constant value of the c_i. In coordinates $(y_1, \ldots, y_m)$ on this neighbourhood $F \circ f$ is just of the form $(x_1(y), \ldots, x_n(y), c_{r+1}, \ldots, c_n)$ where the c_i are constant and this is a smooth function f to N. $\square$

Now for the proof of Theorem 4.3.

Proof: The distribution E on G given by the left-invariant vector fields $\mathfrak{h}$ is integrable as noted above. Let H be the maximal connected integral submanifold through e. Since E is left-invariant left translation gives another integral submanifold so if $h \in H$, $h^{-1}H$ is an integral submanifold. But it passes through $e = h^{-1}h$ so $h^{-1}H = H$ and we deduce that H is closed under multiplication and inversion.

We need to prove that multiplication and inversion are smooth. But $H \times H \subset G \times G$ is smooth and $\mu : G \times G \to G$ is smooth, so putting $L = H \times H$, $N = H$ and $M = G$ in Proposition 4.5 we have the result.

$\square$

5 Global aspects

5.1 Components and coverings

As we saw with $O(n)$, naturally occurring Lie groups are not necessarily connected. It is straightforward to see that a manifold is connected in the topological sense of not containing any sets which are closed and open if and only if it is path-connected.

The connected component G_0 of a Lie group is a subgroup since paths $g(t), h(t)$ from the identity to g, h define a path $g(t)h(t)$ to gh. It is also a *normal subgroup* since $x \mapsto gxg^{-1}$ is a homeomorphism and homeomorphisms take connected components to connected components. Conjugation also takes the identity to the identity so takes G_0 to G_0. The set of components $\pi_0(G)$ is therefore a group: the quotient of G by the normal subgroup G_0. It has the discrete topology (every point is open and closed) and is countable because all our manifolds have a countable basis of open sets.

Examples:

1. The homomorphism $\det : O(n) \to \pm 1$ has kernel $SO(n)$ which is connected, so $\pi_0(O(n)) \cong \mathbf{Z}_2$.

2. If we take $\mathbf{R}^p \oplus \mathbf{R}^q$ and put a positive definite inner product on $\mathbf{R}^p$ and a negative one on $\mathbf{R}^q$ we can write an element of $O(p, q)$ in block form

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

and then $(\operatorname{sgn}\det A, \operatorname{sgn}\det D)$ is a homomorphism to $\mathbf{Z}_2 \times \mathbf{Z}_2$ with connected kernel.

Connectedness is a simple concept. More interesting is whether a Lie group is *simply-connected* i.e. whether any continuous map $f : S^1 \to G$ is contractible to a point by a continuous family $f_t : S^1 \times [0, 1] \to G$. Clearly the circle is not simply connected, but spheres are simply connected hence $SU(2)$ is. When a space is not simply-connected it has non-trivial covering spaces and we encounter this way covering groups.

Definition 18 *A smooth surjective map $\pi : M \to N$ of manifolds is a covering map if each point of N has a neighbourhood U such that $\pi^{-1}(U)$ is the disjoint union of open sets U_i such that $\pi : U_i \to U$ is a diffeomorphism.*

Remark: If $\pi : M \to N$ is a map of *compact* manifolds such that $D\pi$ is an isomorphism at each point, then using the inverse function theorem one can see this is a covering map where there are a finite number of U_is.

For Lie groups we have the following:

Theorem 5.1 *Let $\pi : G \to H$ be a Lie group homomorphism, with H connected, then π is a covering map if and only if $D\pi_e$ is an isomorphism of Lie algebras.*

Examples:

1. We observed that $SU(2)$ and $SO(3)$ have isomorphic 3-dimensional Lie algebras. The adjoint representation $\mathrm{Ad} : SU(2) \to GL(3, \mathbf{R})$ has $SO(3) \subset GL(3, \mathbf{R})$ as image and this is a covering homomorphism. The element $-I \in SU(2)$ acts trivially by conjugation so this is a 2-fold covering with kernel $\pm I$.

2. The group $SL(2, \mathbf{C})$ of 2×2 complex invertible matrices of determinant 1 is a covering of the identity component of $SO(1,3)$, the Lorentz group of special relativity. We can see this by considering the 4-dimensional real vector space V of 2×2 Hermitian matrices $X = X^*$ and the action of $A \in SL(2, \mathbf{C})$ as $X \mapsto AXA^*$. If

$$X = \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix}$$

where a, c are real, then $\det X = ac - b\bar{b} = (a+c)^2/2 - (a-c)^2/2 - b\bar{b}$ is a quadratic form of signature $(1,3)$ and since $\det AXA^* = |\det A|^2 \det X = \det X$, the action preserves the indefinite inner product. Again, $-I \in SL(2, \mathbf{C})$ acts trivially.

Proof:

(i) If π is a covering then it is a local diffeomorphism and so $D\pi_e$ gives an isomorphism of Lie algebras.

(ii) Suppose $D\pi_e$ is an isomorphism then by the inverse function theorem π is a local diffeomorphism at e and by left translation also at any point. At $e \in G$ it maps an open neighbourhood to an open neighbourhood of $e \in H$, which therefore generates the whole group H, hence π is surjective.

By the inverse function theorem $e \in G$ has a neighbourhood W which maps diffeomorphically to a neighbourhood $U \in H$ and therefore contains only one point of $\operatorname{Ker} \pi$, namely e. To get a covering we want, for each $k \in \operatorname{Ker} \pi$ such a neighbourhood mapping diffeomorphically to the same U, perhaps by shrinking U.

Since the map $f(g_1, g_2) = g_1 g_2^{-1}$ is continuous, $f^{-1}(W)$ is open and so contains an open neighbourhood $V \times V$ of (e, e). Take $k \in \operatorname{Ker} \pi$ and suppose $g \in kV \cap V$. Then $g = v_1 = kv_2$ and so $k = v_1 v_2^{-1}$. But by construction this lies in W which only contains the identity, so V and kV are disjoint. Replacing U by $\pi(V)$, we have

$\pi^{-1}(\pi(V))$ expressed as a disjoint union of open sets each mapping diffeomorphically to $\pi(V)$. Each such set is a translate kV. By left translation we get such open sets for all inverse images.

□

The kernel of a covering homomorphism is normal and also discrete (each element is an open set in the induced topology) as we have seen in the proof of the last theorem. Conjugation $k \mapsto gkg^{-1}$ permutes the elements of $\operatorname{Ker}\pi$ but if the covering group G is *connected* it must act trivially, for given a path $g(t)$ to e, $g(t)kg(t)^{-1}$ is a continuous map from $[0,1]$ to $\operatorname{Ker}\pi$ and so maps to one connected component. But $g(1)kg(1)^{-1} = eke = k$. This means that $k \in \operatorname{Ker}\pi$ commutes with all g and so is a subgroup of the *centre* of G.

5.2 From Lie algebras to Lie groups

Every connected manifold M has a special connected covering called the *universal covering* $\tilde{M}$. The universal covering is simply connected, and this characterizes it.

It is defined, given a base point $z \in M$, as a set of equivalence classes of continuous maps $\gamma : [0,1] \to M$ with $\gamma(0) = z$. Two such maps γ, γ' are equivalent if they have the same endpoint $\gamma(1) = \gamma'(1)$ and there is a continuous family γ_t, for $t \in [0,1]$ fixing the end points and such that $\gamma_0 = \gamma, \gamma_1 = \gamma'$. The map $[\gamma] \mapsto \gamma(1)$ is a covering map π of manifolds. It is universal in the sense that if $p : N \to M$ is any other connected covering then there is a covering map $q : \tilde{M} \to N$ with $\pi = p \circ q$. The map q is unique so long as we choose a point $x \in \tilde{M}$ and $y \in N$ mapping to $z \in M$ and require $q(x) = y$.

When G is a Lie group and we take $z = e$, then multiplication of paths $\gamma(t)\gamma'(t)$ defines a product on $\tilde{G}$ and with the identity as the equivalence class of the trivial path $\gamma(t) = e$, we can define inversion by $(\gamma(t))^{-1}$. Thus $\tilde{G}$ is a Lie group and the projection $\pi : \tilde{G} \to G$ a Lie group homomorphism. Moreover given a covering homomorphism $p : H \to G$, if we choose x, y, z as previously to be the identity elements in each group, then the unique $q : \tilde{G} \to H$ is a Lie group homomorphism. Thus each covering group is obtained from $\tilde{G}$ as the quotient group by a subgroup of $\operatorname{Ker}\pi$.

Examples:

1. The universal covering of the circle is the additive real line: $\pi(t) = e^{2\pi i t}$. The kernel is $\mathbf{Z} \subset \mathbf{R}$.

2. The universal covering of $SO(3)$ is $SU(2)$ with the projection defined above and kernel $\pm I$.

3. The universal covering of $SL(2, \mathbf{R})$ is a connected Lie group which is not a subgroup of any $GL(n, \mathbf{R})$. In this case $\operatorname{Ker} \pi \cong \mathbf{Z}$ (this is the fundamental group π_1 of $SL(2, \mathbf{R})$).

Thus far we have associated Lie algebras to Lie groups, but the properties of covering homomorphisms show that any two coverings of G have isomorphic Lie algebras so the association is not one-to-one. However:

Theorem 5.2 *Let G be a simply connected Lie group and H another Lie group, then*

- *there is a one-to-one correspondence between Lie algebra homomorphisms $\mathfrak{g} \to \mathfrak{h}$ and Lie group homomorphisms $G \to H$*
- *if both G and H are simply connected and have isomorphic Lie algebras then they are isomorphic as Lie groups.*

Proof:

(i) If $\psi : \mathfrak{g} \to \mathfrak{h}$ is the given homomorphism of Lie algebras consider its graph

$$\Gamma = \{(x, \psi(x)) \in \mathfrak{g} \oplus \mathfrak{h}\}.$$

This is a Lie subalgebra since ψ is a homomorphism.

Now apply Theorem 4.3 and we get a connected Lie subgroup $S \subset G \times H$ with Lie algebra Γ. Projection from Γ to the first factor is an isomorphism of Lie algebras so from Theorem 5.1 we have a covering homomorphism $\pi : S \to G$. But G is simply connected, hence its own universal covering, which means that it is a diffeomorphism. Then π^{-1} followed by projection onto the second factor in $G \times H$ gives the required Lie group homomorphism.

We already saw, using the exponential map, that a Lie group homomorphism is uniquely determined by the corresponding Lie algebra homomorphism.

(ii) Using the first part for both G and H gives the last part.

□

The outstanding question is whether one can associate a Lie group to any Lie algebra, and this is provided by Ado's theorem which we shall not prove:

Theorem 5.3 (Ado) *For any Lie algebra* V*, there is an injective Lie algebra homomorphism* $V \to \mathfrak{gl}(m, \mathbf{R})$ *for some* m*.*

By Theorem 4.3 this implies there is a connected Lie subgroup G of $GL(m, \mathbf{R})$ with V as Lie algebra. If we pass to the universal cover then we get *Lie's third theorem*:

Theorem 5.4 *There is a one-to-one correspondence between Lie algebras up to isomorphism and simply-connected Lie groups up to isomorphism.*

6 Representations of Lie groups

6.1 Basic notions

Recall that:

Definition 19 *A representation of a Lie group G on V is a Lie group homomorphism $\varphi : G \to \mathrm{Aut}(V)$.*

From Theorem 4.2 we need only assume that the homomorphism is continuous. We can also describe a representation as an action of G on V by linear transformations: $gv = \varphi(g)v$. Some of the following examples entail the action on certain functions on a vector space. The action on functions is $(gf)(x) = f(g^{-1}x)$ for then

$$(h(gf))(x) = gf(h^{-1}x) = f(g^{-1}h^{-1}x) = f(hg)^{-1}x) = ((hg)f)(x).$$

Examples:

1. For each $n \in \mathbf{Z}$ we have a representation U_n of S^1 on $\mathbf{C}$ given by $\varphi(e^{i\theta})z = e^{in\theta}z$. The case $n = 0$ is the *trivial* representation which of course any group has.

