Lie groups and Lie algebras

Wilfried Schmid
Lecture notes by Tony Feng

Spring 2012

Contents

Disclaimer

These are course notes that I wrote for Math 222: Lie Groups and Lie Algebras, which was taught by Wilfried Schmid at Harvard University in Spring 2012. There are, undoubtedly, errors, which are solely the fault of the scribe.

I thank Eric Larson for pointing out some corrections on earlier versions of these notes.

Chapter 1

Geometric preliminaries

1.1 Lie groups

Definition 1.1.1. A *Lie group*, G, is a group endowed with the structure of a C^∞ manifold such that the inversion map

$$\begin{aligned} s : G &\to G \\ g &\mapsto g^{-1} \end{aligned}$$

and multiplication map

$$\begin{aligned} m : G \times G &\to G \\ (g, h) &\mapsto gh \end{aligned}$$

are smooth.

Remark 1.1.2.

1. It suffices to require that the single map $(g, h) \mapsto gh^{-1}$ is smooth. For example, the inversion map is smooth because it is the restriction of this map to $e \times G$.
2. With moderate effort, one can show that requiring all structures to be C^2 is enough to imply smoothness.
3. Historically, Lie defined a Lie group not as above, but as (in modern times) the germ of a Lie group action.

Let us now formally define the notion of "Lie group action."

Definition 1.1.3. Suppose G is a Lie group and M is a C^∞ manifold. An *action* of G on M is a map $A : G \times M \to M$ such that

$$A(g(A(h, m))) = A(gh, m).$$

A more common notation for the action is $g \cdot m$; in these terms, this condition is $g \cdot (h \cdot m) = (gh) \cdot m$.

Definition 1.1.4. Suppose G, H are Lie groups. A *homomorphism* from G to H is a C^∞ map $F : G \to H$ which is also a group homomorphism.

Definition 1.1.5. Suppose G is a Lie group. A subgroup H of G in the algebraic sense is a *Lie subgroup* if $H \subset G$ is a locally (with respect to H) closed submanifold of G.

It is not obvious from this definition that a Lie subgroup is even a Lie group, so we will prove this presently.

Proposition 1.1.6. *A Lie subgroup is a Lie group when equipped with the induced C^∞ structure and the inclusion $H \to G$ is a Lie group homomorphism.*

Proof. First we clarify the meaning of "locally closed submanifold." This means that for every $h \in H$, there exists an open neighborhood U_H of h in H and an open neighborhood U_G of h in G such that

1. $U_H = U_G \cap H$.

2. $U_H \subset U_G$ is a closed submanifold, e.g. looks locally like $\mathbb{R}^k \subset \mathbb{R}^n$.

There is a subtlety here that we only have to consider a subset *in* H; this allows possible strange global behavior such as H actually "bunching up" at some point.

Let $N \subset M$ be a locally closed submanifold. A continuous function f on an open subset $U_N \subset N$ is C^∞ with respect to the induced C^∞ structure on N if and only if for all p in the domain of f, there exist open neighborhoods U_N, U_M of the domain of f such that $f|_{U_N}$ is the restriction of a C^∞ function $\widetilde{F} : U_M \to \mathbb{R}$ to U_N.

We just saw that $H \subset G$ is a submanifold such that the inclusion $H \to G$ is a C^∞ map. It is a subgroup by definition. The map $h \mapsto h^{-1}$ is a C^∞ map to G because it is the restriction of the inversion map on G to H. That tells us that the inversion map is a C^∞ map to G taking values in H. By the definition of the induced C^∞ structure (namely, C^∞ maps to H are those which pull back C^∞ maps from H to C^∞ maps on the domain), this is a C^∞ map to H.

□

Example 1.1.7. The following are examples of Lie groups.

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

2. Giving $\mathbb{Q}$ the discrete topology, $\mathbb{R} \times \mathbb{Q} \subset \mathbb{R}^2, +$ is a Lie subgroup of dimension 1 of $(\mathbb{R}^2, +)$. We recall that there are two topological restrictions on topological spaces that can be manifolds: Hausdorff, and second countable. Therefore, the same construction with $\mathbb{R}$ replacing $\mathbb{Q}$ would be a non-example.

3. $(\mathbb{R}^n/\mathbb{Z}^n, +)$: an n-dimensional torus. Smoothness, etc. follow because they are *local* properties.

4. $\mathrm{GL}_n(\mathbb{R})$, the group of invertible $n \times n$ matrices with real entries under matrix multiplication. This has a Lie group structure from its embedding as an open subset of $\mathbb{R}^{n^2}$ by matrix entries.

5. $\mathrm{GL}_n(\mathbb{C})$, with matrix multiplication.

We can summarize Lie theory as the (shockingly successful) study of *linearization*, whereby one tries to study non-linear structure of a Lie group by the linear structure of its Lie algebra.

Theorem 1.1.8. *Let G be a Lie group, $H \subset G$ a subgroup in the algebraic sense. If H is closed as a subset of G, then H has a unique structure of a Lie subgroup.*

Combining this result with example (4), we get lots of examples, including all the classical Lie groups: $\mathrm{SO}(n), \mathrm{SU}(n), \mathrm{Sp}(n, \mathbb{R}), \mathrm{SO}(p, q), \mathrm{SU}(p, q), \ldots$.

Exercise 1.1.9. (a) Show directly (i.e. not using the preceding theorem) that $\mathrm{SO}(n)$ has a natural Lie group structure, and (b) $\mathrm{SO}(n) \subset \mathrm{GL}(n, \mathbb{R})$ is a Lie subgroup.

1.2 Lie algebras

Suppose K is a field.

Definition 1.2.1. A *Lie Algebra* $\mathfrak{g}$ over K is a vector space over K equipped with a bilinear map called the *Lie bracket*

$$\begin{aligned} \mathfrak{g} \times \mathfrak{g} &\to \mathfrak{g} \\ (a, b) &\mapsto [a, b] \end{aligned}$$

such that

1. $[a, a] = 0$ for all $a \in \mathfrak{g}$.

2. $[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$ for all $a, b, c \in \mathfrak{g}$.

Remark 1.2.2. The definition tells us that the Lie bracket is skew-symmetric, so $[a, b] = -[b, a]$. Moreover, these conditions are equivalent if the characteristic of K is not 2.

Definition 1.2.3. A *homomorphism* of Lie algebras $F : \mathfrak{g} \to \mathfrak{g}$ is a K-linear map which preserves the Lie bracket, i.e.

$$[Fa, Fb] = F[a, b].$$

Definition 1.2.4. A *Lie subalgebra* of $\mathfrak{g}$ is a K-linear subspace $\mathfrak{h} \subset \mathfrak{g}$ such that $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$.

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

Definition 1.2.5. An *ideal* in $\mathfrak{g}$ is a K-linear subspace $\mathfrak{h} \subset \mathfrak{g}$ such that $[\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h}$.

If $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then $\mathfrak{g}/\mathfrak{h}$ has a unique Lie algebra structure so that the quotient map $\mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$ is a homomorphism of Lie algebras. The Lie bracket on the quotient is $[\overline{a}, \overline{b}] = \overline{[a, b]}$ for lifts a, b of $\overline{a}, \overline{b}$.

Example 1.2.6.

1. Suppose A is an associative algebra over K. Then we can turn A into a Lie algebra by defining $[a, b] = ab - ba$. The first condition is obvious, and the second can be checked by explicit computation. In particular, we may take $A = \mathrm{End}(V)$ for some vector space V.

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

In some sense, all Lie algebras arise in this way, though perhaps indirectly.

Example 1.2.7. Let M be a C^∞ manifold, $A = C^\infty(M)$. Then $\mathcal{D}$ is the derivations of $C^\infty(M)$, i.e. the Lie algebra of vector fields.

1.3 Vector fields

Let M be a C^∞ manifold. By a “function” on M we mean either a real-valued function or complex-valued function unless otherwise specified. Recall that a *vector field* on M is a derivation of $C^\infty(M)$, which is the space of smooth functions on M.

Lemma 1.3.1 (“Localization of vector fields.”)**.** *Suppose X is a vector field on M, $p \in M$, and f_1, f_2 are functions on M such that $f_1 = f_2$ on some open neighborhood of p. Then $Xf_1(p) = Xf_2(p)$.*

Proof. We use the existence of “cutoff functions.” Namely, suppose $U_0 \subset M$ is open and $U_1 \subset U_0$ is open such that $\overline{U_1}$ is compact and contained in U_0. Then there exists a function $u \in C^\infty(M)$ such that

(i) $f \equiv 1$ on U_1, and

(ii) $\mathrm{Supp}\, f \subset U_0$.

This follows from the existence of partitions of unity: just take an appropriate open cover (i.e. wherein U_0 is the only open intersecting U_1) and a partition of unity and consider the sum of those parts whose support intersects U_1.

Apply this with U_0, U_1 nested open neighborhoods containing p and $f_1 \equiv f_2$ on U_0. Then $hf_1 \equiv hf_2$, since $f_1 \equiv f_2$ on the support of h. Then $X(hf_1) = X(hf_2)$, so

$$X(hf_i)(p) = h(p)Xf_i(p) + f_i(p)Xh(p).$$

Equating these for $i = 1, 2$, we conclude that $Xf_1(p) = Xf_2(p)$.

□

Corollary 1.3.2. *Suppose X is a vector field on M and $U \subset M$ is open. Then X can be restricted to U.*

Proof. If $f \in C^\infty(U)$, $p \in U$. Choose $h \in C^\infty(M)$ such that $h \equiv 1$ near p and $\operatorname{Supp} h \subset U$. Define $X|_U f = X(hf)(p)$. This is well-defined by the lemma. □

Corollary 1.3.3. *Vector fields can be patched, i.e. given open subset $U_1, U_2 \subset M$ and vector fields X_i on U_i for $i = 1, 2$ such that $X_1|_{U_1 \cap U_2} = X_2|_{U_1 \cap U_2}$, then there exists a unique vector field X on $U_1 \cup U_2$ such that $X|_{U_i} = X_i$.*

Proof. At any $p \in U_i$, define $(Xf)(p)$ by $X_i f|_{U_i}(p)$. □

These two corollaries together tell us that vector fields constitute a *sheaf.*

Definition 1.3.4. A **coordinate neighborhood** in M is a pair $(U, x_1, \ldots, x_n)$ with $U \subset M$ open and $x_1, \ldots, x_n \in C^\infty(M)$ real-valued such that $p \mapsto (x_1(p), \ldots, x_n(p)) \in \mathbb{R}^n$ establishes a diffeomorphism between U and its image in $\mathbb{R}^n$ (which must be open).

Corollary 1.3.5. *Suppose X is a vector field on M and $(U, x_1, \ldots, x_n)$ is a coordinate chart. Then $X|_U$ can be expressed uniquely as*

$$X|_U = \sum_{j=1}^{u} a_j \frac{\partial}{\partial x_j}, \qquad a_j \in C^\infty(M).$$

Conversely, any such expression describes a unique vector field on M.

Proof. Defining $a_j = Xx_j$. Suppose $f \in C^\infty(U)$ and $p \in U$. Then we may write

$$f = f(p) \cdot 1 + \sum_i \frac{\partial f}{\partial dx_i}(p)(x_i - x_i(p)) + h,$$

where h is a linear combination of products $h_1 h_2$ with $h_1(p) = h_2(p) = 0$ (this is just Taylor's theorem). Note that $X1 = 0$ since X is a derivation. On a similar note, $X(h_1 h_2)(p) = h_1(p)Xh_2(p) + h_2(p)Xh_1(p) = 0$. Therefore,

$$(Xf)(p) = \sum_i X\left(\frac{\partial f}{\partial x_i}(p) x_i\right) = \left(\sum a_i \frac{\partial f}{\partial x_i}\right)(p).$$

□

1.4 Tangent and cotangent spaces

Definition 1.4.1. A **tangent vector** at p is a linear map $X_p : C^\infty(M) \to \mathbb{R}$ or $\mathbb{C}$ such that

$$X_p(f_1 f_2) = f_1(p) X_p f_2 + f_2(p) X_p f_1.$$

Corollary 1.4.2. *If X_p is a tangent vector at p and $f \in C^\infty(M)$, then $X_p f$ depends only on the values of f near p (i.e. in any neighborhood of p).*

This tells us that $X_p f$ is well-defined for any C^∞ function f defined on some neighborhood of p; that is, tangent vectors are defined on *germs* of functions at p.

Definition 1.4.3. The vector space (over $\mathbb{R}$ or $\mathbb{C}$ as appropriate) of tangent vectors at p is called the **tangent space** of M at p, and denoted $T_p M$.

For $p \in M$, define

$$\mathscr{I}_p = \{f \in C^\infty(M) \mid f(p) = 0\},$$

i.e. the ideal of functions vanishing at p. Given $f \in C^\infty(M)$, we may write

$$f = f(p) \cdot 1 + f_p, \qquad f_p \in \mathscr{I}_p.$$

Denote by $\mathscr{I}_p^2$ the usual product of ideals. Suppose $X_p \in T_p M$ and $f \in \mathscr{I}_p^2$; by the derivation property, $X_p f = 0$.

Corollary 1.4.4. *The map $X_p \mapsto (f \mapsto X_p f)$, where $X_p \in T_p M$, determines a linear inclusion $T_p M \to (\mathscr{I}_p/\mathscr{I}_p^2)^*$.*

Proof. We have a direct sum decomposition $C^\infty(M) = \mathscr{I}_p \oplus \mathbb{R}$ (or $\mathbb{C}$). Since X_p kills constants and $\mathscr{I}_p^2$, it is uniquely determined by its action on $(\mathscr{I}_p/\mathscr{I}_p^2)$. □

Definition 1.4.5. $\mathscr{I}_p/\mathscr{I}_p^2 \simeq T_p^* M$ is the **cotangent space** to M at p.

In fact, the inclusion in the previous corollary is an isomorphism, which we will prove presently.

Corollary 1.4.6. *If $(U; x_1, \ldots, x_n)$ is a coordinate neighborhood of p, then $\{\frac{\partial}{\partial x_1}|_p, \ldots, \frac{\partial}{\partial x_n}|_p\}$ is a basis of $T_p M$.*

Proof. Completely analogous to Corollary 1.3.5. □

For $f \in C^\infty(M)$ and $p \in M$, define $df|_p \in T^*M$ as the image in $T^*M \simeq \mathscr{I}_p/\mathscr{I}_p^2$ of $f - f(p) \cdot 1$. If $(U; x_1, \ldots, x_n)$ is a coordinate neighborhood of p, then $\{dx_1|_p, \ldots, dx_n|_p\}$ is a basis of $T_p^* M$. This is essentially Taylor's theorem: we can write any function as a constant plus first order terms plus higher order terms about p,

$$f - f(p) = \sum \frac{\partial f}{\partial x_i}(p)(x_i - x_i(p)) + \ldots = \sum \frac{\partial f}{\partial x_i}(p) dx_i.$$

By counting dimensions, we conclude that $T_p^* M = (T_p M)^*$.

Given a vector field X on M and $p \in M$, define $X_p \in T_p M$ by $X_p f = (Xf)(p)$. Therefore, any vector field X is completely determined by $X_p \in T_p M$ for all p. Conversely, given an assignment $p \mapsto X_p \in T_p M$, there exists a vector field X whose values are the X_p if and only if for every $f \in C^\infty(M)$, the function $p \mapsto X_p f$ is smooth. (If this is satisfied, it is clearly a derivation $C^\infty(M) \to C^\infty(M)$ since it is a derivation pointwise.) Since smoothness is a local property, we can check this criterion on coordinate neighborhoods.

1.5 Pullbacks and pushforwards

Suppose $F : M \to N$ is a C^∞ map between manifolds. Then we can define the "pullback" $F^* : C^\infty(N) \to C^\infty(M)$ by $F^*f = f \circ F$.

If $p \in M$, then $F^*\mathscr{I}_{N,f(p)} \subset \mathscr{I}_{M,p}$ and $F^*\mathscr{I}^2_{N,f(p)} \subset \mathscr{I}^2_{M,p}$, so we get a linear map $F^* : T^*_{F(p)}N \to T^*_pM$. Dually, there is a map $F_* : T_pM \to T_{F(p)}N$.

Also, for any $f \in C^\infty(N)$, F^* sends $df|_{F(p)} \in T^*_{F(p)}N \mapsto d(F^*f)|_p \in T^*_pM$. In other words, d computes with pullbacks: $F^*d = dF^*$. This is clear from the interpretation of d as subtracting a constant from the function to put it in the ideal sheaf.

We can pull back functions and differential forms under smooth maps, but a problem arises when we attempt to push vector fields forward. We will address this problem now.

Definition 1.5.1. Let $F : M \to N$ be a smooth map. Two vector fields X_M on M and X_N on N are *F-related* if for every $f \in C^\infty(N)$,

$$F^*(X_N f) = X_M(F^*f).$$

In other words, the following diagram commutes.

$$\begin{tikzcd} C^\infty(N) \arrow[r, "X_N"] \arrow[d, "F^*"] & C^\infty(N) \arrow[d, "F^*"] \\ C^\infty(M) \arrow[r, "X_M"] & C^\infty(M) \end{tikzcd}$$

Yet another equivalent formulation is that $F_*X_M|_p = X_N|_{F(p)}$ for every $p \in M$. You can see the difficulty with pulling back in general: the left hand sides must agree for all p lying over a common point in N.

Exercise 1.5.2. Check that if X_M, X_N and Y_M, Y_N are pairs of F-related vector fields, then $[X_M, Y_M]$ is F-related to $[X_N, Y_N]$.

Note that if F is a diffeomorphism, then for any vector field X on M there exists a unique vector field F_*X on N which is F-related to X. After all, M and N are basically the same manifold. We can define F^*X by the formula above since we know that the pre-image of any point exists and is unique.

Chapter 2

The Lie algebra of a Lie group

2.1 Invariant vector fields

Suppose G is a Lie group, $g \in G$. Define C^∞ maps

$$\begin{aligned} \ell_g : G &\to G \\ h &\mapsto gh \end{aligned} \qquad\qquad \begin{aligned} r_g : G &\to G \\ h &\mapsto hg^{-1}. \end{aligned}$$

These are both diffeomorphisms by the definition of Lie group.

Definition 2.1.1. The **Lie algebra** $\mathfrak{g}$ of G is the Lie algebra of all left-invariant vector fields, i.e. vector fields X such that $(\ell_{g*})X = X$ for all $g \in G$, with the Lie bracket being the commutator.

Note that G-invariant means $\ell_{g*}X_h = X_{gh}$ for all $g, h \in G$. This is equivalent to $\ell_{g*}X_e = X_g$ for all $g \in G$. This gives an inclusion $\mathfrak{g} \to T_eG$ sending $X \mapsto X_e$ (any left-invariant vector field is completely determined by its value at the identity).

Exercise 2.1.2. Show that this map is an isomorphism: $\mathfrak{g} \simeq T_eG$.

Suppose that $F : G \to H$ is a homomorphism of Lie groups. Define $F_* : \mathfrak{g} \to \mathfrak{h}$ as the composition

$$\mathfrak{g} \simeq T_eG \xrightarrow{F_*} T_eH \simeq \mathfrak{h}.$$

Theorem 2.1.3. *The map $F_* : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras.*

Remark 2.1.4. If F_1, F_2 are Lie group homomorphisms that can be composed, then $F_{1*} \circ F_{2*} = (F_1 \circ F_2)_*$. This follows from the analogous statement about pullbacks of functions. Therefore, "Lie algebra" is a *covariant functor* from the category of Lie groups (with Lie group homomorphisms) to Lie algebras.

Proof. Suppose $X, Y \in \mathfrak{g}$. Define $\widetilde{X} \in \mathfrak{h}$ by the requirement that $F_*X_e = \widetilde{X}_e \in T_eH$ and similarly for Y, $\widetilde{Y}$.

We claim that $X, \widetilde{X}$ and $Y, \widetilde{Y}$ are F-related. To see this, suppose $g \in G$. Then

$$\begin{aligned}\widetilde{X}_{F(g)} &= \ell_{F(g)*}\widetilde{X}_e = \ell_{F(g)*}F_*X_e = (\ell_{F(g)} \circ F)_*X_e \\ &= (F \circ \ell_g)_*X_e = F_*\ell_{g*}X_e = F_*X_g,\end{aligned}$$

and similarly for $Y, \widetilde{Y}$.

Therefore, $[X, Y]$ and $[\widetilde{X}, \widetilde{Y}]$ are F-related, which means that for all $f \in C^\infty(H)$,

$$F^*(\widetilde{X}f) = X(F^*f).$$

Applying this to $[\widetilde{X}, \widetilde{Y}] = \widetilde{X}\widetilde{Y} - \widetilde{Y}\widetilde{X}$ we see that

$$[\widetilde{X}, \widetilde{Y}] = \widetilde{[X, Y]}.$$

By the same calculations as before, $[X, Y]$ and $\widetilde{[X, Y]}$ are F-related.

□

2.2 Integral curves

Definition 2.2.1. Suppose M is a manifold, X a vector field on M. An *integral curve* to X is a pair (I, φ) where $I \subset \mathbb{R}$ is an open interval and $\varphi : I \to M$ is a smooth map satisfying

$$\varphi_*\frac{d}{dt}|_{t_0} = X|_{\varphi(t_0)} \text{ for all } t_0 \in I.$$

Equivalently, $\frac{d}{dt}$ and X are φ-related.

Properties of integral curves.

1. If (I, φ) is an integral curve, then so is $(I + a, \varphi_a)$, where $\varphi_a(t) = \varphi(t - a)$.

2. If (I, φ) is an integral curve, then $(s^{-1}I, \varphi(st))$ is an integral curve for sX, for $s \neq 0$. This is actually true for $s = 0$ if we interpret $s^{-1}I = \mathbb{R}$.

3. Suppose $(U; x_1, \ldots, x_n)$ is a coordinate neighborhood and X is a vector field on U with local expression

$$X = \sum a_i \frac{\partial}{\partial x_i}, \qquad a_i \in C^\infty(U).$$

   Then (I, φ), where $\varphi : I \to M$ is given by $t \mapsto (x_1(t), \ldots, x_n(t))$ is an integral curve if and only if $\frac{dx_i}{dt} = a_i((x_i(t)))$.

The first and third facts follow directly from the chain rule. The third also follows from the chain rule:

$$\varphi_*\frac{d}{dt} = \sum \frac{dx_i}{dt}\frac{\partial}{\partial x_i}.$$

Therefore, finding integral curves of vector fields amounts to solving an ODE.

Basic facts on ODE.

Theorem 2.2.2 (Existence of local solutions). *If X is a vector field on M and $p \in M$, then there exists an integral curve (I, φ) with $0 \in I$ and $\varphi(0) = p$.*

Theorem 2.2.3 (Uniqueness). *Suppose (I_1, φ_1) and (I_2, φ_2) are two integral curves for X, and $\varphi_1(t_0) = \varphi_2(t_0)$ for some $t_0 \in I_1 \cap I_2$. Then $\varphi_1 = \varphi_2$ on $I_1 \cap I_2$, and hence there exists a unique integral curve $(I_1 \cup I_2, \varphi)$ with $\varphi|_{U_i} = \varphi_i$.*

In particular, given an integral curve (I, φ) there exists a unique maximal extension, i.e. an integral curve $(\widetilde{I}, \widetilde{\varphi})$ such that $I \subset \widetilde{I}$ and $\widetilde{\varphi}|_I = \varphi$ with $\widetilde{I}$ maximal. This is called a *maximal integral curve.*

Theorem 2.2.4 (C^∞ dependence on initial conditions). *If (I, φ) is an integral curve and $p \in M$, then there exists an open neighborhood U of $p \in M$ and $\epsilon > 0$, and a C^∞ map $\Phi : U \times (-\epsilon, \epsilon) \to M$ such that*

(i) $t \mapsto \Phi(\cdot, t)$ is an integral curve for $t \in (-\epsilon, \epsilon)$.

(ii) $\Phi(q, 0) = q$ for any $q \in U$.

This says that we can choose the interval of an integral curve *locally uniformly* in p, and depending smoothly on p.

Definition 2.2.5. A vector field X on M is *complete* if every maximal integral curve is defined on all of $\mathbb{R}$.

Example 2.2.6.

1. Let $M = \mathbb{R}^2 - \{(0, 0)\}$ and $X = \frac{\partial}{\partial x}$. Then the integral curves are linearly parametrized by lines parallel to the x-axis. An integral curve on the x axis cannot be extended to all of $\mathbb{R}$ since there is a "hole" at $(0, 0)$.
2. The vector field $\frac{d}{dx}$ on $(-1, 1)$ is incomplete. Therefore, composing it with a diffeomorphism $(-1, 1) \simeq \mathbb{R}$ results in an incomplete vector field on $\mathbb{R}$.

Proposition 2.2.7. *Suppose G is a Lie group and $X \in \mathfrak{g}$. Then X is complete.*

Proof. Suppose $\varphi : I \to G$ is an integral curve and $g \in G$. Then $t \mapsto \ell_g \circ \varphi$ is also an integral curve, because X is G-invariant. In particular, if the integral curve through one point is infinitely extendable, then *every* integral curve is infinitely extendable and X is complete.

The idea is simple: if the integral curve is incomplete, then it "runs out of steam" at some finite point. But since G is a Lie group and X is left-invariant, we can always translate it to keep it going a little longer.

Suppose (I, φ) is a maximal integral curve with $0 \in I$ and $\varphi(0) = e$. Suppose that I has an endpoint t_0; without loss of generality, suppose $t_0 > 0$. Choose $\epsilon > 0$

such that $|\epsilon| < t_0$ and choose $t_1 < t_0$ with $|t_1 - t_0| < \frac{\epsilon}{2}$. Then $t \mapsto \varphi(t - t_1)$ is well-defined for $|t_1 - t| < \epsilon$ and takes the value e at $t = t_1$. Therefore, $t \mapsto \varphi(t_1)\varphi(t - t_1)$ is an integral curve agreeing with φ at $t = t_1$ and defined on an interval containing even t_0. $\square$

Suppose X is a complete vector field on M. Define $\Phi : \mathbb{R} \times M \to M$ by

(i) $\Phi(0, p) = p$ for all $p \in M$.

(ii) $t \mapsto \Phi(t, p)$ is an integral curve for all $p \in M$.

Theorem 2.2.8 (Global C^∞ dependence on parameters)**.** $\Phi : \mathbb{R} \times M \to M$ *is a smooth map.*

Proof. By Theorem 2.2.4, Φ is smooth on an open neighborhood of $\{0\} \times M$ (as a function of both arguments), i.e. Φ is smooth near $(0, p_0)$.

The map $t \mapsto \Phi(t, \Phi(s, p))$ is an integral curve in t with value $\Phi(s, p)$ at $t = 0$, as is $t \mapsto \Phi(s + t, p)$. By uniqueness of integral curves, we know that

$$\Phi(t, \Phi(s, p)) = \Phi(s + t, p) \text{ for all } p \in M \text{ and all } s, t \in \mathbb{R}.$$

For $p \in M$, let

$$I_p = \{t \in \mathbb{R} \mid \Phi \text{ is smooth near } (t, p)\}.$$

Then I_p is open and $0 \in I_p$. To prove the theorem, we just need to show that $I_p = \mathbb{R}$ for all p. Suppose this is not the case for some $p_0 \in M$; then the connected component of 0 in I_p has some finite endpoint t_0. Without loss of generality, assume that $t_0 > 0$. Then we have by assumption, Φ is C^∞ (in both arguments) on a neighborhood of $[0, t_0) \times \{p_0\}$.

Pick η satisfying $0 < \eta < 1$. Then

$$\Phi(t_0, p) = \Phi(\eta t_0, \Phi((1 - \eta)t_0, p)).$$

If η is small enough, we can find an open neighborhood U of $\Phi(t_0, p_0)$ such that $p \mapsto \Phi(\eta t_0, p)$ is C^∞ for $p \in U$. Making η smaller if necessary, we can arrange that $\Phi((1 - \eta)t_0, p_0) \in U$ because Φ is continuous in the time variable. Moreover, for p near p_0, $p \mapsto \Phi((1-\eta)t_0, p)$ is C^∞ by the definition of t_0. Therefore, the composition

$$p \mapsto \Phi(\eta t_0, \Phi((1 - \eta)t_0, p)) = \Phi(t_0, p)$$

is C^∞ in p near p_0.

We now show that this implies that $(t, p) \mapsto \Phi(t, p)$ is smooth in a neighborhood of (t_0, p_0), which furnishes the desired contradiction. To see this, observe that

$$\Phi(t, p) = \Phi(t - t_0, \Phi(t_0, p)).$$

By hypothesis, $(t, q) \mapsto \Phi(t - t_0, q)$ is smooth near $t = t_0$ and $q = \Phi(p_0, t_0)$. Since $p \mapsto \Phi(t_0, p)$ is smooth, $\Phi(t, p)$ is the composition of smooth functions

$$(t, p) \mapsto (t, \Phi(t_0, p)) \mapsto \Phi(t - t_0, \Phi(t_0, p)),$$

hence is smooth.

$\square$

2.3 The exponential map

Let G be a Lie group with Lie algebra $\mathfrak{g}$. We showed last time that any $X \in \mathfrak{g}$ is complete. Fixing $X \in \mathfrak{g}$, we get by the previous proposition a smooth map

$$\Phi_X : \mathbb{R} \times G \to G$$

satisfying $\Phi_X(0, g) = g$ and $t \mapsto \Phi_X(t, g)$ is an integral curve for X. Therefore, we get a map

$$\begin{aligned} \Phi : \mathbb{R} \times G \times \mathfrak{g} &\to G \\ (t, g, X) &\mapsto \Phi_X(t, g). \end{aligned}$$

Proposition 2.3.1. *Φ is C^∞ as a function of all variables.*

Before giving the proof, we set up the general framework. Consider the vector field X on a product of manifolds $M \times N$. Say that X is *tangential* to the fibers of projection $\pi_2 : M \times N \to N$ if all the integral curves of X lie in the fibers. Equivalently, in local coordinates $(U; x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)$ such that $x_1, \ldots, x_k$ are local coordinates on M and $x_{k+1}, \ldots, x_n$ are local coordinates on N. The vector field X can be expressed locally as

$$X = \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i}.$$

Then X is tangential to the fibers if and only if $a_j = 0$ for $j > k$. In other words, for each $m \in M$ and $n \in N$

$$X_{(m,n)} \in T_m M \times \{0\} \subset T_m M \oplus T_n N = T_{(m,n)} M \times N.$$

We can view a vector field X of this type as a *family* of vector fields on M, parametrized by $n \in N$. Note that X is complete on $M \times N$ if and only if when considered as a family of vector fields on M parametrized by N, each member is complete as a vector field on M; this is clear from the fact that the integral curve to any point of $M \times N$ stays in a signle fiber, which is isomorphic to M.

Hence a family of complete vector fields with parameters in N determines a smooth map $\widetilde{\Phi} : \mathbb{R} \times M \times N \to M \times N$ by Theorem 2.2.8. Composing with the projection $M \times N \to M$ gives a smooth map $\Phi : \mathbb{R} \times M \times N \to M$.

Proof of Proposition 2.3.4. Let $M = G$ and $N = \mathfrak{g}$ in the above discussion. $\square$

Properties of Φ.

(i) $\Phi(0, g, X) = g$.

(ii) $t \mapsto \Phi(t, g, X)$ is an integral curve for X.

(iii) $\Phi(t, g, X) = g\Phi(t, e, X)$.

(iv) $\Phi(t, g, sX) = \Phi(st, g, X)$.

(v) $\Phi(s, e, X)\Phi(t, e, X) = \Phi(s + t, e, X)$.

Proof. (i) and (ii) are the definition of Φ. (iii), (iv), and (v) follow from the observation that both sides of the respective equalities represent integral curves for X with the same initial conditions (we are using, of course, the fact that X is left-invariant). $\square$

In particular, it follows from the above properties that

$$\Phi(t, g, X) = g\Phi(1, e, tX).$$

So Φ is *completely determined* by the map $\exp : \mathfrak{g} \to G$ defined by

$$\exp X = \Phi(1, e, X).$$

This is the famous *exponential map.*

Properties of the exponential map.

(i) $\exp : \mathfrak{g} \to G$ is smooth.