2. The group $SU(2)$ has a defining representation on $V = \mathbf{C}^2$. We can regard this as an action on linear functions in z_1, z_2. The matrix

$$A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

acts on a linear function $f(z)$ by $f(A^{-1}z)$ and using $A^{-1} = A^*$ this is

$$f \mapsto f(\bar{a}z_1 - bz_2, \bar{b}z_1 + az_2).$$

But this formula makes sense for a homogeneous polynomial of any degree. So if V_n denotes the space of degree n polynomials it has a basis of the $(n+1)$ functions $z_1^n, z_1^{n-1}z_2, \ldots, z_2^n$ and this is a representation space.

Given a representation φ of a Lie group, $D\varphi_e$ is a representation of the Lie algebra, meaning we have a homomorphism $\psi : \mathfrak{g} \to \mathrm{End}\, V$ such that $\psi[X, Y] = \psi(X)\psi(Y) - \psi(Y)\psi(X)$. We find this by differentiating one-parameter subgroups at the identity.

Examples:

1. For the representation of S^1 on $\mathbf{C}$ given by $\varphi(e^{i\theta})z = e^{in\theta}z$ we differentiate with respect to θ at $\theta = 0$ and get the Lie algebra homomorphism $\psi(z) = inz$.

2. For $SU(2)$ we can use the basis

$$X = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad Y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

of the Lie algebra. Then the action of $\exp tX$ is $f(z_1, z_2) \mapsto f(e^{-it}z_1, e^{it}z_2)$ and differentiating at $t = 0$ we get

$$Xf = -iz_1\frac{\partial f}{\partial z_1} + iz_2\frac{\partial f}{\partial z_2}$$

and similarly

$$Yf = -z_2\frac{\partial f}{\partial z_1} + z_1\frac{\partial f}{\partial z_2}, \quad Zf = -iz_2\frac{\partial f}{\partial z_1} - iz_1\frac{\partial f}{\partial z_2}.$$

3. From the previous section, each Lie algebra representation defines a representation of the associated simply-connected Lie group, but not necessarily otherwise. So although $SU(2)$ and $SO(3)$ have isomorphic Lie algebras only the spaces V_n for n even are representations of $SO(3)$ since $-I \in SU(2)$ acts as -1 on odd degree polynomials.

Two representations V, W of G are isomorphic if there is a vector space isomorphism commuting with the two actions of G.

There are some natural operations on representation spaces:

- The direct sum $V \oplus W$ with the action $g(v, w) = (gv, gw)$
- The dual V^*. A linear transformation $A : V \to W$ has a natural dual map $A' : W^* \to V^*$ (the transpose of the matrix if we take a dual basis) and to get an action on V^* from $\varphi : G \to \mathrm{Aut}(V)$ we define for $\xi \in V^*$,

  $$g\xi = \varphi(g^{-1})' : V^* \to V^*.$$

  Concretely this is $(A^{-1})^T$ for the group action and $-B^T$ for the Lie algebra action. Note that since $A'f(v) = f(Av)$ this is the natural action on linear functions.
- The tensor product $V \otimes W$ with action

  $$g\sum_i v_i \otimes w_i = \sum_i gv_i \otimes gw_i$$

We differentiate to get the Lie algebra action

$$X \sum_i v_i \otimes w_i = \sum_i X v_i \otimes w_i + \sum_i v_i \otimes X w_i.$$

- The space $\mathrm{Hom}(V, W)$ of homomorphisms from V to W. Here the group action is

  $$(gA)(v) = g(A(g^{-1}v)).$$

  In fact $\mathrm{Hom}(V, W)$ is canonically isomorphic to $V^* \otimes W$ and so this is a combination of the above.

Definition 20 *A representation V is*

- *reducible if there is a proper subspace $W \subset V$ such that $gw \in W$ for all $w \in W$ and $g \in G$.*
- *it is irreducible if $\{0\}$ and V are the only invariant subspaces*
- *it is completely reducible if it is a direct sum $V = V_1 \oplus V_2 \cdots \oplus V_m$ of irreducible representations.*

Examples:

1. Take $G = \mathbf{R}$ the additive reals and $V = \mathbf{R}^2$ with action $t(x, y) = (x + ty, y)$. Then $y = 0$ is an invariant subspace so V is reducible. It is not completely reducible however for if $y \neq 0$, $(x + ty, y) = \lambda(x + 0y, y)$ implies $\lambda = 1$ and $t = 0$.

2. Any one-dimensional representation is clearly irreducible so the representations U_n of S^1 are.

3. The representations V_n above of $SU(2)$ are irreducible. To see this use the Lie algebra action. Using the basis X, Y, Z write

$$N^+ = \frac{1}{2}(Y + iZ) = z_1 \frac{\partial}{\partial z_2}$$

then

$$N^+(z_1^k z_2^{n-k}) = (n-k)(z_1^{k+1} z_2^{n-k-1}).$$

If $W \subset V_n$ is invariant, and $f \in W$, then W contains $N^+ f, (N^+)^2 f, \ldots$ also. Choose $f \in W$ and let k be the smallest integer such that the coefficient of $z_1^k z_2^{n-k}$ is non-zero. Then $(N^+)^{n-k} f$ is a non-zero multiple of z_1^n, hence $z_1^n \in W$. Now consider

$$N^- = -\frac{1}{2}(Y - iZ) = z_2 \frac{\partial}{\partial z_1}.$$

Applying this to z_1^n we get $nz_1^{n-1}z_2$ and by repetition all the basis vectors, hence $W = V_n$.

4. The product of a polynomial of degree m and one of degree n is of degree $m+n$ and so there is an invariant homomorphism $V_m \otimes V_n \to V_{m+n}$. Since V_{m+n} is irreducible the image is the whole space and so if $m, n > 0$ there is a non-zero kernel which is an invariant subspace of $V_m \otimes V_n$ which is therefore reducible.

5. Restrict V_n to $S^1 \subset SU(2)$ acting as $(z_1, z_2) \mapsto (e^{i\theta}z_1, e^{-i\theta}z_2)$ then by looking at the basis vectors $z_1^k z_2^{n-k}$ we can see that it is completely reducible:

$$V_n| = U_n \oplus U_{n-2} \oplus \cdots \oplus U_{-n}.$$

We have this useful result:

Proposition 6.1 (Schur's lemma) *A G-invariant homomorphism $A : V \to W$ between two irreducible representations is either an isomorphism or is zero. If $V = W$ and these are complex vector spaces then A is scalar multiplication by $\lambda \in \mathbf{C}$.*

Proof: Since V is irreducible, $\operatorname{Ker} A = V$ or 0 so A is either zero or injective. But W is irreducible so the image is either 0 or W.

If $V = W$ and the field is $\mathbf{C}$, then A has an eigenvalue λ and so $A - \lambda I$ is an invariant endomorphism of V. By the first part it has to be zero since $A - \lambda I$ is not invertible. $\square$

Remark: It is convenient to work over the complex numbers and only as a secondary issue discuss whether a representation is real. A *real structure* on a complex vector space V is an antilinear involution T, i.e. $T(u + v) = T(u) + T(v), T(\lambda u) = \bar{\lambda}T(u)$ and $T^2 = 1$. The fixed point set of T is a real vector space U and V is naturally its complexification $U \otimes_{\mathbf{R}} \mathbf{C}$ with T becoming complex conjugation. So a complex representation which commutes with such a T is actually a real representation.

As an example consider $(z_1, z_2) \mapsto (\bar{z}_2, -\bar{z}_1)$ acting on $\mathbf{C}^2$. This is antilinear but squares to -1. However its action on polynomials of even degree gives a real structure preserved by $SU(2)$. So the V_{2m}, which we already saw are representations of $SO(3)$, are all real in this sense.

Proposition 6.2 *Over the complex numbers every irreducible representation of an abelian Lie group is one-dimensional.*

Proof: If G is abelian the map $v \mapsto gv$ for a fixed g commutes with G and hence by Schur's Lemma is a non-zero scalar λ_g. So multiples of a fixed vector v form an invariant subspace and hence by irreducibility the whole space. $\square$

If V is a real representation of an abelian group which is completely reducible then we can complexify $V \otimes_{\mathbf{R}} \mathbf{C}$ and from the proposition write

$$V \otimes_{\mathbf{R}} \mathbf{C} = V_1 \oplus V_2 \oplus \cdots \oplus V_n$$

as a direct sum of one-dimensional representations: $g \in G$ acts as a scalar $\lambda(g) \in \mathbf{C}^*$ on each one. Reality means the action commutes with an antilinear involution T: complex conjugation. So T maps V_i into some V_j and the action $\lambda(g)$ on V_i is paired with $\bar{\lambda}(g)$ on V_j. Since ± 1 is the only real subgroup of the unit complex numbers, if G is connected it must act by a complex scalar or the identity. So V is a direct sum of a trivial representation of dimension k, say, and a sum of 2-dimensional real vector spaces where the action is

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

An example is the real defining 3-dimensional representation of $SO(3)$ restricted to the abelian subgroup of rotations about $(0, 0, 1)$. It has the trivial representation with multiplicity 1 (the space $(0, 0, z)$) and an irreducible 2-dimensional real representation on $(x, y, 0)$. The eigenvalues $e^{\pm i\theta}$ of a rotation have complex conjugate eigenspaces which are irreducible one-dimensional complex representations.

6.2 Integration on G

Suppose a representation space V has a positive-definite Hermitian inner product $\langle v, w\rangle$ and the action of G preserves it so that $\langle gv, gw\rangle = \langle v, w\rangle$, then g acts as unitary transformations, and the representation can be viewed as a homomorphism $\varphi : G \to U(n)$ where $\dim V = n$. More importantly, if $W \subset V$ is an invariant subspace, so is its orthogonal complement $W^{\perp}$. For a finite group the existence of such an inner product is straightforward by averaging over the group elements. In other words, choose any inner product (.) and define

$$\langle v, w\rangle = \sum_{g \in G} \frac{1}{|G|}(gv, gw)$$

and since the sum of positive-definite Hermitian inner products is still positive definite we have the required inner product. This means that any representation is completely reducible – pick an invariant subspace $W \subset V$ and write $V = W \oplus W^{\perp}$ and repeat for W and $W^{\perp}$ until it has no invariant subspaces.

For a *compact* Lie group we can do the same thing but we need to integrate over the group rather than sum, so we need to know how to integrate functions on a manifold. Recall the change of variables formula in a multiple integral:

$$\int f(y_1, \ldots, y_n) dy_1 dy_2 \ldots dy_n = \int f(y_1(x), \ldots, y_n(x)) |\det \partial y_i / \partial x_j| dx_1 dx_2 \ldots dx_n.$$

Let $\Omega(X_1, \ldots, X_n)$ be an alternating multilinear form on the tangent space T_a of a manifold of dimension n. So it is linear in each factor and changes sign if we interchange two variables. If we write $X_i = \sum A_{ij} E_j$ for a basis $E_1, \ldots, E_n$ so that X_i is the ith row of the matrix A, then $\det A$ is an example of such a form and in fact all such forms are multiples of this. If $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$ are two local coordinate systems they each give a basis for the tangent space and these are related by

$$\left(\frac{\partial}{\partial x_i}\right)_a = \sum_j \frac{\partial y_j}{\partial x_i}(a) \left(\frac{\partial}{\partial y_j}\right)_a.$$

So

$$\Omega\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right) = \left(\det \frac{\partial y_i}{\partial x_j}\right) \Omega\left(\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_n}\right).$$

Apart from the absolute value this is how multiple integrals transform.