(ii) $t \mapsto g \exp(tX)$ is the integral curve to X with the initial point g at $t = 0$.

(iii) $\exp(sX) \exp(tX) = \exp((s + t)X)$. Equivalently,

$$\exp(X + Y) = \exp(X) \exp(Y) \text{ if } X, Y \text{ are linearly dependent.}$$

Proposition 2.3.2 (Universality of the exponential map.)**.** *Suppose $F : G \to H$ is a Lie group homomorphism. Then the following diagram commutes.*

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "F_*"] \arrow[d, "\exp"] & \mathfrak{h} \arrow[d, "\exp"] \\
G \arrow[r, "F"'] & H
\end{tikzcd}$$

Exercise 2.3.3. Prove this proposition.

Proposition 2.3.4. *The composition*

$$\mathfrak{g} \simeq T_0\mathfrak{g} \stackrel{exp_*}{\to} T_eG \simeq \mathfrak{g}$$

is the identity.

Exercise 2.3.5. Prove this proposition.

Proposition 2.3.6. *For $g \in G$ and $X \in \mathfrak{g}$, $f \in C^\infty(G)$,*

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

Exercise 2.3.7. Prove this proposition.

Corollary 2.3.8. *The map* $\exp : \mathfrak{g} \to G$ *is locally a diffeomorphism near* $0 \in \mathfrak{g}$.

Proof. By Proposition 2.3.4, the derivative of the exponential map is the identity at the origin. The result then follows directly from the inverse function theorem. □

Any manifold is the countable union of its connected components. Note that connectedness and path-connectedness are equivalent notions for a manifold, since a manifold is locally path-connected.

Let G be a Lie group and G^0 the connected component containing the identity of G; we call this the *identity component.* Note that

$$(G^0)^2 = \{g_1g_2 \mid g_i \in G^0\} \qquad \text{and} \qquad (G^0)^{-1} = \{g^{-1} \mid g \in G^0\}$$

are connected and contain e (since we can go from a product to another product by a product of connecting paths), hence equal to G^0. This tells us that G^0 is an open, closed subgroup.

Also for $g \in G$, gG^0g^{-1} is a connected subgroup of G containing e. Therefore, the identity component is a *normal* subgroup. Now suppose $U \subset G^0$ is a neighborhood of the identity. We can arrange that $U = U^{-1}$ by replacing U with $U \cap U^{-1}$ if necessary. Then

$$G_1 = \bigcup_{k=1}^{\infty} U^k$$

is open (since U^k is a union of translates of opens, hence open) and a subgroup of G^0. Therefore G_1 is closed, because its complement is a union of its cosets. Since G^0 was connected, it must be the case that $G_1 = G^0$.

We thus observe that if U is a neighborhood of the identity in a Lie group G, then U generates G^0 in the algebraic sense. We have proved:

Theorem 2.3.9. *Let G be a Lie group with Lie algebra $\mathfrak{g}$.*

(i) $\exp(\mathfrak{g})$ *generates the identity component of* G *in the algebraic sense.*

(ii) Let $F : G \to H$ *be a Lie group homomorphism. Then the restriction of* F *to* G^0 *is completely determined by* $F_* : \mathfrak{g} \to \mathfrak{h}$.

Proof. (i) follows from the fact that exp is a local diffeomorphism from $\mathfrak{g}$ to G, hence contains a neighborhood of $e \in G$. (ii) follows from the preceding observation and the universality of the exponential map, which tells us that knowledge of F_* completely determines F restricted to the image of the exponential map. □

The general theme is that the Lie algebra recognizes everything about the Lie group, except in two cases

1. it doesn't know about other connected components, and
2. it can't distinguish covering spaces.

2.4 The group law in terms of the Lie algebra

If the Lie algebra knows "everything" about G, then it should know about the group operation. This is the subject of the Campbell-Baker-Hausdorff formula, which is our next topic.

Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. We have seen that there exist neighborhoods V_0 containing 0 in $\mathfrak{g}$ and U_e containing e in G such that $\exp : V \simeq U_e$. Let $\log : U_e \to V_0$ denote the inverse.

By shrinking V_0 (and thus U_e), we can arrange that V_0 is $\star$-shaped about 0. Recall that this means that for any $p \in V_0$, the line between p and 0 lies in V_0. By further shrinking V_0 if necessary, we may also assume that $V_0 = -V_0$. By passing to a yet smaller neighborhood V, we assume that there exists a neighborhood U of $e \in G$ such that $\log : U \to \mathfrak{g}$ is smooth and well-defined, and $(\exp V)^2 \subset U$. To summarize, we have an open neighborhood $V \ni 0 \in \mathfrak{g}$ such that

(i) $V = -V$

(ii) V is $\star$-shaped about 0.

(iii) There is a neighborhood U of $e \in G$ such that log is well defined and smooth on U, and $(\exp V)^2 \subset U$.

That allows us to define a C^∞ map $M : V \times V \to \mathfrak{g}$ sending $M(X,Y) = \log(\exp X \exp Y)$. This M is the group multiplication transferred to the Lie algebra via the exponential map. Note the following properties of this map:

1. $M(X,0) = X$ and $M(0,Y) = Y$.
2. $M(X,-X) = 0$.

3. $M(-X, -Y) = -M(Y, X)$.

   This follows from the fact that $(\exp X \exp Y)^{-1} = (\exp Y)^{-1}(\exp X)^{-1}$.

4. $M(X, M(Y, Z)) = M(M(X, Y), Z)$ if X, Y, Z is sufficiently small (for everything to lie in V).

   This is associativity of the group law.

Proposition 2.4.1. *For $X, Y \in \mathfrak{g}$,*

$$\frac{\partial^2}{\partial s \partial t} M(sX, tY)|_{s=t=0} = \frac{1}{2}[X, Y].$$

Proof. For $g \in$ and $f \in C^\infty(G)$,

$$Xf(g) = \frac{\partial}{\partial s} f(g \exp(sX))|_{s=0}.$$

Applying this to Yf, we have $XYf(g) = \frac{\partial}{\partial s} Yf(g \exp(sX))|_{s=0}$. Now applying the same reasoning to Yf, we find that this is

$$\frac{\partial^2}{\partial s \partial t} f(g \exp(sX) \exp(tY))|_{s=t=0}.$$

Therefore,

$$\begin{aligned}
[X, Y]f(e) &= (XY - YX)f(e) \\
&= \frac{\partial^2}{\partial s \partial t} \left(f(\exp(sX)\exp(tY)) - f(\exp(tY)\exp(sX))\right)_{s=t=0} \\
&= \frac{\partial^2}{\partial s \partial t} \left(f \circ \exp(M(sX, tY)) - f \circ \exp(M(tY, sX))\right)|_{s=t=0} \\
&= \frac{\partial^2}{\partial s \partial t} \left(f \circ \exp(M(sX, tY)) - f \circ \exp(-M(-sX, -ty))\right)|_{s=t=0}.
\end{aligned}$$

By setting s and t to zero, we see that the Taylor expansion of $M(sX, tY)$ is

$$M(sX, tY) = sX + tY + stC + (\text{terms of order} \geq 3)$$

where $C = \frac{\partial^2}{\partial s \partial t} M(sX, tY)|_{s=t=0}$. We substitute the terms of order up to 2 into the expression above to obtain

$$= \frac{\partial^2}{\partial s \partial t} \left[f \circ \exp(sX + tY + stC) + f \circ \exp(sX + tY - stC)\right]_{s=t=0}.$$

Replace $f \circ \exp$ by φ, so φ is a C^∞ function defined on V. Since $\exp_*$ is the identity map, the left hand side of the original equation is $[X, Y]\varphi(0)$. In other words,

$$\begin{aligned}
[X, Y]\varphi(0) &= \frac{\partial^2}{\partial s \partial t} \left[\varphi(sX + tY + stC) - \varphi(sX + tY - stC)\right]_{s=t=0} \\
&= 2\frac{\partial^2}{\partial s \partial t} \varphi(sX + tY + stC)|_{s=t=0} \\
&= 2\frac{\partial^2}{\partial s \partial t} \varphi(M(sX, tY))|_{s=t=0}.
\end{aligned}$$

Therefore, $M(X,Y) = X + Y + \frac{1}{2}[X,Y] + \dots$. $\square$

Corollary 2.4.2. *Multiplication on G determines the Lie bracket $[\cdot,\cdot]$ on $\mathfrak{g}$.*

Chapter 3

Lie algebras

3.1 The Lie algebra $\mathfrak{gl}(V)$

Let V be a finite dimensional real or complex vector space, $\mathrm{GL}(V)$ the group of invertible endomorphisms of V. This is open in $\mathrm{End}(V)$. With the induced structure, $\mathrm{GL}(V)$ becomes a Lie group of V, and a *complex* Lie group if V is a vector space over $\mathbb{C}$.

Definition 3.1.1. A **complex Lie group** is a group G endowed with the structure of a complex manifold such that multiplication and inversion are holomorphic maps.

In this case, $\mathfrak{g}$ is the Lie algebra of left-invariant *holomorphic* vector fields. This is the "holomorphic part" of the complexification of the Lie algebra of G considered as a *real* Lie group, essentially because of the Cauchy-Riemann equations. We will elaborate on this discussion later; we just want to indicate how the proceeding discussion applies to both real and complex vector spaces. Note also that in the complex case, $\exp, \log$, and M are holomorphic maps.

Turn V into a normed vector space, i.e. pick a norm $||\cdot|| : V \to k$ satisfying

(i) $||av|| = |a|||v||$ for scalars a and vectors v.

(ii) $||u+v|| \leq ||u|| + ||v||$.

(iii) $||v|| \geq 0$ and equality holds only if $v = 0$.

For $T \in \mathrm{End}(V)$, define the "operator norm"

$$||T|| = \sup_{||v|| \leq 1} ||Tv||.$$

Note that $||S \circ T|| \leq ||S||\ ||T||$.

Define $\mathrm{Exp} : \mathrm{End}(V) \to \mathrm{End}(V)$ by

$$\mathrm{Exp}(T) = \sum_{k=0}^{\infty} \frac{1}{k!} T^k.$$

This converges uniformly on bounded subsets of $\mathrm{End}(V)$. For formal reasons,

$$\mathrm{Exp}(S+T) = \mathrm{Exp}(S)\ \mathrm{Exp}(T)$$

provided that S, T commute, because absolutely convergent series can be re-ordered. In particular, $\mathrm{Exp}(-T) = \mathrm{Exp}(T)^{-1}$, so Exp lands in $\mathrm{GL}(V)$. By the usual arguments, convergent power series can be differentiated term-by-term. So

$$\frac{\partial}{\partial t}\mathrm{Exp}(tT) = \mathrm{Exp}(tT)\cdot T.$$

Therefore, $t \mapsto \mathrm{Exp}(tT)$ is tangential to the left-invariant (holomorphic) vector field whose value at e is T. Since this property also characterizes the exponential map, we must have $\mathrm{Exp} = \exp$. This shows:

Proposition 3.1.2. *The exponential map for $GL(V)$ (in both the real and complex cases) is given by*

$$\exp T = \sum_{k=0}^{\infty} \frac{1}{k!} T^k.$$

What about the logarithm? We have

$$\frac{d}{dw}\log(1-w) = -\frac{1}{1-w} = -\sum_{k=0}^{\infty} w^k.$$

Therefore,

$$\log z = -\sum_{k=1}^{\infty} \frac{(1-z)^k}{k}.$$

The map $\log : \mathrm{GL}(V) \to \mathrm{End}(V)$ which is defined near $1_V \in \mathrm{GL}(V)$ is given by

$$\log g = -\sum_{k=1}^{\infty} \frac{1}{k}(1-g)^k$$

for formal reasons concerning power series.

We now compute the Lie bracket for $\mathfrak{gl}(V) := \mathrm{Lie}(\mathrm{GL}(V))$. If

$$\begin{aligned} g = \exp S \exp T &= \left(1 + S + \frac{1}{2}S^2 + \dots\right)\left(1 + T + \frac{1}{2}T^2 + \dots\right) \\ &= \left(1 + S + T + \frac{1}{2}S^2 + \frac{1}{2}T^2 + ST + \dots\right) \end{aligned}$$

where we omit all terms of order ≥ 3, because they are not relevant to the calculation. Then

$$\begin{aligned} M(S,T) = \log g &= g - 1 - \frac{1}{2}(1-g)^2 + \dots \\ &= S + T + \frac{1}{2}S^2 + \frac{1}{2}T^2 + ST - \frac{1}{2}\left(S^2 + T^2 - ST - TS\right) \\ &= S + T + \frac{1}{2}[S,T] + \dots \end{aligned}$$

This discussion proves the following proposition.

Proposition 3.1.3. *The Lie bracket on End(V), viewed as the Lie algebra of GL(V), is*

$$[S, T] = ST - TS.$$

How should we view this? If we have an arbitrary Lie group, the Lie algebra can't really detect things like covering spaces and different connected components. However, modulo these considerations, all Lie groups are captured as subgroups of $\mathrm{GL}(V)$. In other words, every connected Lie group is a covering of some Lie group sitting inside a subgroup of $\mathrm{GL}(V)$.

3.2 The adjoint representation

Suppose G is a Lie group with Lie algebra $\mathfrak{g}$. For $g, h \in G$, we have

$$\ell_g \circ r_g = ghg^{-1},$$

i.e. conjugation by g. So we get a group homomorphism $G \to \mathrm{Aut}(G)$ sending $g \mapsto \ell_g \circ r_g$.

Define $\mathrm{Ad} : G \to \mathrm{End}(\mathfrak{g})$ by $\mathrm{Ad}\,(g) = (\ell_g \circ r_g)_*$. Since $(\ell_g \circ r_g) \circ (\ell_{g^{-1}} \circ r_{g^{-1}})$ is the identity, we have $\mathrm{Ad}\,(g^{-1}) = (\mathrm{Ad}\,)^{-1}$. So Ad lands in $\mathrm{Aut}(\mathfrak{g}) \subset \mathrm{GL}(\mathfrak{g})$. We already know that $\mathrm{GL}(\mathfrak{g})$ is a Lie group with Lie algebra $\mathrm{End}(\mathfrak{g})$.

Definition 3.2.1. We set $\mathrm{ad} := \mathrm{Ad}_* \colon \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$. Note that this is a Lie algebra homomorphism by the Jacobi identity.

We cal Ad the *adjoint representation* and ad the *infinitesimal adjoint representation*.

Proposition 3.2.2. $(\mathrm{ad}\ X)(Y) = [X, Y]$.

Proof. By definition of the exponential map, $Y \in \mathfrak{g} \simeq T_eG$ is the tangent vector at $t = 0$ to the curve $t \mapsto \exp(tY)$. Therefore, $\mathrm{Ad}g(Y)$ is the tangent vector at $t = 0$ to $t \mapsto g \exp(tY)g^{-1}$. Hence $\mathrm{ad}X(Y)$ is the tangent vector at $s = 0$ to the curve $\mathrm{Ad}(\exp(sX))Y$.

Therefore, for any smooth f near $e \in G$, $\mathrm{ad}X(Y) \in \mathfrak{g} \simeq T_eG$, and

$$\begin{aligned}
(\mathrm{ad}X(Y))f(e) &= \frac{\partial^2}{\partial s\partial t} f(\exp(sX)\exp(tY)\exp(-sX))|_{s=t=0} \\
&= \frac{\partial^2}{\partial s\partial t} f\circ\exp(M(sX, M(tY, -sX)))|_{s=t=0} \\
&= \frac{\partial^2}{\partial s\partial t} f\circ\exp(sX + M(tY,-sX) + \frac{1}{2}[sX, M(tY,-sX)] + \ldots)|_{s=t=0} \\
&= \frac{\partial^2}{\partial s\partial t} f\circ\exp\big(sX + tY - sX + \frac{1}{2}[tY, -sX] \\
&\qquad\qquad + \frac{1}{2}[sX, tY - sX + \ldots]\big)|_{s=t=0} \\
&= \frac{\partial^2}{\partial s\partial t} f\circ\exp\left(tY + st[X,Y] + \ldots\right)_{s=t=0}.
\end{aligned}$$

Letting $\varphi = f\circ\exp$, we have φ is smooth near 0 in $\mathfrak{g}$ since $\exp_* = 1$ (and the inverse function theorem). In terms of φ,

$$\begin{aligned}
((\mathrm{ad}X(Y))\varphi)(0) &= \frac{\partial^2}{\partial s\partial t}\varphi(tY + st[X,Y] + \ldots)|_{s=t=0} \\
&= \varphi([X,Y]) = ([X,Y]\varphi)(0).
\end{aligned}$$

$\square$

Corollary 3.2.3. ad X *for* $X \in \mathfrak{g}$ *is a derivation of the Lie algebra* $\mathfrak{g}$*, i.e.*

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

With $\mathrm{ad}X = [X,\cdot]$, this condition is equivalent to the Jacobian identity.

A derivation of the Lie algebra should be thought of as an "infinitesimal automorphism," and ad X for some $X \in \mathfrak{g}$ should be thought of as an "infinitesimal inner automorphism."

3.3 Poincaré's formula

Suppose $X \in \mathfrak{g}$; then $\mathfrak{g} \simeq T_X\mathfrak{g}$. The exponential map induces

$$\mathfrak{g} \simeq T_X\mathfrak{g} \xrightarrow{\exp_*|_X} T_{\exp X}G.$$

On the other hand,

$$\mathfrak{g} \simeq T_eG \xrightarrow{(\ell_{\exp X})_*} T_{\exp X}G.$$

Since $\ell_{\exp(-X)} = (\ell_{\exp X})^{-1}$, the map $(\ell_{\exp X})_*$ is invertible and threading across the diagram gives a linear map $\mathfrak{g} \to \mathfrak{g}$

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "\cong"] \arrow[d, dotted] & T_X\mathfrak{g} \arrow[r, "\exp_*|_X"] & T_{\exp X}G \arrow[d, equal] \\
\mathfrak{g} & T_eG \arrow[l, "\cong"'] & T_{\exp X}G \arrow[l, "(\ell_{\exp X})_*^{-1}"']
\end{tikzcd}$$

One might ask, what is this map?

Theorem 3.3.1 (Poincare's formula for the differential of the exponential map)**.**

$$\exp_*|_X = (\ell_{\exp X})_* \circ \frac{1 - e^{-\,\mathrm{ad}\, X}}{\mathrm{ad}\ X}.$$

Note that

$$\frac{1 - e^{-z}}{z} = 1 - \frac{z}{2} + \dots,$$

an entire holomorphic function. So we can think of the expression $\frac{1-e^{-\mathrm{ad}X}}{\mathrm{ad}X}$ as a convergent power series in the operator $\mathrm{ad}X$.

Remark 3.3.2*.* This formula is quite useful! By contrast, the exact coefficients of the Campbell-Baker-Hausdorff formula rarely comes up in practice.

Proof. By the preceding discusson, there exists $D(X) \in \mathrm{End}(\mathfrak{g})$ such that

$$\exp_*|_X = (\ell_{\exp X})_* \circ D(X). \tag{3.3.1}$$

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "\cong"] \arrow[d, dotted, "D(X)"] & T_X\mathfrak{g} \arrow[r, "\exp_*|_X"] & T_{\exp X}G \arrow[d, equal] \\
\mathfrak{g} & T_eG \arrow[l, "\cong"'] & T_{\exp X}G \arrow[l, "(\ell_{\exp X})_*^{-1}"']
\end{tikzcd}$$

The map $D : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ is a smooth function of X.

Exercise 3.3.3*.* Prove this.

Define $F(t) = tD(tX)$, which is then a smooth function of t with values in $\mathrm{End}(\mathfrak{g})$. We claim that

$$F'(t) = 1_{\mathfrak{g}} - \mathrm{ad}\,(X) \circ F(t).$$

Supposing this for the moment, define

$$\widetilde{F}(t) = \frac{1 - e^{-t\,\mathrm{ad}\ X}}{\mathrm{ad}\ X} := (1 - e^{-t\,\mathrm{ad}\ X})(\mathrm{ad}\ X)^{-1}.$$

Then

$$\widetilde{F}'(t) = e^{-t\mathrm{ad}X} = 1 - \widetilde{F}(t) \circ \mathrm{ad}(X) = 1 - \mathrm{ad}X \circ \widetilde{F}(t).$$

Also, $F(0) = \widetilde{F}(0) = 0$. Therefore, $F = \widetilde{F}$, which is the desired conclusion.

Now we prove the claim. First we give a more concrete interpretation of (3.3.1). Applying it to Y, we see that for f a smooth function near $\exp X$ in G, we have

$$\frac{\partial}{\partial t} f(\exp(X)\exp(tY))|_{t=0} = (D(X)Y)f(\exp X). \tag{3.3.2}$$

To elaborate, let $Y \in \mathfrak{g}$. We are concerned with its image in $T_{\exp X}G$, which is a functional on $C^\infty(M)$; what is the value of its image at $f \in C^\infty(M)$? Along the map described by the right hand side of (3.3.1), it is just $((D(X)Y)f)(\exp X)$, absorbing the $(\ell_{\exp X})_*$ into the formula. Now consider the map described by the left hand side of (3.3.1). We know from the homework that

$$\frac{\partial}{\partial t} f(\exp(X + tY))|_{t=0} = Yf(\exp X),$$

the reasoning being that $t \mapsto \exp X \exp(tY)$ is an integral curve to Y at $\exp X$.

Since the ambiguity between left-invariant vector fields and elements of the Lie algebra will be confusing, for this proof only we use the notation $r(Y)$ for the left invariant vector field generated by $Y \in \mathfrak{g}$. In these terms, (3.3.2) says

$$(r(Y)f)(g) = \frac{d}{du} f(g \exp uY)|_{u=0}.$$

We used the notation $r(Y)$ because right translation commutes with left translation, so "infinitesimal right translation is a left-invariant vector field."

Recall that we wanted to prove that for $F(t) = tD(tX)$,

$$F'(t) = 1 - \mathrm{ad}\ X \circ F(t).$$

So by the preceding equations, this has something to do with

$$\frac{\partial^2}{\partial s \partial t} f(\exp(t(X + sY)))|_{s=0}$$

Let's calculate it in two different ways.

1. By a homework problem,

$$\frac{\partial}{\partial t} f(\exp(t(X + sY))) = (r(X + sY)f)(\exp(t(X + sY))).$$

   Now (3.3.2) tells us exactly how to differentiate this expression with respect to s. Doing so, we find that

$$\frac{\partial}{\partial s}(*)|_{s=0} = r(Y)f(\exp tX) + r(D(tX)(tY))(r(X)f)(\exp(tX)).$$

2. Reasoning as above, we first compute

$$\frac{\partial}{\partial s} f(\exp(t(X+sY)))|_{s=0} = r(tD(tX)Y)f(\exp(tX)).$$

Then differentiating with respect to t, we find that

$$\frac{\partial}{\partial s}(*)|_{s=0} = r(\frac{d}{dt}(tD(tX))Y)f(\exp tX) + r(X)r(tD(tX)Y)f(\exp tX).$$

Comparing the two expressions, we conclude that

$$r(Y)f(\exp tX) = r(\frac{d}{dt}(tD(tX)Y))f(\exp tX) + r([X, tD(tX)Y])f(\exp tX).$$

Since this holds for all f, we conclude that

$$Y = \frac{d}{dt}(tD(tX))Y + [X, tD(tX)Y].$$

Hence

$$\frac{d}{dt}(tD(tX)) = 1_{\mathfrak{g}} - \mathrm{ad}X \circ tD(tX).$$

□

That was a struggle...

- *Wilfried Schmid*

We re-iterate Poincaré's formula for emphasis:

$$\boxed{\exp_*|_X = (\ell_{\exp X})_* \circ \frac{1-e^{-\mathrm{ad}X}}{\mathrm{ad}X}.}$$

Now let's manipulate this identity.

$$\begin{aligned}
\exp_*|_X &= (r_{\exp(-X)})_* \circ (r_{\exp(X)})_* \circ (\ell_{\exp(X)})_* \circ \frac{1-e^{-\mathrm{ad}X}}{\mathrm{ad}X} \\
&= (r_{\exp(-X)})_* \circ \mathrm{Ad}\exp X \circ \frac{1-e^{-\mathrm{ad}X}}{\mathrm{ad}X} \\
&= (r_{\exp(-X)})_* \circ \exp(\text{ad X}) \circ \frac{1-e^{-\mathrm{ad}X}}{\mathrm{ad}X} \quad \text{(universality of exp)} \\
&= (r_{\exp(-X)})_* \circ \frac{e^{\mathrm{ad}X}-1}{\mathrm{ad}X}.
\end{aligned}$$

We have shown the right-translation analogue:

Corollary 3.3.4. *We have*

$$\exp_*|_X = (r_{\exp(-X)})_* \circ \frac{e^{\mathrm{ad}\ X}-1}{\mathrm{ad}\ X}.$$

3.4 The Campbell-Baker-Hausdorff formula

Now put a linear norm on $\mathfrak{g}$. Then there exists $C > 0$ such that

$$||[X,Y]|| \leq C||X||\ ||Y||$$

Choose an open neighborhood V of 0 in $\mathfrak{g}$ such that

(i) $V = -V$

(ii) V is $\star$-shaped about 0.

(iii) $X, Y \in V \implies ||M(X,Y)|| \leq \frac{2\pi}{C}$. (This implies that $||\mathrm{ad}M(X,Y)|| < 2\pi$.)

With this choice, for all $X, Y \in V$, $h_j(\mathrm{ad}M(X,Y))$ is well-defined.

Define

$$h_1(z) = \frac{z}{e^z - 1} = \left(\frac{e^z - 1}{z}\right)^{-1}$$
$$h_2(z) = \frac{z}{1 - e^{-z}} = \left(\frac{1 - e^{-z}}{z}\right)^{-1}.$$

As power series in z, h_1 and h_2 have radius of convergence 2π. Therefore, for $||X|| < \frac{2\pi}{C}$, both $h_1(\mathrm{ad}X)$ and $h_2(\mathrm{ad}(X))$ are well-defined. So with this choice of V, $h_j(\mathrm{ad}M(X,Y))$ is well-defined for all $X, Y \in V$.

Since V is open, for X and Y in V there exists $\epsilon > 0$ such that

$$|t| < 1 + \epsilon \implies tX, tY \in V.$$

Define $F(t) = M(tX, tY)$. Then the map $F : (-1-\epsilon, 1+\epsilon) \to \mathfrak{g}$ is smooth.

Lemma 3.4.1. *For $X, Y \in V$ and $|s|, |t| < 1 + \epsilon$ we have the following identities:*

$$\frac{\partial}{\partial s} M(sX, tY) = h_1(\mathrm{ad}\ M(sX, tY))X$$
$$\frac{\partial}{\partial t} M(sX, tY) = h_2(\mathrm{ad}\ M(sX, tY))Y$$

Proof. Suppose f is a C^∞ function defined near $\exp sX \exp tY$. Consider

$$\frac{\partial}{\partial t} f(\exp sX \exp tY) = \frac{\partial}{\partial u} f(\exp sX \exp tY \exp uY)|_{u=0}.$$

Thinking of Y as a left-invariant vector field, this is equal to

$$Yf(\exp sX \exp tY).$$

Just by the chain rule, we have

$$\frac{\partial}{\partial t} f(\exp sX \exp tY) = \underbrace{\exp_* |_{M(sX,tY)} \underbrace{\frac{\partial}{\partial t} M(sX,tY)}_{\in \mathfrak{g} \simeq T_e G}}_{\in T_{\exp M} G} f.$$

Now applying the formula for the differential of the exponential map, we get

$$\Big((\ell_{\exp sX \exp tY})_* \circ \underbrace{\frac{1-e^{-\mathrm{ad} M(sX,tY)}}{\mathrm{ad} M(sX,tY)} \frac{\partial}{\partial t} M(sX,tY)}_{\in \mathfrak{g} \simeq T_e G} \Big) f.$$

Considering $\frac{1-e^{-\mathrm{ad} M(sX,tY)}}{\mathrm{ad} M(sX,tY)} \frac{\partial}{\partial t} M(sX,tY)$ as a left-invariant vector field, this is the same as

$$\left(\frac{1-e^{-\mathrm{ad} M(sX,tY)}}{\mathrm{ad} M(sX,tY)} \frac{\partial}{\partial t} M(sX,tY) \right) f(\exp sX \exp tY).$$

That tells us that

$$Y = \frac{1-e^{-\mathrm{ad} M(sX,tY)}}{\mathrm{ad} M(sX,tY)} \frac{\partial}{\partial t} M(sX,tY) = h_2(\mathrm{ad} M(sX,tY))^{-1} \frac{\partial}{\partial t} M(sX,tY).$$

as left-invariant vector fields. We conclude that

$$\frac{\partial}{\partial t} M(sX,tY) = h_2(\mathrm{ad} M(sX,tY))Y.$$

The other formula is obtained the same way. Actually, we have to identify $\mathfrak{g}$ in that case with right-invariant vector fields, or consider the "anti-automorphism" $g \mapsto g^{-1}$. In any case, there is no real difference between left and right. □

Corollary 3.4.2. *With the notation above,*

$$F'(t) = h_1(\mathrm{ad}\ F(t))X + h_2(\mathrm{ad}\ F(t))Y.$$

Proof. $F'(t) = \frac{\partial}{\partial t} M(tX,tY)$. Apply Lemma 3.4.1, using the chain rule. □

So far we only know that F is a C^∞ function, but we can write down its Taylor series

$$F(t) \text{``}=\text{''} \sum_{k=1}^{\infty} t^k M_k(X,Y).$$

Differentiating and applying the corollary,

$$\sum_{k=1}^{\infty} k t^{k-1} M_k(X,Y) = h_1\left(\sum_{\ell=1}^{\infty} t^\ell \mathrm{ad} M_\ell(X,Y) \right) X + h_2\left(\sum_{\ell=1}^{\infty} t^\ell \mathrm{ad} M_\ell(X,Y) \right) Y. \tag{3.4.1}$$

Now,

$$h_1(z) = \frac{z}{e^z - 1} = \frac{z}{z + \frac{z^2}{2} + \frac{z^3}{6} + \dots}$$
$$= \frac{1}{1 + \frac{z}{2} + \dots} = 1 - \frac{z}{2} + \dots,$$

and similarly
$$h_2(z) = \frac{z}{1 - e^{-z}} = \frac{z}{z - \frac{z^2}{2} + \dots} = 1 + \frac{z}{2} + \dots.$$

So $h_1(0) = h_2(0) = 1$ and $h_1'(0) = -\frac{1}{2}$, $h_2'(0) = \frac{1}{2}$. What is the first order term? On the right hand side of (3.4.1), we get

$$M(tX, tY) = t(X + Y) + \dots$$

Now we want to identify coefficients of t^k. First considering only the linear term of h_1, we get

$$(k+1)M_{k+1}(X, Y) = \frac{1}{2}[X, M_k(X, Y)] - \frac{1}{2}[Y, M_k(X, Y)] + \dots$$

Since the series inside the arguments of the h_j have no constant term, the other contributions to t^k come from linear combinations of terms

$$\mathrm{ad} M_{j_1}(X, Y) \mathrm{ad} M_{j_2}(X, Y) \dots \mathrm{ad} M_{j_n}(X, Y)$$

with $1 \leq j_1, j_2, \dots, j_n < k$ such that $j_1 + \dots + j_n = k$.

Definition 3.4.3. A *Lie monomial of degree* k in variables X, Y is a well-formed expression in the terms of X, Y, and bracket.

A *Lie polynomial homogeneous of degree* k in X and Y is a linear combination of linear monomials in X, Y of degree k.

Example 3.4.4. The expression $[X, [Y, X]]$ is a Lie monomial of degree 3. The expression $[[X, Y], [X, [X, Y]]]$ is a Lie monomial of degree 5.

Theorem 3.4.5 (Campbell-Baker-Hausdorff)**.** *There exist universal Lie polynomials $M_k(X, Y)$, homogeneous of degree k, with the following property: if G is a Lie group with Lie algebra $\mathfrak{g}$, then there exists a constant $R(C)$ depending on C but not otherwise on $G, \mathfrak{g}$ such that for $X, Y \in \mathfrak{g}$ satisfying $||X||, ||Y|| < R(C)$, the series*

$$M(X, Y) = \sum_{k=1}^{\infty} M_k(X, Y)$$

converges absolutely and uniformly on compact subsets of $\{||X||, ||Y|| < R(C)\}$ to the function $M(X, Y)$ satisfying

$$\exp M(X, Y) = \exp X \exp Y.$$

Remark 3.4.6. Before embarking on the proof, we give some historical perspective.