This is at one point $a \in M$ but if we want a smoothly varying form this is the definition:

Definition 21 *A differential form of degree m on a manifold M is an alternating multilinear function of vector fields $\Omega(X_1, \ldots, X_m) \in C^{\infty}(M)$.*

When $m = n$, if f is a function with support in a coordinate neighbourhood then we can define the integral of the n-form $f\Omega$ and this will be independent of coordinates so long as the sign of $\det \partial y_i / \partial x_j$ is positive. For a Lie group this certainly holds for we can take a non-zero multilinear form ω on T_e and by translation extend it to all of G so that for a basis of left-invariant vector fields $X_1, \ldots, X_n$ we have

$$\Omega(X_1, \ldots, X_m) = \omega((X_1)_e, \ldots, (X_n)_e)$$

If we restrict to local coordinates $x_1, \ldots, x_n$ for which

$$\Omega\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right) > 0$$

then we have a consistent notion of integration on each coordinate neigbourhood. A general function can be written on a compact manifold as a finite sum of smooth functions supported in coordinate neighbourhoods and then we can define integration of functions in a coordinate-independent way.

The form described above is left-invariant and on a compact manifold its integral is finite, so we can normalize it so that the integral of Ω is 1. It is also right-invariant because the adjoint action on left-invariant vector fields induces an action on the one-dimensional space of degree n ($n = \dim T_e$)) multilinear maps, a homomorphism from G to $\mathbf{R}^*$. But ± 1 is the only compact subgroup of the multiplicative group of non-zero reals and if G is connected the action must be trivial.

Example: For the circle $x \mapsto e^{2\pi i x}$ for $x \in [0,1]$, we can take $\Omega = dx$, its value on the vector field d/dx is 1 and its integral is 1.

From now on, in performing integrals we shall omit Ω and just write

$$\int_G f(g).$$

Then given a representation V we can always find a G-invariant inner product by choosing any inner product $(\ .\ )$ and defining

$$\langle v, w\rangle = \int_G (gv, gw).$$

For an irreducible representation the inner product is unique up to multiplication by a positive real number. This follows from Schur's lemma since two inner products are related by $(u,v)_1 = (Au,v)_2$ where $A = A^*$ with respect to $(\ .\ )_2$. But if both are G-invariant so is A and so by Schur it is a multiple of the identity.

Differential forms transform naturally under a diffeomorphism F which is denoted $F^*\Omega$. The integral of the differential form $f\Omega$ transforms as

$$\int_M (f \circ F) F^*\Omega = \pm \int_M f\Omega.$$

depending on whether F preserves or reverses the orientation. The form defines a *measure* which is insensitive to orientation: a positive function will always have a positive integral.

Since L_h and R_h preserve our choice of Ω for G this means that for a fixed $h \in G$,

$$\int_G f(hg) = \int_G f(gh) = \int_G f(g).$$

Example: For the circle this is just the statement that

$$\int_0^1 f(x+a)dx = \int_0^1 f(x)dx$$

for a periodic function with period 1.

Inversion $g \mapsto g^{-1}$ acts as -1 on the tangent space at the identity which introduces a sign $(-1)^{\dim G}$ in the transform of the differential form Ω but the associated measure satisfies

$$\int_G f(g^{-1}) = \int_G f(g).$$

The assumption above was that f is a real-valued function but it could equally be vector valued. So consider a representation space V and for $v \in V$ the function $g \mapsto gv$ on G. Then

$$h\int_G gv = \int_G hgv = \int_G gv$$

so the integral is an invariant vector. On the other hand if u is an invariant vector

$$\int_G gu = \int_G u = u$$

since the integral of 1 is 1. So

$$P(v) = \int_G gv$$

is a projection onto the subspace $V^G \subset V$ of invariant vectors. It is also an orthogonal projection ($P = P^*$) since using an invariant inner product,

$$\langle Pv, w\rangle = \int_G \langle gv, w\rangle = \int_G \langle v, g^{-1}w\rangle$$

since G acts as unitary transformations. But using the invariance of the integral under $g \mapsto g^{-1}$ this is

$$\int_G \langle v, gw\rangle = \langle v, Pw\rangle.$$

6.3 Characters and orthogonality

Recall the trace of a matrix $\operatorname{tr} A = \sum_i A_{ii}$. It has the property that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, $\operatorname{tr} A^T = \operatorname{tr} A$. This has a more invariant interpretation if we use the canonical isomorphism between $\operatorname{End} V$ and $V^* \otimes V$. The (linear extension of) the map $\xi \otimes v \mapsto \xi(v)$ is the trace.

Definition 22 *The character of a representation $\rho : G \to \mathrm{Aut}(V)$ is the function $\chi_V(g) = \mathrm{tr}(\rho(g))$.*

Here are some properties of the character for a compact group, which is evidently a smooth function on G.

- $\chi_V(e) = \dim V$ since $\mathrm{tr}\, I_V = \dim V$
- $\chi_V(hgh^{-1}) = \chi_V(g)$ since $\mathrm{tr}\, \rho(hgh^{-1}) = \mathrm{tr}(\rho(h)^{-1}\rho(h)\rho(g)) = \mathrm{tr}(\rho(g))$
- If V, W are isomorphic as representations then they have the same character. This is again invariance of trace under conjugation.
- $\chi_{V\oplus W} = \chi_V + \chi_W$
- $\chi_{V\otimes W} = \chi_V \chi_W$. To see this note that if $v_1, \ldots, v_m$ is a basis for V and $w_1, \ldots, w_n$ a basis for W then $v_i \otimes w_j$ is a basis for the tensor product. So considering a linear map of the form

$$\sum_{ij} c_{ij} v_i \otimes w_j \mapsto \sum_{ij} c_{ij} Av_i \otimes Bw_j$$

summing the diagonal terms for basis vectors $v_i \otimes w_1$ gives $\mathrm{tr}\, AB_{11}$ and continuing gives $\mathrm{tr}\, A\, \mathrm{tr}\, B$.
- $\chi_{V^*}(g) = \chi_V(g^{-1})$. This is because $\mathrm{tr}(A^{-1})^T = \mathrm{tr}(A^{-1})$. But also the representation is unitary and for a unitary matrix $\mathrm{tr}\, A^{-1} = \mathrm{tr}\, \bar{A}^T = \mathrm{tr}\, \bar{A} = \overline{(\mathrm{tr}\, A)}$ and so also $\chi_{V^*}(g) = \overline{\chi_V(g)}$.

We see immediately from the previous section that

$$\int_G \chi_V(g) = \dim V^G$$

since this is the trace of the orthogonal projection P onto the invariant subspace.

We can define an L^2 inner product on smooth functions on G by integration and then we have:

Theorem 6.3 *For two representations V, W of a compact Lie group G*

$$\langle \chi_V, \chi_W \rangle = \dim \mathrm{Hom}^G(V, W).$$

In particular, if V and W are inequivalent irreducible representations, their characters are orthogonal.

Proof: Using $\mathrm{Hom}(V,W) \cong V^* \otimes W$ and the properties of characters above we see that

$$\chi_{V^*\otimes W} = \chi_{V^*}\chi_W = \overline{\chi_V}\chi_W.$$

Integrating the right hand term gives $\langle \chi_W, \chi_V \rangle$. Integrating the first gives $\dim \mathrm{Hom}_G(V,W)$.

If V, W are inequivalent then by Schur's lemma $\mathrm{Hom}^G(V,W) = 0$. In fact if $V = W$ then Schur says that $\dim \mathrm{Hom}^G(V,V) = 1$ and so χ_V has norm 1.

□

It is not just characters which satisfy orthogonality relations:

Theorem 6.4 *Let V, W be irreducible representations of a compact Lie group G and take $v_1, v_2 \in V$, $w_1, w_2 \in W$. Then*

$$\int_G \langle gv_1, v_2 \rangle \overline{\langle gw_1, w_2 \rangle}$$

vanishes if V and W are inequivalent and if $V \cong W$ equals

$$\frac{1}{\dim V}\langle v_1, w_1 \rangle \langle v_2, w_2 \rangle.$$

If $v_1, \ldots, v_m$ is a unitary basis for V then $\langle gv_i, v_j \rangle$ is the (i,j) entry in the matrix representing the action of G so the theorem tells us that the matrix entries of irreducible representations form an orthonormal set in $L^2(G)$.

Proof: Consider the term $\overline{\langle gw_1, w_2 \rangle}$ in this expression. Since a Hermitian inner product is antilinear in the second factor, for each $w \in W$, $\overline{\langle w, u \rangle}$ is complex linear in u and defines an element $\xi_w \in W^*$. In different words, a Hermitian inner product gives an antilinear isomorphism from W to its dual. Since G preserves the inner product, $\overline{\langle gw_1, w_2 \rangle} = g\xi_{w_1}(w_2)$ where g is the dual action.

Now

$$\int_G gv_1 \otimes g\xi_{w_1}$$

is the orthogonal projection P onto the invariant subspace of $V \otimes W^* = \mathrm{Hom}(W,V)$. But by Schur's lemma if V, W are irreducible this is zero. If $V = W$ then the invariant part consists of scalar multiples of the identity so for $A \in \mathrm{End}\, V$ the orthogonal projection onto multiples of I is

$$P(A) = \frac{1}{\dim V}(\mathrm{tr}\, A)I.$$

Now $\mathrm{tr}(v \otimes \xi) = \xi(v)$ so $P(v_1 \otimes \xi_{w_1}) = \xi_{w_1}(v_1)I/\dim V = \overline{\langle w_1, v_1\rangle} I/\dim V = \langle v_1, w_1\rangle I/\dim V$.

Evaluating $gv_1 \otimes g\xi_{w_1} \in V \otimes V^*$ on $\xi_{v_2} \otimes w_2$ is

$$\langle gv_1, v_2\rangle\langle w_2, gw_1\rangle = \langle gv_1, v_2\rangle\overline{\langle gw_1, w_2\rangle}$$

and so

$$\int_G \langle gv_1, v_2\rangle\overline{\langle gw_1, w_2\rangle} = \frac{1}{\dim V}\langle v_1, w_1\rangle \,\mathrm{tr}(\xi_{v_2} \otimes w_2) = \frac{1}{\dim V}\langle v_1, w_1\rangle\langle v_2, w_2\rangle.$$

□

If $v_1, \ldots, v_m$ is an orthonormal basis for V and $w_1, \ldots, w_n$ for W then $\langle gv_i, v_j\rangle$ is the (i, j) matrix coefficient for g acting on V and similarly for W. The theorem then says first of all that if V, W are inequivalent irreducible representations the matrix coefficients, as functions of G, are orthogonal. And if $V = W$ we have

$$\int_G \langle gv_i, v_j\rangle\overline{\langle gv_k, v_\ell\rangle} = \frac{1}{\dim V}\langle v_i, v_k\rangle\langle v_j, v_\ell\rangle.$$

The right hand side is zero unless $i = k$ and $j = \ell$ which means that the matrix coefficients $\langle gv_i, v_j\rangle$ are orthogonal.

For the circle, the matrix coefficients of irreducible representations are the functions $e^{in\theta}, n \in \mathbf{Z}$ and these form a complete orthonormal basis for the Hilbert space of all L^2 functions on S^1, or equivalently periodic functions on $[0, 2\pi]$. The same is true of any compact Lie group.

Remark: The proof we shall give is for a compact subgroup of $GL(n, \mathbf{R})$: a *matrix group*. In fact any compact subgroup is embedded in a general linear group, but this is a consequence of the theorem below given a general proof. Such a proof can be found in many texts (or online) using some functional analysis. For any compact Lie group, the image of Ad is of course a matrix group and Ado's theorem (which we did not prove either) tells us that some quotient of the universal covering is a matrix group.

If $G \subset GL(n, \mathbf{R})$ then it is a compact submanifold of $\mathbf{R}^{n^2}$. From point set topology any continuous function on G can be extended to a continuous function on $\mathbf{R}^{n^2}$ and the Weierstrass approximation theorem says that continuous functions there can be approximated by polynomials, so continuous functions on G can be approximated by polynomials in the entries of the $n \times n$ matrix.

Call the representation space for $G \subset GL(n, \mathbf{R})$ V (which may decompose as a sum of irreducibles) then the matrix coefficients for V are the linear polynomials and for $V \otimes V$ the homogeneous quadratic ones etc. So we see that the matrix coefficients for a *countable* collection of irreducible representations are dense in the continuous functions on G with the uniform norm. Then using the orthogonality of Theorem 6.4, by choosing orthogonal bases $\{v_1, \dots, v_{n_m}\}$ for each irreducible V_m and normalizing we get a complete orthonormal basis in $L^2(G)$.

Rather than dealing with an orthogonal sequence this result, the *Peter-Weyl theorem* is better stated as

Theorem 6.5 (Peter-Weyl) *Let G be a compact Lie group. Then*

$$L^2(G) \cong \hat{\bigoplus} V_L \otimes V_R^*,$$

the L^2-completion of the direct sum over all finite-dimensional irreducible representations of G, where V_L denotes the left action on functions and V_R the right action.

Proof: The matrix coefficients are the functions $f(g) = \langle gv_i, v_j \rangle$ as the v_i range through an orthonormal basis of V and the V run through (representatives of equivalence classes of) irreducible representations.

Consider the functions $\langle gv_i, v_j \rangle$. For fixed v_j this is a representation of G. Varying j shows that it has multiplicity $\dim V$ in $L^2(G)$. More invariantly we define an action of $G \times G$ on the space spanned by these functions

$$(h, k)\langle gv_i, v_j \rangle = \langle k^{-1}ghv_i, v_j \rangle$$

which gives the term $V_L \otimes V_R^*$. $\square$

Remark:

1. The theorem above holds for any compact group and in particular for a finite group. There $L^2(G)$ has dimension $|G|$ and is the *regular representation.*

2. The character χ_V of an irreducible representation V is the function on G given by

$$\chi_V(g) = \sum_1^n \langle gv_i, v_i \rangle.$$

In terms of the Peter-Weyl theorem this is the identity $I_V \in \mathrm{Hom}(V_R, V_L)$ on that factor.

3. The Hilbert space $L^2(G)$ above of complex functions is clearly the complexification of the space of real functions, even though the representations V may be complex. But they appear as $\mathrm{End}\, V \subset L^2(G)$ and $T(A) = A^*$ is an antilinear involution which commutes with the G-action, so these are real representations.

7 Maximal tori

7.1 Abelian subgroups

The eigenspaces of a unitary matrix A are orthogonal. This means we can find an orthonormal basis such that A is diagonal. Another way of saying this is to consider the set of diagonal matrices in $U(n)$: $\operatorname{diag}(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_n})$. This is a Lie subgroup isomorphic to the product of n copies of S^1: the torus T^n. The argument above shows that any element in $U(n)$ is conjugate to an element in T^n.

This is particularly relevant when considering the character χ_V of a representation V. Since $\chi_V(hgh^{-1}) = \chi_V(g)$, the function χ_V is determined by its restriction to T^n. Orthogonality of characters implies that the character uniquely determines V (up to equivalence of course) so this function on T^n determines V. Not only that, but since T^n is abelian, V can be expressed as a direct sum of one-dimensional invariant subspaces. Considering the actions of the separate factors, on each irreducible the action is given by a scalar of the form $\exp i(m_1\theta_1 + m_2\theta_2 + \cdots + m_n\theta_n)$ for $m_k \in \mathbf{Z}$. In other words V is determined by a collection of integer vectors. This is a general feature which we shall investigate next.

Definition 23 *A torus $T \subset G$ is a Lie subgroup T isomorphic to a product of circles. A maximal torus is maximal under the inclusion of tori.*

Tori, being the image of a compact group, are closed in G and hence embedded Lie subgroups. They are connected and so (using the exponential map) any proper inclusion $T \subset T'$ implies that $\dim T' > \dim T$. It follows that any torus is contained in a maximal one. A maximal torus is maximal among connected abelian subgroups A, for the closure of A is a torus.

Definition 24 *Let $T \subset G$ be a maximal torus. The Weyl group W of T is the quotient of the normalizer by the normal subgroup T*

$$W = N(T)/T = \{g \in G : gTg^{-1} = T\}/T$$

This definition seems to depend on the choice of maximal torus but we shall prove that in fact all maximal tori are conjugate so that the Weyl group (up to isomorphism) is independent of choice.

Example:

1. For $SO(3)$ a maximal torus is the circle subgroup of rotations by θ about an axis given by a unit vector $\mathbf{u}$. Conjugating by a rotation which takes $\mathbf{u}$ to $-\mathbf{u}$ takes θ to $-\theta$. The Weyl group is $\mathbf{Z}_2$.

2. Return to $U(n)$. A matrix which commutes with all diagonal matrices is diagonal so $\mathrm{diag}(e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_n})$ is a maximal torus. Let $h \in T$ have distinct diagonal entries. Then since conjugation leaves the eigenvalues unchanged $g \in N(T)$ must permute the entries. But permutation of the elements of an orthonormal basis is a unitary transformation π. It follows that $\pi^{-1}g$ leaves fixed the open set in T with distinct entries and by continuity every point. But then it commutes with all diagonals and so lies in T. Hence the Weyl group $N(T)/T$ is isomorphic to the symmetric group S_n.

Proposition 7.1 *The Weyl group is a finite group.*

Proof: The normalizer $N(T)$ acts on T and its Lie algebra, commuting with the exponential map. It thus preserves the kernel of exp which from Theorem 3.4 consists of integer multiples of basis vectors $v_1, \ldots, v_n$. Since this is discrete and the action is continuous an element connected to the identity acts trivially. Let N_0 be the component of the identity. It is a connected Lie subgroup which contains T and acts trivially on T by conjugation, so there is a one parameter subgroup which commutes with T, but this contradicts the maximality of T amongst connected abelian subgroups, so $N_0 = T$. Then $W = N(T)/T = N(T)/N_0$ is a compact group with the discrete topology and hence is finite.

□

7.2 Conjugacy of maximal tori

We shall prove that maximal tori are conjugate by using the map

$$F : G/T \times T \to G$$

(where G/T is the space of cosets of T) defined by $F(gT, t) = gtg^{-1}$. This is well-defined because replacing g by gs for $s \in T$ gives $gsts^{-1}g^{-1} = gtg^{-1}$. Clearly the image consists of elements conjugate to an element in T, so if we can prove F is surjective then we conclude that every element h is conjugate to an element of T.

Example: Consider $G = SO(3)$ and $T = S^1 \subset G$ the circle subgroup of rotations about the unit vector $\mathbf{k}$. To each coset gT associate the unit vector $g\mathbf{k}$ and this gives an identification of G/T with the unit sphere $S^2 \subset \mathbf{R}^3$. If $h = gtg^{-1}$ then $hg\mathbf{k} = gtg^{-1}g\mathbf{k} = g\mathbf{k}$ so the unit vector is the axis of rotation of h and t is the angle of rotation. But a rotation of θ about $\mathbf{u}$ is the same as a rotation of $-\theta$ about $-\mathbf{u}$, so the inverse image of a general point in $SO(3)$ is two points in $S^2 \times S^1$. If $h = I$, however, the inverse image is $S^2 \times \{e\}$.

This example reveals the general features: first that $G/T \times T$ and G are compact manifolds of the same dimension, and secondly that the smooth map F is a covering space on an open subset. We shall show this in general and use a theorem in the theory of manifolds: *If a smooth map F between two compact, connected, orientable manifolds has non-zero degree, then it is surjective.* A proof of this can be found in

https://people.maths.ox.ac.uk/hitchin/hitchinnotes/manifolds2012.pdf

but we will sketch the idea below.

The proper way to treat this is to use the theory of differential forms and the exterior derivative, but we have not introduced that here. However, we have a distinguished n-form on G, so we will work with a compact manifold M with an everywhere non-vanishing form Ω, whose existence is the definition of being orientable. We will call a differential n-form ω on a manifold of dimension n *exact* if there is a vector field X such that

$$\omega = \mathcal{L}_X \Omega$$

where we have taken the Lie derivative of Ω. Recall that diffeomorphisms act on n-forms and integrating X to a local one-parameter subgroup of diffeomorphisms we have

$$\mathcal{L}_X \Omega = \frac{\partial}{\partial t} \varphi_t^* \Omega |_{t=0}.$$

But since φ_t is a diffeomorphism connected to the identity the integral of $\varphi_t^* \Omega$ over a compact manifold M is constant (effectively just a change of variables in the integration). Thus the integral of an exact form is zero. In $\mathbf{R}^n$ if

$$X = \sum_i a_i \frac{\partial}{\partial x_i} \qquad \Omega \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \right) = f$$

then

$$\mathcal{L}_X \Omega = Xf + f \sum \frac{\partial a_i}{\partial x_i} = \operatorname{div} fX$$

so locally, if f has compact support, this is the divergence theorem. More importantly is the global converse theorem, which we do not prove here:

Proposition 7.2 *Two differential n-forms on a compact, connected, oriented manifold of dimension n differ by a divergence if their integrals are the same.*

Now differential forms, unlike vector fields, transform not just via diffeomorphisms but for arbitrary smooth maps $F : M \to N$. This is the pull-back $F^*\Omega$ of a form on N defined pointwise by

$$(F^*\Omega)_x(X_1, \ldots, X_n) = \Omega_{F(x)}(DF_x(X_1), \ldots, DF_x(X_n))$$

for tangent vectors X_i at $x \in M$. Then one can show that the pull-back of an exact form is exact: if $\tilde{\Omega}$ is a non-vanishing form on M then the required vector field Y on M is defined by

$$\tilde{\Omega}_x(Y, X_1, \ldots, X_{n-1}) = \Omega_{F(x)}(X, DF_x(X_1), \ldots, DF_x(X_{n-1})).$$

Now if $F : M \to N$ is a smooth map of compact manifolds which is not surjective, its image is compact and hence closed, so its complement is non-empty and open. We can then take an n-form $f\Omega$ on N where f has support in some small open set of the complement and such that the integral is 1. Then $F^*(f\Omega) = 0$ because f vanishes on the image of F. On the other hand, if we have a point x over which the map is a covering map, and we take a form $g\Omega$ of integral one in a small neighbourhood U, then the integral of $F^*(g\Omega)$ is a sum of terms over the finite number of open sets $U_1, \ldots, U_k$ whose union is $F^{-1}(U)$. Since $F : U_i \to U$ is a diffeomorphism the integral is ± 1 on each U_i depending on whether $F^*\Omega$ is a positive or negative multiple of $\tilde{\Omega}$. The degree is defined to be the sum of these terms, which is clearly an integer. If all signs are positive it counts the number of inverse images.

But since $f\Omega$ and $g\Omega$ have the same integral on N, they differ by an exact form, and hence so do $F^*(f\Omega)$ and $F^*(g\Omega)$ in which case $F^*(g\Omega)$ has the same integral on M as $F^*(f\Omega) = 0$. So if the degree is non-zero, there is a contradiction and the map is surjective.

To apply this to our situation we have three tasks:

1. Show that G/T is an orientable manifold

2. Find an element $g \in G$ over which F is a covering map

3. Calculate the orientations at the finite number of points $F^{-1}(g)$.

The space G/T is the set of cosets gT, so any point can be obtained by the left action of some g on the identity coset $[e] \in G/T$. In order to define a manifold structure it suffices to define a chart in a neighbourhood of the identity and transport it around by the left action. So let $\mathfrak{t} \subset T_eG$ be the tangent space of T at $e \in G$ and using the invariant inner product take its orthogonal complement $\mathfrak{t}^\perp$. Then the exponential map restricted to this gives, by the inverse function theorem, a submanifold of a neighbourhood U of $e \in G$ which intersects each coset in UT in one point. Using these charts (and the fact that T is a compact group to provide the Hausdorff condition) we give G/T the structure of a compact manifold of dimension $\dim G - \dim T$.