1. The first proof (about 1885) was due to Friedrich Schur (not the usual Schur). This was the first proof standing up to modern levels of rigor.
2. Campbell treated the problem purely formally (not worrying about convergence).
3. The second proof was given in around 1890 by Poincaré.
4. Baker came later, around 1902.
5. Hausdorff (1906) gave the modern formulation, in terms of Lie polynomials, with arguments essentially the same as ours.

Proof of Theorem 3.4.5. At this point, the only issue left is that of convergence. The differential equation

$$F'(t) = h_1(\mathrm{ad}F(t))X + h_2(\mathrm{ad}F(t))Y. \tag{3.4.2}$$

is a "holomorphic ordinary differential equation," since we know that $h_1(\mathrm{ad}F(t))$ and $h_2(\mathrm{ad}F(t))$ are holomorphic on the relevant region. There is a general theorem about the existence of solutions in such a situation, precisely by looking at and bounding coefficients of power series.

Set $\mathfrak{g}_{\mathbb{C}} := \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{g}$, a complex Lie algebra, and regard the ODE (3.4.2) as taking values in $\mathbb{C}$ with t as a complex parameter. Then (3.4.2) is a holomorphic ODE. By generalities on ODE, a solution certainly exists in a neighborhood of 0. Moreover, the point of going to the complex analytic context is that the radius of convergence of a complex-analytic function at a point is always the distance to the nearest singularity. So there exists $r > 0$, potentially depending on X and Y, such that the series

$$F(t) := \sum t^k M_k(X, Y)$$

has radius of convergence r. We need a nice estimate on r in terms of X and Y. If $r = \infty$, we are done. Otherwise, suppose that $r < \infty$. Recall that h_1, h_2 have Taylor series with radius of convergence 2π, so by the existence and uniqueness of solutions of ODE,

$$\sup_{|t|<r} ||\,\mathrm{ad}\ F(t)|| \geq 2\pi$$

since the differential equation (3.4.2) is sensible as long as $||\,\mathrm{ad}\ F(t)|| < 2\pi$. By our choice of C,

$$\sup_{|t|<r} ||F(t)|| \geq \frac{2\pi}{C}.$$

So there exists r_1 with $0 < r_1 < r$ such that

$$\sup_{|t|\leq r_1} ||F(t)|| = \frac{\pi}{C}$$

since the sup of $||F(t)||$ on disks depends continuously on the radius of the disk, and must assume any value less than $\frac{2\pi}{C}$. Because h_1, h_2 have Taylor series with radius of convergence 2π,

$$D = \max\{\sum |a_k|\pi^k, \sum |b_k|\pi^k\} < \infty$$

where $h_1(t) = \sum a_k t^k$ and $h_2 = \sum b_k t^k$. Then for $|t| < r_1$,

$$||\,\text{ad } F(t)|| \leq \pi \implies ||F'(t)|| \leq D(||X|| + ||Y||).$$

So

$$||F(t)|| \leq \int_0^{|t|} ||F'\left(r\frac{t}{|t|}\right)||\,dr \leq Dr_1(||X|| + ||Y||) \text{ for } |t| \leq r_1.$$

Hence

$$\sup_{|t|\leq r_1} ||F(t)|| = \frac{\pi}{C} \leq Dr_1(||X|| + ||Y||).$$

So $r > r_1 \geq \frac{\pi}{CD}\frac{1}{||X||+||Y||}$. For $||X||$ and $||Y||$ sufficiently small (depending only on C and D), we have $r > 1$, and thus

$$M(X,Y) = F(1) = \sum M_k(X,Y).$$

□

Chapter 4

Geometry of Lie groups

4.1 Vector bundles

Let M be a C^∞ manifold of dimension n. Everything we say will apply equally well to the complex analytic context, replacing C^∞ by "holomorphic."

Definition 4.1.1. A rank k *vector bundle* over M with real (resp. complex) fibers consists of the following data:

- A manifold E.
- A surjective smooth map $p : E \to M$.
- The structure of a k-dim real (resp. complex) vector space on $E_x := p^{-1}(x)$ for each $x \in M$, such that x has an open neighborhood U fitting into a commutative diagram

$$\begin{tikzcd} p^{-1}(U) \arrow[r, "\approx"] \arrow[d, "p"] & U \times \mathbb{R}^k \arrow[d, "\pi"] \\ U \arrow[r, "\approx", "\mathrm{id}"'] & U \end{tikzcd}$$

  which preserves the vector space structure (resp. replacing $\mathbb{R}^k$ by $\mathbb{C}^k$). This is called a *local trivialization* of the vector bundle.

Definition 4.1.2. A *section* of E over U is a C^∞ map $s : U \to p^{-1}U$ such that $p \circ s = 1_U$.

Sections can be added or multiplied by a scalar, so the sections over U comprise a vector space.

Definition 4.1.3. A *frame* for E over U is a k-tuple of sections $(s_1, \dots, s_k)$ which at each $x \in U$ provide a basis for E_x.

A local trivialization for the vector bundle over U is the same thing as a local frame over U. This is left as an exercise.

Example 4.1.4. Let TM be the tangent bundle of M and T^*M be the cotangent bundle. These are vector bundles: if $(U; x_1, \ldots, x_n)$ is a coordinate neighborhood, then $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\}$ is a local frame for TM and $\{dx_1, \ldots, dx_n\}$ is a local frame for T^*M. Sections of the tangent bundle are vector fields; sections of the cotangent bundle are 1-forms.

Definition 4.1.5. Let $p : E \to M$ be a vector bundle of rank k. A *rank ℓ sub-bundle of E* is a rank ℓ vector bundle $p_F : F \to M$ together with a commutative diagram

$$\begin{tikzcd} F \arrow[r, hook] \arrow[d, "p_F"] & E \arrow[d, "p"] \\ M \arrow[r, "\approx", "\mathrm{id}"'] & M \end{tikzcd}$$

such that the inclusion $F \to E$ is an injective linear map on each fiber. Then each $x \in M$ has an open neighborhood U with local frame $s_1, \ldots, s_k$ for E such that $s_1, \ldots, s_\ell$ is a local frame for F.

4.2 Frobenius' Theorem

Consider a rank k sub-bundle $E \subset TM$. For each $x \in M$, E_x is a k-dimensional subspace of T_xM.

Definition 4.2.1. An *integral submanifold* for E is a k-dimensional submanifold $S \subset M$ such that for each $x \in S$, $T_xS = E_x \subset T_xM$.

Definition 4.2.2. A tangent sub-bundle $E \subset TM$ is *involutive* if for any two vector fields X, Y on an open $U \subset M$ taking values in E, the commutator $[X, Y]$ also takes values in E.

Remark 4.2.3. The more common terminology for tangent sub-bundle is "distribution."

That is, a vector field is a section of the tangent bundle. It takes values in E if its value over $x \in M$ lies in $E_x \subset T_xM$. The bundle E is involutive if for any two such vector fields taking values in E, so does their Lie bracket.

Definition 4.2.4. The tangent sub-bundle E is *integrable* if each point in M has a coordinate neighborhood $(U; x_1, \ldots, x_n)$ such that $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k}$ constitute a local frame for E.

Note that in this situation,

1. The set $\{x_{k+1} = c_{k+1}, \ldots, x_n = c_n\}$ where the c_i are constant, are the integral submanifolds for E.

2. E is involutive, because
$$\left[\sum_{j=1}^k a_j \frac{\partial}{\partial x_j}, \sum_{j=1}^k b_j \frac{\partial}{\partial x_j}\right] = \sum_{i,j=1}^k a_j \frac{\partial b_i}{\partial x_j} \frac{\partial}{\partial x_i} - \sum_{i,j=1}^k b_j \frac{\partial a_i}{\partial x_j} \frac{\partial}{\partial x_i}.$$

Theorem 4.2.5 (Frobenius)**.** *A tangent sub-bundle is integrable if and only if it is involuble.*

As remarked above, it is obvious that integrable implies involuble; the substance of the theorem is the other direction.

Example 4.2.6. Suppose E is a rank one sub-bundle. Given any point $m \in M$, there exists an open neighborhood U and a vector field X on U which is a frame for E on U. The values of X must be non-zero at every point of U (by definition of local frame). Then any two vector fields Y_1, Y_2 defined near m and taking values in E can be expressed as $Y_i = f_i X$.

$$[Y_1, Y_2] = (f_1 X f_2 - f_2 X f_1) X$$

So in the rank one case, *involutivity is automatic.* According to the theorem, E should be integrable. Let's try to see this.

By the theory of ODEs, after shrinking U if necessary, we can construct a map

$$\Phi : (-\epsilon, \epsilon) \times U \to M$$

such that for all $u \in U$, $t \mapsto \Phi(t, u)$ is an integral curve of X and $\Phi(0, u) = u$. After shrinking U more if necessary, we can introduce coordinates $x_1, \ldots, x_n$ on U such that $x_j(m) = 0$. Then after making a linear change of coordinates, we can arrange that

$$X_m = \frac{\partial}{\partial x_1}\Big|_m.$$

Define a C^∞ map from a neighborhood V of $0 \in \mathbb{R}^n$ to M by $(y_1, \ldots, y_n) \mapsto \Phi(y_1; 0, y_2, \ldots, y_n)$. It is easy to see that the differential of this map at 0 is invertible at m. By the inverse function theorem, this map can be inverted locally near 0, which tells us that near 0, we can regard $y_1, \ldots, y_n$ as coordinates. But in terms of these coordinates, we have locally $\frac{\partial}{\partial y_1} = X$ since Φ defines an integral curve for X.

Let's recap at a conceptual level what we did here. By a linear change of coordinates, we arranged coordinates so that $X = \frac{\partial}{\partial x_1}$ at a single point. Then using the theory of integral curves, we showed that this must hold *locally.* We summarize our work in the theorem below.

Theorem 4.2.7 ("Local structure of a non-zero vector field")**.** *Let M be a manifold, $m \in M$, and X a vector field defined near m such that $X_m \neq 0$. Then there exist local coordinates $x_1, \ldots, x_n$ defined near m such that $X = \frac{\partial}{\partial x_1}$ locally.*

This proves the rank one case of Frobenius' theorem. The general case proceeds by induction on the rank, and the induction step is a clever reduction to the rank one case. (See e.g. Warner.)

Example 4.2.8. Now consider the rank 2 case on a neighborhood of 0 in $\mathbb{R}^3$. There exists a local frame which, up to linear change of coordinates, can be written as follows.

$$X_1 = \frac{\partial}{\partial x_1} + a(x)\frac{\partial}{\partial x_3} \dots$$
$$X_2 = \frac{\partial}{\partial x_2} + b(x)\frac{\partial}{\partial x_3} \dots$$

where $a(0) = b(0) = 0$. Then

$$[X_1, X_2] = \left(\frac{\partial b}{\partial x_1} - \frac{\partial a}{\partial x_2}\right)\frac{\partial}{\partial x_3}.$$

The term $\frac{\partial b}{\partial x_1} - \frac{\partial a}{\partial x_2}$ is the obstruction to involutivity, so generically a rankd 2 tangent sub-bundle of $T\mathbb{R}^3$ is not involutive, hence not integral.

Example 4.2.9. Let G be a Lie group with Lie algebra $\mathfrak{g}$ and $\mathfrak{h} \subset \mathfrak{g}$ a Lie subalgebra. Define a tangent sub-bundle $E_{\mathfrak{h}} \subset TG$ as follows:

$$(E_{\mathfrak{h}})_g = \ell_{g*}\mathfrak{h} \subset T_gG \subset \mathfrak{g} = T_eH.$$

Is this a tangent sub-bundle, i.e. is it spanned locally by a frame?

Yes; this is true even globally. Let $Y_1, \dots, Y_k$ be a basis of $\mathfrak{h} \subset \mathfrak{g}$, and view them as left-invariant vector fields; then $(Y_1, \dots, Y_k)$ is a global frame for $E_{\mathfrak{h}}$. Then $[Y_h, Y_j] \in \mathfrak{h} \implies E_{\mathfrak{h}}$ is involutive. (It suffices to check involutivity on the level of frames, since any vector fields can be written as a $C^\infty(M)$ linear combination of these.)

4.3 Foliations

Proposition 4.3.1. *Let M be a connected manifold, $E \subset TM$ an integrable rank k tangent subbundle. Given any point $m \in M$, there exists a unique maximal connected integral submanifold which contains m.*

Let $(U; x_1, \dots, x_n)$ be a coordinate neighborhood such that $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_k}$ is a frame for E on U. If $S \cap U \neq 0$, then by definition $S \cap U$ must be contained in a countable union (because S is second-countable) of submanifolds of the form

$$\{x \in U \mid x_{k+1} = c_{k+1}, \dots, x_n = c_n\}$$

By maximality, it must contain all of the slices that it meets. The issue to be established is second-countability.

We establish some (non-standard) terminology.

Definition 4.3.2. A coordinate neighborhood $(U; x_1, \dots, x_n)$ is *adapted* in E if

1. the image of U in $\mathbb{R}^n$ is $(-1,1)^n$
2. $\frac{\partial}{\partial x_1} \dots, \frac{\partial}{\partial x_k}$ is a frame for $E|_U$.
3. there exists a coordinate neighborhood $(\widetilde{U}; \widetilde{x}_1, \dots, \widetilde{x}_n)$ such that (a) $\overline{U}$ is compact and contained in $\widetilde{U}$, (b) $\widetilde{x}_j|_U = x_j$, and (c) $\frac{\partial}{\partial \widetilde{x}_1}, \dots, \frac{\partial}{\partial \widetilde{x}_k}$ is a frame for E over $\widetilde{U}$.

Note that M, being second-countable, can be covered by a countable number of adapted coordinate neighborhoods $U_1, U_2, \dots$.

Definition 4.3.3. A *slice* of an adapted coordinate neighborhood $(U; x_1, \dots, x_n)$ is a set of the form

$$\{(x_1, \dots, x_n) \mid x_{k+1} = c_{k+1}, \dots, x_n = c_n\}.$$

So "slice" means maximal connected integral submanifold of $E|_U$. The slices in U are parametrized by $(x_{k+1}, \dots, x_n) \in (-1,1)^{n-k}$.

We shall call a continuous (not necessarily C^1) path $\gamma : [0,1] \to M$ *tangential to* E if it satisfies the following condition:

(*) Let U be an adapted coordinate neighborhood of $\gamma(t_0) \in U$, $t_0 \in [0,1]$. Then for t near t_0, $\gamma(t)$ lies in the same slice as $\gamma(t_0)$.

This is equivalent to the weaker condition

(**) Given any $t_0 \in [0,1]$ there exists an adapted coordinate neighborhood U of $\gamma(t_0)$ such that for t near t_0, $\gamma(t)$ lies in the same slice as $\gamma(t_0)$

because locally, integral submanifolds are unique. There is just no way to jump from one slice to another.

Lemma 4.3.4. *Let U_α, U_β be adapted coordinate neighborhoods and S a slice of U_α. Then $S \cap U_\beta$ is contained in a finite union of slices in U_β.*

Proof. Suppose not, i.e. suppose that there exists a sequence $\{p_k\}$ in $S \cap U_\beta$ such that for $k \neq \ell$, p_k and p_ℓ lie in different slices of U_β. But $\overline{S}$ is compact, so the sequence $\{p_k\}$ converges to some $p_\infty \in \overline{S}$.

Recall that U_α has compact closure in $\widetilde{U}_\alpha$ and $\overline{S}$ is contained in a slice of $\widetilde{U}_\alpha$, say $\widetilde{S}$. Furthermore, $p_\infty \in U_\beta \subset \widetilde{U}_\beta$ so p_∞ lies in a slice of U_β, say $\widetilde{S}_\beta$.

But integral submanifolds for E are locally unique, in the sense that $\widetilde{S} \cap \widetilde{S}_\beta$ contains an open neighborhood (with respect to $\widetilde{S}$) of p_∞ that is also contained in $\widetilde{S}_\beta$. So the limiting p_k cannot all be in distinct slices. $\square$

To prove the existence of a maximal connected integral submanifold passing through a given point $m \in M$, choose $U_0, U_1, \ldots$ with $m \in U_0$. Let γ be a path tangential to E, with $\gamma(0) = m$. By compactness of $[0,1]$, there exists a partition $0 < t_0 < \ldots, < t_N = 1$ of $[0,1]$ such that for $0 < j \leq N$, $\gamma([t_i, t_{i+1}]) \subset U_{\ell_j}$, so in particular $\gamma(1) \in U_{\ell_N}$. By the lemma, there is only a finite number of possible slices for U_{ℓ_N} on which $\gamma(1)$ can lie. The possibilities for these slices is completely determined by the datum of $U_0 = U_{\alpha_1} = \ldots = U_{\alpha_N}$.

So to prove existence of a maximal connected integral submanifold S, consider all finite chains $U_{\alpha_0} = U_0, U_{\alpha_1}, \ldots, U_{\alpha_N}$ such that $U_{\alpha_{j-1}} \cap U_{\alpha_j} \neq \emptyset$. Then take successive unions of slices, starting with the slice of U_0 on which M lies, then dictated by intersections of $U_{\alpha_{j-1}}$ and U_{α_j}. The union will be a union of a countable number of slices of the U_ℓ, connected by construction.

To summarze, we have proved:

Theorem 4.3.5. *Let M be a manifold, $E \subset TM$ an integrable tangent sub-bundle, $m \in M$. Then there exists a unique maximal connected integral submanifold passing through m.*

Theorem 4.3.6. *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and $\mathfrak{h} \subset \mathfrak{g}$ a Lie subalgebra. Then there exists a unique connected Lie subgroup $H \subset G$ whose Lie algebra, considered as as subalgebra of $\mathfrak{g}$, coincides with $\mathfrak{h}$.*

Proof. Define $E_\mathfrak{h} \subset TG$ as before, $(E_\mathfrak{h})_g = \ell_{g*}\mathfrak{h}$. By the previous theorem, there exists a unique maximal integral submanifold H which contains e. Suppose $h \in H$; we need to check that $hH \subset H$. Then $\ell_{h^{-1}}H$ is also an integral submanifold because $E_\mathfrak{h}$ is invariant under left translation. Therefore, $\ell_{h^{-1}}H$ is an integral submanifold containing e and is therefore contained in H by maximality. So $H \subset G$ is a subgroup.

This shows that multiplication $H \times H \to G$ is C^∞ and takes values in H. By the discussion of adapted coordinate neighborhoods, given $h \in H$ there exist open neighborhoods U of h in G and U_H of h in H such that $U_H \to U$ is a closed embedding with U_H connected. Therefore, any C^∞ map from N to G which takes values in H is smooth as a map $N \to H$. So $H \subset G$ is a connected Lie subgroup.

The Lie algebra of H, as a subspace of $\mathfrak{g} \simeq T_eG$ is by construction $T_eH = (E_\mathfrak{h})_e = \mathfrak{h}$. To establish uniqueness, recall that

$$\begin{tikzcd} \mathfrak{h} \arrow[r, "\exp"] \arrow[d, hook] & H \arrow[d, hook] \\ \mathfrak{g} \arrow[r, "\exp"] & G \end{tikzcd}$$

is a commutative diagram. Let U be an open neighborhood of $0 \in G$. Then $\exp(U \cap \mathfrak{h})$ contains an open neighborhood of e in H, and we saw earlier than any neighborhood of e in a connected Lie group generates the whole group in the algebraic sense. Since this discussion applies equally well to any potential other connected Lie subgroup H' with Lie algebra $\mathfrak{h}$, we have the desired uniqueness of H. $\square$

As a corollary, we see that H can also be described as the smallest subgroup of G, in the algebraic sense, containing $\exp(U \cap \mathfrak{h})$. In particular, this gives a bijection between connected subgroups of G and Lie subalgebras of $\mathfrak{g}$.

4.4 Lie subgroups

Let M be a manifold. We discuss various notions of "submanifolds" of M.

Immersion. S a manifold equipped with a map $\varphi : S \to M$ such that φ_* is locally 1-1 for every $s \in S$. It follows from the inversion function theorem that φ is locally injective.

This is a very weak notion. For instance, consider the map $\mathbb{R} \to \mathbb{R}^2$ sending $t \mapsto (\cos t, \sin t)$. This is an immersion, but it is from from being injective.

Injective Immersion. S a manifold equipped with a *globally injective* immersion $\varphi : S \to M$. This is the notion of submanifold that we have been using.

One can get pathological examples where the topology on S is incompatible with the subspace topology from M. For instance, the inclusion $\mathbb{R}\times\mathbb{Q} \subset \mathbb{R}^2$ or the "figure 8" immersion.

Embedding. This is an injective immersion where the intrinsic topology of S must agree with the induced (subspace) topology. For example, the inclusion of an open interval in $\mathbb{R}^2$.

Closed embedding. An embedding in which $\varphi(S) \subset M$ is closed.

However, for submanifolds which are *Lie groups*, many of these notions collapse.

Theorem 4.4.1. *For a Lie subgroup $H \subset G$, the following are equivalent.*

1. *H is closed as a subset of G.*

2. *The intrinsic and induced topologies agree.*

3. *$\iota : H \hookrightarrow G$ is a closed embedding.*

Example 4.4.2. $\mathbb{R} \times \mathbb{Q} \hookrightarrow \mathbb{R}^2$ is a Lie subgroup which is not a closed embedding.

The submanifold $(0,1) \to (0,1) \times (0,1)$ looking like a σ with a tail at the boundary of the square is closed but does not have the induced topology.

The open interval in $\mathbb{R}^2$ has the induced topology but is not closed.

Proof. By definition, the first two conditions together are the third. So we must show that in the setting of a Lie subgroup $H \subset G$, (1) $\iff$ (2). Both (1) and (2) are local (with respect to G) properties. So it suffices to prove (1') $\iff$ (2'):

(1') For any open neighborhood U of $g \in G$, $H \cap U$ is closed in $G \cap U$

(2') For any open neighborhood U of $g \in G$, the inclusion $U \cap H \subset H$ is a homeomorphism onto the image.

We know that any open neighborhood of $g \in G$ contains an adapted (to $E_{\mathfrak{h}}$) coordinate neighborhood. So we may assume that $U \simeq (-1,1)^n$ is an adapted coordinate neighborhood and P the set of slices, $P \simeq (-1,1)^{n-k}$ and $U \simeq (-1,1)^k \times P$ (these are all topological isomorphisms, i.e. homeomorphisms). Let $P_H \subset P$ denote the set of slices that are contained in $U \cap H$. We know that $P_H \subset P$ is at most countable.

For the *induced* topology, we have the homeomorphism $U \cap H \simeq (-1,1)^k \times P_H$. On the other hand, for the *intrinsic* topology we have the homeomorphism $U \cap H \simeq (-1,1)^k \times P_H$ when P_H is given the discrete topology.

This reduces the problem to showing that for any adapted U, $P_H \subset P$ is closed if and only if P_H is discrete. Suppose $P_H \subset P$ is not closed; then there exists an adapted U such that P_H is not closed with this U. Then there exists a sequence $p_n \in P_H$ converging to some $p \in P$ where $p \notin P_H$.

The $p_n, p \in P$ correspond to h_n, h with $h_n \simeq 0 \times p_n$ and $h \simeq 0 \times p$. For N large and all $n \geq N$, $h_N^{-1} h_n$ will lie in a particular adapted coordinate neighborhood of the identity. Now, $h_N^{-1} h_n \in H$ but $h_N^{-1} h \notin H$, so $h_N^{-1} h_n$ lie in slices of $U_e \cap H$ which accumulate at the slice of e. Since by construction the h_n are "infinitely far apart" in H, this tells us that P_H is not discrete.

Now suppose that there exists U such that P_H is not discrete. By left translation, we may arrange that U is an adapted neighborhood of the identity and the slice of e in U is not isolated in P_H. Translating again, we see that no slice is isolated. If P_H were closed, then for $p \in P_H$ the set $P_H - \{p\}$ would be dense and open in P_H. So

$$\bigcap_{p \in P_H} (P_H \setminus \{p\})$$

is a countable intersection of open dense subsets, which is empty. This contradicts the Baire category theorem if P_H were a complete metric space in its induced topology. Hence P_H is *not* closed. $\square$

Lemma 4.4.3. *Suppose that $H \subset G$ is a closed Lie group. Then G/H, equipped with the quotient topology, is Hausdorff.*

The fact that H is closed implies that the identity coset in G/H is closed, so every point of G/H is closed, i.e. G/H has the "T1" axiom. It is a general fact that for topological groups, this implies "T2." However, we will go through the details.

Proof. Suppose $g \in G$ and $g \notin H$. Then g has an open neighborhood which does not intersect H. We can expression this open neighborhood as $U_1 g$, where U_1 is an open neighborhood of e. Now we may choose neighborhoods U_2, U_3 of e such that $U_2^2 \subset U_1$ (because multiplication is continuous) and $U_2 = U_2^{-1}$(by replacing U_2 with $U_2 \cap U_2^{-1}$ if necessary) and $U_3 \subset U_2$ and $gU_3g^{-1} \subset U_2$.

If $gU_3H \cap U_3H = \emptyset$, then these two opens descend to open neighborhoods of gH and eH in G/H which are disjoint, as had to be shown. If not, there exists some $x \in gU_3H \cap U_3 = gU_3g^{-1}gH \cap U_3 \subset U_2gH \cap U_2$. This gives an element $h \in H$ lying in $U_2^{-1}U_2g \subset U_1g$, which is a contradiction. □

Theorem 4.4.4. *Keeping the notation above, let $p : G \to G/H$ denote the quotient map. There exists a unique structure of a C^∞ manifold on G/H which is compatible with the quotient topological structure, such that for a continuous function f defined on an open subset $U \subset G/H$, $f \in C^\infty(U)$ if and only if $p^*f \in C^\infty(p^{-1}U)$. Furthermore, this structure has the following properties.*

- *$p : G \to G/H$ and the action map $G \times G/H \to G/H$ are smooth.*
- *p_* induces a linear isomorphism $\mathfrak{g}/\mathfrak{h} \simeq T_{eH}(G/H)$.*
- *Each point in G/H has an open neighborhood U over which $p : p^{-1}(U) \to U$ has a C^∞ section.*

Proof. Uniqueness is obvious: to know the notion of C^∞ fnction on open subsets is to know the manifold structure. Assuming that G/H equipped with this structure is a manifold, $p : G \to G/H$ is C^∞ because the pullbacks of C^∞ functions are C^∞.

The fact that $G \times G \overset{m}{\to} G$ is C^∞ is C^∞, since the pullback of a C^∞ function on G/H is a function on $G \times G/H$ whose pullback to $G \times G$ is right H-invariant and smooth.

We shall see, in the proof of existence of the smooth structure, that p_* is surjective and $\dim(G/H) = \dim G - \dim H$. Now, $T_{eH}G/H \simeq \mathfrak{g}/\ker p_*$. The fact that C^∞ functions on G/H are right H-invariant implies that $\mathfrak{h} \subset \ker p_* : T_eG \to T_{eH}(G/H)$. By dimension reasons, $\mathfrak{h} = \ker p_*$.

To complete the proof, we must construct the manifold structure and local sections near the identity coset (because left translation by any $g \in$ is a diffeomorphism) which relates local sections at the identity coset to sections near gH.

To establish the manifold structure, we need to show that locally G/H, with the presumed smooth structure, is diffeomorphic to an open set in $\mathbb{R}^n$. Again, it suffices to do this for an open neighborhood of e. Note that this is an entirely *local* question.

As before, let $E_\mathfrak{h}$ denote the left-invariant tangent sub-bundle of TG whose fiber at e is $\mathfrak{h}$. Let U be an adapted coordinate neighborhood of e. We may suppose that the coordinates are centered at e, i.e. $x_j(e) = 0$ for all j. In the following, we will shrink U and and tacitly scale the coordinates so that the shrunk U is an adapted coordinate neighborhood (recall that adapted coordinates have as their image $(-1,1)^n$). As before, we have $U \simeq P \times S_0$ where $P = (-1,1)^{n-k}$ and S_0 is the slice through $e \simeq 0$, isomorphic to $(-1,1)^k$.

Note that each slice is contained in a coset gH (by left invariance of the tangent sub-bundle $\mathfrak{h}$, integral manifolds are preserved). After shrinking U, we can arrange that $U \cap H = S_0$. Shrinking further if necessary, we claim that we may arrange that no two distinct slices in U lie in the same coset gH. If this were not true, there would exist sequences p_n' and p_n'' lying in distinct slices in P both converging to 0 and lying in the same coset. Let g_n' and g_n'' be the group elements to $(p_n', 0)$ and $(p_n'', 0)$. Then $g_n' \to e$, and $g_n'' \to e$, but $g_n'H = g_n''H$. Then $(g_n'')^{-1}g_n' \to e$ and $(g_n'')^{-1}g_n'H = H$, so $(g_n'')^{-1}g_n' \in S_0$ for $n \gg 0$. So we get convergence $(g_n'')^{-1}g_n' \to e$ in H, which implies that $g_n'' \to g_n'$ in H, but this is impossible because they are "far apart" in H (the different slices having the discrete topology).

It must then be the case that $p^{-1}(P) = U \simeq P \times H$. That implies $p(U) \simeq P \simeq (-1,1)^{n-k}$ and hence $p(U)$ is isomorphic to an open subset of $\mathbb{R}^{n-k}$, the identification preserving the smooth structure. The functions on $p(U)$ are the right H-invariant functions on $P \times H$, i.e. functions on P.

The section $s : (-1,1)^{n-k} \simeq P \to U$ sending $x \mapsto (x, 0)$ is smooth by construction. Then $p : G \to G/H$ "looks like" the projection $U \simeq (-1,1)^{n-k} \times (-1,1)^k \to (-1,1)^{n-k}$. $\square$

Theorem 4.4.5. *Let G be a Lie group, $H \subset G$ a subgroup in the algebraic sense such that H as a subset of G is closed. Then H has a natural, unique structure of a Lie subgroup.*

Proof. Let $\mathfrak{g}$ be the Lie algebra of G. Define $\mathfrak{h} = \{X \in \mathfrak{g} \mid \exp tX \in H \text{ for all } t\}$.

Lemma 4.4.6. $\mathfrak{h} \subset \mathfrak{g}$ *is a linear subspace.*

Proof. From the defninion it is immediate that $\mathfrak{h}$ is closed under scalar multiplicaiton, so it suffices to show that $\mathfrak{h}$ is closed under addition. Define $M : V \times V \to \mathfrak{g}$ as before, i.e. multiplication on G pulled back to $\mathfrak{g}$ near 0 via exp. Suppose $X, Y \in \mathfrak{h}$ and $t \in \mathbb{R}$. For $n \gg 0$,

$$\frac{t}{n}X, \frac{t}{n}Y \in V.$$

Then

$$M\left(\frac{t}{n}X, \frac{t}{n}Y\right) = \frac{t}{n}(X+Y) + O(\frac{1}{n^2}).$$

Then $\exp(t(X+Y))$ is the limit as $n \to \infty$ of

$$\begin{aligned}\exp\left(nM\left(\frac{t}{n}X, \frac{t}{n}Y\right)\right) &= \left(\exp M\left(\frac{t}{n}X, \frac{t}{n}Y\right)\right)^n \\ &\approx \left(\exp\left(\frac{t}{n}X\right)\exp\left(\frac{t}{n}Y\right)\right)^n \in H\end{aligned}$$

The fact that $H \subset G$ is closed implies that $\exp(t(X+Y)) \in H$ for all t. $\square$

Let $\mathfrak{h}' \subset \mathfrak{g}$ be a linear complement, i.e. $\mathfrak{g} \simeq \mathfrak{h} \oplus \mathfrak{h}'$.