The action of $g \in G$ gives an isomorphism from the tangent space at $[e]$ to the tangent space at $[g]$, but this is only well-defined up to the right action of T on g. This is the adjoint action of T on $\mathfrak{t}^\perp$, so only properties of this vector space which are T-invariant can be propagated around G/T. For example, there are no invariant vectors since T is a maximal connected abelian subgroup so unlike G itself we don't have invariant vector fields. The inner product restricted to $\mathfrak{t}^\perp$ is invariant, and an inner product defines a multilinear form $\Omega(E_1, \dots, E_n) = 1$ for an oriented orthonormal basis $E_1, \dots, E_n$. Since T is connected it cannot alter the orientation so this defines a non-vanishing form on G/T which is therefore oriented.

2. In the example of $SO(3)$ we could take a general element of the torus T to get a covering and here we do the same, but interpret "general" as being a generator t in the sense that the closure of the subgroup $\{t^n; n \in \mathbf{Z}\}$ is T. If the kernel of the exponential map for T consists of $\mathbf{Z}$-linear combinations of $v_1, \dots, v_k$ then $t = \exp(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n)$ is a generator if $\{1, c_1, c_2, \dots, c_n\}$ are linearly independent over the rational numbers. One property of this choice is the following:

Proposition 7.3 *If $t \in T$ is a generator then there are $|W|$ points in the inverse image of F, where W is the Weyl group of T.*

Proof: If $F(gT, s) = gsg^{-1} = t$ then $s^n = g^{-1}t^n g$ and so $g^{-1}Tg \subset T$. Thus $g \in N(T)$ and so there is one coset gT for each $g \in N(T)/T = W$. The cosets are just the orbit of T under the left action of $N(T)$. □

To prove that F is a covering map we need its derivative at these $|W|$ points, and this is necessary for the orientation question too.

3. Define $\tilde{F} : G \times T \to G$ by $\tilde{F}(g, t) = gtg^{-1}$. This is a lift of the map F. By left translation we want to compute the derivative at (g, t) in terms of the geometry at the identity (e, e). This means first translating on $G \times T$ from (e, e) to (g, t), applying $\tilde{F}$ and then left multiplying by $\tilde{F}(g, t)^{-1}$. It is easiest to calculate imagining it as a matrix group: the derivative of gtg^{-1} is $\dot{g}tg^{-1} + g\dot{t}g^{-1} - gtg^{-1}\dot{g}g^{-1}$ and left translation by $(gtg^{-1})^{-1} = gt^{-1}g^{-1}$ gives

$$gt^{-1}g^{-1}\dot{g}tg^{-1} + gt^{-1}\dot{t}g^{-1} - \dot{g}g^{-1}$$

so if $X = g^{-1}\dot{g}$ is a left-invariant vector field on G and $Y = t^{-1}\dot{t}$ on T this is

$$\mathrm{Ad}(gt^{-1})(X) + \mathrm{Ad}(g)(Y) - \mathrm{Ad}(g)(X) = \mathrm{Ad}(g)(\mathrm{Ad}(t^{-1})X + Y - X).$$

For the derivative of F we want to restrict X to the subspace $\mathfrak{t}^{\perp}$ and then this is a linear map from $T_eG = \mathfrak{t}^{\perp} \oplus \mathfrak{t}$ to itself.

For a covering map we want this to be invertible. Since $\mathrm{Ad}(g)$ is invertible this will fail to be invertible only if $\mathrm{Ad}(t^{-1})$ has an eigenvalue $+1$ on $\mathfrak{t}^{\perp}$. But as remarked above that does not hold since T is maximal abelian. Now $\mathrm{Ad}(g)$ is orthogonal so its determinant is ± 1 for each of the $|W|$ points g_i over t. However, the g_i are all connected to the identity so $\det \mathrm{Ad}(g_i) = 1$ and hence $\det \mathrm{Ad}(g_i)(\det(\mathrm{Ad}(t^{-1}) - I))$ is the same non-zero number for each point and thus the degree is non-zero. We deduce therefore that F is surjective and hence

Theorem 7.4 *Every element $g \in G$ is contained in a maximal torus and all such tori are conjugate.*

The last part of the theorem follows by taking t' to be a generator of a maximal torus T'. Then $t' = gtg^{-1}$ for $t \in T$ and taking powers and the closure it follows that $T' \subset gTg^{-1}$ but by maximality this must be equality.

Remark: The power of this type of argument is that by investigating the equation $F(x) = g$ for a general point g we deduce the existence of a solution for any point. For example, suppose g is in the centre $Z(G)$, the subgroup which commutes with everything. Then by the theorem it lies in a maximal torus T, but $hgh^{-1} = g$ so it lies in all conjugates, hence in the intersection of all maximal tori. Conversely, since every element is in a maximal torus, which is abelian, every element commutes with the intersection, which is therefore equal to the centre. We saw earlier that a covering homomorphism $G \to H$ is given by $H = G/\Gamma$ for $\Gamma \subset Z(G)$. It follows that the maximal tori of H are quotients of those of G by Γ.

For example $\mathrm{diag}(e^{i\theta}, e^{-i\theta})$ is the maximal torus of $SU(2)$ which contains the centre $\pm I$. The adjoint representation $SU(2) \to SO(3)$ is a covering homomorphism in which $-I$ acts trivially. The group $SO(3)$ has no centre.

From the theorem it follows that, up to conjugation, there is a single Weyl group W and the dimension of T, called the *rank*, is an invariant of G. There is one further point:

Proposition 7.5 *The dimension of G/T is even.*

Proof: The tangent space at $[e] \in G/T$ is $\mathfrak{t}^\perp$ and this is a real representation of the compact abelian group T. We saw that $\mathrm{Ad}(t^{-1})$ has no eigenspace with eigenvalue $+1$. Nor does it have one with -1, for if so $\mathrm{Ad}(t^2)$ would have a $+1$ eigenvalue. But t^2 is also a generator since if $\{1, c_1, c_2, \dots, c_n\}$ are linearly independent over the rationals so are $\{1, 2c_1, 2c_2, \dots, 2c_n\}$. It follows that the real irreducible subspaces have dimension 2 and hence $\dim \mathfrak{t}^\perp$ is even. $\square$

7.3 Roots

In the proof of Proposition 7.5 we used the fact that $\mathfrak{t}^\perp$ is a direct sum of irreducible real 2-dimensional representation spaces for T, or homomorphisms $T \to SO(2)$. So the tangent space at e, or equivalently the Lie algebra, is a direct sum

$$\mathfrak{g} = \mathfrak{t} \oplus \bigoplus_a \mathfrak{g}_a.$$

Let $k = \dim \mathfrak{t}$ be the rank of G. Choosing an orientation on $\mathbf{R}^2$ allows us to identify the rotation action on $\mathfrak{g}_a$ as $e^{2\pi i\theta_a}$ where, for $(e^{2\pi i x_1}, e^{2\pi i x_2}, \dots, e^{2\pi i x_k}) \in T$, we have $\theta_a = m_1 x_1 + \cdots + m_k x_k$ for integers m_i. Choosing the opposite orientation gives $-\theta_a$.

Definition 25 *The linear forms $\pm\theta_a \in \mathfrak{t}^*$ are called the roots of G.*

Example:

1. For $U(n)$ we have seen that T consists of the diagonal matrices and the Lie algebra is the space of skew Hermitian matrices, so $B_{ij} = -\bar{B}_{ji}$. The 1-dimensional complex space of entries in the upper triangular (i, j)-place with $i < j$ is a real 2-dimensional space $\mathfrak{g}_a$ since it is acted on by T as $e^{ix_i}e^{-ix_j}$. Thus the roots are $x_i - x_j$, $i \neq j$.

2. For $SU(n)$ the only difference is that since $\det A = 1$ the maximal torus is given by $\sum_i x_i = 0$. In particular for $SU(2)$ the roots are $\pm(x_1 - x_2) = \pm 2x_1$.

3. If $G = SO(2m)$ then the maximal torus consists of m 2×2 blocks down the diagonal each of the form

$$\begin{pmatrix} \cos 2\pi x_i & \sin 2\pi x_i \\ -\sin 2\pi x_i & \cos 2\pi x_i \end{pmatrix}.$$

It acts on the space of skew-symmetric matrices broken up into 2×2 blocks, with the $i < j$ place a real 2×2 matrix acted on the left by a rotation by $2\pi x_i$ and on the right by $-2\pi x_j$. This is a tensor product $V_i \otimes V_j$ and complexifying and looking at the eigenspaces the four one-dimensional representations are $e^{2\pi i(\pm x_i \pm x_j)}$. Thus the roots are $\pm x_i \pm x_j$ for $i < j$.

Given a root θ_a the kernel of the homomorphism $T \to S^1$ is a subtorus of dimension $k-1$ which acts trivially on $\mathfrak{g}_a$. So take a one-parameter subgroup with tangent $X \in \mathfrak{g}_a$. Its closure together with the kernel generates a connected abelian subgroup. In fact since the maximum dimension is k it will be automatically closed (there is a geometric reason for this which will become evident below). It follows that points on a *root hyperplane* defined by $\theta_a(x) = 0$ are tangent to many maximal tori.

Conversely, suppose that $t \in T$ is not contained in the kernel of any of the root homomorphisms. This means its action on each $\mathfrak{g}_a$ is non-trivial. But this is the condition in the proof of Theorem 7.4 for the derivative of F to be an isomorphism and F to be a covering map in a neighbourhood. Moreover this neighbourhood contains generators of T, so as in the proof of Proposition 7.3 the cosets in the fibre is an orbit of $N(T)$.

Hence if t is contained in another maximal torus hTh^{-1} then $t = hsh^{-1}$ for some $s \in T$ which means (hT, s) is in $F^{-1}(t)$ but then $hT = gT$ for $g \in N(T)$, so $h = gu$, $u \in T$ and $h \in N(T)$. This means $hTh^{-1} = T$.

Hence if t is not contained in the kernel of any of the root homomorphisms it lies in a unique maximal torus. The root planes are thus the tangent spaces of the intersection of other maximal tori with T.

Example: In $U(n)$ the root planes are $x_i - x_j = 0$, so a matrix A lies in a unique maximal torus if its eigenvalues $e^{2\pi i x_i}$ are distinct.

The subgroup $N(T) \subset G$ acts on $\mathfrak{g}$ via the adjoint representation preserving $\mathfrak{t} \subset \mathfrak{g}$, but T acts trivially on itself by conjugation, so the quotient $W = N(T)/T$, the Weyl group, acts. If $\rho : T \to \mathrm{Aut}(\mathfrak{g})$ is the adjoint action then $\rho(gtg^{-1}) = \mathrm{Ad}(g)\rho(t)\,\mathrm{Ad}(g)^{-1}$, for $g \in N(T)$, is an equivalent representation thus it permutes the irreducible components of the action and so permutes the roots.

The discussion above shows that if $X \in \mathfrak{t}$ does not lie in any root hyperplanes then its orbit under W has $|W|$ elements and so the action on this open subset is free. In particular W embeds as a finite subgroup of linear automorphisms of $\mathfrak{t}$. It also preserves the inner product, so it embeds in $O(k)$.

Example:

1. For $SU(3)$ the rank $k = 2$ and $W = S_3$ so this symmetric group embeds as a subgroup of $O(2)$: it can only be the dihedral group of rotations and reflections of an equilateral triangle.

2. For $SU(4)$ we have $S_4 \subset O(3)$, the symmetries of the cube.

If $X \in \mathfrak{t}$ lies in a root hyperplane then there are elements of W which leave it fixed. We prove the following:

Theorem 7.6 *The orthogonal reflection in each root hyperplane lies in the Weyl group.*

Since reflection in a hyperplane has one eigenvalue -1 and the rest 1, its determinant is -1 and so W never lies in $SO(k)$.

Proof: The 2-dimensional root space $\mathfrak{g}_a$ has a positive definite inner product: choose an orthonormal basis X_1, X_2. Then $X_1 \pm iX_2$ is acted on by T as $e^{\pm 2\pi i\theta_a(x)}$ and so $[X_1, X_2] = i[X_1 + iX_2, X_1 - iX_2]/2 = \alpha$ is invariant by T and so lies in $\mathfrak{t}$. It is non-zero, for otherwise the root hyperplane together with X_1, X_2 span an abelian subalgebra of $\mathfrak{g}$ of dimension $k+1$ which contradicts maximality.