Lemma 4.4.7. *If U is a sufficiently small open neighborhood of $0 \in \mathfrak{g}$, then*

$$\exp(U \cap \mathfrak{h}) = \exp(U) \cap H.$$

Proof. Note that $\exp \mathfrak{h} \subset H$ by definition. Therefore, $\exp(V \cap \mathfrak{h}) \subset H \cap \exp V$ for any open neighborhood V of 0 in $\mathfrak{g}$. If the lemma were false, we could find a sequence $h_n \in H$ converging to e such that $h_n \notin \exp \mathfrak{h}$. Let $\mathfrak{q} \subset \mathfrak{g}$ be a linear complement to $\mathfrak{h}$. Then the map

$$\begin{aligned} \mathfrak{q} \oplus \mathfrak{h} &\to G \\ (X, Y) &\to \exp X \exp Y \end{aligned}$$

is a local diffeomorphism of a neighborhood of 0 in $\mathfrak{q} \oplus \mathfrak{h}$ into a neighborhood of e in G. Eventually, the h_n lie in the image of this map, so their pre-images have non-zero $\mathfrak{q}$-component. This says that there exist sequences $\{X_n\} \subset \mathfrak{q}$ and $\{Y_n\} \subset \mathfrak{h}$ such that $\exp X_n \exp Y_n = h_n$, $X_n \to 0$, $Y_n \to 0$, and $X_n \neq 0$ for all n.

Introduce a linear norm $|| \cdot ||$ on $\mathfrak{g}$ and write $X_n = t_n \widetilde{X_n}$ where $||\widetilde{X}_n|| = 1$. Passing to a subsequence if necessary, we may assume that $\widetilde{X}_n \to \widetilde{X} \in \mathfrak{q}$. We must necessarily have $||\widetilde{X}|| = 1$. Let $t \in \mathbb{R}$ be fixed. We know that $t_n \to 0$, so we may choose $k_n \in \mathbb{Z}$ such that $t - t_n < k_n t_n \leq t$. Then $k_n t_n \to t$, so

$$\exp(k_n X_n) = \exp(k_n t_n \widetilde{X}_n) \to \exp(t\widetilde{X}).$$

But on the other hand, $\exp(k_n X_n) = (\exp X_n)^{k_n}$, so

$$\exp(k_n X_n) = (\exp X_n)^{k_n} \to \exp t\widetilde{X}.$$

So $\exp(t\widetilde{X}) \in H$. Since this holds for all t, it must be the case that $\widetilde{X} \in \mathfrak{h}$. Hence $\widetilde{X} \in \mathfrak{h} \cap \mathfrak{q} = \{0\}$. But $\widetilde{X}$ is a unit vector, so this is impossible. $\square$

Near the origin, exp is a diffeomorphism. By the preceding lemma, e has an open neighborhood U_e such that $H \cap U_e$ is a closed, embedded submanifold of U_e. By translation, every $h \in H$ has an open neighborhood U_h with the same property.

Since H is closed, for any $g \in G - H$ there exists an neighborhood U_g of g such that $U_g \cap H = \emptyset$. Therefore, $H \hookrightarrow G$ is a closed embedding of a submanifold. Hence the multiplication map $H \times H \to G$ and the inversion map $H \to G$ are restrictions to $H \times H$ (resp. H) of smooth maps from $G \times G$ (resp. G), hence are smooth maps. Moreover, they take values in H, hence they are C^∞ as maps into H (precisely because this is a smooth embedding; we gave counterexamples in the other cases). $\square$

4.5 Lie's Theorems

Theorem 4.5.1. *Suppose $\Phi : G \to H$ is a homomorphism of Lie groups such that on the level of Lie algebras, $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ is injective. Then*

(i) $\ker \Phi \subset G$ *is discrete.*

(ii) If Φ is onto, then $\Phi : G \to H$ is a covering map.

Proof. $\ker \Phi$ is a closed, normal Lie subgroup (e.g. by the theorem we just proved). By a homework problem, $G/\ker \Phi$ has a natural Lie group structure. The map $\Phi : G \to H$ factors through $G/\ker \Phi \to H$ by the definition of the C^∞ structure. Therefore, $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ factors through $\mathfrak{g}/\mathrm{Lie}(\ker \Phi) \to \mathfrak{h}$. Since Φ_* is injective, $\ker \Phi$ must have dimension 0, which implies that $\ker \Phi$ has the discrete topology. This establishes (i).

Note that we have implicitly used the Baire category theorem; it is an exercise to give a proof not using this theorem.

Φ factors through $G/\ker \Phi \to H$. If Φ is onto, then this is an isomorphism of Lie groups. Also, $G/\ker \Phi$ has the same Lie algebras as G, so $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ is a linear isomorphism. Locally near the identity, Φ is a diffeomorphism, so there exist open neighborhoods $U_{G,0}$ of $e_G \in G$ and $U_{H,0}$ of $e_H \in H$ such that $U_{G,0} \cap \ker \Phi = \{e\}$ and Φ restricts to a diffeomorphism $U_{G,0} \to U_{H,0}$.

Now let U_G be an open neighborhood of e such that $U_G = U_G^{-1}$ and $U_G^2 \subset U_{G,0}$. Let $U_H = \Phi(U_G)$. Then $\Phi : U_G \to U_H$ is a diffeomorphism. We claim that for $g \in \ker \Phi$ with $g \neq 0$, then $U_G g \cap U_G = \emptyset$ (otherwise, we would have a non-identity element $g \in U_G^{-1} U_G \subset U_{G,0}$, which is impossible by construction). By translation, $U_G g_1 \cap U_G g_2 = \emptyset$ whenever g_1, g_2 are distinct elements of $\ker \Phi$. This implies that

$$\Phi^{-1}(U_H) = \bigcup_{g \in \ker \Phi} U_G g,$$

i.e. the pre-image of U_H is a disjoint union of neighborhoods of the form $U_G g$ such that $\Phi : U_G g \to U_H$ is a diffeomorphism for any $g \in \ker \Phi$.

By left translation, we obtain an analogous description for suitably small neighborhoods of any given $g \in G$ and $U_{\Phi(g)}$ of $\Phi(g) \in H$, hence Φ is a covering map. $\square$

Corollary 4.5.2. *Suppose $\Phi : G \to H$ is a Lie group homomorphism such that $\Phi_* : \mathfrak{g} \to \mathfrak{h}$ is an isomorphism. If G, H are connected then Φ is a covering map. In particular, if H is also simply connected, Φ is an isomorphism.*

Proof. The point is to show surjectivity. By the universality of exp, the image of Φ contains an open neighborhoods of the identity in H. Since H is connected, this open neighborhood generates H, so Im $\Phi = H$. By the previous theorem, $\Phi : G \to H$ is a covering map.

If H is simply connected, then obviously Φ must be a homeomorphism. $\square$

Corollary 4.5.3. *Let G, H be Lie groups with Lie algebras $\mathfrak{g}, \mathfrak{h}$ and $\varphi : \mathfrak{g} \to \mathfrak{h}$ a Lie algebra homomorphism. Suppose G and H are connected and G is simply connected. Then there exists a unique Lie group homomorphism $\Phi : G \to H$ such that $\Phi_* = \varphi$.*

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

$$\{(X, \varphi(X)) \mid X \in \mathfrak{g}\} \subset \mathfrak{g} \oplus \mathfrak{h}$$

Then we know that there exists a closed Lie subgroup $\Gamma \subset G \times H$ with this Lie algebra. Let p_1, p_2 be the two projection maps $G \times H \to G$ and $G \times H \to H$. Well, p_{1*} restrict to the graph of φ is 1-1 and onto, so $p_1 : \Gamma \to G$ is a covering map. Since G is simply connected, this is an isomorphism. So $p_1^{-1} : G \to \Gamma$ is a well-defined Lie group homomorphism. Define $\Phi = p_2 \circ p_1^{-1}$; then $\Phi_* = \varphi$ by construction. □

Corollary 4.5.4. *Let G, H be Lie groups and $\Phi : G \to H$ a homomorphism of groups in the algebraic sense. If Φ is continuous, it is C^∞, hence a Lie group homomorphism.*

Proof. Let Γ be the graph of Φ,

$$\Gamma = \{(g, \Phi(g)) \mid g \in G\} \subset G \times H.$$

This is a subgroup in the algebraic sense because Φ is a homomorphism, and a closed subset since Φ is continuous. Hence this is a Lie subgroup of $G \times H$. By construction, $p_1|_\Gamma : \Gamma \to G$ is 1-1 and onto, hence an isomorphism. By definition of Γ, $\Phi = p_2 \circ p_1^{-1}$. □

Proposition 4.5.5. *A discrete normal subgroup of a connected Lie group G is central (lies in the center).*

Proof. Suppose $h \in H$ is not the identity. The map $g \mapsto ghg^{-1}$ is continuous, so given an open U_h of h, there exists an open neighborhood U of $e \in G$ such that $g \in U \implies ghg^{-1} \in U_h$. We can do this for some U_h such that $U_h \cap H = h$. By normality, $ghg^{-1} \in H \cap U_H$, so $ghg^{-1} = h$. That tells us that an open neighborhood of $e \in G$ commutes with h, but such a neighborhood generates G, so all of G commutes with h.

A more concise way to say this: the map $g \mapsto ghg^{-1}$ sends G to a discrete set, hence is constant. □

Note that subgroups of this type arise as kernels of Lie group homomorphisms $\Phi : G \to H$ such that Φ_* is 1-1 and G is connected. In particular, this applies to Lie group shomomorphisms $\Phi : G \to H$ that are covering maps, with G connected.

Suppose G is a connected Lie group and $\widetilde{G}$ the universal cover.

Theorem 4.5.6. *There exists a natural structure of a Lie group on $\widetilde{G}$ such that the covering map $\widetilde{G} \to G$ is a Lie group homomorphism.*

Exercise 4.5.7. Prove this theorem.

Exercise 4.5.8. Let $\mathfrak{g}$ be a finite dimensional Lie algebra over $\mathbb{R}$. Does there exist a Lie group G with Lie algebra $\mathfrak{g}$?

An approach to this problem is through Ado's theorem:

Theorem 4.5.9 (Ado)**.** *Any such $\mathfrak{g}$ can be realized as a Lie subalgebra of End$(\mathbb{R}^N)$ for some N.*

Combined with our work above, this tells us the answer is yes.

Chapter 5

The Universal Enveloping Algebra

5.1 Construction of the universal enveloping algebra

Let K be an arbitrary field, $\mathfrak{g}$ a Lie algebra over K. We define

$$\bigotimes \mathfrak{g} = \bigoplus_{k=0}^{\infty} \mathfrak{g}^{\otimes k}.$$

This is a graded associative algebra over K with unit, freely generated by $\mathfrak{g}^{\otimes 1} = \mathfrak{g}$. These conditions characterize $\bigotimes \mathfrak{g}$ as an algebra.

Similarly, we define

$$S(\mathfrak{g}) = \bigoplus_{k=0}^{\infty} S^k \mathfrak{g},$$

where $S^k \mathfrak{g}$ is the k^{th} symmetric power of $\mathfrak{g}$. This is a commutative algebra over K with unit, freely generatd by $S^1 \mathfrak{g} = \mathfrak{g}$, which uniquely characterizes $S(\mathfrak{g})$.

Now define

$$U(\mathfrak{g}) = \bigotimes \mathfrak{g} / (\{XY - YX - [X, Y] \colon X, Y \in \mathfrak{g}\}).$$

This is the tensor algebra modded out by the two sided ideal generated by relations of the form $XY - YX - [X, Y]$. The composition

$$\mathfrak{g} = \mathfrak{g}^{\otimes 1} \to U(\mathfrak{g})$$

is a Lie algebra homomorphism when $U(\mathfrak{g})$ is considered as a Lie algebra via commutators.

We have constructed $U(\mathfrak{g})$ to have the following properties:

- it is an associative algebra over K with unit (*),
- it is filtered (**),
- it is equipped with a linear map $i : \mathfrak{g} \to U_1(\mathfrak{g})$ which is a Lie algebra homomorphism via commutators, generated by $i(\mathfrak{g})$.

(**) Let $U_k(\mathfrak{g})$ denote the image of $\bigoplus_{\ell \leq k} \mathfrak{g}^{\otimes \ell}$. Then we have a filtration

$$U_0(\mathfrak{g}) \subset U_1(\mathfrak{g}) \subset \ldots \subset U_k(\mathfrak{g}) \ldots \subset_{k=0}^{\infty} U_k(\mathfrak{g}) = U(\mathfrak{g})$$

satisfying $U_k(\mathfrak{g})U_\ell(\mathfrak{g}) \subset U_{k+\ell}(\mathfrak{g})$.

(*) Note that

$$\bigotimes \mathfrak{g} = \underbrace{K \cdot 1}_{\mathfrak{g}^{\otimes 0}} \oplus \underbrace{\mathfrak{g} \otimes \mathfrak{g}}_{\bigoplus_{k=1}^{\infty} g^{\otimes k}} .$$

The second summand is called the *augmentation ideal.* Note that the ideal we are quotienting out by lies in the augmentation ideal, and hence never harms the constants. Hence $U_0(\mathfrak{g}) = K \cdot 1$.

Definition 5.1.1. The pair $(U(\mathfrak{g}), i)$ is called the *universal enveloping algebra* of $\mathfrak{g}$.

5.2 Poincaré-Birkhoff-Witt

Question. Is $i : \mathfrak{g} \to U_1(\mathfrak{g})$ injective? Well, there isn't anything that is transparently in the kernel, but it also isn't necessarily obvious that the relations can't somehow kill something at degree 1. In fact, it *is* injective, as we shall eventually prove.

Observe that for each k, "$U_k(\mathfrak{g})$ is commutative modulo $U_{k-1}(\mathfrak{g})$." In other words, the commutator of two elements in $U_k(\mathfrak{g})$ differ by an element of $U_{k-1}(\mathfrak{g})$. This is because the elements of $i(\mathfrak{g})$ in $U_2(\mathfrak{g})$ commute modulo $i(\mathfrak{g})$. For instance, in $U_3(\mathfrak{g})$ the elements XYZ, YZX, XZY, etc. all commute up to elements in $U_2(\mathfrak{g})$ (note that this is a different meaning of "commute" from what is sometimes used elsewhere). Consider the associated graded algebra to $U(\mathfrak{g})$:

$$\mathrm{gr}(U(\mathfrak{g})) := \bigoplus_{k=0}^{\infty} U_k(\mathfrak{g})/U_{k-1}(\mathfrak{g}).$$

This is a graded associative algebra over K with unit, generated by $i(\mathfrak{g})$. It is commutative by the remarks above.

Hence by the universality of $S(\mathfrak{g})$, the map linear map $i : \mathfrak{g} \to U_1(\mathfrak{g})$ induces a unique algebra homomorphism $S(\mathfrak{g}) \to \mathrm{gr}(U(\mathfrak{g}))$ of commutative algebras with unit. Moreover, this map is surjective by construction, since $\mathfrak{g}$ generates $\mathrm{gr}(U(\mathfrak{g}))$.

Theorem 5.2.1 (Poincaré-Birkhoff-Witt)**.** *This morphism of algebras is an isomorphism:*

$$S(\mathfrak{g}) \simeq \mathrm{gr}(U(\mathfrak{g})).$$

Corollary 5.2.2. *$i : \mathfrak{g} \to U_1(\mathfrak{g})/U_0(\mathfrak{g})$ is an isomorphism, hence $i : \mathfrak{g} \to U(\mathfrak{g})$ is injective.*

This tells us how to think of $U(\mathfrak{g})$: it is the *universal* associative algebra which contains $\mathfrak{g}$ as a Lie sub-algebra with the Lie bracket of commutators.

Before embarking on the proof, we observe some more consequences. Suppose $\dim \mathfrak{g} < \infty$, and let $X_1, \ldots, X_n$ be a basis of $\mathfrak{g}$ over K.

Corollary 5.2.3. *$\{X_1^{j_1} \ldots X_n^{j_n} \mid 0 \leq j_1, \ldots, j_n < \infty\}$ is a basis for $U(\mathfrak{g})$.*

Proof. The subset of elements above with $j_1 + \ldots + j_n = k$ obviously comprise a basis for $S^k(\mathfrak{g})$, hence they push forward to a basis for U_k/U_{k-1}. □

What if $\mathfrak{g}$ is infinite-dimensional? Similar considerations apply, but one has to be more careful about stating the result.

Corollary 5.2.4. *Let $\mathfrak{h} \subset \mathfrak{g}$ be a Lie subalgebra. Then $U(\mathfrak{h})$ is a subalgebra of $U(\mathfrak{g})$.*

Proof. Choose the basis of $\mathfrak{g}$ to contain a basis for $\mathfrak{h}$ as a subset, and then use the form of solutions in Corollary 5.2.3. □

Corollary 5.2.5. *Suppose $\mathfrak{h}_1, \mathfrak{h}_2$ are Lie subalgebras of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$ as vector spaces. Then $U(\mathfrak{g}) \simeq U(\mathfrak{h}_1) \otimes_K U(\mathfrak{h}_2)$ as left $U(\mathfrak{h}_1)$-modules and right $U(\mathfrak{h}_2)$-modules.*

Proof. This is clear by considering bases: hoose the basis of $\mathfrak{g}$ so that the first things are a basis of $\mathfrak{h}_1$ and the second things are a basis of $\mathfrak{h}_2$. Any basis element in the form of Corollary 5.2.3 may be expressed as a tensor of two such elements for $\mathfrak{h}_1$ and $\mathfrak{h}_2$. □

Now suppose that K has characteristic 0. Define

$$\begin{aligned} \widetilde{s}_k : \mathfrak{g}^{\otimes k} &\to U_k(\mathfrak{g}) \\ Y_1 \otimes \ldots \otimes Y_k &\mapsto \frac{1}{k!} \sum_{\sigma \in S_k} Y_{\sigma(1)} \ldots Y_{\sigma(k)}. \end{aligned}$$

We initially define this for basis elements only, and then extend by linearity to the whole tensor product. Note that $\widetilde{s}_k$ factors through $S(\mathfrak{g})$, hence induces

$$s_k : S^k(\mathfrak{g}) \to U_k(\mathfrak{g}).$$

This map s_k is called the "symmetrization map." Note that this definition doesn't depend on PBW and, in fact, even without PBW it is clear that Im S_k spans $U_k(\mathfrak{g})/U_{k-1}(\mathfrak{g})$. What PBW tells us is that this is not only a surjection, but an isomorphism.

Corollary 5.2.6. *The symmetrization map* $s = \bigoplus_{k=0}^{\infty} s_k$ *defines a vector space isomorphism* $S(\mathfrak{g}) \simeq U(\mathfrak{g})$.

5.3 Proof of Poincaré-Birkhoff-Witt

We shall prove PBW in the following situation: $\mathfrak{g}$ is the Lie algebra of a Lie group, so in particular $K = \mathbb{R}$ or $\mathbb{C}$, and $\dim \mathfrak{g} < \infty$. Granted Ado's theorem, this furnishes a proof for arbitrary real or complex Lie algebras.

We know that the map from the symmetric algebra to the universal enveloping algebra is surjective; we want it to be injective. To do this, we want to show that the universal enveloping algebra is as big as we expect: that there are no unexpected relations. We achieve this by finding a space (the space of functions on G) on which it acts sufficiently nontrivially.

We first set up some preliminaries. Let M by a C^∞ manifold of dimension n. To an open set $U \subset M$, define $\mathcal{D}(U)$ to be the algebra of linear differential operators on U with C^∞ coefficients of finite order. Let $\mathcal{D}_k(U)$ be the space of differential operators of order $\leq k$.

Let $(U; x_1, \ldots, x_n)$ be a coordinate neighborhood. Then

$$\mathcal{D}_k(U) = \left\{ \sum_{\ell=0}^{k} \sum_{\substack{0 \leq j_1, \ldots, j_n < \infty \\ j_1 + \ldots + j_n = \ell}} a_{j_1 \ldots j_n} \frac{\partial^{j_1}}{\partial x_1^{j_1}} \frac{\partial^{j_2}}{\partial x_2^{j_2}} \cdots \frac{\partial^{j_n}}{\partial x_n^{j_n}} \mid a_{j_1 \ldots j_n} \in C^\infty(M) \right\}.$$

Define

$$\sigma_k \left(\sum_{\ell=0}^{k} \sum_{\substack{0 \leq j_1, \ldots, j_n < \infty \\ j_1 + \ldots + j_n = \ell}} a_{j_1 \ldots j_n} \frac{\partial^{j_1}}{\partial x_1^{j_1}} \frac{\partial^{j_2}}{\partial x_2^{j_2}} \cdots \frac{\partial^{j_n}}{\partial x_n^{j_n}} \right) = \sum_{j_1 + \ldots + j_n = k} a_{j_1 \ldots j_n} \frac{\partial^{j_1}}{\partial x_1^{j_1}} \frac{\partial^{j_2}}{\partial x_2^{j_2}} \cdots \frac{\partial^{j_n}}{\partial x_n^{j_n}}.$$

This is called the k^{th} *principal symbol* (it picks out the "homogeneous" piece of degree k). It could depend on the choice of coordinates, but since differential operators of degree k commute with multiplication by smooth functions *modulo* differential operators of degree $k-1$, this becomes well defined as a linear map on the quotient:

$$\sigma_k : \mathcal{D}_k(U)/\mathcal{D}_{k-1}(U) \to C^\infty(U; S^k TM)$$

where $S^k TM$ is the k^{th} symmetric power of TM, i.e. the vector bundle over M where the fiber at x is $S^k(T_x M)$. To see that $S^k TM$ is a vector bundle, note that there exist local frames, e.g.

$$\left\{ \frac{\partial^{j_1}}{\partial x_1^{j_1}} \frac{\partial^{j_2}}{\partial x_2^{j_2}} \cdots \frac{\partial^{j_n}}{\partial x_n^{j_n}} \mid j_1 + \ldots + j_n = k \right\}$$

over $(U; x_1, \ldots, x_n)$. So in fact, the principal symbol induces an *isomorphism*

$$\sigma_k : \mathcal{D}_k(U)/\mathcal{D}_{k-1}(U) \simeq C^\infty(U; S^k TM), \tag{5.3.1}$$

at least when U is a coordinate neighborhood. In fact, this holds for any open subset $U \subset M$, and in particular $U = M$, by the usual partition of unity arguments.

Let G be a Lie group. Then G acts on $\mathcal{D}(G)$ by left (and also right) translation. For $D \in \mathcal{D}(G)$ and $f \in C^\infty(G)$, define

$$(\ell_g D)(f) = \ell_g^*(D(\ell_{g^{-1}}^* f)).$$

G also acts by left translation on $S^k TM$, and the isomorphism (5.3.1) is equivariant with respect to this action.

Now let $\mathcal{D}^\ell(G)$ denote the associative algebra of left-invariant differential operators. Then σ_k induces an isomorphism

$$\sigma_k : \mathcal{D}_k^\ell(G)/\mathcal{D}_{k-1}^\ell(G) \simeq C^\infty(G, S^k TG)^G$$

(The notation $C^\infty(G, S^k TG)^G$ means the G-invariants). This is clear from geometric reasons because we can patch differential operators using partitions of unity. This is a *model* for PBW. Composing with evaluation at e, we get an element of $S^k T_e G = S^k \mathfrak{g}$. Since this was obtained by restricting the isomorphism σ_k to a subspace, this is injective. As we indicated above, a partition of unity argument shows that it is in fact an isomorphism.

Identifying $\mathfrak{g}$ with left-invariant vector fields on G induces a map $\mathfrak{g} \to D_1^\ell(G)/D_0^\ell(G)$. But $D_0^\ell(G)$ are left-invariant degree 0 operators, i.e. constant functions. This induces a map $U(\mathfrak{g}) \to \mathcal{D}^\ell(G)$ since the relations in $U(\mathfrak{g})$ were precisely constructed from the model of vector fields. Furthermore, this map preserves degree, so we get for each graded component k maps:

$$\begin{tikzcd}
\mathcal{D}_k^\ell(G)/\mathcal{D}_{k-1}^\ell(G) \arrow[rr, hook, "\sigma_k"] & & S^k\mathfrak{g} \arrow[dl, "\mathrm{gr}"] \\
& U_k(\mathfrak{g})/U_{k-1}(\mathfrak{g}) \arrow[ul] &
\end{tikzcd}$$

We claim that this is a commutative diagram. This is clear for $k = 0$ and $k = 1$, since when $k = 0$ everything is the space of constants, and when $k = 1$ everything is the space of left-invariant vector fields. Furthermore, these are compatible with composition of differential operators and multiplication in $S(\mathfrak{g}), U(\mathfrak{g})$, so we deduce the commutativity for all k.

But since this is a commutative *triangle*, all the maps are automatically isomorphisms.

Corollary 5.3.1. *$U(\mathfrak{g}) \simeq \mathcal{D}^\ell(G)$, i.e. $U(\mathfrak{g})$ is the algebra of left-invariant differential operators on G.*

To give a conceptual review of the proof, we wanted to show that the universal enveloping algebra is really big, with no unexpected relations. We achieve this by finding a space (the space of functions on G) on which it acts sufficiently nontrivially. Another variant of this argument runs as follows. Viewing $U(\mathfrak{g})$ as differential operators on the space of functions on G, we can analyze its action on the space of power series which vanish to order $k-1$, evaluated at 0. This is equivalent to picking the k^{th} degree term, i.e. the k^{th} principal symbol.

Chapter 6

Representations of Lie groups

6.1 Haar Measure

We need the notion of Haar measure in discussing the representation theory of Lie groups. This is fairly transparent on Lie groups, but to satisfy the "general education" requirement we will discuss Haar measure in greater generality.

Let G be a locally compact Hausdorff topological group.

Theorem 6.1.1 (Haar)**.** *There exists a nontrivial, regular, left-invariant Borel measure on G. It is unique up to scaling and satisfies $m(U) > 0$ for every nonempty open Borel set U.*

We clarify some of the terms in this theorem.

- A *Borel set* (in a locally compact Hausdorff space) is an element of the σ-ring (closed under set differences and countable unions) generated by the compact subsets.

- A *Borel measure* is defined on Borel sets and finite on compact sets.

- A Borel measure is *regular* if for every Borel measurable set S,

$$\begin{aligned} m(S) &= \sup\{m(C) \mid C \text{ compact } \subset S\}, \\ &= \inf\{m(U) \mid U \text{ open, Borel set } \supset S\} \end{aligned}$$

G acts on measures by right translation $S \mapsto Sg^{-1} \mapsto m(Sg^{-1})$. Let m be the Haar measure, normalized in some way. Then for $g \in G$, $r(g)m$ is again the Haar measure, possibly with a different normalization. So

$$r(g)m = \Delta(g)m \text{ for some } \Delta(g) \in \mathbb{R}_{>0}.$$

In fact, Δ is a homomorphism $G \to \mathbb{R}_{>0}$.

Exercise 6.1.2. Show tha Δ is continuous.

Recall that Haar measure is left-invariant by definition.

Definition 6.1.3. G is **unimodular** if Haar measure is also right-invariant, i.e. $\Delta \equiv 1$.

So Δ is called the "unimodular function."

Note that compact groups do not admit nontrivial continuous homomorphisms into $\mathbb{R}_{>0}$, since if anything not equal to 1 is in the image, we can generate arbitrarily big elements of the image through multiplication, but the image must be compact.

Corollary 6.1.4. *Compact groups are unimodular.*

By definition, compact groups have finite Haar measure. The general convention is to normalize the Haar measure of compact groups so that $m(G) = 1$.

If G is a Lie group, there obviously exists a left-invariant non-zero differential form of top degree, unique up to scaling. This defines Haar measure.

6.2 Representations

Let G be a locally compact Hausdorff group, V a topological vector space which is

(a) finite-dimensional over $\mathbb{R}$ or $\mathbb{C}$, or

(b) real or complex Hilbert space, or

(c) real or complex Banach space.

Let $\mathrm{Aut}(V)$ be the group of bounded linear maps from V to itself which have bounded inverses. This is a group, but we shall *not* topologize it unless $\dim V < \infty$.

Definition 6.2.1. A *representation* of G on V is a homomorphism

$$\pi : G \to \mathrm{Aut}(G)$$

which is continuous in the following sense: the action map

$$\begin{aligned} G \times V &\to V \\ (g, v) &\mapsto \pi(g)v \end{aligned}$$

is continuous.

Definition 6.2.2. If V is a Hilbet space, a representation π of G on V is called *unitary* if $\pi(g)$ for $g \in G$ is a unitary operator (i.e. preserves the inner product).

Example 6.2.3. Let $L^2(G)$ denote the space of L^2 (with respect to Haar measure) Borel measurable functions[1]. Then left translation defines a homomorphism

$$\ell : G \to \mathrm{Aut}(L^2(G)).$$

This preserves the inner product. We claim that its action is continuous, making ℓ a unitary representation of G on $L^2(G)$. This is called the *left regular representation.* It is important that we use the weaker notion of continuity here, rather than continuity of ℓ itself.

Exercise 6.2.4. Check the claim.

Example 6.2.5. By left translation, $(\mathbb{R}, +)$ acts on $L^\infty(\mathbb{R})$. This is *not* continuous. Why? Consider the characteristic function χ for $[0, 1]$. If we translate slightly, we get a function which differs in norm from χ by 1.

Let $\pi : G \to \mathrm{Aut}(V)$ be a homomorphism, $u, v \in V$ and $g, h \in G$. Then

$$\begin{aligned}
\pi(g)u - \pi(h)v &= (\pi(g)u - \pi(h)u) + \pi(h)(u - v) \\
&= \pi(g)(1 - \pi(g^{-1}h))u + \pi(g)\pi(g^{-1}h)(u - v).
\end{aligned}$$

So π is continuous if and only if it satisfies:

(i) $G \times V \to V$ is continuous at $(e, 0)$.

(ii) For any $v \in V$, the map $g \mapsto \pi(g)v$ is continuous (equivalently, continuous at e).

The first condition ensure that both terms in

$$\pi(g)(1 - \pi(g^{-1}h))u + \pi(g)\pi(g^{-1}h)(u - v)$$

are small. The second term only seems to control the first term, a priori. It is clear that (i) $\implies$ (ii), but in fact we will show that (ii) $\implies$ (i).

Lemma 6.2.6. *In the notation above, (ii) $\implies$ (i).*

Proof. We apply the "uniform boundedness principle": if V is a Banach space and B a set of bounded linear operators on V, then $\{Tv \mid T \in B\}$ is bounded for all $v \in V$ if and only if $\{||T|| \mid T \in B\}$ is bounded.

Apply this to the set

$$\{\pi(g) \mid g \in C\}$$

where C is a compact neighborhood of e. Then (ii) $\implies$ $\{\pi(g)v \mid g \in C\}$ is bounded for every $v \in V$ because it is the image of a compact set under a continuous map. Hence the boundedness principle implies that $\{||\pi(g)|| \mid g \in C\}$ is bounded. Substituting the bounds into the second equation above, we get what we want. □

[1] equivalently, the closure in the L^2 norm of the compactly supported continuous functions.

Definition 6.2.7. A homomorphism $\pi : G \to \mathrm{Aut}(V)$ is *strongly continuous* if $g \mapsto \pi(g)v \in V$ is continuous for every $v \in V$.

We have shown that strong continuity implies continuity (note that strong continuity is, formally, a weaker notion than continuity).

On the other hand, a homomorphism $\pi : G \to \mathrm{Aut}(V)$ is *weakly continuous* if the map

$$g \mapsto \langle \tau, \pi(g)v \rangle \in \mathbb{R} \text{ or } \mathbb{C}$$

is continuous if for all $v \in V, \tau \in V^*$. It is a fact, which we shall not prove, that weak continuity implies strong continuity.

6.3 Schur orthogonality relations

Definition 6.3.1. Suppose $\pi : G \to \mathrm{Aut}(V)$ is a representation. We say that π is *irreducible* if there does not exist a nontrivial closed subspace $V_1 \subset V$ which is invariant under G, i.e. $\pi(g)V_1 \subset V$ for all $g \in G$.

Suppose π is a unitary representation of G on a Hilbert space V, $V_1 \subset V$ a closed, invariant subspace. Then $V_1^\perp$ is also also a closed invariant subspace, giving a decomposition $V = V_1 \oplus V_1^\perp$. If in addition $\dim V < \infty$, then we say (π, V) is *completely reducible,* i.e. a finite direct sum of irreducible representations.

Lemma 6.3.2. *let (π, V) be a finite-dimensional representation of a compact Lie group G. Then it can be made unitary, i.e. we can put an inner product on V which makes π unitary.*

Proof. This is the usual "averaging trick." Let $(\cdot, \cdot)_0$ be an arbitrary inner product and define

$$(u, v) = \int_G (\pi(g)u, \pi(g)v)_0 \, dg.$$

□

Corollary 6.3.3. *Finite dimensional representations of compact groups are completely reducible.*

Proof. This is the usual argument: if there is an irreducible subspace, then the orthogonal complement is irreducible, etc. □

Henceforth, we consider only representations over $\mathbb{C}$; this can be used to deduce the desired results for $\mathbb{R}$.

Lemma 6.3.4 (Schur)**.** *Let (π_1, V_1) and (π_2, V_2) be finite-dimensional, irreducible representations. Let $T : V_1 \to V_2$ be a G-invariant linear map. Then*

(a) $T = 0$ unless $(\pi_1, V_1) \overset{T}{\simeq} (\pi_2, V_2)$, and

(b) if $(\pi_1, V_1) = (\pi_2, V_2)$, then T is a multiple of the identity.

Proof. $\ker T$ and $\operatorname{Im} T$ are invariant subspace, hence either 0 or the whole space, which implies (a). If $(\pi_1, V_1) = (\pi_2, V_2)$, then every eigenspace for T is G-invariant, which implies (b).

Remark 6.3.5. We used algebraic closedness of $\mathbb{C}$ to deduce that T has an eigenvalue. This certainly fails if the ground field is not algebraically closed.

□

Corollary 6.3.6. *With $(\pi_1, V_1) = (\pi_2, V_2)$ as before, let $h : V_1 \times V_2 \to \mathbb{C}$ be a Hermitian pairing. Then*

(a) $h = 0$ unless $(\pi_1, V_1) \simeq (\pi_2, V_2)$.

(b) If $(\pi_1, V_1) = (\pi_2, V_2)$, then h is a real multiple of a G-invariant inner product.

Proof. Let $(\cdot, \cdot)$ be a G-invariant (hermitian) inner product on V_1. Then $(\cdot, \cdot)$ defines a G-invariant *conjugate*-linear isomorphism $V_1 \simeq V_1^*$. Similarly, h defines a G-invariant conjugate-linear map $V_2 \to V_1^*$. Then the second linear map composed with the inverse of the first defines a G-invariant $\mathbb{C}$-linear map $V_2 \to V_1$, which by Schur's Lemma is either 0 or an isomorphism, in which case it must be multiplication by a constant, which must be real because h is Hermitian. □

Corollary 6.3.7. *If (π, V) is a finite-dimensional irreducible representation, then there exists only one G-invariant product, up to scaling.*

Now suppose (π, V) is a finite-dimensional irreducible representation with $(\cdot, \cdot)$ a G-invariant inner product. An inner product on V determines an inner product on V^*, by declaring the dual basis to an orthonormal basis of V to be an orthonormal basis of V^*. It is left as an exercise to show that this doesn't depend on the choice of orthonormal basis for V. The inner product on V^* induced in this way is G-invariant by construction. The two inner products scale inverse proportionally, so there exists a distinguished inner product on $V \otimes V^*$, which will clearly be G-invariant, which is invariant under scaling the inner product on V.

We have $V \otimes V^* \simeq \operatorname{End}(V)$, defined by sending $u \otimes u^* \mapsto T_{u \otimes u^*}$, where

$$T_{u \otimes u^*}(v) = \langle u^*, v \rangle u.$$

Of course, one has to ensure that this is well-defined; this follows from observing that this map is bilinear in u and u^*. So we can and shall describe the distinguished inner product on $\operatorname{End}(V)$, rather than $V \otimes V^*$. Under these identifications, we define the *Hilbert-Schmidt inner product*

$$(T_1, T_2)_{\mathrm{HS}} = \operatorname{tr}(T_2^* T_1).$$

In particular, $(T, T)_{\mathrm{HS}}$ is the sum of the squares of the absolute values of the matrix entries of T.

Let $\widehat{G}$ be the set of isomorphism classes of finite-dimensional, irreducible, unitary representations of G. (Note that the qualifier "unitary," is somewhat extraneous, since we can make any irreducible representation unitary in a unique way up to scaling. It is much less trivial that irreducible and unitary imply finite-dimensionality, and this depends on compactness of G). For each $i \in \widehat{G}$, choose a concrete representative (π_i, V).

Choose an inner product which rescales the canonical Hilbert-Schmidt inner product on $V \otimes V^*$ by $\frac{1}{\dim V}$, so that

$$(1_V, 1_V) = \frac{1}{\dim V}(1_V, 1_V)_{\mathrm{HS}} = 1.$$

Define $\Phi_i : V_i \otimes V_i^* \to C(G)$ by

$$\Phi_i(v \otimes v^*)(g) = \langle v^*, \pi_i(g^{-1})v\rangle$$

Note that Φ_i is $G \times G$-invariant with respect to $\pi_i \otimes \pi_i^*$ on $V_i \otimes V_i^*$ and $\ell \times r$ on $C(G)$ because

$$\begin{aligned}
(\ell(g_1)r(g_2)\Phi(v \otimes v^*))(g) &= \Phi(v \otimes v^*)(g_1^{-1}gg_2) \\
&= \langle v^*, \pi_i(g_2^{-1}g^{-1}g_1)v\rangle \\
&= \langle v^*, \pi_i(g_2)^{-1}\pi_i(g^{-1})\pi_i(g_1)v\rangle \\
&= \langle \pi_i^*(g_2)v^*, \pi_i(g^{-1})\pi_i(g_1)v\rangle \\
&= \Phi_i((\pi_i(g_1) \otimes \pi_i^*(g_2))(v \otimes v^*))(g)
\end{aligned}$$

Proposition 6.3.8 (Schur orthogonality relations). *Keeping the notation above,*

(i) $\operatorname{Im} \Phi_i \perp \operatorname{Im} \Phi_j$ *in* $L^2(G)$ *if* $i \neq j$.

(ii) $\Phi_i : V_i \otimes V_i^* \to C(G) \to L^2(G)$ *is an isometry.*

Proof. Define $S : V_1 \times V_1^* \times V_2 \times V_2^* \to \mathbb{C}$ by

$$S(v_1, v_1^*, v_2, v_2^*) = (\Phi_i(v_1 \otimes v_1^*), \Phi_i(v_2 \otimes v_2^*))_{L^2(G)}.$$

Note that S is

- quadri-linear over $\mathbb{R}$,
- $\mathbb{C}$-linear in the first and last variables (basically because it is conjugated twice)
- conjugate-linear in the second and third variables,
- and G-invariant.

Then for fixed v_1^*, v_2^*, this becomes a G-invariant Hermitian pairing from $V_1 \times V_2 \to \mathbb{C}$. By the first corollary to Schur's Lemma, this implies (i).

Now suppose that $V_i = V_j$ (we've been operating under the notation $i = 1, j = 2$). Then

$$S(v_1, v_1^*, v_2, v_2^*) = h(v_2^*, v_1^*)(v_1, v_2).$$

where $h : V_i^* \times V_i^* \to \mathbb{C}$ is a Hermitian pairing, hence some (real) constant multiple of any one. But then $h(v_2^*, v_1^*) = \lambda(v_2^*, v_1^*)_{V_i^*}$ for some $\lambda \in \mathbb{R}$. We must show that $\lambda = \frac{1}{\dim V_i}$. The point is that we want to show that S is the product of the inner products on V_i and V_i^*, which is the Hilbert-Schmidt inner product.

To establish this, choose an orthonormal basis $\{v_\alpha\}$ of V_i and the dual basis $\{v_\beta^*\}$ of V_i^*. Then $\{v_\alpha \otimes v_\beta^*\}$ is an orthonormal basis for $(\cdot,\cdot)_{\mathrm{HS}}$ on $V_i \otimes V_i^*$. Note that as linear functions on V_i,

$$v \mapsto \langle v_\beta^*, v\rangle$$

coincides with $v \mapsto (v, v_\beta)$, by definition of the dual basis. Now we calculate

$$\begin{aligned}
\sum_{\alpha,\beta} (\Phi_i(v_\alpha \otimes v_\beta^*), \Phi_i(v_\alpha \otimes v_\beta^*))_{L^2(G)} &= \sum_{\alpha,\beta} \int_G \langle v_\beta^*, \pi_i(g^{-1})v_\alpha\rangle \overline{\langle v_\beta^*, \pi_i(g^{-1})v_\alpha\rangle}\, dg \\
&= \sum_{\alpha,\beta} \int_G (\pi_i(g^{-1})v_\alpha, v_\beta)(v_\beta, \pi_i(g^{-1})v_\alpha)\, dg \\
&= \sum_{\alpha} \int_G (\pi_i(g^{-1})v_\alpha, \pi_i(g^{-1})v_\alpha)\, dg \\
&= \dim V
\end{aligned}$$

On the other hand, the left hand side is

$$\sum_{\alpha,\beta} \lambda(v_\alpha, v_\alpha^*)(v_\beta, v_\beta^*) = \lambda(\dim V)^2.$$

This shows that $\lambda = \frac{1}{\dim V}$. □

6.4 The Peter-Weyl Theorem

We just saw that

$$\bigoplus_{i \in \widehat{G}} \Phi_i : \bigoplus_{i \in \widehat{G}} V_i \otimes V_i^* \to L^2(G)$$

is a $G \times G$-invariant (with respect to $\pi_i \otimes \pi_i^*$ and $\ell \times r$) isometry (with respect to the renormalized inner products on the $V_i \otimes V_i^*$). Hence we can take the completion:

$$\Phi = \widehat{\bigoplus_{i \in \widehat{G}}} \Phi_i : \widehat{\bigoplus}_{i \in \widehat{G}} V_i \otimes V_i^* \to L^2(G).$$

Theorem 6.4.1 (Peter-Weyl). *This is a Hilbert space isomorphism of unitary $G \times G$ representations:*

$$L^2(G) \simeq \widehat{\bigoplus}_{i \in \widehat{G}} V_i \otimes V_i^*$$

Example 6.4.2. When G is compact and abelian, its irreducible representations are all one-dimensional. Then

$$\widehat{G} \simeq \mathrm{Hom}(G, \mathbb{C}^*).$$

Then the Peter-Weyl Theorem tells us that

$$L^2(G) \simeq \widehat{\bigoplus}_{\chi_i \in \mathrm{Hom}(G, \mathbb{C}^*)} \mathbb{C}\chi_i$$

For $i \in \widehat{G}$, we had defined

$$\Phi_i : V_i \otimes V_i^* \to C(G) \subset L^2(G)$$

and taking the completion,

$$\widehat{\Phi} = \widehat{\bigoplus_{i \in \widehat{G}}} \Phi_i : \widehat{\bigoplus_{i \in \widehat{G}}} V_i \otimes V_i^* \to L^2(G).$$

By Schur orthogonality, this is a $G \times G$-invariant isometry. In fact, the Peter-Weyl theorem asserts that it is an isomorphism.

Proof. We only need to show that $A = \Phi(\bigoplus_{i \in \widehat{G}} V_i \otimes V_i^*)$ is dense in $L^2(G)$, or even dense in $C(G)$ in the sup norm (since $C(G)$ is dense in $L^2(G)$).

Observe that Im $\Phi_i \subset C(G)$ is the the linear span of matrix coefficients of $\pi_i^*(g)$, or equivalently the linear span of matrix coefficients of $\pi_i(g^{-1})$, or equivalently the linear span of the complex conjugates of $\pi_i(g)$. Why? Let $\{v_\alpha\}$ be an orthonormal basis of V_i and $\{v_\beta^*\}$ be the dual basis. Then $\operatorname{im} \Phi_i$ is the linear span of

$$\begin{aligned} g \mapsto \langle v_\beta^*, \pi_i(g^{-1}) v_\alpha \rangle &= \langle \pi_i^*(g) v_\beta^*, v_\alpha \rangle \\ &= (\pi_i(g^{-1}) v_\alpha, v_\beta) \\ &= (v_\alpha, \pi_i(g) v_\beta) \\ &= \overline{(\pi_i(g) v_\beta, v_\alpha)}. \end{aligned}$$

(The $\langle \cdot, \cdot \rangle$ is the dual pairing and $(\cdot, \cdot)$ is the inner product on V.)

Definition 6.4.3. Suppose (π, V) is an infinite-dimensional representation. A vector $v \in V$ is said to be *G-finite* if it is contained in a finite dimensional G-invariant subspace.

Proposition 6.4.4. *For $f \in L^2(G)$, the following are equivalent.*

(a) $f \in A$,

(b) f is $G \times G$-finite with respect to $\ell \times r$,

(c) f is G-finite with respect to ℓ,

(d) f is G-finite with respect to r,

(e) f is a linear combination of matrix coefficients of finite-dimensional representations.

Proof. $(a) \implies (b)$: By definition, the image of $V_i \otimes V_i^*$ under Φ is $G \times G$-invariant, hence if $f \in A$ then f is a finite linear combination of $G \times G$-invariant finite-dimensional subspaces, hence is $G \times G$ invariant.

Clearly (b) implies (c) and (d).

$(e) \implies (a)$: since finite-dimensional representations are completely reducible, and the matrix coefficients of a direct sum are sums of matrix coefficients, f is a linear combination of matrix coefficients of *irreducible* finite-dimensional representations.

The remaining directions are (c) implies (e) and (d) implies (e), which are essentially similar.

We now set up more preliminaries for the proof. Let (π, V) be a representation of G on a Hilbert or Banach space. Then for $\varphi \in C(G)$, define $\pi(\varphi) \in \mathrm{End}(V)$ by

$$\pi(\varphi)v = \int_G \varphi(g)\pi(g)v\, dg,$$

noting that $\varphi(g)\pi(g)$ is a continuous V-valued function. This is bounded because $\pi : G \times V \to V$ is continuous and G is compact, hence $\sup ||\pi|| < \infty$. So

$$||\pi(\varphi)|| \leq \sup |\varphi(g)| \sup ||\pi(g)||.$$

Now apply this for $(\pi, V) = (\ell, L^2(G))$. Then

$$\ell(\varphi)f = \int_G \varphi(g)\ell(g)f dg \qquad \varphi \in C(G), f \in L^2(G)$$

so

$$(\ell(\varphi)f)(g) = \int_G \varphi(h)f(h^{-1}g)\, dh.$$

Notation. Denote by $\varphi * f$ the function whose value at g is

$$\int_G \varphi(h)f(h^{-1}g)\, dh = \int_G \varphi(gh)f(h^{-1})dh.$$

This is called *convolution* (notice the similarity to the Fourier transform!). Note that

$$\begin{aligned}(r(\varphi)f)(g) &= \left(\int_G \varphi(h)r(h)f\,dh\right)(g)\\ &= \int_G \varphi(h)f(gh)\,dh\\ &= \int_G f(gh)\check{\varphi}(h^{-1})\,dh\\ &= (f * \check{\varphi})(g).\end{aligned}$$

where $\check{\varphi}(g) = \varphi(g^{-1})$. Then

$$\varphi * f(g) - \varphi * f(g_1) = \int_G (\varphi(gh) - \varphi(g_1 h))f(h^{-1})\,dh.$$

We conclude that

1. $|\varphi * f| \leq \sup|\varphi(g)| \cdot ||f||_2$.
2. $\varphi * f$ is continuous for $\varphi \in C(G)$ and $f \in L^2(G)$.
3. $\sup|\varphi * f(g)| \leq \sup|\varphi| \cdot ||f||_2$.
4. The "modulus of continuity" of $\varphi * f$ depends only on φ and $||f||_2$.

Why? (1) is clear. (2) For g close to g_1, $\varphi(gh) - \varphi(g_1 h)$ is small; but since G is compact, this can be made uniformly small. (3) is also clear from the formula, and (4) is similar to (2). (Note that this is true even for the L^1 norm on f)

Of course, all of this holds equally well for $f * \varphi$.

Now we can finally prove that (c) $\implies$ (e). Let $V \subset L^2(G)$ be a finite-dimensional $\ell(G)$-invariant subspace. We want to show that V is the linear span of matrix coefficients of finite-dimensional representations. Without loss of generality, we may suppose that V is $\ell(G)$-irreducible. Suppose $f \in V$, and let $\{f_k\}$ be an orthonormal basis for V and $\{f_k^*\}$ the dual basis.

$$\begin{aligned}\ell(g^{-1})f &= \sum_i (\ell(g^{-1})f, f_i)f_i\\ &= \sum_{i,j}(f, f_j)(\ell(g^{-1})f_j, f_i)f_i\\ &= \sum_{i,j}(f, f_j)\langle f_i^*, \ell(g^{-1})f_j\rangle f_i\\ &= \sum_{i,j}(f, f_j)\underbrace{\langle \ell(g)^* f_i^*, f_j\rangle}_{\text{matrix coeff of } (\ell^*, V^*)} f_i\end{aligned}$$

We claim that $V \subset C(G)$. Assuming this for the moment, we may evaluate the preceding identity at e to find that

$$f(g) = \sum_{i,j} (f, f_j) f_i(e) \langle \ldots \rangle.$$

This expresses f as a finite linear combination of matrix coefficients, as desired.

Now it suffices to prove the claim. The actions ℓ and r commute, so if $\varphi \in C(G)$, then

$$r(\varphi)V \subset V.$$

So $r(\varphi)V$ is $\ell(G)$-invariant and consists of continuous functions (a convolution with continuous functions is coninuous). If the claim were false, we would have to conclude (by irreducibility) that for every $\varphi \in C(G)$ and every $r(\varphi)f = f * \check{\varphi} \equiv 0$. But

$$r(\varphi)f(e) = \int f(h)\varphi(h^{-1})\, dh = (f, \check{\varphi})$$

so $V \perp \{\check{\varphi} \mid \varphi \in C(G)\} = C(G)$. But $C(G)$ is dense in L^2, which is a contradiction. $\square$

Now to prove the Peter-Weyl theorem, it suffices to show that A is dense in $C(G)$, because $C(G)$ is dense in $L^2(G)$. Because of the proposition, it suffices to show that the linear span of finite dimensional ℓ-invariant subspaces of $L^2(G)$ is dense in $C(G)$.

Lemma 6.4.5. *For $\varphi \in C(G)$, $r(\varphi) \in End(L^2(G))$ is a compact operator, i.e. the $r(\varphi)$ image of any bounded set has compact closure in $L^2(G)$. Furthermore, if $\check{\varphi} = \varphi$, then $r(\varphi)$ is self-adjoint, i.e.*

$$(r(\varphi)f_1, f_2) = (f_1, r(\varphi)f_2).$$

Proof. Suppose $F \subset L^2(G)$ is a bounded set. Then we know that the set of functions

$$\{r(\varphi)f \mid f \in F\}$$

lies in $C(G)$, is bounded, and is equicontinuous. By Ascoli's theorem, this set has a compact closure in $C(G)$. Since $C(G)$ is dense in $L^2(G)$, it has compact closure in $L^2(G)$.

For $f_1, f_2 \in L^2(G)$ and $\varphi \in C(G)$,

$$\begin{aligned}
(r(\varphi)f_1, f_2) &= \int_{G\times G} f_1(gh)\overline{f_2(g)}\varphi(h)\, dg\, dh \\
&= \int_{G\times G} f_1(g)\overline{f_2(gh^{-1})}\varphi(h)\, dgdh \\
&= \int_{G\times G} f_1(g)\overline{f_2(gh)}\check{\varphi}(h^{-1})\, dgdh \\
&= (f_1, r(\check{\varphi})f_2)
\end{aligned}$$

□

Now suppose $\varphi \in C(G)$ is real-valued and $\varphi = \check{\varphi}$. Then $r(\varphi)$ is compact and self-adjoint. By the spectral theorem for such operators, the eigenvalues of $r(\varphi)$ are real and discrete except at 0 (i.e. if λ_n is a convergent sequence of distinct eigenvalues, then $\lambda_n \to 0$). Moreover, the eigenspaces corresponding to non-zero eigenvalues are finite-dimensional.

What do we know about Im $r(\varphi)$? We claim that Im $r(\varphi) \cap A$ is dense in Im $r(\varphi)$, because Im $r(\varphi) \cap \{0 - \text{eigenspace}\}^{\perp}$ is dense in Im $r(\varphi)$. Indeed, the eigenspaces are ℓ-invariant (because left and right translation commute), and finite-dimensional, hence lie in A.

Now let φ_n run through an approximate identity satisfying $\check{\varphi}_n = \varphi_n$ for all n, i.e. $\varphi_n \geq 0$ and $\int_G \varphi_n(g)\, dg = 1$ and such that supp $\varphi_n \downarrow \{e\}$. (If G doesn't satisfy first countability, we need to use a net instead of a sequence). For any $f \in C(G)$,

$$r(\varphi_n) f \to f$$

(this is the point of "approximate identity.")

The conclusion is that any $f \in L^2(G)$ can be approximated arbitrarily well by $r(\varphi_n)\widetilde{F}$ with $\widetilde{F} \in C(G)$, therefore arbitrarily well by $r(\varphi_n)\widetilde{f}$, $\widetilde{f} \in L^2(G)$, hence by $\widetilde{f} \in A$. □

Chapter 7

Compact Lie groups

7.1 Compact tori

Suppose T is a connected, compact, abelian Lie group. By Campbell-Baker-Hausdorff, the map

$$\exp : \mathfrak{t} \to T$$

is a homomorphism of Lie groups. Let $L \subset \mathfrak{t}$ be the kernel. We know:

1. $L \subset \mathfrak{t}$ is closed and discrete,

2. $\mathfrak{t}/L$ is compact,

It is a fact from algebra that any finitely generated, torsion-free $\mathbb{Z}$-module is free, so $L \subset \mathfrak{t}$ is a lattice, i.e. the $\mathbb{Z}$-linear span of an $\mathbb{R}$-basis.

We quickly sketch the proof. Let $\{v_1, \ldots, v_r\} \subset L$ be a maximal $\mathbb{R}$-independent subset and L_0 be the $\mathbb{Z}$-linear span of $\{v_1, \ldots, v_r\}$. Then $\mathbb{R} \otimes_{\mathbb{Z}} L_0/L_0 \simeq \mathbb{R}^r/\mathbb{Z}^r$, and hence has a compact fundamental domain. By (i), $\{v_1, \ldots, v_r\}$ must be maximal among any $\mathbb{Z}$-linearly independent subsets, since the multiples of any other element would otherwise have an accumulation within the fundamental domain, and L/L_0 must be finite or else it would again have an accumulation point in the fundamental domain. Therefore, L has a $\mathbb{Z}$-basis and (ii) compactness of $\mathfrak{t}/L$ implies that the $\mathbb{Z}$-basis is an $\mathbb{R}$-basis.

The conclusion is that $T \simeq \mathbb{R}^n/\mathbb{Z}^n \simeq (\mathbb{R}/\mathbb{Z})^n$ is a *torus.*

By Schur's Lemma, all irreducible finite-dimensional representations of T are one-dimensional. Therefore,

$$\widehat{T} = \mathrm{Hom}(T, \mathbb{C}^*) = \mathrm{Hom}(T, S^1)$$

(since T is compact, its image is compact). This is called the *character group.*

Remark 7.1.1. In the context of abelian groups, "character" normally means "homomorphism to $\mathbb{C}^*$" as equivalently "one-dimensional representation." This is not so for non-abelian groups, as will be explored in the homework exercises.

7.2 Weight decomposition

Let (π, V) be a finite-dimensional representation of T. Then by complete reducibility, π is a direct sum of one-dimensional representations. We know $T \simeq \mathfrak{t}/L$, so we can lift $\pi\colon T \to S^1$ to a homomorphism $\mathfrak{t} \to S^1$, which must be purely imaginary in order to take values in S^1. That implies that

$$\widehat{T} \simeq \Lambda = \{\lambda \in i\mathfrak{t}^* \mid \langle \lambda, L\rangle \subset 2\pi i\mathbb{Z}\}.$$

This identification sends λ to e^λ, where

$$e^\lambda(\exp X) = e^{\langle \lambda, X\rangle}.$$

By complete reducibility, we have a decomposition

$$V = \bigoplus_{\lambda\in\Lambda} V^\lambda$$

where

$$\begin{aligned} V^\lambda &= \{v \in V \mid \pi(t)v = e^\lambda(t)v \text{ for all } t \in T\} \\ &= \{v \in V \mid \pi_*(X)v = \langle \lambda, X\rangle v \text{ for all } X \in \mathfrak{t}\}. \end{aligned}$$

We say that

- λ is a *weight* of π if $V^\lambda \neq 0$, and
- $\dim V^\lambda$ is the *multiplicity* of the weight, and
- V^λ is the λ-*weight space.*

Observation: if (π_1, V_1) and (π_2, V_2) are two finite-dimensional representations of T and $\lambda \in \Lambda$, then

$$(V_1 \oplus V_2)^\lambda = V_1^\lambda \oplus V_2^\lambda$$

and

$$(V_1 \otimes V_2)^\lambda = \bigoplus_{\mu\in\Lambda} V_1^\mu \otimes V_2^{\lambda-\mu}$$

and

$$(V_1^*)^\lambda = (V_1^{-\lambda})^*.$$

Proposition 7.2.1. *Suppose that G is a connected, compact Lie group and $T \subset G$ is a connected Lie subgroup with Lie algebra $\mathfrak{t}$. Then the following are equivalent:*

1. *T is a torus, and is maximal among tori in G,*
2. *T is abelian, and is maximal among connected abelian Lie subgroups,*

3. $\mathfrak{t}$ *is abelian, and is maximal among abelian subalgebras.*

Proof. (1) $\implies$ (2). Suppose T is a maximal torus and A is connected and abelian, $T \subset A$. Then the closure of A is abelian, closed (hence compact), and connected, therefore a torus. By maximality, $A = T$.

(2) is equivalent to (3) because of the connection between connected Lie subgroups and Lie subalgebras, and abelian Lie groups and abelian Lie algebras.

It is clear that (2) implies (1). □

In particular, if $G \neq \{e\}$, then G must contain a maximal torus T of positive dimension.

Our general strategy for analyzing Lie groups is to understand the maximal torus. Being abelian, they are easy to understand, and they turn out to be quite large in G.

Consider a particular compact, connected Lie group G. Fix $T \subset G$ a maximal torus. We shall see later that T is unique up to conjugacy. We note a systematic change of notation: for any Lie groups $G, H, \ldots$ we shall denote by $\mathfrak{g}_{\mathbb{R}}, \mathfrak{h}_{\mathbb{R}}$ the Lie algebras of $G, H, \ldots$ and $\mathfrak{g} = \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{g}_{\mathbb{R}}, \mathfrak{h} = \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{h}_{\mathbb{R}}$, i.e. the complexifications.

Let $T \simeq \mathfrak{t}_{\mathbb{R}}/L$,

$$\widehat{T} = \Lambda = \{\lambda \in i\mathfrak{t}_{\mathbb{R}}^* \mid \langle \lambda, L \rangle \subset 2\pi i \mathbb{Z}\}.$$

L is called the "unit lattice" and Λ the "weight lattice." Moreover,T acts on $\mathfrak{g}$ Ad.

Definition 7.2.2. A *root* of $(\mathfrak{g}, \mathfrak{t})$ is a non-zero weight of T acting on $\mathfrak{g}$.

Let $\Phi \subset \Lambda - \{0\}$ be the set of roots (it is finite). We write $\Phi = \Phi(\mathfrak{g}, \mathfrak{t})$ when there is any ambiguity.

Proposition 7.2.3. *With the notation above,*

$$\mathfrak{g} = \mathfrak{t} \oplus \left(\bigoplus_{\alpha \in \Phi} \mathfrak{g}^{\alpha} \right).$$

Proof. We know that $\mathfrak{g} = \mathfrak{g}^0 \oplus \left(\bigoplus_{\alpha\in\Phi} \mathfrak{g}^{\alpha} \right)$. Clearly $\mathfrak{t} \subset \mathfrak{g}^0$ because T is abelian, hence acts trivially on its own Lie algebra. Since $\Lambda \subset i t_{\mathbb{R}}^*$,

$$\overline{\mathfrak{g}^{\alpha}} = \mathfrak{g}^{-\alpha},$$

and in particular $\overline{\mathfrak{g}^0} = \mathfrak{g}^0$ (complex conjugation with respect to $\mathfrak{g}_{\mathbb{R}} \subset \mathfrak{g}$). Hence if $t \neq \mathfrak{g}^0$, there must exist $X \in \mathfrak{g}^0 \cap \mathfrak{g}_{\mathbb{R}}$ not in $\mathfrak{t}_{\mathbb{R}}$. Then $t_{\mathbb{R}} \oplus \mathbb{R}X$ is abelian, contradicting maximality. □

Corollary 7.2.4. $\overline{\mathfrak{g}^{\alpha}} = \mathfrak{g}^{-\alpha}$ *and* $\Phi = -\Phi$.

We shall say that g^{α} is the *α-root space.*

Remark 7.2.5. We shall not talk about the "multiplicity of a root" because we shall see later that $\mathfrak{g}^{\alpha} = 1$ for all $\alpha \in \Phi$.

Let (π, V) be a finite-dimensional representation of G. Then, when we consider the action of T on V,

$$V = \bigoplus_{\mu \in \Lambda} V^{\mu}.$$

Observe that the action map $\mathfrak{g} \times V \to V$ is G-invariant (with respect to $\mathrm{Ad} \times \pi$ on the left and π on the right). Hence, if $\alpha \in \Phi$ is a weight, then

$$\pi_*(\mathfrak{g}^\alpha)V^\mu \begin{cases} \subset V^{\mu+\alpha} & \mu+\alpha \text{ a weight} \\ = 0 & \text{otherwise} \end{cases}$$

Proposition 7.2.6. *Suppose $\alpha, \beta \in \Phi$. Then*

$$[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] \begin{cases} \subset \mathfrak{g}^{\alpha+\beta} & \alpha+\beta \in \Phi \\ \subset \mathfrak{t} & \alpha+\beta = 0 \\ 0 & \textit{otherwise} \end{cases}.$$

Proof. This follows from the Jacobi identity:

$$[t, [\mathfrak{g}^\alpha, \mathfrak{g}^\beta]] + [\mathfrak{g}^\alpha, [\mathfrak{g}^\beta, t]] + [\mathfrak{g}^\beta, [t, \mathfrak{g}^\alpha]] = 0.$$

□

Definition 7.2.7. We define $W = N_G(T)/T$ to be the *Weyl group.*

We will use $W = W(G, T)$ when there is any ambiguity. This is a *finite* group, which acts via conjugation and Ad on $T, \mathfrak{t}_{\mathbb{R}}, \mathfrak{t}, L, \Lambda, \Phi$. Why finite? Note that $N_G(T)$ acts continuously on L, which is discrete, so $N_G(T)^0$ acts trivially. Therefore, $N_G(T)^0 \subset Z(T)^0 = T$ because T is maximal among connected abelian subgroups. So

$$W = N_G(T)/T = N_G(T)/N_G(T)^0$$

is 0-dimensional and compact, hence finite.

7.3 The maximal torus

Lemma 7.3.1. *Suppose A is a compact abelian Lie group such that A/A^0 is cyclic (A^0 is a torus). Then A is "singly topologically generated", i.e. there exists $a \in A$ such that*

$$\{a^n \mid n \in \mathbb{Z}\}$$

is dense in A.

Proof. We know $A^0 \simeq \mathbb{R}^n/\mathbb{Z}^n$. Let

$$p : \mathbb{R}^n \to \mathbb{R}^n/\mathbb{Z}^n = A^0$$

be the projection, $\{U_k \mid k \in \mathbb{N}\}$ a countable neighborhood base for A^0. We claim that there exists a sequence

$$Q_0 \supset Q_1 \supset Q_2 \supset \dots$$

of compact n-dimensional rectangles with non-empty interior and integers $m_k, k \geq 0$ such that

$$p(m_k Q_k) \subset U_k.$$

Suppose this for the moment. Then by compactness, there exists $a_0 \in \bigcap p(Q_k)$. For each k, $a_0^{m_k} \in U_k$, so the cyclic group generated by a_0 meets every U_k so its closure is A_0.