Since T preserves $\mathfrak{g}_\alpha$ this means that $\{X_1, X_2, \alpha\}$ spans a 3-dimensional non-abelian Lie subalgebra. Its elements are skew-adjoint with respect to the inner product so ad defines an isomorphism to $\mathfrak{so}(3)$. There is thus a corresponding connected Lie subgroup G_a of G which is a covering of $SO(3)$. But the double covering of $SO(3)$ is $SU(2)$ which is simply connected so the Lie subgroup is either $SU(2)$ or $SO(3)$ which are compact and hence the subgroup is embedded.

Take $Y \in \operatorname{Ker} \theta_a$, then $[Y, X_1] = [Y, X_2] = 0$ since Y acts trivially on $\mathfrak{g}_a$. And $[\alpha, Y] = 0$ because $Y, \alpha \in \mathfrak{t}$ which is abelian. Hence G_a fixes the hyperplane.

Consider $(Y, \alpha) = (Y, [X_1, X_2]) = (Y, \operatorname{ad}(X_1)X_2)$. Since the inner product is bi-invariant, $\operatorname{ad}(X_1)$ is skew adjoint and so

$$(Y, \operatorname{ad}(X_1)X_2) = -(\operatorname{ad}(X_1)Y, X_2) = -([X_1, Y], X_2)$$

which vanishes since $[X_1, Y] = 0$. Hence α is orthogonal to the root hyperplane.

There is a rotation in $SO(3)$ which takes α to $-\alpha$ and a corresponding element $g \in G_a$. This fixes pointwise $\operatorname{Ker} \theta_a$ and preserves $\operatorname{Ker} \theta_a \oplus \mathbf{R}\alpha = \mathfrak{t}$ and so lies in $N(T)$. Consequently it is an action of an element of order 2 in the Weyl group. $\square$

Remark: Every one-parameter subgroup in $SU(2)$ or $SO(3)$ is a circle and embedded. The maximal tori which intersect the root plane are then generated by the $(k-1)$-dimensional kernel of $\rho_a : T \to S^1$ and one of the 2-parameter family of circle subgroups of G_a.

8 Representations and maximal tori

8.1 The representation ring

We have observed that given representations V, W we can form the direct sum $V \oplus W$ and the tensor product $V \otimes W$. We can also write $mV = V \oplus V \oplus \cdots \oplus V$ for a positive integer m. If we now take equivalence classes $[V]$ of representations and allow negative values of m we get a ring $R(G)$ called the *representation ring*. Since an equivalence class is determined by its character χ_V and $\chi_{V\oplus W} = \chi_V + \chi_W$, and $\chi_{V\otimes W} = \chi_V \chi_W$ this is also the *character ring* consisting of finite integer combinations of the functions χ_V. We take complex representations here.

Example: For the circle S^1, every irreducible representation is of the form U_n with character $e^{in\theta}$. So any representation is a sum of these and $R(S^1) = \mathbf{Z}[t, t^{-1}]$, the ring of finite Laurent series in t where t is the basic representation U_1 with character $e^{i\theta}$.

Restricting a representation ρ of G to a maximal torus T, it splits into one-dimensional irreducibles with character of the form $\exp 2\pi i(a_1x_1 + \cdots + a_mx_m)$ and these linear maps $a_1x_1 + \cdots + a_mx_m$ are called the *weights* of the representation. The roots are the weights for the adjoint representation.

The representation ring $R(T)$ for $T = S^1 \times S^1 \times \cdots \times S^1$ is, from the example above, the Laurent polynomials in k variables where $k = \dim T$: $R(T) = \mathbf{Z}[t_1^{\pm 1}, t_2^{\pm 1}, \ldots, t_k^{\pm 1}]$. Just as the Weyl group permutes the roots, it also permutes the weights of ρ and so the character of a representation restricted from G must be invariant under W.

Proposition 8.1 *The restriction map $R(G) \to R(T)^W$, to the fixed part under the Weyl group, is injective.*

Proof: By orthogonality of characters a representation is determined by its character χ, and the character is conjugation-invariant so since every element is conjugate to one in T, χ is determined by its restriction to T. The conjugation action of $N(T)$ leaves χ invariant so its restriction to T must be invariant under the Weyl group. $\square$

Example: If $G = SU(2)$, $T = \mathrm{diag}(e^{i\theta}, e^{-i\theta})$ and the Weyl group is $S_2 = \mathbf{Z}_2$ taking θ to $-\theta$ (a rather simple case of reflection!) or t to t^{-1}. So $R(T)^W$ is the ring of Laurent polynomials of the form

$$\pi = m_0 + m_1(t + t^{-1}) + \cdots + m_n(t^n + t^{-n}). \tag{4}$$

The character χ_n of the irreducible representation V_n in Section 6.1 is

$$t^n + t^{-n} + t^{n-2} + t^{-(n-2)} + \dots$$

so we can rewrite π as a linear combination of irreducible characters:

$$\pi = m_0 + m_1\chi_1 + m_2(\chi_2 - \chi_0) + \dots + m_n(\chi_n - \chi_{n-2}).$$

The character of any irreducible representation can therefore be written in this form. By orthogonality of characters, it must therefore be one of the V_n. The character of any genuine representation is a sum of non-negative multiples of χ_n: a Laurent polynomial (4) where $m_i - m_{i+2} > 0$.

Since

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

π can also be written as a polynomial in a single variable $t + t^{-1}$ so algebraically $R(T)^W = \mathbf{Z}[t + t^{-1}]$.

In the example, we saw that the restriction map $R(G) \to R(T)^W$ was surjective and hence an isomorphism. This is a general fact though we shall not prove it here. Identifying the irreducible characters in the representation ring is in general far more difficult than the case of $SU(2)$. We shall demonstrate the isomorphism in another example next.

8.2 Representations of $U(n)$

The Weyl group for $U(n)$ is the symmetric group S_n acting on an n-dimensional torus. Any polynomial in $t_1, \dots, t_n$ which is symmetric is itself a polynomial in the *elementary symmetric functions* σ_m: the coefficient of x^m in the expansion of $(x-t_1)(x-t_2)\dots(x-t_n)$. The representation ring $R(T) = \mathbf{Z}[t_1^{\pm}, \dots, t_n^{\pm}]$ and involves the negative powers, but we can convert such a Laurent polynomial into a genuine polynomial by multiplying by a power of $\sigma_n = t_1 t_2 \dots t_n$. If we can find representations of $U(n)$ with character $\sigma_1, \sigma_2, \dots, \sigma_n, \sigma_n^{-1}$ then we will have an isomorphism $R(U(n)) \cong R(T)^W$ and

$$R(T)^W \cong \mathbf{Z}[\sigma_1, \sigma_2, \dots, \sigma_n, \sigma_n^{-1}]$$

as a ring.

If V is the n-dimensional vector space which defines $U(n)$ we shall use the *exterior powers* $\Lambda^m V$ for $0 \leq m \leq n$. This is the algebra which lies behind the differential forms which we used in discussing integration, but now using a complex vector space. The easiest way to define $\Lambda^m V$ is the dual space of the vector space of alternating multilinear functions $M(v_1, \ldots, v_m)$ on V. So $\Lambda^1 V$ is the dual space of V^* which is canonically V itself. Given vectors $v_1, \ldots, v_m \in \Lambda^m V$ we define a linear function of multilinear forms M by

$$v_1 \wedge v_2 \wedge \cdots \wedge v_m(M) \stackrel{\text{def}}{=} M(v_1, \ldots, v_m).$$

This expression is linear in each v_i and the alternating property of M means it changes sign if any two v_i are interchanged. A general element in $\Lambda^m V$ is a linear combination of terms like this. In particular, if $v_1, \ldots, v_n$ is a basis for V, a basis for $\Lambda^m V$ is provided by $v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_m}$ where $i_1 < i_2 < \cdots < i_m$. A matrix $A \in U(n)$ acts as

$$A \sum_{i_1 < \cdots < i_m} a_{i_1 i_2 \ldots i_m} v_{i_1} \wedge v_{i_2} \wedge \ldots v_{i_m} = \sum_{i_1 < \cdots < i_m} a_{i_1 i_2 \ldots i_m} A v_{i_1} \wedge A v_{i_2} \wedge \ldots A v_{i_m}$$

Taking the maximal torus as the diagonal matrices with respect to the unitary basis $v_1, \ldots, v_n$, the character of this representation is the elementary symmetric function σ_m. When $m = n$, $\Lambda^n V$ is one-dimensional since if $w_i = \sum_j A_{ij} v_j$, then

$$w_1 \wedge w_2 \wedge \cdots \wedge w_n = \det A v_1 \wedge v_2 \wedge \cdots \wedge v_n.$$

Its dual space has character σ_n^{-1}. Hence the generators of $R(T)^W$ are all characters of representations of $U(n)$ and so $R(U(n)) \cong R(T)^W$.

These representations are irreducible because if $U \subset \Lambda^m V$ is an invariant subspace, it is invariant by T and hence a sum of one-dimensional representations given by the weights. But the weights are all distinct so some $v_{i_1} \wedge v_{i_2} \wedge \ldots v_{i_m} \in U$ and the character χ_U contains the monomial $t_{i_1} t_{i_2} \ldots t_{i_m}$. But it is invariant under the symmetric group and so contains all such terms and is the character of $\Lambda^m V$.

Example: The adjoint representation has weights which are the roots $\pm(x_i - x_j)$and n trivial weights, so its character is

$$\chi_{Ad} = n + \sum_{i \neq j} t_i t_j^{-1}.$$

Multiply by $\sigma_n = t_1 t_2 \ldots t_n$ and we get $n\sigma_n + 2\sigma_{n-1}$ so in $R(T)^W$ we have

$$\chi_{Ad} = n + 2\frac{\sigma_{n-1}}{\sigma_n}.$$

Elements of the group $SU(n)$ have determinant 1 and the determinant is the action on $\Lambda^n V$ which is therefore the trivial representation and $t_1 \dots t_n = 1$, which is the equation of its maximal torus. In this case the generators σ_n, σ_n^{-1} are removed. In particular when $n = 2$ we recover the description of $R(SU(2))$ above.

The representation ring contains the characters of all representations. An additive basis is provided by the irreducible ones but the ring by itself does not identify these. This requires the further study of roots, weights and the Weyl group which is beyond this course. By the same token, it is difficult to determine the product in the ring in terms of a basis of irreducibles. This is an important issue: given two irreducible representations V, W, how does $V \otimes W$ break up as a sum of irreducible components?

We can answer this for $SU(2)$, since we know that the V_n, the representation on homogeneous polynomials $f(z_1, z_2)$ of degree n, give all the irreducibles.

Proposition 8.2 (Clebsch-Gordan) *The tensor product of the irreducible representations V_m, V_n (where $m \geq n$) of $SU(2)$ decomposes as*

$$V_m \otimes V_n = V_{m+n} \oplus V_{m+n-2} \oplus \dots \oplus V_{m-n}.$$

Proof: The character of V_n is $t^n + t^{n-2} + \dots + t^{-n}$ or

$$\chi_{V_n} = \frac{t^{-n}(1 - t^{2n+2})}{(1 - t^2)}.$$

Hence

$$\chi_{V_m \otimes V_n} = \chi_{V_m} \chi_{V_n} = \frac{t^{-m-n}(1 - t^{2m+2})}{(1 - t^2)}(1 + t^2 + \dots + t^{2n})$$

and collecting terms the numerator is

$$t^{-m-n}(1 - t^{2m+2n+2}) + t^{-m-n+2}(1 - t^{2m+2n}) + \dots$$

□

8.3 Integration on T

The argument for injectivity $R(G) \subset R(T)^W$ was based on orthogonality of characters. This means by integration over G. But since $\chi_V(ghg^{-1}) = \chi_V(h)$ the function itself is uniquely determined by its restriction to T, so there should be a formula for integrating a function of this form (a *class function*) over G in terms of an integral over T.