Choose a generator for A/A_0, $a_1 \in A$, and we may assume that $a_1^m \in A^0$ for some $m > 0$. Then $a_1^m a_0^{-1} \in A_0 = A_0^m$. So

$$a_0 = a_1^m a_2^m \qquad a_2 \in A_0,$$

and we may replace a_1 by $a_1 a_2$ and thereby arrange that $a_1^m = a_0$. Then a_1 is a topological generator for A, since it generates a_0 and meets every connected component.

It only remains to prove the claim. We proceed by induction. Choose $Q_0 \subset [0,1]^n$ so that $p(Q_0) \subset U_0$ and set $m_0 = 1$. For the inductive step, suppose $Q_0, \dots, Q_{k-1}$ and $m_0, \dots, m_{k-1}$ have been so chosen. There exists $m_k \in \mathbb{Z}$ and $\ell \in \mathbb{Z}^n$ such that

$$m_k Q_{k-1} + \ell \supset [0,1]^n.$$

Then we can choose Q_k as required so that its projection lies inside U_k. $\square$

Let G be a connected compact Lie group, $T \subset G$ the maximal torus. We know that

$$T_{eT}(G/T) \simeq \mathfrak{g}_\mathbb{R}/\mathfrak{t}_\mathbb{R}$$

We saw that

$$\mathfrak{g} \simeq \mathfrak{t} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}^\alpha,$$

and $\overline{\mathfrak{g}^\alpha} = \mathfrak{g}^{-\alpha}$, so $\dim_\mathbb{C} \mathfrak{g}/\mathfrak{t}$ is even (the representations come in pairs!). We conclude that G/T is a connected, compact, even dimensional manifold. We claim that it is also orientable. Why? We can choose an orientation at the identity coset and propagate it around using the action of G. However, there is an ambiguity which comes from the action of T itself on G/T; since T is connected and this is orientation-preserving at e, it must be so everywhere.

Suppose $g \in G$. We get an action

$$\ell(g) : G/T \to G/T$$

Observe that $\ell(g)$ has a fixed point if and only if g is conjugate to an element of T. This is just algebra:

$$\ell(g)g_1T = g_1T \iff gg_1T = g_1T \iff gg_1 \in g_1T \iff g \in g_1Tg_1^{-1}.$$

To prove the existence of fixed points, one can use the Lefschetz fixed point formula.

Lefschetz fixed point formula. Suppose X is an oriented, connected, compact manifold and $F : X \to X$ is a smooth map. The Lefschetz number of F is defined to be

$$L(F) = \sum_{p=0}^{\dim X} \operatorname{tr}(F^* : H^p(X, \mathbb{R}) \to H^p(X, \mathbb{R})).$$

This is a homotopy invariant and $L(1_\chi) = \chi(X)$.

Theorem 7.3.2 (Lefschetz)**.** *Suppose F has only isolated many fixed points. Then one can attach integers n_j to each x_j, namely the intersection number of the graph of F with the diagonal in $X \times X$ at (x_j, x_j) such that*

$$L(F) = \sum n_j.$$

The intersection numbers are defined as follows: near each intersection, the manifolds can be deformed to be transverse. Then the intersection number is just defined to be ± 1 according to the orientations.

If the intersection is already known to be transverse,

$$\det(1 - F_*) : T_{x_j}X \to T_{x_j}X$$

is non-zero, and $n_j = \operatorname{sign}\det(1 - F_*)$. If F has no fixed points, then the Lefschetz number is 0.

Theorem 7.3.3. *For every $g \in G$,*

$$\ell_g : G/T \to G/T$$

has a fixed point. Moreover, $L(\ell(g)) = \#W$.

Corollary 7.3.4. *We have the following.*

1. *Every conjugacy class meets T, i.e. every $g \in G$ is conjugate to some $t \in T$.*

2. *Any two maximal tori are conjugate.*

3. $\exp : \mathfrak{g}_{\mathbb{R}} \to G$ *is surjective.*

4. *Let* $S \subset G$ *be a torus,* $g \in Z_G(S)$*. Then there exists a maximal torus in* G *which contains both* S *and* g*.*

5. *With* S *as before,* $Z_G(S)$ *is connected, and is the union of the maximal tori which contain* S*.*

6. *Every maximal torus is maximal among abelian subgroups of* G *(but not every maximal abelian subgroup is a torus).*

Proof. (1) Follows from the observation that g is conjugate to an element of T if and only if ℓ_g has a fixed point.

(2) Follows from the existence of a topological generator in tori. A topological generator for T_1 is conjugate to an element of T, which must be a topological generator.

(3) This follows from (1), the map Ad, and the corresponding statement for tori.

(4) Let $\widetilde{S}$ be the smallest closed Lie subgroup of G containing S and g. Then $\widetilde{S}$ is abelian, since the smallest group containing S and g is abelian, hence its closure is as well. Then $\widetilde{S}^0$ is a torus and $\widetilde{S}/\widetilde{S}^0$ is generated by g. $\widetilde{S}$ has a topological generator by the lemma proved last time. By (1), it is conjugate to an element of a maximal torus.

(5) Clear from (4).

□

Proof of Theorem 7.3.3. The fact that G is connected implies that $L(\ell(g))$ is independent of the choice of g, and therefore must equal $L(1_{G/T}) = \chi(G/T)$, which is non-zero. Therefore, it suffices to find some $g \in G$ such that

$$L(\ell(g)) = \#W.$$

Fix a topological generator $t_0 \in T$. Then $gT \in G_0/T$ is a fixed point of $\ell(t_0)$, if and only if $t_0 \in gTg^{-1} \iff T = gTg^{-1} \iff g \in N_G(T)$. So

$$W = N_G(T)/N_G(T)^0 = N_G(T)/T \subset G/T$$

is the set of fixed points of $\ell(t_0)$.

However, we are not done yet because we must check that the local multiplicities are 1. Recall that in the case of a transverse intersections, these should be $\operatorname{sign}(1 - \det \operatorname{Ad} t_0)$.

Observe that $W = N_G(T)/T$ acts on G/T by right translation: for $u \in N_G(T)$,

$$r(u)gT = gTu^{-1} = gu^{-1}uTu^{-1} = gu^{-1}T.$$

This action commutes with left translation, in particular with $\ell(t_0)$. So we have the commutative diagram

$$\begin{tikzcd}
T_{uT}G/T \arrow[r, "\ell(t_0)_*"] \arrow[d, "r(u)_*"] & T_{uT}G/T \arrow[d, "r(u)_*"] \\
T_{eT}G/T \arrow[r, "\ell(t_0)_*"] & T_{eT}G/T
\end{tikzcd}$$

So all the local Lefschetz numbers of $\ell(t_0)$ agree. At the identity coset,

$$T_{eT}G/T \simeq \mathfrak{g}/\mathfrak{t} \simeq \bigoplus_{\alpha\in\Phi} \mathfrak{g}^\alpha.$$

By naturality, $\ell(t_0)_*$ corresponds to $\mathrm{Ad}(t_0)$ because $\ell(t_0)gT = t_0gT = t_0gt_0^{-1}T$, so if $g = \exp X$ then we have

$$\ell(t_0)(\exp X)T = \exp(\mathrm{Ad}(t_0)X)T.$$

Then

$$\det(1 - \mathrm{Ad}(t_0)) : \mathfrak{g}/\mathfrak{t} \to \mathfrak{g}/\mathfrak{t} = \prod_{\alpha\in\Phi} (1 - e^\alpha(t_0))^{\dim \mathfrak{g}_\alpha}.$$

We shall soon see that $\dim \mathfrak{g}_\alpha = 1$, but it is not necessary right now. Since things occur in complex conjugates, this is

$$\prod_{\alpha\in\Phi} |1 - e^\alpha(t_0)|^{\dim \mathfrak{g}_\alpha}$$

If $e^\alpha(t_0) = 1$, then $t_0 \in \ker e^\alpha$ hence cannot generate T, so each term in the product above is 1. $\square$

7.4 Trace forms

Let G be a connected compact Lie group, T a maximal torus. We now know that T is unique up to conjugacy.

Definition 7.4.1. A symmetric bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ is called *Ad-invariant* if

$$B(\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y) = B(X, Y)$$

for all $g \in G$, $X, Y \in \mathfrak{g}$. On the infinitesimal level (i.e. set $g = \exp tZ$ and differentiate), this says

$$B([Z, X], Y) + B(X, [Z, Y]) = 0.$$

Since G is connected, the infinitesimal statement implies the global statement, and the converse is obvious.

Let (π, V) be a finite dimensional representation of G. Define

$$B_\pi : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$$
$$B_\pi(X, Y) = \operatorname{tr} \pi_*(X)\pi_*(Y)$$

(first for $X, Y \in \mathfrak{g}_\mathbb{R}$, then complexify).

Lemma 7.4.2. *B_π is symmetric, bilinear, Ad-invariant, and defined over $\mathbb{R}$ (i.e. real if X, Y are real), negative semi-definite, and the radical of $B_\pi = \ker \pi_*$ (the radical being the set of x such that $B(x, -) = 0$).*

Proof. Symmetric and bilinear are obvious. Suppose $X, Y \in \mathfrak{g}_\mathbb{R}$ and $g \in G$. Then

$$\begin{aligned} B_\pi(\operatorname{Ad}(g)X, \operatorname{Ad}(g)Y) &= \operatorname{tr} \pi_*(\operatorname{Ad}(g)X)\pi_*(\operatorname{Ad}(g)Y) \\ &= \operatorname{tr} \pi(g)\pi_*(X)\pi(g^{-1})\pi(g)\pi_*(Y)\pi(g^{-1}) \\ &= \operatorname{tr} \pi_*(X)\pi_*(Y). \end{aligned}$$

This shows Ad-invariance. Now suppose $X \in \mathfrak{g}_\mathbb{R}$. Then, up to conjugacy, $X \in \mathfrak{t}_\mathbb{R}$. The weight space decomposition diagonalizes the action of the maximal torus. So

$$B_\pi(X, X) = \operatorname{tr} \pi_*(X)^2 = \sum_{\mu \in \Lambda} \langle \mu, X \rangle^2 \dim V^\mu$$

But we know that $X \in \mathfrak{t}_\mathbb{R}$ and $\mu \in i\mathfrak{t}_\mathbb{R}^*$, so $\langle \mu, x \rangle \in i\mathbb{R}$. Hence $B_{\pi(X,X)} < 0$ unless $\langle \mu, X \rangle = 0$ for all weights μ of π i.e. unless $\pi_*(X) = 0$. So the radical of B_π is the (complexification of the) kernel of $\pi_* : \mathfrak{g}_\mathbb{R} \to \operatorname{End}(V)$, in other words the kernel of $\pi_* : \mathfrak{g} \to \operatorname{End}(V)$.

Finally, note that it takes real values on $\mathfrak{g}_\mathbb{R}$ by polarization. □

B_π is called the *trace form* of π. If $\pi = \operatorname{Ad}$, it is the *Killing form.* Note that

$$\bigcap_{i \in \widehat{G}} \ker \pi_i = \{e\}$$

by Peter-Weyl. Therefore, on the infinitesimal level,

$$\bigcap_{i \in \widehat{G}} \ker \pi_{i*} = \{0\}.$$

Therefore, there exists a finite-dimensional representation (π, V) such that $\ker \pi_* = \{0\}$. In other words, there exists a finite-dimensional (π, V) such that G acts faithfully. Fix $B = B_\pi$ for some such (π, V). We know that B is symmetric, bilinear, defined over $\mathbb{R}$, Ad-invariant, and negative definite on $i\mathfrak{g}_\mathbb{R}$. In particular, B induces a positive definite inner product $(,)_*$ on $i\mathfrak{t}_\mathbb{R}$, and by duality on $i\mathfrak{t}_\mathbb{R}^* \supset \Lambda \supset \Phi$. This inner product is W-invariant.

7.5 Representations of $SU(2)$

As usual, let G be a connected, compact Lie group and $T \subset G$ a maximal torus.

Definition 7.5.1. Suppose (π, V) is a finite-dimensional representation of G. Then the *character* of π is

$$\chi_\pi(g) = \operatorname{tr} \pi(g)$$

Note that χ_π is conjugation-invariant, since the trace is invariant under conjugation. The conjugates of T cover G, so π is determined up to isomorphism by $\chi_\pi|_T$.

Let the weight space decomposition for V be

$$V = \bigoplus_{\mu \in \Lambda} V^\mu$$

Then $\chi_\pi|_T = \sum_\mu \dim V^\mu e^\mu$. Hence the integers $\dim V^\mu$, for $\mu \in \Lambda$, determine π up to isomorphism.

We now study in depth the following special case: $G = SU(2)$, i.e. unitary 2×2 matrices of determinant 1.

$$G = \left\{ \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \mid |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

This is described by two complex parameters and one real condition, hence $\boxed{\dim G = 3.}$ What does this tell us about the maximal torus? It can't have dimension 3, since G is non-abelian. It can't have dimension 2, because G/T has even dimension. It has dimension at least 1, hence exactly 1. In fact, we have a representation

$$T = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \mid |\alpha| = 1 \right\}.$$

Topologically we have $SU(2) \approx S^3$ so $SU(2)$ is simply-connected. Any other Lie group with the same Lie algebra would be covered by $SU(2)$. The fundamental group of such a space would be a discrete normal subgroup, hence contained in the center of $SU(2)$, which one can easily check is $\{\pm 1\}$.

Note that $n = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ normalizes T, so

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

Since n determines a nontrivial element of the Weyl group which acts faithfully on $\mathfrak{t}_\mathbb{R} \simeq \mathbb{R}$, we know that $W = \mathbb{Z}/2\mathbb{Z}$ generated by $\operatorname{Ad} u$.

For classical groups, $\mathfrak{g}$ is the real Lie algebra. So

$$\mathfrak{su}(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & -a \end{pmatrix} \mid b \in \mathbb{C}, a \in i\mathbb{R} \right\}.$$

Hence the complexified Lie algebra consists of all 2×2 complex matrices of trace 0. A basis for this is

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Now, iH is "real" with respect to the structure of $\mathfrak{su}(2)$. Let $^-$ denote complex conjugation with respect to $\mathfrak{su}(2)$, so $\bar{H} = -H$, $\bar{E_+} = -E_-$. Finally,

$$\begin{aligned} [H, E_+] &= 2E_+ \\ [H, E_-] &= -2E_- \\ [E_+, E_-] &= H. \end{aligned}$$

Let

$$\mathfrak{t}_{\mathbb{R}} = \left\{ \begin{pmatrix} ix & 0 \\ 0 & -ix \end{pmatrix} \mid x \in \mathbb{R} \right\}$$

so

$$i\mathfrak{t}_{\mathbb{R}} = \left\{ \begin{pmatrix} x & 0 \\ 0 & -x \end{pmatrix} = xH \mid x \in \mathbb{R} \right\} \simeq \mathbb{R}.$$

Then the unit lattice (i.e. points whose exponential is trivial) is

$$L = \{2\pi i n H \mid n \in \mathbb{Z}\}$$

and $\Lambda \subset i\mathfrak{t}_{\mathbb{R}}^* \simeq \mathbb{R}$, so $\Lambda \simeq \mathbb{Z} \simeq \widehat{T}$. Under this identification,

$$n \mapsto \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \mapsto \alpha^n \right\}$$

When $1 \in \mathbb{Z} \simeq \Lambda$ is evaluated on H, we get 1. Since H acts on E_+ by 2 and E_- by -2, we see that the roots of $\mathfrak{su}_2$ are

$$\Phi(\subset \Lambda \simeq \mathbb{Z}) = \{\pm 2\}.$$

(Recall that Φ was the set of weights for the representation of T on $\mathfrak{g}$.) The finite dimensional representations of $SU(2)$ are completely reducible. For any irreducible representation (π, V), we have a decomposition

$$V = \bigoplus_{n \in \mathbb{Z}} V^n,$$

as representations of T. From now on, we shall denote the action of X in the Lie algebra on V not by $\pi_*(X)$, but simply by juxtaposition. Then for $v \in V^n$, $Hv = nv$. So

$$E_+ V^n \subset V^{n+2} \qquad E_- V^n \subset V^{n-2}$$

Now let n be the "highest weight" of π. Pick $v_n \in V^n$ non-zero, and define $v_{n-2k} = E_-^k v_n$. Then $v_{n-2k} \in V^{n-2k}$. We have

$$\begin{aligned} Hv_{n-2k} &= (n-2k)v_{n-2k} \\ E_- v_{n-2k} &= v_{n-2k-2}. \end{aligned}$$

What about $E_+E_-v_{n-2k}$?

Lemma 7.5.2.

$$E_+v_{n-2k} = k(n+1-k)v_{n-2k+2}.$$

Proof. Proceed by induction. We know that for $k=0$, $E_+v_n \in V^{n+2} = 0$. For the induction step,

$$\begin{aligned} E_+v_{n-2(k+1)} &= E_+E_-v_{n-2k} \\ &= (H + E_-E_+)v_{n-2k} \\ &= (k(n+1-k) + (n-2k))v_{n-2k} \\ &= (k+1)(n-k)v_{n-2k} \end{aligned}$$

□

Hence the linear span of v_{n-2k} is invariant under the action of $\mathfrak{su}(2)$, hence invariant under the action of $SU(2)$. Dropping the 0 terms, we must get a basis for V.

W acts on Λ, hence also on the weight space decomposition of any representation. We saw that the generator flips the sign, so $-n$ must be the *lowest* weight. The conclusion is that

$$\boxed{V \text{ has basis } \{v_{n-2k} \mid 0 \le k \le n\}.}$$

We have already determined the action of the complexified Lie algebra of $\mathfrak{su}(2)$. Since $SU(2)$ is simply connected, any finite-dimensional representation of the Lie algebra lifts to $SU(2)$.

Note that the raising/lifting identities also determine an irreducible representation of the Lie algebra. This proves:

Theorem 7.5.3. *For each $n \ge 0$, there exists an irreducible representation of $SU(2)$ of dimension $n+1$, unique up to isomorphism. It has weights $n, n-2, \dots, -n$. Furthermore, $E_+ : V^k \to V^{k+2}$ is an isomorphism unless $k = n$, in which case it is 0; and $E_- : V^k \to V_{k-2}$ is an isomorphism unless $k = -n$, in which case it is 0.*

The representation on $\mathbb{C}^2$ is just the obvious action of $SU(2)$ on $\mathbb{C}^2$. What about the three-dimensional representation? Representations are functorial; how can we go functorially from $\mathbb{C}^2$ to $\mathbb{C}^3$?

The answer is by taking the symmetric square! In fact, the irreducible representation of dimension $n+1$ is the n^{th} symmetric power of the 2-dimensional representation.

Corollary 7.5.4. *Let (π, V) be a finite-dimensional, possibly irreducible representation of $SU(2)$. Then all weights are integral, and*

$$E_+ V^{-n} \xrightarrow{1-1} V^{-n+2} \text{ for } n \geq 1$$
$$E_- V^{n} \xrightarrow{1-1} V^{n-2} \text{ for } n \geq 1.$$

For $v_0 \in V^0$, $E_+ v_0 = 0 \iff E_- v_0 = 0$.

Proof. Decompose V into a direct sum of irreducibles and apply the results above. $\square$

Chapter 8

Root systems

8.1 Structure of roots

Let G be a connected, compact Lie group and T its maximal torus. We had constructed a bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ which was symmetric, Ad-invariant, non-degenerate, and negative definite on $\mathfrak{g}_{\mathbb{R}}$ (it is the "trace form" associated to the adjoint representation).

By restriction of B to $i\mathfrak{t}_{\mathbb{R}}$ plus duality, we get a positive definite inner product $(\cdot,\cdot)$ on $i\mathfrak{t}_{\mathbb{R}}$ and on $i\mathfrak{t}_{\mathbb{R}}^*$ which are W-invariant.

Lemma 8.1.1. *Suppose $\alpha, \beta \in \Phi$. Then*

1. $B(\mathfrak{g}^\alpha, \mathfrak{g}^\beta) = 0$ *unless* $\alpha + \beta = 0$.
2. $B(\mathfrak{g}^\alpha, \mathfrak{t}) = 0$.
3. $B : \mathfrak{g}^\alpha \times \mathfrak{g}^{-\alpha} \to \mathbb{C}$ *is non-degenerate.*
4. *There exists* $\widetilde{H}_\alpha \in i\mathfrak{t}_{\mathbb{R}}$ *such that for any* $E_\alpha \in \mathfrak{g}^\alpha$,

$$[E_\alpha, E_{-\alpha}] = B(E_\alpha, E_{-\alpha})\widetilde{H}_\alpha$$

Proof. (1) B is Ad-invariant, but

$$B(\mathrm{Ad}g(\mathfrak{g}^\alpha), \mathrm{Ad}g(\mathfrak{g}^\beta)) = e^{\alpha+\beta} B(\mathfrak{g}^\alpha, \mathfrak{g}^\beta).$$

If $\alpha + \beta \neq 0$, we must get 0.

(2) is clear since we must have $\alpha = 0$, but $\mathfrak{g}^0 = t$ is abelian.

(1) and (2) and the non-degeneracy of B imply (3).

We know $[\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}] \subset \mathfrak{t}$. By non-degeneracy, $[E_\alpha, E_{-\alpha}]$ is completely determined by the action of B for $H \in \mathfrak{t}$,

$$B([E_\alpha, E_{-\alpha}], H) = B([H, E_\alpha], E_{-\alpha}) = \langle \alpha, H \rangle B(E_\alpha, E_{-\alpha}).$$

So $[E_\alpha, E_{-\alpha}] \in (\alpha^\perp)^\perp$, which is one-dimensional. Let's clarify what we mean here. The first $\perp$ picks out the hyperplane in $\mathfrak{t}$ on which $\alpha \in \mathfrak{t}^*$ vanishes. The second $\perp$ is with respect to the inner product induced by B on $\mathfrak{t}$. Hence $[E_\alpha, E_{-\alpha}]$ is completely determined by $B(E_\alpha, E_{-\alpha})$. □

Taking complex conjugates with respect to $\mathfrak{g}_\mathbb{R}$ sends $\mathfrak{g}^\alpha$ to $\mathfrak{g}^{-\alpha}$, hence we may replace $\overline{E_\alpha}$ by $E_{-\alpha}$, etc. to get

$$[E_{-\alpha}, E_\alpha] = B(E_\alpha, E_{-\alpha})\overline{\widetilde{H}_\alpha}.$$

That shows that $\widetilde{H}_\alpha = -\overline{\widetilde{H}_\alpha}$, i.e. $\widetilde{H}_\alpha \in i\mathfrak{t}_\mathbb{R}$ is purely imaginary.

Recall that B defined a negative-definite bilinear form on $\mathfrak{g}_\mathbb{R}$, hence an inner product on $i\mathfrak{g}_\mathbb{R}$. Define

$$\boxed{H_\alpha = \frac{2}{(\alpha,\alpha)}\widetilde{H}_\alpha.}$$

Choose $E_\alpha \in \mathfrak{g}^\alpha$ non-zero and define $E_{-\alpha} = -\overline{E}_\alpha$. B is negative definite on $\mathfrak{g}_\mathbb{R}$ so $B(X, \overline{X}) < 0$ for $X \in \mathfrak{g}$ non-zero. Therefore, $B(E_\alpha, E_{-\alpha}) > 0$, and in particular $\overline{E_\alpha}$ spans the line in $\mathfrak{g}_\alpha$ pairing non-degenerately with E_α.

Recale both E_α and $E_{-\alpha}$ by the same positive factor so that

$$\boxed{B(E_\alpha, E_{-\alpha}) = \frac{2}{(\alpha,\alpha)}.}$$

while preserving $E_{-\alpha} = -\overline{E}_\alpha$. We have the relations

$$[E_\alpha, E_{-\alpha}] = B(E_\alpha, E_{-\alpha})\widetilde{H}_\alpha.$$
$$[H_\alpha, E_\alpha] = \langle \alpha, H_\alpha\rangle E_\alpha$$

since $H_\alpha \in \mathfrak{t}_\mathbb{R}$. We want to compute this quantity $\langle \alpha, H_\alpha \rangle$, so pair with $E_{-\alpha}$.

$$\langle \alpha, H_\alpha\rangle B(E_\alpha, E_{-\alpha}) = B([H_\alpha, E_\alpha], E_{-\alpha}) = B([E_\alpha, E_{-\alpha}], H_\alpha)$$

Inserting the normalizations we made above, we conclude that

$$\langle \alpha, H_\alpha\rangle \frac{2}{(\alpha,\alpha)} = B(H_\alpha, H_\alpha)$$
$$\langle \frac{\alpha}{||\alpha||}, \frac{H_\alpha}{||H_\alpha||}\rangle = \frac{1}{2}||\alpha|| \cdot ||H_\alpha||.$$

By construction, $\frac{H_\alpha}{||H_\alpha||}$ is a unit vector in $(\alpha^\perp)^\perp$ (see above for a discussion of what this means). Therefore, we must have

$$\langle \frac{\alpha}{||\alpha||}, \frac{H_\alpha}{||H_\alpha||}\rangle = \pm 1.$$

From the equation we see that the sign must be $+1$, so

$$1 = \frac{1}{2}||\alpha||||H_\alpha||.$$

Therefore, we have

$$\boxed{\langle \alpha, H_\alpha \rangle = \frac{1}{2}\left(||\alpha||||H_\alpha||\right)^2 = 2.}$$

Summary. With our normalization, we have the relations

$$\begin{aligned}
[H_\alpha, E_\alpha] &= 2E_\alpha \\
[H_\alpha, E_{-\alpha}] &= -2E_{-\alpha} \\
[E_\alpha, E_{-\alpha}] &= H_\alpha
\end{aligned}$$

and $\overline{H}_\alpha = -H_\alpha$, $\overline{E}_\alpha = -E_{-\alpha}$.

Define $\varphi_\alpha : \mathfrak{sl}(2, \mathbb{C}) \to \mathfrak{g}$ by

$$\begin{aligned}
\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &\mapsto H_\alpha \\
\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} &\mapsto E_\alpha \\
\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} &\mapsto E_{-\alpha}
\end{aligned}$$

Then by the data in the conclusion, φ *is a Lie algebra homomorphism.* By construction, for $X \in \mathfrak{sl}(2, \mathbb{R})$ we have

$$\varphi(\overline{X}) = \overline{\varphi(X)},$$

where the complex conjugation is respect to $\mathfrak{su}(2)$ and $\mathfrak{g}_\mathbb{R}$, respectively.

Let $\mathfrak{s}_\alpha = \varphi_\alpha(\mathfrak{sl}(2, \mathbb{C}))$. We shall see later that the $\dim \mathfrak{g}^\alpha = 1$. Therefore, $\mathfrak{s}_\alpha = \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha} \oplus [\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}]$. This shows that everything depends only on α, not on the particular choice of E_α, etc.

Now, $SU(2)$ is simply connected with center $\{\pm 1\}$. Therefore, φ_α lifts to a Lie group homomorphism

$$\widetilde{\varphi_\alpha} : SU(2) \to G$$

Let $S_\alpha \subset G$ be the image. The kernel is a discrete central subgroup, hence a subgroup of $\{\pm 1\}$. Therefore, it must be the case that

$$S_\alpha \simeq SU(2) \text{ or } SU(2)/\{\pm 1\}$$

Theorem 8.1.2. *Let $\alpha, \beta \in \Phi$. Then*

1. $\dim \mathfrak{g}^\alpha = 1$, *and*

2. $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$ *provided that* $\alpha + \beta \in \Phi$.

3. *The only roots proportional to* α *are* $\pm\alpha$.

4. *There exist integers* $p, q \in \mathbb{Z}$ *such that*

$$\{k \in \mathbb{Z} \mid \beta + k\alpha \in \Phi \cup \{0\}\} = \{k \in \mathbb{Z} \mid p \leq k \leq q\}.$$

   Moreover, $p + q = -2\frac{(\alpha,\beta)}{(\alpha,\alpha)}$.

Proof. Let S be the set of integers k such that $\beta + k\alpha \in \Phi \cup \{0\}$. Define

$$\mathfrak{g}(S) = \bigoplus_{k\in S} \mathfrak{g}^{\beta+k\alpha} \oplus \begin{cases} \mathbb{C}H_\alpha & 0 = \beta + k\alpha \text{ for some } k \in S, \\ 0 & \text{otherwise} \end{cases}.$$

We claim that $[\mathfrak{s}_\alpha, \mathfrak{g}(S)] \subset \mathfrak{g}(S)$. This is clear because bracketing any summand above with E_α lands in another thing of this form, i.e. $[E_\alpha, \mathfrak{g}^{\beta+k\alpha}] \subset \mathfrak{g}^{\beta+(k+1)\alpha}$. Via φ_α, we get a representation of $\mathfrak{sl}(2,\mathbb{C})$ on $\mathfrak{g}(S)$. What are the weights for this representation?

$$\langle k\alpha, H_\alpha\rangle + \langle \beta, H_\alpha\rangle = 2k + \frac{(\beta,\alpha)}{(\alpha,\alpha)}\langle \alpha, H_\alpha\rangle = 2k + 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}.$$

because $H_\alpha \in (\alpha^\perp)^\perp$. These must be integers by our classification of finite-dimensional $\mathfrak{su}_2$-representations, so we must have

$$2\frac{(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}.$$

We know that in any finite dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ maps the weight spaces corresponding to to strictly negative weights injectively to the subsequent weight space. Then

$$[E_\alpha, \mathfrak{g}^{-\alpha}] \subset \mathbb{C}H_\alpha.$$

This is injective, hence $\dim \mathfrak{g}^{-\alpha} = 1$.

The conclusion is that the weight spaces are one dimensional with all weights having the same parity. Therefore, $\mathfrak{g}(S)$ is irreducible under the action of $\mathfrak{s}_\alpha$.

In an irreducible finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ maps the n-weight space isomorphically to the $n+2$ weight space unless one of two spaces is zero. Therefore,

$$[E_\alpha, \mathfrak{g}^\beta] \simeq \mathfrak{g}^{\beta+\alpha}$$

if $\beta + \alpha$ is a root (β is a root by assumption). This proves the second assertion.

Now suppose $\alpha, \beta = c\alpha$ for $c \in \mathbb{R}$ are roots and $\beta \neq \pm\alpha$. Interchanging α and β or negating α if necessary, we may assume that $0 < c < 1$. From earlier, we know that

$$2\frac{(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z} \implies c = \frac{1}{2}.$$

Now apply part (2) with $\alpha/2$ in place of α and β:

$$[\mathfrak{g}^{\alpha/2}, \mathfrak{g}^{\alpha/2}] = \mathfrak{g}^{\alpha}.$$

Since these are one-dimensional spaces,

$$[\mathbb{C}E_{\alpha/2}, \mathbb{C}E_{\alpha/2}] = \mathfrak{g}^{\alpha}$$

but the left hand side of the equation is obviously 0. This establishes (3).

The set of weights of a finite dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$ is symmetric about the origin, and is an uninterrupted string of integers of the same parity. Therefore,

$$S = \{p \leq k \leq q\} \text{ for some } p, q \in \mathbb{Z}.$$

So it must be the case that

$$2p + 2\frac{(\beta,\alpha)}{(\alpha,\alpha)} = -\left(2q + 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}\right)$$

so

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

□

By definition of S, it contains 0, so $p \leq 0 \leq q$.

Corollary 8.1.3. *Suppose α, β are roots and $\alpha \neq \pm\beta$.*

1. *If $(\alpha,\beta) > 0$ then $\beta - \alpha \in \Phi$.*
2. *If $(\alpha,\beta) < 0$ then $\beta + \alpha \in \Phi$.*
3. *If $(\alpha,\beta) = 0$ then $p + q = 0$, so either $\alpha + \beta$ and $\alpha - \beta$ are both roots, or neither is a root.*
4. *If $\alpha \pm \beta \notin \Phi$, then $(\alpha,\beta) = 0$.*

Suppose α, β are roots and $\alpha \neq \pm\beta$. If $\alpha \pm \beta \notin \Phi$, then one says that α, β are "strongly orthogonal." In this situation,

$$[\mathfrak{s}_\alpha, \mathfrak{s}_\beta] = 0,$$

i.e. $\mathfrak{s}_\alpha$ and $\mathfrak{s}_\beta$ commute. In fact, this is an if and only if condition.

Suppose now that (π, V) is an irreducible, finite-dimensional representation of G. Suppose $\alpha \in \Phi$ and λ is a weight of π. Define

$$S = \{k \in \mathbb{Z} \mid \lambda + k\alpha \text{ is a weight.}\}$$

Corollary 8.1.4. *In the notation above,*

$$S = \{k \in \mathbb{Z} \mid p \le k \le q\}$$

for some $p + q \in \mathbb{Z}$, *and*

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

Proof. Let

$$V(S) = \bigoplus_{k \in S} V^{\lambda + k\alpha}.$$

Then $\mathfrak{s}_\alpha V(S) \subset V$, and the rest of the proof goes through as before. ☐

Corollary 8.1.5. *In the notation above,*

1. *If* $(\alpha, \lambda) > 0$ *so then* $\lambda - \alpha \in \Phi$.
2. *If* $(\alpha, \lambda) < 0$ *then* $\lambda + \alpha \in \Phi$.
3. *If* $(\alpha, \lambda) = 0$ *then* $p + q = 0$, *so either* $\alpha + \lambda$ *and* $\alpha - \lambda$ *are both roots, or neither is a root.*

In $SU(2)$, consider

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

We saw that this normalizes the diagonal maximal torus in $SU(2)$ and acts on it by sending

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \mapsto \begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{pmatrix}.$$

Now, $\mathfrak{t}$ normalizes $\mathfrak{s}_\alpha$ and T normalizes S_α, and

$$\ker\{e^\alpha : T \to \mathbb{C}^*\}$$

centralizes S_α. Define

$$n_\alpha := \varphi \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

This normalizes T and inverts $\begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{pmatrix}$. On the level of Lie algebras, $\mathrm{Ad}\, n_\alpha$ is the identity on $\{X \in \mathfrak{t} \mid \langle \alpha, X \rangle = 0\}$ and sends H_α to $-H_\alpha$. Dually, on $i\mathfrak{t}^*_{\mathbb{R}}$, $\mathrm{Ad}\, n_\alpha$ is (orthogonal) reflection about $\alpha^\perp$.

Definition 8.1.6. There exists an element $s_\alpha \in W$, with W viewed as a group acting on $i\mathfrak{t}^*_{\mathbb{R}}$, which is reflection about α (namely, $s_\alpha = \mathrm{Ad}\, n_\alpha$).

8.2 Weyl Chambers

For each root α, we constructed a subgroup $S_\alpha \subset G$ with complexified Lie algebra

$$\mathfrak{s}_\alpha = \mathfrak{g}^\alpha \oplus \mathfrak{g}^{-\alpha} \oplus [\mathfrak{g}^\alpha, \mathfrak{g}^{-\alpha}],$$

so $S_\alpha \simeq SU(2)$ or $SU(2)/\{\pm 1\}$. S_α is normalized by T and commutes with

$$\ker\{e^\alpha : T \to \mathbb{C}^*\}.$$

We constructed $s_\alpha \in W$, $s_\alpha = \operatorname{Ad} n_\alpha$ the image of $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in SU(2)$, such that s_α is the reflection about $\alpha^\perp$: for $\mu \in i\mathfrak{t}_\mathbb{R}^*$,

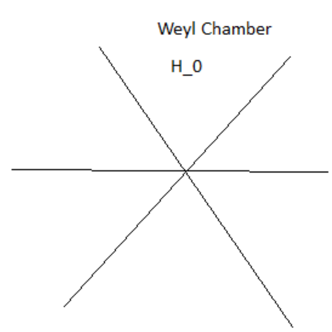
$$s_\alpha \mu = \mu - 2\frac{(\mu, \alpha)}{(\alpha, \alpha)}\alpha.$$

We saw that $2\frac{(\mu,\alpha)}{(\alpha,\alpha)}\alpha \in \mathbb{Z}$, and has significance concerning the roots of the form $\mu + k\alpha$.

The point is that $SU(2)$ commutes with $\ker e^\alpha$. So s_α fixes the kernel of e^α and acts as -1 on the other piece of $\mathfrak{t}$ (or inversion on the other piece of T).

Definition 8.2.1. An element $H \in i\mathfrak{t}_\mathbb{R}$, respectively $\mu \in i\mathfrak{t}_\mathbb{R}^*$, is *regular* if $\langle \alpha, H \rangle \neq 0$ (resp. $(\alpha, \mu) \neq 0$), for all $\alpha \in \Phi$.

The picture is that the regular set in $i\mathfrak{t}_\mathbb{R}$ (resp. $i\mathfrak{t}_\mathbb{R}^*$) is the complement of a finite number of hyperplanes, which slice up the vector space into wedges.



Definition 8.2.2. An *open Weyl chamber* in $i\mathfrak{t}_\mathbb{R}$ (resp. $i\mathfrak{t}_\mathbb{R}^*$) is a connected component of the regular set.

A *Weyl chamber* is the closure of an open Weyl chamber.

We can think of $i\mathfrak{t}_{\mathbb{R}} \simeq i\mathfrak{t}_{\mathbb{R}}^*$ under the inner product, sending the regular sets to each other. Therefore, the Weyl chambers in $i\mathfrak{t}_{\mathbb{R}}$ correspond bijectively to those in $i\mathfrak{t}_{\mathbb{R}}^*$ via this isomorphism.

Choose a regular $H_0 \in i\mathfrak{t}_{\mathbb{R}}$ and define

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

Lemma 8.2.3. *Keeping the notation above,*

(i) $\Phi = \Phi^+ \sqcup \Phi^-$.

(ii) For $\alpha, \beta \in \Phi^+$, *if* $\alpha + \beta \in \Phi$ *then* $\alpha + \beta \in \Phi^+$.

(iii) Φ^+ *depends only on the Weyl Chamber in which* H_0 *lies, not on the particular choice of* H_0.

(iv) The Weyl chamber C^+ *in which* H_0 *lies is characterized by*

$$\mathit{int}(C^+) = \{H \in i\mathfrak{t}_{\mathbb{R}} \mid \langle \alpha, H \rangle > 0 \textit{ for all } \alpha \in \Phi^+\}.$$

$$C^+ = \{H \in i\mathfrak{t}_{\mathbb{R}} \mid \langle \alpha, H \rangle \geq 0 \textit{ for all } \alpha \in \Phi^+\}.$$

Proof. (i) and (ii) are obvious. For (iii), observe that if we can connected H_0 to some other regular H by a path in this regular set, then by the intermediate value theorem

$$\langle \alpha, H_0 \rangle > 0 \iff \langle \alpha, H_1 \rangle > 0.$$

Therefore,

$$\mathrm{int}(C^+) \subset \{H \in i\mathfrak{t}_{\mathbb{R}} \mid \langle \alpha, H \rangle > 0 \text{ for all } \alpha \in \Phi^+\}.$$

If the two sets were not equal, there would exist $H_1 \notin \mathrm{int}(C^+)$ such that

$$\langle \alpha, H_1 \rangle > 0 \text{ for all } \alpha \in \Phi^+.$$

By convexity, the entire line connecting H_1 and H_0 is in the regular set, furnishing a contradiction. This proves (iv). □

An obvious observation that we used in the proof is:

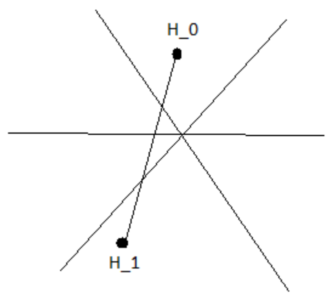
Corollary 8.2.4. *The open Weyl chambers are convex.*

Theorem 8.2.5. W *is generated, as a group, by the* s_α *for* $\alpha \in \Phi$. *Moreover, it acts simply transitively on the set of Weyl chambers.*

(Recall that a simply transitive action is one in which no element but the identity fixes anything, i.e. the stabilizers are all trivial).

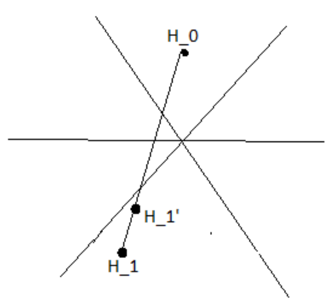
Proof. The Weyl group W acts on Φ, hence on the regular set, hence also on the Weyl chambers. Also, $s_\alpha \in W$ for every $\alpha \in \Phi$. Suppose C^+ and C_1 are two Weyl chambers with $H_0 \in \mathrm{int}(C^+)$, and choose $H_1 \in \mathrm{int}(C_1)$.

Moving H_1 slightly, we can arrange that the straight line segment ℓ from H_0 to H_1 does not cross any intersection of two or more root hyperplanes (think about the codimension).



Let N denote the number of hyperplanes $\alpha^\perp$ crossed by ℓ, and let $N(C^+, C_1)$ denote the minimum possible N over all possible choices of H_0, H_1. Then $N(C^+, C_1)$ measures the "distance" between C^+ and C_1; in particular, $N(C^+, C_1) = 0$ if and only if $C^+ = C_1$.

Now choose H_0 and H_1 such that $N = N(C_1^+, C_1)$. Replace H_1 by a "new" H_1 which is very close to one root hyperplane $\alpha^\perp$ but not comparably close to any other (traverse the line ℓ from H_0 to H_1, and choose the new H_1 to be just after the last crossing of a hyperplane).



Then $s_\alpha H_1$ is regular (just the reflection of H_1 across across the very close hyperplane $\alpha^\perp$) and lies in a different Weyl chamber, namely $s_\alpha C_1$. By construction,

$N(C^+, s_\alpha C_1) < N(C^+, C_1)$. Repeating this argument, we find $\alpha_1, \ldots, \alpha_n \in \Phi$ such that

$$N(C^+, s_{\alpha_1} \ldots s_{\alpha_n} C_1) = 0$$

i.e. $C^+ = s_{\alpha_1} \ldots s_{\alpha_n} C_1$. This shows that the subgroup of W generated by the s_α for $\alpha \in \Phi$ acts transitively on the set of Weyl chambers.

To complete the proof, we must show that if $w \in W$ and $wC^+ = C^+$, then $w = e$. Supposing that w fixes C^+, choose $\widetilde{H}_0 \in \text{int}(C^+)$. By convexity,

$$\frac{1}{|W|} \sum_{k=1}^{|W|} w^k \widetilde{H}_0 = H_0 \in \text{int}(C^+)$$

and $wH_0 = H_0$. Choose $n \in N_G(T)$ such that $w = \text{Ad } n$. Let S be the closure of $\{\exp(itH_0) \mid t \in \mathbb{R}\} \subset T$ (since $H_0 \in i\mathfrak{t}_\mathbb{R}$). Then S is connected, compact, abelian, and contained in T, hence it is a torus. Also $\text{Ad } n(H_0) = H_0$, therefore $n \in Z_G(S)$, which is connected.

Since $T \supset S$ the complexified Lie algebra $\mathfrak{z}$ of $Z_G(S)$ contains $\mathfrak{t}$, hence has a weight decomposition

$$\mathfrak{t} \oplus (\bigoplus_{\alpha \in B} \mathfrak{g}^\alpha) \qquad B \subset \Phi.$$

By construction, for each $\alpha \in B$ we must have $e^\alpha = 1$ on S, hence $\langle \alpha, H_0 \rangle = 0$ (since $Z_G(S) = Z_\mathfrak{g}(H_0)$, but there are no such α since H_0 is in the interior of a Weyl chamber, so $B = \emptyset$. Therefore $\mathfrak{z} = \mathfrak{t}$, hence $Z_G(S)^0 = T$, and by connectedness $Z_G(S) = T$, so $n \in T$. Therefore, $w = \text{id}$.

□

With the original choices of H_0, Φ^+,

$$\text{int}(C^+) = \{H \in i\mathfrak{t}_\mathbb{R} \mid \langle \alpha, H \rangle > 0 \forall \alpha \in \Phi^+\}.$$

Φ^+ is a "system of positive roots" and C^+ is called the "dominant Weyl Chamber."

Definition 8.2.6. $\alpha \in \Phi^+$ is a *simple root* if α cannot be expressed as $\alpha = \beta + \gamma$ with $\beta, \gamma \in \Phi^+$.

Of course, this notion depends on the choice of dominant Weyl chamber C^+ which was used to define Φ^+. Let $\Psi \subset \Phi^+$ denote the set of simple roots.

Theorem 8.2.7. *Keeping the notation above,*

(a) For $\alpha, \beta \in \Psi$ such that $\alpha \neq \beta$, then $(\alpha, \beta) \leq 0$, and if equality holds then α and β are strongly orthogonal ($\alpha \pm \beta$ are both not roots).

(b) Ψ is a vector space basis of the $\mathbb{R}$-linear span of Φ.

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

(d) *Every $\alpha \in \Phi$ is W-conjugate to a simple root.*

(e) *W is generated by s_α with $\alpha \in \Psi$.*

Proof. (a) Suppose $\alpha, \beta \in \Psi$ are distinct simple roots with $(\alpha, \beta) > 0$. From the discussion last time, we know that $\beta - \alpha \in \Phi$. Interchange the roles of α and β if necessary so that $\alpha - \beta = \gamma \in \Phi^+$. Then $\alpha = \beta + \gamma$, which is a contradiction.

In particular, this argument shows that $\alpha - \beta \notin \Phi$. Therefore, if $(\alpha, \beta) = 0$ then they must be strongly orthogonal (again, look at the corollary from last time: for orthogonal roots, either both their sum and difference are roots or neither is).

(b),(c) Now suppose that $\alpha \in \Phi^+$. If $\alpha \notin \Psi$, then there exist positive roots $\alpha', \alpha'' \in \Phi^+$ such that $\alpha = \alpha' + \alpha''$. Observe that

$$\langle \alpha, H_0 \rangle = \langle \alpha', H_0 \rangle + \langle \alpha'', H_0 \rangle$$

with all terms positive. Since there are only finitely many positive roots, Φ is a discrete set, and this process must eventually terminate, in which we have expressed α as a integer combination of simple roots. Therefore, Ψ spans the $\mathbb{R}$-linear span of Φ. We also get (c) after showing linear independence of the simple roots.

Suppose that Ψ is not linearly dependent over $\mathbb{R}$. Say

$$\sum_{j=1}^{r} c_j \alpha_j = 0$$

where $\Psi = \{\alpha_1, \ldots, \alpha_r\}$ and the c_j real and not all zero. Renumbering the simple roots and multiplying the identity by -1 if necessary, we may assume that $c_1, \ldots, c_k > 0$ and $k \geq 1$, while $c_{k+1}, \ldots, c_r \leq 0$. Then

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

so taking the inner product gives something non-negative:

$$0 \leq || \sum_{j \leq k} c_j \alpha_j ||^2 = \sum_{j \leq k < i} c_j(-c_i)(\alpha_j, \alpha_i)$$

but $c_j > 0$, $-c_i \geq 0$, and $(\alpha_j, \alpha_i) \leq 0$. That forces

$$\sum_{j \leq k} c_j \alpha_j = 0 \implies \langle \sum_{j \leq k} c_j, H_0 \rangle = 0$$

which is a contradiction since the α_j were positive roots and $c_j \geq 0$ in this range.

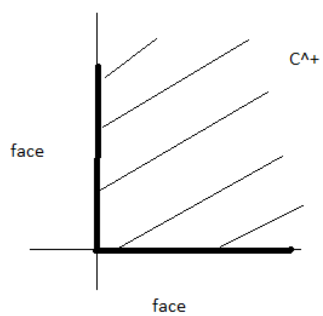
By (a)-(c), we have now characterized

$$C^+ = \{H \in i\mathfrak{t}_{\mathbb{R}} \mid \langle \alpha, H\rangle \geq 0 \text{ for all } \alpha \in \Psi\}.$$

Note that $i\mathfrak{t}_{\mathbb{R}}$ is the direct sum of the central part and the linear span of the duals of the roots

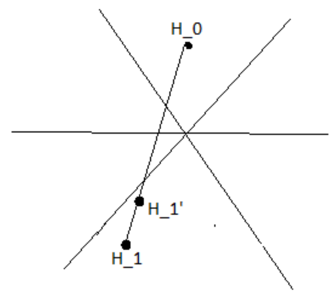
$$\mathfrak{t} = \mathfrak{z} \oplus (\mathfrak{t} \cap [\mathfrak{g}, \mathfrak{g}]),$$

and C^+ is a "partial quadrant" in the latter summand. In particular, every face of C^+ (i.e. any codimension 1 component of ∂C^+ is contained in exactly one of the hyperplanes $\alpha^\perp$ for $\alpha \in \Psi$). Conversely, every one of these hyperplanes contains a face of C^+ since the simple roots impose independent conditions.



Let α be a root. Then the hyperplane $\alpha^\perp$ contains a face of some Weyl chamber C_1. We know that $C_1 = wC^+$ for some $w \in W$ since W acts simply transitively on the set of Weyl chambers. So $(w\alpha)^\perp$ contains a face of C^+, so $(w\alpha)^\perp = \beta^\perp$ for $\beta \in \Psi$. By independence, $w_\alpha = \pm\beta$. But $s_\alpha \alpha = -\alpha$, so $\beta = w\alpha$ or $\beta = ws_\alpha\alpha$. This establishes (d).

To prove (e), recall the proof of the generation of W by the s_α, $\alpha \in \Phi$. Given any two Weyl chambers C_1, C_2 we defined $N(C_1, C_2)$ with $N(C_1, C_2) \geq 0$ and $N(C_1, C_2) = 0 \iff C_1 = C_2$. We had seen that there exists $\alpha \in \Phi$ such that $\alpha^\perp$ is a face of C_1 and $N(s_\alpha C_1, C_2) < N(C_1, C_2)$,



Apply this with $C_1 = C^+$: there exists a simple root such that $N(s_\alpha C^+, C_2) < N(C^+, C_2)$. W acts on the chambers, so $N(s_\alpha C^+, C_2) = N(C^+, s_\alpha C_2)$. Repeating this argument, we eventually arrive at $C^+ = s_{\alpha_n} \dots s_{\alpha_1} C_2$ with $\alpha_j \in \Psi$, which establishes 5.

□

We have almost completed the ingredients for the classification of the compact Lie groups. As a consequence of the theorem, we see that knowledge of Ψ as an abstract set, along with the data $||\alpha||$, $\alpha \in \Psi$, and the inner products (α, β) with $\alpha, \beta \in \Psi$, completely determine Φ as a subset of the inner product space spanned by Φ over $\mathbb{R}$.

Now suppose $\alpha, \beta \in \Phi$ such that $\alpha \neq \pm\beta$. Switching the roles of α and β if necessary, we may suppose that $||\alpha|| \geq ||\beta||$.

Theorem 8.2.8. *In this situation, if α and β are not orthogonal then*

$$||\alpha||^2 = k||\beta||^2 \qquad k = 1, 2, 3$$

and

$$(\cos \angle(\alpha, \beta))^2 = \frac{(\alpha, \beta)^2}{||\alpha||^2||\beta||^2} = \frac{k}{4}.$$

Proof. We know that

$$2\frac{(\alpha, \beta)}{||\alpha||^2}, 2\frac{(\alpha, \beta)}{||\beta||^2} \in \mathbb{Z}.$$

Taking the product,

$$4\frac{(\alpha, \beta)^2}{||\alpha||^2||\beta||^2} \in \mathbb{Z}$$

but α and β are neither proportional nor perpendicular, so $\frac{(\alpha,\beta)^2}{||\alpha||^2||\beta||^2} = \frac{k}{4}$ where $k = 1, 2,$ or 3. Since $||\alpha|| \geq ||\beta||$, the first term in the first equation is the smaller integer, hence

$$\left|2\frac{(\alpha,\beta)}{||\alpha||^2}\right| = 1.$$

Straightforward manipuulations of this imply what we want. □

Corollary 8.2.9. *Suppose α, β are distinct simple roots and $(\alpha, \beta) \neq 0$. Then*

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

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

This potentially allows us to reconstruct the root system from the knowledge of a couple of roots.

Chapter 9

Classificiation of compact Lie groups

9.1 Classification of semisimple Lie algebras

Definition 9.1.1. A Lie algebra $\mathfrak{g}$ over a field k is *simple* if it contains no proper ideal and has dimension at greater than 1.

Note that this rules out abelian Lie algebras, since every subalgebra of a Lie algebra is an ideal.

Definition 9.1.2. A Lie algebra $\mathfrak{g}$ over a field k is *semisimple* if it is a direct sum of commuting simple ideals.

Definition 9.1.3. A Lie group G is *simple* (resp. semisimple) if its Lie algebra $\mathfrak{g}_{\mathbb{R}}$ is simple (resp. semisimple) as a Lie algebra over $\mathbb{R}$.

Proposition 9.1.4. *Suppose G is a connected, compact, simple Lie group. Then $\mathfrak{g}$ is simple over $\mathbb{C}$.*

This says that $\mathfrak{g}_{\mathbb{R}}$ is "absolutely simple" (simple when tensored up with the algebraic closure).

Proof. Suppose not; then $\mathfrak{g}$ contains a simple ideal $\mathfrak{g}_1$. Then $\mathfrak{g}_1 \neq \overline{\mathfrak{g}_1}$ because $\mathfrak{g}_{\mathbb{R}}$ is simple. We claim that $\mathfrak{g} = \mathfrak{g}_1 \oplus \overline{\mathfrak{g}_1}$. Indeed, if $\mathfrak{g}_1 \oplus \overline{\mathfrak{g}_1}$ isn't the whole space, then its intersection with $\mathfrak{g}_{\mathbb{R}}$ is a proper ideal (because the Galois group of $\mathbb{C}/\mathbb{R}$ interchanges the two factors). Note also that $[\mathfrak{g}_1, \overline{\mathfrak{g}_1}] = 0$.

Define $J : \mathfrak{g} \to \mathfrak{g}$ by

$$JX = \begin{cases} iX & X \in \mathfrak{g}_1 \\ -iX & X \in \overline{\mathfrak{g}_1} \end{cases}.$$

Then $J = \overline{J}$, so

$$J : \mathfrak{g}_{\mathbb{R}} \to \mathfrak{g}_{\mathbb{R}}$$

with $J^2 = -1$. Suppose $u, v \in \mathfrak{g}_{\mathbb{R}}$. Then

$$u = u_1 + \bar{u}_1, v = v_1 + \bar{v}_1 \qquad u_1, v_1 \in \mathfrak{g}_1$$

so

$$[Ju, v] = [J(u_1 + \bar{u}_1), v_1 + \bar{v}_1] = [iu_1 - i\bar{u}_1, v_1 + \bar{v}_1] = i[u_1, v_1] - i[\bar{u}_1, \bar{v}_1]$$

Observe that this coincides with

$$J[u, v] = J[u_1 + \bar{u}_1, v_1 + \bar{v}_1] = J([u_1, v_1] + [\bar{u}_1, \bar{v}_1]) = i[u_1, v_1] - i[\bar{u}_1, \bar{v}_1].$$

So $J : \mathfrak{g}_{\mathbb{R}} \to \mathfrak{g}_{\mathbb{R}}$ with $J^2 = -1$ and

$$[Ju, v] = J[u, v] = [u, Jv].$$

Turn $\mathfrak{g}_{\mathbb{R}}$ into a Lie algebra over $\mathbb{C}$ by defining $iX = JX$.

What are we doing? Turning a finite dimensional real vector space into a complex vector space amounts to giving a linear map that squares to -1. We have just proved that the complex structure is compatible with the Lie bracket. We had seen earlier that this implies that G is a complex Lie group. By assumption, it is compact and non-abelian, which is impossible (see homework). $\square$

On the homework, we saw that if G is a connected, compact Lie group then G is semisimple if and only if Z_G is finite. If so, then there exist connected compact simple Lie groups $G_1, \ldots, G_n$ with a finite covering map $G_1 \times \ldots \times G_n \to G$. So, to understand connected, compact, semisimple Lie groups, it suffices to understand connected, compact, simple Lie groups and their centers.

As usual, let G be a connected compact Lie group and T a torus. Observation: G is semisimple if and only if the $\mathbb{R}$-linear span of Φ is $i\mathfrak{t}_{\mathbb{R}}^*$, which is the case if and only if Ψ is an $\mathbb{R}$-basis of $i\mathfrak{t}_{\mathbb{R}}^*$.

Definition 9.1.5. Φ is *irreducible* if Φ cannot be expressed as a disjoint union

$$\Phi = \Phi_1 \sqcup \Phi_2$$

with $\Phi_i = -\Phi_i$ non-empty, and $\Phi_1 \perp \Phi_2$.

Definition 9.1.6. Ψ is *irreducible* if Ψ cannot be expressed as

$$\Psi = \Psi_1 \sqcup \Psi_2$$

with Ψ_i non-empty and $\Psi_1 \perp \Psi_2$.

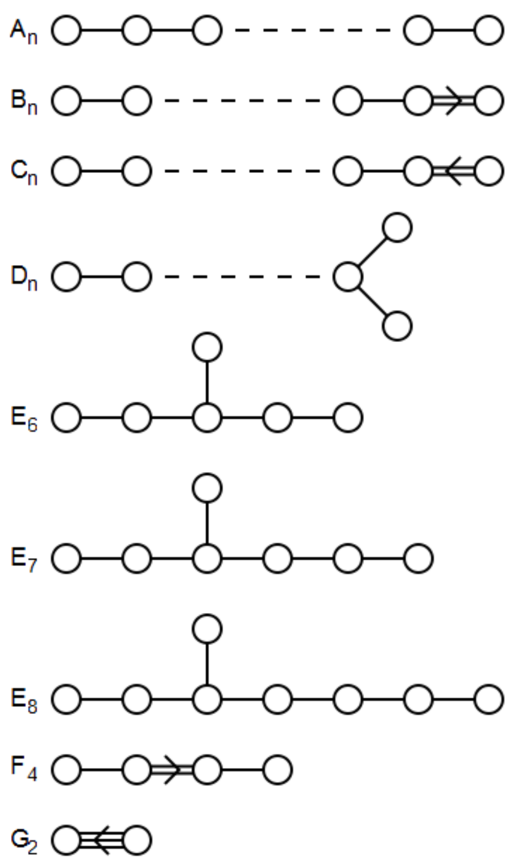
Another observation: if G is semisimple, then G is simple if and only if Φ is irreducible, which is the case if and only if Ψ is irreducible. This is again straightforward, and its proof is left as an exercise.

Now suppose that G is connected, compact, and simple. The *Dynkin diagram* of G is the graph constructed as follows:

- The vertices are the simple roots.
- $\alpha_1, \alpha_2 \in \Psi$ are connected by k edges if $(\alpha_1, \alpha_2) \neq 0$ and $||\alpha_1|| = k||\alpha_2||^2$ or vice versa, and are not connected if $(\alpha_1, \alpha_2) = 0$ (equivalently strongly orthogonal). If $k \neq 1$, the edge is directed from the longer root to the shorter.

By the observation, the Dynkin diagram must be connected, since the connected components correspond to reducible pieces.

Theorem 9.1.7. *The only possible Dynkin diagrams of connected, compact, simple Lie groups are:*



Each of these does correspond to a connected, compact, simple Lie group, which is unique up to covering.

Proof sketch. The first part follows from some elementary analysis using inner products, etc. For instance, one quickly realizes that cycles are impossible, and then only one pair of nodes can have multiple edges.

A_n is $SU(n+1)$, $B_n = SO(2n+1)$, $C_n = Sp(2n) \cap U(4n)$, and $D_n = SO(2n)$. These are the classical groups.

Why the uniqueness? From the Dynkin diagram we can reconstruct the roots and root system (see remarks above). For simple α, take $E_\alpha, E_{-\alpha}$, and H_α as above, and normalize the scaling. Additional roots are obtained by bracketing these for simple roots.

Compact semisimple Lie groups are adjoing groups (the image under the adjoint representation). The last question concerns the centers; we will address this next time.

□

9.2 Simply-connected and adjoint form

The homework, together with the discussion last time, implies that to "understand" all connected, compact Lie groups, it suffices to "understand" the coverings of simple, connected, compact Lie groups.

Suppose G is connected, compact, and semisimple, $T \subset G$ a maximal torus and L, Λ as usual. We define

$$G_{\mathrm{Ad}} := G/Z_G$$

which is isomorphic to the image of G under the adjoint representation (something commutes with a neighborhood of 0 if and only if it commutes with everything, since a neighborhood of 0 generates everything). Obviously, $Z_G \subset T$ is discrete, hence finite. Similarly, we have $T_{\mathrm{Ad}} = T/Z_G$, etc. Identifying $\mathfrak{t}_{\mathbb{R}}/L \simeq T$, this maps to $T_{\mathrm{Ad}} \simeq \mathfrak{t}_{\mathbb{R}}/L_{\mathrm{Ad}}$ with kernel L_{Ad}/L, so $Z_G \simeq L_{\mathrm{Ad}}/L$.

Then

$$\begin{aligned} L_{\mathrm{Ad}} &= \{H \in \mathfrak{t}_{\mathbb{R}} \mid \exp H \in Z_G\} \\ &= \{H \in \mathfrak{t}_{\mathbb{R}} \mid \exp H \in \ker \mathrm{Ad}\} \\ &= \{H \in \mathfrak{t}_{\mathbb{R}} \mid e^{\alpha} \exp H = 1 \text{ for all } \alpha\} \\ &= \{H \in \mathfrak{t}_{\mathbb{R}} \mid \langle \alpha, H\rangle \in 2\pi i\mathbb{Z} \text{ for all } \alpha \in \Phi\} \\ &= \{H \in \mathfrak{t}_{\mathbb{R}} \mid \langle \alpha, H\rangle \in 2\pi i\mathbb{Z} \text{ for all } \alpha \in \Psi\} \end{aligned}$$

We have seen that

$$\begin{aligned} \Lambda &\overset{(*)}{\subset} \{\mu \in i\mathfrak{t}_{\mathbb{R}} \mid 2\frac{(\alpha,\mu)}{(\alpha,\alpha)} \in \mathbb{Z} \text{ for all } \alpha \in \Phi\} \\ &= \{\mu \in i\mathfrak{t}_{\mathbb{R}} \mid \langle \mu, H_\alpha\rangle \in \mathbb{Z} \text{ for all } \alpha \in \Phi\} \\ &= \{\mu \in i\mathfrak{t}_{\mathbb{R}} \mid \langle \mu, H_\alpha\rangle \in \mathbb{Z} \text{ for all } \alpha \in \Psi\} \\ &:= \Lambda_{\mathrm{sc}} \end{aligned}$$

The inclusion $(*)$ is actually an equality for G simply connected, and depends on knowing that every $\mu \in \Lambda$ is the weight of some finite dimensional irreucible representation. We'll get to this later.

If Λ_{sc} were the weight lattice of some covering group of G, then the corresponding unit lattice would be $L_{\mathrm{sc}} =$ the $\mathbb{Z}$-linear span of $2\pi i H_\alpha$ for $\alpha \in \Psi$.

Theorem 9.2.1. *With the notation above,*

1. *$\pi_1(G)$ is finite.*
2. *Let G_{sc} be the universal covering. Then its unit lattice is L_{sc} and its weight lattice is Λ_{sc}. In particular, $\pi_1(G_{\mathrm{Ad}}) \simeq Z_{G_{sc}} \simeq L_{\mathrm{Ad}}/L_{sc}$.*

Sketch. This is only a sketch; a more detailed argument will be posted online.