In fact the formula comes from our discussion in 7.2 of the map $F : G/T \times T \to G$ together with the properties of root hyperplanes in 7.3. Recall that when $g \in G$ lies in a unique maximal torus the map F in a neighbourhood of $F^{-1}(g)$ is a covering transformation of degree $|W|$ and so there is a dense open set of G over which F is a covering. In particular, a set of full measure, so we can restrict our integral to this set.

Further, because F is a local diffeomorphism, the integral of an n-form transforms via the determinant of DF, relative to the invariant n-forms on G and $G/T \times T$. But we calculated this, $forg = t$, as

$$(\det(\mathrm{Ad}(t^{-1}) - I))$$

acting on $\mathfrak{t}^{\perp}$ although we omitted to determine the sign of this.

However, t preserves each root space $\mathfrak{g}_a$ and acts via the root θ_a so the determinant is a product of terms

$$\det \begin{pmatrix} \cos 2\pi\theta_a(x) - 1 & \sin 2\pi\theta_a(x) \\ -\sin 2\pi\theta_a(x) & \cos 2\pi\theta_a(x) - 1 \end{pmatrix} = 2(1 - \cos 2\pi\theta_a(x))$$

and, since $\theta_a(x) \neq 0$ because x is not on the root hyperplane, this is positive.

The degree of the covering is $|W|$ so the integral over $F^{-1}(U)$ for a coordinate neighbourhood U in G is $|W|$ times the integral over U.

Now suppose $f : G \to \mathbf{C}$ is a class function, and consider $f \circ F : G/T \times T \to G$. This is $f(gT, t) = f(gtg^{-1}) = f(t)$. For fixed t this is constant on G/T. Applying Fubini's theorem we get

Proposition 8.3 (Weyl integration formula) *If f is a class function,*

$$\int_G f(g) = \frac{1}{|W|} \int_T \det(\mathrm{Ad}(t^{-1}) - I)|_{G/T} f(t).$$

Example: For $SU(2)$ with roots $\pm 2x$, we have

$$\det(\mathrm{Ad}(t^{-1}) - I)|_{G/T} = 2(1 - \cos 4\pi x) = -(t - t^{-1})^2$$

if $t = e^{2\pi i x}$. Then since the Weyl group has order 2, the Weyl integration formula gives

$$\int_{SU(2)} f(g) = \int_0^1 (1 - \cos 4\pi x) f(x) dx.$$

We can test orthogonality of the characters of V_m, V_n this way by replacing this by the contour integral around a circle C. Since $\chi_{V_m} = e^{2\pi imx} + \cdots + \ldots e^{-2\pi imx}$ is real

$$\int_{SU(2)} \chi_{V_m}\bar{\chi}_{V_n} = \int_{SU(2)} \chi_{V_m}\chi_{V_n} = -\frac{1}{2}2\pi i \int_C (t-t^{-1})^2 \frac{t^{-m-n}(1-t^{2n+2})(1-t^{2m+2})}{(1-t^2)^2} t^{-1} dt$$

which is by Cauchy's residue theorem the residue at $t = 0$ of

$$-\frac{1}{2t^3}(t^{-m-n} - t^{-m+n+2} - t^{m-n+2} + t^{m+n+4})dt$$

and this vanishes unless $m = n$ in which case it is 1.

9 Simple Lie groups

9.1 The Killing form

So far we have worked on a compact Lie group with a bi-invariant positive definite inner product on the Lie algebra constructed by averaging. There are choices: at one extreme is the torus where any left-invariant inner product is bi-invariant. But for other groups there is a natural choice. Since $\operatorname{tr} AA^T = \sum_{i,j} A_{ij}^2$, $\operatorname{tr} AB^T$ is a positive definite inner product on the space of all $n \times n$ real matrices so for $SO(n)$, whose Lie algebra consists of the skew-symmetric matrices, $\operatorname{tr} A^2 = -\operatorname{tr} AA^T$ is negative-definite and, since $\operatorname{tr} PXP^{-1} = \operatorname{tr} X$, is invariant under the adjoint action.

Every compact Lie group has a homomorphism $\mathrm{Ad} : G \to SO(\mathfrak{g})$ and so we can use this natural inner product on $SO(\mathfrak{g})$ to give:

Definition 26 *The Killing form on a Lie algebra $\mathfrak{g}$ is the symmetric bilinear form*

$$(X, Y) = \operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y).$$

For a compact Lie group G this is negative definite by the observation above so long as $\operatorname{ad} : \mathfrak{g} \to \operatorname{End}(V)$ has zero kernel. The centre $Z(G)$ acts trivially by conjugation and so acts trivially under Ad, and so its Lie algebra lies in this kernel. Conversely if $\operatorname{ad} X = 0$ the exponential of tX acts trivially on $\mathfrak{g}$ and hence on a neighbourhood of the identity in G so if G is connected it lies in the centre.

Example:

1. For $SL(2, \mathbf{R})$, which is not compact, the Killing form is non-degenerate but not positive definite. In fact the adjoint representation is a covering map $SL(2, \mathbf{R}) \to SO(2, 1)$ and the form has signature $(2, 1)$.

2. For $SO(n)$ the Killing form is $(n - 2) \operatorname{tr} XY$ which is negative definite for $n > 2$.

3. The Killing form for $U(n)$ is not negative definite since elements $e^{i\theta} I$ lie in the centre. It is negative definite for $SU(n)$ and is $2n \operatorname{tr} XY$.

9.2 Ideals and simplicity

An abstract group is *simple* if it has no nontrivial normal subgroups. If $H \subset G$ is a connected normal Lie subgroup of a Lie group then its Lie algebra $\mathfrak{h}$ is preserved by the adjoint action of G, and so for $Y \in \mathfrak{h}$ and $X \in \mathfrak{g}$, $[X, Y] \in \mathfrak{h}$.

Definition 27 *An ideal in a Lie algebra $\mathfrak{g}$ is a subspace V such that for $Y \in V$ and $X \in \mathfrak{g}$, $[X, Y] \in V$.*

With an Ad-invariant inner product on $\mathfrak{g}$, the endomorphism $\operatorname{ad} X$ is skew adjoint. It follows that if V is preserved by all transformations of the form $\operatorname{ad} X$, so is its orthogonal complement $V^\perp$ which is therefore an ideal also.

Definition 28 *A Lie algebra is simple if it is non-abelian and has no non-trivial ideals. It is semi-simple if it is a direct sum of simple Lie algebras.*

For a connected Lie group G, the ideals are the invariant subspaces for the adjoint representation, and when G is compact we have complete reducibility of any representation so $\mathfrak{g}$ splits as a direct sum of trivial representations and non-trivial irreducible ones. Thus if there are no trivial ones, the Lie algebra is semisimple.

The connected Lie group corresponding to an ideal in $\mathfrak{g}$ is a normal subgroup i.e. preserved by conjugation. A trivial ideal is the Lie algebra of a connected Lie subgroup in the centre of G.

Definition 29 *A connected Lie group is simple if it is non-abelian and has no non-trivial connected normal Lie subgroups. It is semisimple if its Lie algebra is semisimple.*

Note that a simple Lie group may have a discrete centre which is of course a normal subgroup.

Although a semisimple Lie algebra is a direct sum of simple Lie algebras, Theorem 5.2 shows that the corresponding statement about semisimple Lie groups only holds for simply-connected ones.

Examples:

1. The group $U(n)$ is not semisimple because it has a 1-dimensional centre: the scalar matrices $e^{i\theta}$.

2. The case of $SU(n)$ is simple because the adjoint representation is irreducible. In fact since the roots $x_i - x_j$ are permuted by the Weyl group, they are the weights of a single representation so any complementary representation must be acted on trivially by the maximal torus. But these are diagonal matrices and, unless they are scalars (which is impossible if the trace is zero) they can be conjugated to be non-diagonal.

3. The centre of $SU(n)$ consists of scalars ωI where $\omega^n = 1$, the cyclic group $\mathbf{Z}_n$. So take the diagonal copy of $\mathbf{Z}_n \subset \mathbf{Z}_n \times \mathbf{Z}_n$ and the quotient $SU(n) \times SU(n)/\mathbf{Z}_n$. This group is semisimple but not a product.

The result that the Lie algebra of a compact group is a direct sum of simple ones does not require a positive definite invariant inner product. In particular it can be used for groups for which the Killing form is nondegenerate.

Theorem 9.1 *If the Lie algebra* $\mathfrak{g}$ *contains no abelian ideals and admits an invariant (possibly indefinite) inner product then it is semisimple.*

Proof: Let $\mathfrak{m}$ be a minimal ideal. The span $[\mathfrak{m}, \mathfrak{m}]$ of commutators $[X, Y]$ is again an ideal, contained in $\mathfrak{m}$. It cannot be zero, since $\mathfrak{m}$ would then be abelian, so it must be $\mathfrak{m}$. Take its orthogonal complement $\mathfrak{m}'$ which is an ideal. Note that for an indefinite inner product there exist null vectors, so a vector space can intersect non-trivially its orthogonal complement.

If $\mathfrak{m} \cap \mathfrak{m}' = 0$ we have $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{m}'$, a direct sum of Lie algebras. Moreover any abelian ideal in either factor is an ideal in $\mathfrak{g}$. The inner product is nondegenerate on each factor too if it is nondegenerate on the sum. So we can continue, reducing dimension each time.

If $\mathfrak{m} \cap \mathfrak{m}' \neq 0$ it must be $\mathfrak{m}$ by minimality. Together with $\mathfrak{m} = [\mathfrak{m}, \mathfrak{m}]$ this means that the inner product is zero on $\mathfrak{m}$.

Take $A = \sum_i [B_i, C_i] \in \mathfrak{m}$ where $B_i, C_i \in \mathfrak{m}$. Then if $X \in \mathfrak{g}$

$$(A, X) = \sum_i ([B_i, C_i], X) = \sum_i (B_i, [C_i, X]) = 0$$

since $[C_i, X] \in \mathfrak{m}$ as $\mathfrak{m}$ is an ideal. But this contradicts the nondegeneracy of the inner product. $\square$

With this approach, by Lie's third theorem one can reduce the classification of universal covers of compact Lie groups to the discussion of simple Lie groups.

9.3 Classification

Simply-connected compact simple Lie groups are classified. They are:

- The special unitary group $SU(n)$
- The double cover $Spin(n)$ of the special orthogonal group $SO(n)$

- The quaternionic unitary group $Sp(n)$. The notation here is a little ambiguous (Sp = symplectic). The reason is that the *complexification* of the Lie algebra is $\mathfrak{sp}(2n, \mathbf{C})$, the Lie algebra of the non-compact group of complex $2n \times 2n$ matrices leaving fixed a nondegenerate skew symmetric form.

- The exceptional groups G_2, F_4, E_6, E_7, E_8. The subscript denotes the rank of the group.

The group $Sp(n)$ is one we haven't discussed. The quaternions $\mathbf{H}$ consist of real linear combinations of $1, i, j, k$ where i, j, k satisfy the defining relations

$$i^2 = j^2 = k^2 = ijk = -1.$$

So $q = x_0 + ix_1 + jx_2 + kx_3$ is a quaternion. Multiplication is associative but not commutative. The quaternionic conjugate is $\bar{q} = x_0 - ix_1 - jx_2 - kx_3$ and $\overline{pq} = \bar{q}\bar{p}$.

The group $Sp(n)$ consists of $n \times n$ matrices A with quaternionic entries such that $A\bar{A}^T = I$. When $n = 1$ this is the 3-sphere $x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1$ which we have seen as $SU(2)$. In fact in low dimensions there are special isomorphisms among the simply-connected groups:

$$Sp(1) \cong SU(2) \quad Spin(4) \cong Sp(1) \times Sp(1) \quad Spin(5) \cong Sp(2) \quad Spin(6) \cong SU(4).$$

More invariantly, if we consider $\mathbf{H}^n$ as a quaternionic vector space by right multiplication by $\mathbf{H}$, then $Sp(n)$ is the group of orthogonal matrices commuting with this action. Another even more invariant way to say this is to pick out the i in the quaternions and then regard j as an antilinear map J such that $J^2 = -1$. In Section 6.1 we saw that the action of $SU(2)$ on the representation space V_n commuted with such an action when n was odd. This means that V_{2m-1} corresponds to a homomorphism $SU(2) \to Sp(m)$.

The exceptional groups have a rather more complicated description, especially E_8. It is often said that if you want to prove something about Lie groups by case-by-case treatment using the classification, by the time you get to E_8 you can see the general proof. However G_2 is more amenable and we shall finally discuss this.

9.4 The group G_2

The "exceptional" group G_2 is in some respects not exceptional and has more in common with the "classical" groups like $SO(n)$. To justify this, consider $O(n)$ defined as the subgroup of $GL(n, \mathbf{R})$ which preserves the positive definite bilinear form

$(x,y) = x_1y_1 + \cdots + x_ny_n$. An invertible matrix $A \in GL(n, \mathbf{R})$ acts on all such bilinear forms by $B(x,y) \mapsto B(Ax, Ay)$. Positive definiteness is an open condition and any positive definite form can be transformed to (x,y) by this action. So the *orbit* of B in the space of symmetric bilinear forms (which has dimension $n(n+1)/2$) is open, and the stabilizer is a Lie group of dimension $n^2 - n(n+1)/2 = n(n-1)/2$ conjugate to $O(n)$.

Instead of symmetric bilinear forms we can consider alternating trilinear forms $T(x,y,z)$. In fact every Lie algebra has one $(X, [Y,Z])$ but these are very special. In a 7 dimensional space the picture is more like the symmetric forms: there are open orbits of the 49-dimensional Lie group $GL(7, \mathbf{R})$ on the 35-dimensional space $(\Lambda^3\mathbf{R}^7)^*$ of trilinear forms T and the stabilizer of a point has dimension $49 - 35 = 14$. This will be the group G_2. To show this we need to pick a representative form T and show that its stabilizer is a compact subgroup of dimension 14.

We take an orthonormal basis $e_1, \ldots, e_7$ of $\mathbf{R}^7$ and set

$$T = e_7(e_1e_2 + e_3e_4 + e_5e_6) + e_1e_3e_5 - e_1e_4e_6 - e_2e_3e_6 - e_2e_4e_5.$$

This terminology means that evaluating T on three of these vectors is the coefficient in the above expression, taking account of the alternating property, so $1 = T(e_7, e_1, e_2) = -T(e_1, e_7, e_2)$ etc.

Definition 30 *The Lie group G_2 is the identity component of the stabilizer in $SO(7)$ of the alternating trilinear form T.*

We shall then prove

Proposition 9.2 *The form T lies in an open orbit of $GL(7, \mathbf{R})$ in the vector space of alternating trilinear forms and the stabilizer in $GL(7, \mathbf{R})$ of any form in this open set is conjugate to G_2.*

Proof: The Lie algebra action of a matrix A is

$$T(Ax, y, z) + T(x, Ay, z) + T(x, y, Az).$$

First we shall find a 14-dimensional Lie subalgebra of $\mathfrak{so}(7)$ which leaves T invariant, then show that the connected Lie group corresponding to it is the stabilizer of T in $GL(7, \mathbf{R})$.

1. For each $i \neq j$ let $E^{(ij)}$ denote the skew symmetric matrix taking e_i to e_j and e_j to $-e_i$ and zero on all other basis vectors. Consider $E^{(71)}$. It transforms T to

$$e_1(e_3e_4 + e_5e_6) - e_7(e_3e_5 + e_4e_6).$$

But $E^{(36)} + E^{(45)}$ transforms T to

$$-2e_1(e_3e_4 + e_5e_6) + 2e_7(e_3e_5 + e_4e_6)$$

and so $2E^{(71)} + E^{(36)} + E^{(45)}$ lies in the Lie algebra of the stabilizer. Continuing with $E^{(7i)}$ for $1 \leq i \leq 6$ we have a 6-dimensional space which leaves T invariant and to find the other elements we need only restrict to linear combinations of $E^{(ij)}$ for $i, j \neq 7$. This is the Lie subalgebra of $\mathfrak{so}(6)$ leaving fixed e_7.

If e_7 is fixed and T is fixed, then so are the alternating 2-form $e_1e_2 + e_3e_4 + e_5e_6$ and the 3-form $e_1e_3e_5 - e_1e_4e_6 - e_2e_3e_6 - e_2e_4e_5$ which make up T. The skew bilinear form B defined by the first is $B(x, y) = (Ix, y)$ where $Ie_1 = e_2, Ie_2 = -e_1$ et. So $e_i + ie_2, e_3 + ie_4, e_5 + ie_6$ form a basis for $\mathbf{R}^6$ as a complex vector space $\mathbf{C}^3$. The elements in $SO(6)$ which commute with I are orthogonal and complex linear hence in $U(3)$. The 3-form can be written

$$e_1e_3e_5 - e_1e_4e_6 - e_2e_3e_6 - e_2e_4e_5 = \mathrm{Re}(e_1 + ie_2)(e_3 + ie_4)(e_5 + ie_6)$$

which lies in the complex 1-dimensional space $\Lambda^3\mathbf{C}^3$. The group $U(3)$ acts on this by the complex determinant $e^{i\theta}$, so if the real part is preserved the group must lie in $SU(3)$.

Thus $SU(3)$ lies in the stabilizer of T and the Lie algebra $\mathfrak{g}_2$ is $\mathfrak{su}(3) \oplus \mathbf{R}^6$ which has dimension 14.

2. Now suppose U is any trilinear form on $\mathbf{R}^7$. We shall define an associated inner product. This involves the exterior product of two forms. Given M of degree p and N of degree q the exterior product $M \wedge N$ is defined as

$$M \wedge N(x_1, \ldots, x_{p+q}) = \frac{1}{p!q!} \sum_{\sigma \in S_{p+q}} M(x_{\sigma(1)}, \ldots, x_{\sigma(p)}) N(x_{\sigma(p+1)} \ldots, x_{\sigma(p+q)}).$$

If α, β are two linear forms then $\alpha \wedge \beta = -\beta \wedge \alpha$ and more generally $M \wedge N = (-1)^{pq} N \wedge M$.

For a vector $x \in \mathbf{R}^7$ define a 2-form B_x by $B_x(y, z) = U(x, y, z)$, Then consider

$$B_x \wedge B_y \wedge U.$$

Since 2-forms commute this is symmetric in x, y and takes values in the 1-dimensional space of 7-linear forms in 7 dimensions. This is almost an inner product, except it takes values in a one-dimensional space on which $A \in GL(7, \mathbf{R})$ acts as $\det A$. So a representative matrix B transforms like $B \mapsto ABA^T \det A$ and so $\det B \mapsto (\det A)^9 B$.

Then $B/(\det B)^{1/9}$ defines an inner product if $\det B \neq 0$. This may not be positive definite.

3. Now consider T. The given positive definite inner product for which $e_1, \ldots, e_7$ is an orthonormal basis is the one determined as above. In fact since T is invariant by $SU(3)$ this is an invariant inner product on $\mathbf{R}^6 \oplus \mathbf{R}$. As a real representation of $SU(3)$ $\mathbf{R}^6$ is irreducible so it can only be a scalar multiple of the standard one so it suffices to check that e_7 and e_1 have the same length in the canonical inner product. In this case

$$B_{e_7} = e_1e_2 + e_3e_4 + e_5e_6$$

and so $B_{e_7} \wedge B_{e_7} = 6e_1e_2e_3e_4e_5e_6e_7$. And

$$B_{e_1} = -e_7e_2 + e_3e_5 - e_4e_6$$

which gives $B_{e_1} \wedge B_{e_1} = 6e_1e_2e_3e_4e_5e_6e_7$. These have the same sign and value so the inner product is positive definite and (up to a factor) the given one.

We conclude from this that the connected component of the stabilizer of T under the action of $GL(7, \mathbf{R})$ lies in $SO(6)$ and so from the first part is a 14-dimensional compact Lie group. The map $A \mapsto T(Ax, Ay, Az)$ from $GL(7, \mathbf{R})$ to the 35-dimensional space of trilinear forms is smooth. The derivative at the identity has, as we have seen, a 14-dimensional kernel and so since $49 - 14 = 35$ has a surjective derivative. By the inverse function theorem it maps to an open set and so the orbit of T is open. $\square$

We can now consider properties of G_2 as a Lie group, using the subgroup $SU(3)$.

- We have seen that the Lie algebra $\mathfrak{g}_2 = \mathfrak{su}(3) \oplus \mathbf{R}^6$ and the $\mathbf{R}^6$ is the defining 3-dimensional complex vector space for $SU(3)$. Its complexification is the representation $V \oplus \bar{V}$ in the notation of Section 8.2. So the maximal torus T of $SU(3)$ acts on $\mathbf{R}^6$ with weights $\pm x_1, \pm x_2, \pm x_3$. Since none of these is zero, T is the maximal torus of G_2 which is therefore a group of rank 2. Moreover the roots are (where $x_1 + x_2 + x_3 = 0$):

$$\pm(x_1 - x_2), \pm(x_2 - x_3), \pm(x_3 - x_1), \pm x_1, \pm x_2, \pm x_3.$$

- The Weyl group is generated by reflections about these six root planes and is the dihedral group of symmetries in $O(2)$ of a regular hexagon.

- The group G_2 is *simple*: any ideal in $\mathfrak{g}_2$ is a representation space for $SU(3)$ so (since $SU(3)$ is simple) if G_2 had an ideal it would be $\mathfrak{su}(3)$ and also its orthogonal complement $\mathbf{R}^6$. But $\mathbf{R}^6$ is not closed under Lie bracket. We could check this directly but since it has no zero weights, as a representation of $SU(3)$, all the Lie brackets would vanish. But then we would have a 6-dimensional abelian subalgebra which contradicts the rank calculation.

- The group G_2 has the same maximal torus as the subgroup $SU(3)$ but the Weyl group consists of the symmetries of the hexagon instead of the triangle: it has the extra symmetry of multiplication by -1. This acts on $R(T)$ by the involution $(t_1, t_2, t_3) \mapsto (t_1^{-1}, t_2^{-1}, t_3^{-1})$. We saw that $R(T)^W$ for $SU(3)$ was $\mathbf{Z}[\sigma_1, \sigma_2]$ where

$$\sigma_1 = t_1 + t_2 + t_3 \qquad \sigma_2 = t_1t_2 + t_2t_3 + t_3t_1 \qquad t_1t_2t_3 = 1.$$

So the extra involution takes σ_1 to $\sigma_2/\sigma_3 = \sigma_2$ and $\sigma_1 + \sigma_2, \sigma_1\sigma_2$ are generate the invariant subring.

Now the 7-dimensional representation of G_2 splits as $1 \oplus V \oplus \bar{V}$ as a representation of $SU(3)$ with character

$$1 + t_1 + t_2 + t_3 + t_1^{-1} + t_2^{-1} + t_3^{-1} = 1 + \sigma_1 + \sigma_2$$

and the adjoint representation

$$2 + t_1t_2^{-1} + \cdots + t_1 + t_2 + t_3 + t_1^{-1} + t_2^{-1} + t_3^{-1}.$$

But

$$\sigma_1\sigma_2 = (t_1 + t_2 + t_3)(t_1^{-1} + t_2^{-1} + t_3^{-1})$$

so the character of the adjoint representation is

$$\sigma_1\sigma_2 - 1 + \sigma_1 + \sigma_2.$$

We see therefore that $R(G_2) = R(T)^W = \mathbf{Z}[\sigma_1 + \sigma_2, \sigma_1\sigma_2]$ is a polynomial ring on two generators.