Define

$$G_{\mathrm{reg}} := \{g \in G \mid \dim Z_G(g) = \operatorname{rank} G\},$$

where $\operatorname{rank} G$ is the dimension of the maximal torus. Then G_{reg} is a conjugation invariant. Then

$$\begin{aligned} G - G_{\mathrm{reg}} &= \{g \in G \mid \dim Z_G(g) > \operatorname{rank} G\} \\ &= \{g \in G \mid \dim \ker(1 - \mathrm{Ad}\ g) > \operatorname{rank} G\} \end{aligned}$$

since $\ker(1-\mathrm{Ad}\ g) = Z_G(g)$. This implies that $G - G_{\mathrm{reg}}$ is a real analytic subvariety of G, i.e. cut out by real analytic equations. By invariance under conjugation, and the fact that every element of G is conjugate to an element of T,

$$G - G_{\mathrm{reg}} = \{gtg^{-1} \mid t \in T - T_{\mathrm{reg}}, g \in G, \text{“}g \in G/T''\}$$

where $T_{\mathrm{reg}} = G_{\mathrm{reg}} \cap T$ and “$g \in G/T''$ means that we need only take coset reprentatives of G/T (representatives of the same coset will give the same gtg^{-1}). The centralizer of $t \in T$ contains T, and is bigger if and only if $e^\alpha(t) = 1$ for some α.

$$T - T_{\mathrm{reg}} = \{t \in T \mid e^\alpha(t) = 1 \text{ for some } \alpha \in \Phi\}.$$

Then

$$G - G_{\mathrm{reg}} = \bigcup_\alpha \{gtg^{-1} \mid e^\alpha(t) = 1, \text{“}g \in G/TS_\alpha''\}$$

where S_α is defined as in our proof of the structure theorem for roots, since any such t will commute with s_α as well, and T normalizes S_α. The condition $e^\alpha(t) = 1$ imposes a codimension-one condition, and TS_α has dimension $\dim T + 2$ since S_α has dimension 3 but intersects T in a one-dimensional piece. So $\dim(G - G_{\mathrm{reg}}) \leq (\dim T - 1) + \dim G - (\dim T + 2) \leq \dim G - 3$. In fact, this is an equality.

What does cutting out a codimension 3 thing mean? Cutting out codimension 1 disconnects. Cutting out codimension 2 doesn't disconnect, but changes the fundamental group. Codimension 3 doesn't change the fundamental group, since homotopies occur on a surface, and we can homotope things to be transverse to the codimension 3 things.

Therefore, the inclusion

$$G_{\text{reg}} \hookrightarrow G$$

induces an isomorphism on π_1. Consider the map

$$\begin{aligned} G/T \times T_{\text{reg}} &\to G_{\text{reg}} \\ (gT, t) &\mapsto gtg^{-1}. \end{aligned}$$

Facts:

1. This is a covering map with covering group W, which acts on G/T by right translation and on T_{reg} by conjugation.

2. T_{reg} is connected.

Now,

$$\pi_1(G/T) \simeq \pi_1(G/T \times \{\text{pt}\}) \subset \pi_1(G/T \times T_{\text{reg}}).$$

Since the map is a covering map, the induced map on fundamental groups is injective, so

$$\pi_1(G/T \times T_{\text{reg}}) \to \pi_1(G_{\text{reg}}) \simeq \pi_1(G)$$

is injective. Let $g(s)$ represent a loop in G/T, i.e.

$$g(1)T = g(0)T$$

and $s \mapsto g(s)$ is continuous. Then the image loops in $\pi_1(G)$ is

$$s \mapsto g(s)t_0 g(s)^{-1} \qquad t_0 \in T_{\text{reg}}.$$

Note that we couldn't choose $t_0 = e$ because t_0 must be regular. However, this can be contracted in G by letting $t_0 \to e$. The conclusion is that $\pi_1(G/T) = 0$.

The map $G \to G/T$ is a fibration with fiber bundle T, so

$$\pi_1(T) \to \pi_1(G) \to \pi_1(G/T)$$

is exact. Since the last term is 0,

$$\pi_1(T) \to \pi_1(G)$$

is surjective. But $T = \mathfrak{t}_{\mathbb{R}}/L$, so we get a surjection

$$L \twoheadrightarrow \pi_1(G).$$

That tells us that $\pi_1(G)$ is finitely generated and abelian.

What we've seen is that if we pass to a covering, the unit lattice shrinks and the weight lattice grows. But the inclusion $(*)$ that we haven't proven yet shows that the weight lattice is "bounded above."

This already implies $\pi_1(G)$ is finite because Λ is "bounded above." (We know that $\pi_1(G)$ is finitely generated and abelian.)

Consider the map $SU(2) \to G$, with image S_α.

$$\begin{tikzcd} \text{“}L \text{ of } SU(2)\text{''} \arrow[r] \arrow[d] & L \arrow[d] \\ \pi_1(SU(2)) \arrow[r] & \pi_1(G) \end{tikzcd}$$

We know that $SU(2) \simeq S^3$ is simply connected, and the path which represents the generator $\begin{pmatrix} 2\pi i & 0 \\ 0 & -2\pi i \end{pmatrix}$ in the unit lattice of $SU(2)$ is

$$s \mapsto \begin{pmatrix} e^{2\pi i s} & 0 \\ 0 & e^{-2\pi i s} \end{pmatrix} \to G$$

the composite map being $s \mapsto \exp(2\pi i s H_\alpha)$.

Conclusion. Under the surjection $L \to \pi_1(G)$, the $2\pi i H_\alpha$ all go to e. So $\pi_1(G)$ is a quotient of L/L_{sc}, in particular finite.

To complete the proof, we must show that this is an isomorphism when $G = G_{\mathrm{sc}}$, the universal cover of G. This will be posted online. $\square$

We now establish the inclusion (*) from earlier. Suppose $\mu \in \Lambda$, i.e. μ lifts to a character

$$e^\mu : T \to \mathbb{C}^*.$$

Consider

$$\{f \in C(G) \mid f(gt) = e^\mu(t) f(g) \text{ for all } t \in T\}.$$

This is non-zero because locally the fibration $G \to G/T$ is a product, i.e. there are local sections. So we can take a function on G/T with compact support contained within the appropriate open set, and extend to the pre-image in G by multiplying by e^μ on the T factor. By Peter-Weyl, the L^2 closure of this space is

$$\widehat{\bigoplus_{i \in \widehat{G}}} V_i \otimes (V_i^*)^\mu$$

so $(V_i^*)^\mu$ cannot be zero for all $i \in \widehat{G}$, hence μ is a weight of V_i^*.

Chapter 10

Representations of compact Lie groups

10.1 Theorem of the highest weight

Now suppose that G is connected and compact (not necessarily semisimple, so we could have a center of positive dimension). Let $T \subset G$ be the maximal torus, C^+ a dominant Weyl chamber. Thus we have Φ^+ and Ψ. Let (π, V) be an irreducible, finite-dimensional representation of G.

Theorem 10.1.1 ("Theorem of the highest weight")**.** *The following conditions on a weight λ of π are equivalent.*

(1) $\lambda + \alpha$, for $\alpha \in \Phi^+$, is not a weight.

(2) There exists a non-zero $v_\lambda \in V^\lambda$ such that

$$E_\alpha v_\lambda = 0 \text{ for all } \alpha \in \Phi^+.$$

(3) Any weight of π can be expressed as

$$\lambda - \alpha_1 - \alpha_2 - \ldots - \alpha_N$$

with $\alpha_1, \ldots, \alpha_N \in \Phi^+$ (repetitions allowed).

A weight with these properties exists and is unique. It is called the "highest weight" of π, and it has the following properties.

1. $\dim V^\lambda = 1$.

2. $(\lambda, \alpha) \geq 0$ *for every $\alpha \in \Phi^+$, i.e. λ is dominant.*

3. *λ determines π up to isomorphism.*

Remark 10.1.2. We shall see later that every dominant $\lambda \in \Lambda$ is the highest weight of some irreducible representation. Therefore,

$$\widehat{G} \simeq \{\text{dominant } \lambda \in \Lambda\}.$$

Proof. (1) $\implies$ (2). This is because $E_\alpha V^\lambda \subset V^{\lambda+\alpha}$ (in fact, we get the result for *any* such v_λ).

(2) $\implies$ (3). Define

$$\mathcal{L}_+ = \text{linear span of } \mathfrak{t} \text{ and } \mathfrak{g}^\alpha, \alpha \in \Phi^+$$

$$\mathcal{M}_- = \bigoplus_{\alpha \in \Phi^+} \mathfrak{g}^{-\alpha}.$$

Both are Lie subalgebras of $\mathfrak{g}$, so

$$\mathfrak{g} = \mathcal{L}_+ \oplus \mathcal{M}_-$$

(a vector space direct sum). By PBW,

$$U(\mathfrak{g}) = U(\mathcal{M}_-)U(\mathcal{L}_+).$$

Suppose $v_\lambda \in V^\lambda$ is non-zero. By irreducibility,

$$V = U(\mathfrak{g})v_\lambda = U(\mathcal{M}_-)U(\mathcal{L}_+)v_\lambda.$$

Now,

$$\mathfrak{t}v_\lambda = \begin{cases} \mathbb{C}v_\lambda & \lambda \neq 0 \\ 0 & \lambda = 0 \end{cases}$$

so if $E_\alpha v_\lambda = 0$ then for all $\alpha \in \Phi^+$,

$$U(\mathfrak{g})v_\lambda = U(\mathcal{M}_-)\underbrace{U(\mathcal{L}_+)v_\lambda}_{\mathbb{C}v_\lambda} = U(\mathcal{M}_-)v_\lambda.$$

This is the linear span of $E_{-\alpha_1}, E_{-\alpha_2}, \ldots, E_{-\alpha_N}v_\lambda$ for all possible N-tuples of positive roots. This proves (3).

Also, $U(\mathcal{M}_-)v_\lambda = U(\mathcal{M}_-)(\mathcal{M}_- \oplus \mathbb{C})v_\lambda$. The summand coming from $U(\mathcal{M}_-)\mathcal{M}_-$ has weight strictly less than λ, so $\dim V^\lambda = 1$.

(3) $\implies$ (1). Suppose $\lambda + \alpha$ is a weight for some $\alpha \in \Phi^+$. By (3), $\lambda + \alpha = \lambda - \alpha_1 - \ldots - \alpha_N$, so therefore,

$$\alpha + \alpha_1 + \ldots + \alpha_N = 0.$$

Apply this to some regular $H_0 \in C^+$:

$$\langle \alpha + \alpha_1 + \ldots + \alpha_N, H_0 \rangle = 0,$$

which is a contradiction.

For uniqueness, suppose for the sake of contradiction that λ_1 and λ_2 both satisfy properties (1)-(3). Then

$$\lambda_1 = \lambda_2 - \alpha_1 \ldots - \alpha_N \qquad \alpha_i \in \Phi^+$$

$$\lambda_2 = \lambda_1 - \beta_1 - \ldots - \beta_M \qquad \beta_j \in \Phi^+$$

Equating these, we find that

$$\alpha_1 + \ldots + \alpha_N + \beta_1 + \ldots + \beta_M = 0$$

which contradicts positivity by evaluating on any H_0 in the dominant Weyl chamber.

To prove existence, suppose that μ_0 is a weight of π. If μ_0 satisfies (1), set $\lambda = \mu_0$. If not, we may choose a positive $\alpha \in \Phi^+$ such that $\mu_1 = \mu_0 + \alpha_1$ is a weight. If μ_1 satisfies (1), set $\lambda = \mu_1$; if not, continue. The process must stop because V is finite dimensional, so there are only finitely many weights. Eventually, we will produce a $\lambda = \mu_k$ which satisfies (1).

Let λ be as in (1)-(3). For $\alpha \in \Phi^+$, $\lambda + \alpha$ is not a weight. Hence $(\lambda, \alpha) \geq 0$ (by the $p + q$ thing), which verifies dominance.

The last thing to verify is that λ determines π up to isomorphism. As before, define subalgebras of $\mathfrak{g}$

$$\mathcal{L}_+ = \text{linear span of } \mathfrak{t} \text{ and } \mathfrak{g}^\alpha, \alpha \in \Phi^+$$

$$\mathcal{M}_- = \bigoplus_{\alpha \in \Phi^+} \mathfrak{g}^{-\alpha}.$$

So $\mathfrak{g} = \mathcal{L}_+ \oplus \mathcal{M}_-$. Define the $U(\mathfrak{g})$ module M_λ, where

$$M_\lambda = U(\mathfrak{g})/\text{left ideal generated by } \mathfrak{g}^\alpha, H - \langle \lambda, H \rangle 1 \qquad \alpha \in \Phi^+, H \in \mathfrak{t}.$$

This is called the "Verma module of highest weight λ." Let m_λ be the image of $1 \in M_\lambda$ under the quotient. Note that $m_\lambda \neq 0$. Why? We had expressed

$$U(\mathfrak{g}) = U(\mathcal{M}_-)U(\mathcal{L}_+).$$

We are essentially dividing out by $U(\mathcal{L}_+)$: the relations that lie in the augmentation ideal (the $\mathfrak{g}^\alpha$) obviously can't kill 1, since they raise degrees; the other relations take the constants to the Lie algebra of the torus, which doesn't intersect $\mathfrak{g}^\alpha$ for $\alpha \in \Phi^+$, hence also can't kill 1.

By definition, $\mathfrak{t}$ acts on m_λ by λ and the $\mathfrak{g}_\alpha$ for $\alpha \in \Phi^+$ kill m_λ, so

$$U(\mathcal{L}_+)m_\lambda = \mathbb{C}m_\lambda.$$

Then we claim that

$$\begin{aligned} U(\mathcal{M}_-) &\simeq M_\lambda \\ X &\mapsto Xm_\lambda \end{aligned}$$

as vector spaces. This is because M_λ is obviously generated by m_λ as a $\mathfrak{g}$ module (this is true even before quotienting), and $U(\mathfrak{g}) = U(\mathcal{M}_-)U(\mathcal{L}_+)$, so

$$M_\lambda \simeq U(\mathfrak{g})m_\lambda \simeq U(\mathcal{M}_-)U(\mathcal{L}_+)m_\lambda \simeq U(\mathcal{M}_-)m_\lambda.$$

We next claim that $\mathfrak{t}$ acts finitely, i.e. every $m \in M_\lambda$ lies in a finite dimensional, $\mathfrak{t}$-stable subspace. Indeed, for $X \in U(\mathcal{M}_-)$, $H \in \mathfrak{t}$,

$$\begin{aligned} HXm_\lambda &= (HX - XH + XH)m_\lambda \\ &= \operatorname{ad} H(X)m_\lambda + XHm_\lambda \\ &= \operatorname{ad} H(X)m_\lambda + \langle \lambda, H\rangle Xm_\lambda. \end{aligned}$$

So H preserves the degree (by which we mean the filtered components of the filtration). This also implies that $\mathfrak{t}$ acts semisimply (since it acts semisimply on the $\mathfrak{g}^\alpha$), with eigenvalues of the form $\lambda - \alpha_1 - \ldots - \alpha_N$, $\alpha_i \in \Phi^+$ (the λ coming from the second term in the equation above, and the other stuff from the first term.)

In particular, the weight λ occurs with multiplicity 1. Let $N \subset M_\lambda$ be a proper $U(\mathfrak{g})$ submodule. Then $\mathfrak{t}$ acts on N, with weights as above except λ cannot be a weight for N since m_λ generates all of M_λ.

So M_λ contains a unique maximal, proper, $U(\mathfrak{g})$ submodule (i.e. not M_λ, but conceivably zero), namely the linear span of all proper submodules. Therefore, M_λ has a unique irreducible, simple $U(\mathfrak{g})$ quotient. If V is an irreducible representation with highest weight λ, then note that the ideal defining M_λ annihilates V^λ by (2), so we get a surjective homomorphism of $U(\mathfrak{g})$ modules $M_\lambda \to V$ sending $m_\lambda \mapsto v_\lambda$. Hence V is this unique irreducible, simple quotient.

□

10.2 The Weyl character formula

Let G be a compact, connected Lie group, and T, Φ^+, Ψ be as usual. Define

$$\rho = \frac{1}{2}\sum_{\alpha\in\Phi^+} \alpha \in i\mathfrak{t}^*_{\mathbb{R}}.$$

2ρ is a sum of roots, hence $2\rho \in \Lambda_{\mathrm{Ad}}$, i.e. the $\mathbb{Z}$-linear span of Ψ.

Lemma 10.2.1. *We have*

1. *For $w \in W$, $w\rho - \rho \in \Lambda_{\mathrm{Ad}}$.*

2. $2\frac{(\rho,\alpha)}{(\alpha,\alpha)} = 1$ *for every* $\alpha \in \Psi$.

Proof. Choose $\alpha \in \Psi$ and let $\beta \in \Phi^+$, $\beta \neq \alpha$. Then

$$s_\alpha \beta = \beta - 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}\alpha \in \Phi.$$

We know that when β is expressed as a linear combination of simple roots, all coefficients are non-negative integers. Therefore, the same must be true of $s_\alpha\beta$ because some coefficient in the expression for $s_\alpha\beta$ is positive. That says

$$s_\alpha(\Phi^+ - \{\alpha\}) = \Phi^+ - \{\alpha\}.$$

(Obviously, $s_\alpha\alpha = -\alpha$.) So

$$\begin{aligned} s_\alpha\rho &= s_\alpha\left(\frac{1}{2}\sum_{\beta\in\Phi^+,\beta\neq\alpha}\beta + \frac{1}{2}\alpha\right) \\ &= \frac{1}{2}\sum_{\beta\in\Phi,\beta\neq\alpha}\beta - \frac{\alpha}{2} \\ &= \rho - \frac{\alpha}{2} - \frac{\alpha}{2} = \rho - \alpha. \end{aligned}$$

So we've found that $s_\alpha\rho = \rho - \alpha$. On the other hand, we know that

$$s_\alpha\rho = \rho - 2\frac{(\rho,\alpha)}{(\alpha,\alpha)}\alpha$$

so we can conclude that

$$2\frac{(\rho,\alpha)}{(\alpha,\alpha)} = 1 \text{ for all } \alpha \in \Psi.$$

This also proves that $s_\alpha\rho \in \rho + \Lambda_{\mathrm{Ad}}$ and s_α generate W, establishing (1). $\square$

The lemma implies that $\rho \in \Lambda_{\mathrm{sc}}$, since

$$\Lambda_{\mathrm{sc}} = \{\mu \in i\mathfrak{t}^*_{\mathbb{R}} \mid 2\frac{(\mu,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z} \text{ for all } \alpha \in \Psi\}.$$

In the following, we could take one of two courses:

(a) either use the fact that Λ_{sc} is the weight lattice of the universal covering of G, or

(b) go to a 2-fold cover of T so that $e^{2\rho} : T \to \mathbb{C}^*$ has a square root, namely e^ρ from the twofold cover. Then keep track of computations to see that they make sense.

We will do the second. Define the *formal* expression

$$\Delta = \prod_{\alpha \in \Phi^+} (e^{\alpha/2} - e^{-\alpha/2}) = \underbrace{e^\rho}_{\text{defined up to } \pm 1 \text{ on } T} \underbrace{\prod_{\alpha \in \Phi^+} (1 - e^{-\alpha})}_{\text{well-defined on } T}.$$

Weyl integration formula. Normalize the Haar measure on G, T to have total measure 1, so the "quotient measure" on G/T also has this property (the quotient measure satisfies the property that the measure of a set in the quotient is the measure of its pre-image). Then for $f \in C(G)$,

$$\boxed{\int_G f\, dg = \frac{1}{\# W} \int_T \int_{G/T} |\Delta(t)|^2 f(gtg^{-1})\, dg\, dt.}$$

Corollary 10.2.2. *If f is conjugation invariant, then*

$$\int_G f\, dg = \frac{1}{\# W} \int_T |\Delta(t)|^2 f(t)\, dt.$$

Weyl character formula. Let (π, V) be an irreducible representation for G of highest weight λ and χ_λ its character. Then

$$\boxed{\chi_\lambda|_T = \frac{\sum_{w \in W} \epsilon(w) e^{w(\lambda+\rho)}}{\Delta}.}$$

Moreover, each dominant $\lambda \in \Lambda$ is the highest weight of some finite dimensional irreducible (π, V).

Definition 10.2.3. For $w \in W$, we define

$$\epsilon(w) = \det\{w : \mathfrak{t}_\mathbb{R} \to \mathfrak{t}_\mathbb{R}\} = \pm 1.$$

The value is ± 1 because the Weyl group is finite dimensional, hence the determinant is a root of unity; but this is a map of real vector spaces.

Remark 10.2.4. Some observations:

1. We have $w(\lambda + \rho) = w\lambda + w\rho = w\lambda + w\rho - \rho + \rho$. Obviously $w\lambda \in \Lambda$, and Lemma 10.2.1 tells us that $w\rho - \rho \in \Lambda$, so we deduce that $w(\lambda + \rho) \in \Lambda + \rho$. Therefore, $e^{w(\lambda+\rho)}$ has an ambiguity in sign, but it is the same ambiguity as $\Delta = e^\rho \prod_{\alpha \in \Phi^+} (1 - e^{-\alpha})$, so the quotient is well-defined.

2. Observe that $\chi_\lambda|_T$ is a class function and T meets all conjugacy classes, so this completely determines χ.

3. A priori, the expression on the right hand side of the Weyl character formula might have "poles." However, this can only happen if $e^{-\alpha} = 1$, so the function is well-defined on
$$T_{\text{reg}} = \{t \in T \mid e^{\alpha}(t) \neq 1 \forall \alpha \in \Phi\},$$
which is open and dense in T.

10.3 Proof of Weyl character formula

For $i \in \widehat{G}$, let χ_i be the character. We know that $\{\chi_i \mid i \in \widehat{G}\}$ is an orthonormal basis of the space of L^2 conjugation-invariant functions. Define

$$\varphi_i = \chi_i|_T \cdot \Delta.$$

By the Weyl integration formula,

$$\int_T \varphi_i \overline{\varphi_j}\, dt = \#W \int_G \chi_i \overline{\chi_j}\, dg = \begin{cases} 0 & i \neq j \\ \#W & i = j \end{cases}.$$

Also, if f is an L^2 conjugation invariant function and $\varphi = f|_T \cdot \Delta$, then

$$\int_T \varphi \overline{\varphi}_i\, dt = 0 \text{ for all } i \in \widehat{G} \implies f = 0.$$

Fix $i \in \widehat{G}$ and let $\lambda \in \Lambda$ be the highest weight. We know that

$$(\lambda, \alpha) > 0 \text{ for all } \alpha \in \Phi^+$$
$$(\rho, \alpha) > 0 \text{ for all } \alpha \in \Phi^+$$

so $(\lambda + \rho, \alpha) > 0$ for all $\alpha \in \Phi^+$, i.e. $\lambda + \rho$ lies in the interior of the dominant Weyl chamber in $i\mathfrak{t}_{\mathbb{R}}^*$. In particular, the terms $w(\lambda + \rho)$ for $w \in W$ are distinct because W acts freely on the interiors of the Weyl chambers.

We obviously have the identity

$$\varphi_i = \left(\sum_{\mu \in \Lambda} \dim V^{\mu} e^{\mu} \right) e^{\rho} \prod_{\alpha \in \Phi^+} (1 - e^{-\alpha}).$$

The highest weight appears in the left term, with multiplicity one. So we can write this as
$$\varphi_i = e^{\lambda+\rho} + \text{ (an integral linear combination of terms) } e^{\mu+\rho}$$
where $\mu = \lambda -$ sum of positive roots. In particular, the $e^{\lambda+\rho}$ term certainly doesn't get canceled.

Note that $\chi_i|_T$ is W-invariant, since W acts on the torus by conjugation. We had seen that for any *simple* root $\alpha \in \Psi$, $s_\alpha(\Phi^+ - \{\alpha\}) = \Phi^+ - \{\alpha\}$ and $s_\alpha(\alpha) = -\alpha$ (i.e. s_α preserves all the positive roots except α, which it negates). Therefore

$$s_\alpha \Delta = s_\alpha \prod_{\beta \in \Phi^+} \left(e^{\beta/2} - e^{-\beta/2}\right) = -\Delta$$

because exactly one factor changes sign. This implies that Δ is "W-alternating" in the sense that $w\Delta = \epsilon(W)\Delta$ for any $w \in W$. Then φ_i is W-alternating, so

$$\varphi_i = \sum_{w \in W} \epsilon(w) e^{w(\lambda+\rho)} + \substack{\text{integral linear combination of other terms} \\ \text{with no cancellation}}.$$

By orthogonality,

$$\int_T |\varphi_i|^2 \, dt = \#W + \int |\text{other terms}|$$

but we also know that

$$\int_T |\varphi_i|^2 \, dt = \#W.$$

Therefore, the other terms must vanish. This proves the Weyl character formula: if χ is the character of the irreducible representation with highest weight λ, then

$$\chi|_{T_{\text{reg}}} = \frac{\sum_{w \in W} \epsilon(w) e^{w(\lambda+\rho)}}{\prod_{\alpha \in \Phi^+} (e^{\alpha/2} - e^{-\alpha/2})}.$$

This can also be written as

$$\chi|_{T_{\text{reg}}} = \frac{\sum_{w \in W} \epsilon(w) e^{w(\lambda+\rho)-\rho}}{\prod_{\alpha \in \Phi^+} (1 - e^{-\alpha})},$$

which shows it to be well-defined.

Applying this to the trivial representation, where the highest weight is $\lambda = 0$, we obtain what is called the *Weyl denominator formula*.

Corollary 10.3.1 (Weyl denominator formula)**.**

$$\prod_{\alpha \in \Phi^+} (e^{\alpha/2} - e^{-\alpha/2}) = \sum_{w \in W} \epsilon(w) e^{w\rho}.$$

Remark 10.3.2. The Weyl character formula also implies the following statement which we previously proved by algebraic means: the highest weight of an irreducible representation π determines π up to isomorphism.

We had stated this proposition earlier.

Proposition 10.3.3. *If $\lambda \in \Lambda$ is dominant, i.e. $(\lambda, \alpha) \geq 0$ for each $\alpha \in \Phi^+$, then λ is the highest weight of some irreducible representation.*

Proof. Let λ_i be the highest weight of π_i, $\lambda \neq \lambda_i$ for all $i \in \widehat{G}$. Define

$$\varphi = \sum_{w \in W} \epsilon(w) e^{w(\lambda+\rho)}$$

There's no cancellation here because the $w(\lambda+\rho)$ are all distinct (as mentioned at the beginning of the lecture). Also, φ is W-alternating, hence $\frac{\varphi}{W}$ is W-invariant. Then there exists a unique W-invariant (conjugation-invariant, since the group of deck transformations for the covering in the section below is precisely W) function $f \in C(G_{\text{reg}})$ such that $f|_{T_{\text{reg}}} = \frac{\varphi}{\Delta}$. Since $G - G_{\text{reg}}$ has measure 0, we can regard f as an element of $L^2(G)$.

By the Weyl integration formula,

$$\begin{aligned}
\int_G f \overline{\chi}_i \, dg &= \frac{1}{\#W} \int_T f \Delta \overline{\chi_i \Delta} \, dt \\
&= \frac{1}{\#W} \int \sum_{w,v \in W} \epsilon(w) e^{w(\lambda+\rho)} \epsilon(v) e^{-v(\lambda_i+\rho)} \, dt \\
&= 0 \text{ by orthogonality.}
\end{aligned}$$

□

10.4 Proof of the Weyl integration formula.

We know that the map

$$\begin{aligned}
G/T \times T_{\text{reg}} &\xrightarrow{F} G_{\text{reg}} \\
(gT, t) &\mapsto gtg^{-1}
\end{aligned}$$

is a covering map with group of deck transformations equal to $W = N_G(T)/T$. Therefore,

$$\begin{aligned}
\int_G f(g) \, dg &= \frac{1}{\#W} \int_{G/T} \int_T F^* f \, dt \, dg \\
&= \frac{1}{\#W} \int_{G/T} \int_T |\det F_*| f(gtg^{-1}) \, dt \, dg
\end{aligned}$$

Left multiplication by g gives a map

$$T_{gT} G/T \leftarrow T_{eT} G/T \simeq \mathfrak{g}/\mathfrak{t} \simeq \bigoplus_{\alpha \in \Phi} \mathfrak{g}^\alpha$$

and $T_t T \simeq \mathfrak{t}$ (this is the tangent space to T at t). So

$$T_{(gT,t)}(G/T \times T_{\text{reg}}) \simeq \bigoplus_{\alpha \in \Phi} \mathfrak{g}^\alpha \oplus \mathfrak{t} \simeq \mathfrak{g} \simeq T_{gtg^{-1}} G_{\text{reg}}.$$

These identifications we made are a little sketchy, but the measure is left invariant, etc. so things are OK.

Suppose $X \in \bigoplus_{\alpha\in\Phi} \mathfrak{g}^\alpha \cap \mathfrak{g}_\mathbb{R}$ and let $f \in C^\infty(G)$. We want to compute F_*; let's trace through the identifications made above.

$$F(g\exp(sX)T, t) = g\exp(sX)t\exp(-sX)g^{-1} = g\exp(sX)\exp(-s\,\mathrm{Ad}\ tX)tg^{-1}.$$

So

$$\begin{aligned} F_*(X,0)f(gtg^{-1}) &= \frac{\partial}{\partial s} f(g\exp(sX)\exp(-s\,\mathrm{Ad}\ tX)tg^{-1})|_{s=0} \\ &= \underbrace{\ell(-\mathrm{Ad}\ g(1-\mathrm{Ad}\ t)X)}_{\text{left infinitesimal translation by ...}}\ f(gtg^{-1}) \end{aligned}$$

Recall that if $f \in C^\infty(G)$ and $X \in \mathfrak{g}_\mathbb{R}$, then

$$r(X)f(g) = \frac{\partial}{\partial s} f(g\exp(sX))|_{s=0}.$$

and

$$\ell(X)f(g) = \frac{\partial}{\partial s} f(\exp(-sX)g)|_{s=0}.$$

For $Y \in \mathfrak{t}_\mathbb{R}$,

$$F(gT, \exp(sY)t) = g\exp(sY)tg^{-1}.$$

Hence

$$F_*(0,Y)f(gtg^{-1}) = \frac{\partial}{\partial s} f(g\exp(sY)tg^{-1})|_{s=0} = \ell(-\mathrm{Ad}\ g(Y))f(gtg^{-1}).$$

That means that

$$\mathrm{Ad}\ g^{-1}\circ F_* : \begin{cases} X \in \bigoplus_{\alpha\in\Phi}\mathfrak{g}^\alpha & \mapsto (1-\mathrm{Ad}\ t)X \\ Y \in \mathfrak{t} & \mapsto Y \end{cases}$$

so

$$\begin{aligned} \det(\mathrm{Ad}\ g^{-1}\circ F_*(t)) &= \prod_{\alpha\in\Phi}(1-e^\alpha(t)) \\ &= \prod_{\alpha\in\Phi^+}(1-e^\alpha(t))(1-e^{-\alpha}(t)) \\ &= |\Delta(t)|^2. \end{aligned}$$

10.5 Borel-Weil

Suppose (π, V) is irreducible of highest weight λ. The theorem of the highest weight implies that

$$V^{\mathfrak{N}_+}(= \text{ space of } \mathfrak{N}_+\text{-invariant vectors}) = V^{\lambda}.$$

where $\mathfrak{N}_+ = \bigoplus_{\alpha\in\Phi^+} \mathfrak{g}^{\alpha}$ (for groups, we look at the elements that map to the identity. For Lie algebras, we look at things that get mapped to 0). This is an equality of vector spaces with an action of T (or $\mathfrak{t}$) since T normalizes $\mathfrak{N}_+$. Also let

$$\mathfrak{N}_- = \bigoplus_{\alpha\in\Phi^+} \mathfrak{g}^{-\alpha}.$$

Dually we have

$$(V^*)^{\mathfrak{N}_-} = (V^*)^{-\lambda}.$$

By Peter-Weyl,

$$V = \{f \in \widehat{\bigoplus_{i\in\widehat{G}} V_i} \otimes (V_{i^*}^{\mathfrak{N}_-})^{-\lambda}\}$$

i.e. all functions f on G on which T acts via right translation by $e^{-\lambda}$. What this says is:

$$V \simeq \left\{ f \in C^{\infty}(G) \mid \substack{r(X)f=0 \text{ for all } X\in\mathfrak{N}_- \\ f(gt)=e^{-\lambda}(t)f(g)\forall t\in T} \right\}.$$

Facts: G/T can be turned into a complex manifold such that

(a) G acts on G/T holomorphically.

(b) The holomorphic tangent space to G/T at eT is isomorphic to

$$\mathfrak{N}_+ \subset \bigoplus_{\alpha\in\Phi} \mathfrak{g}^{\alpha} \simeq T_{\mathrm{et}} G/T$$

(Note that $\mathfrak{N}_+$ is in the complexified Lie algebra.)

(c) If $\lambda \in \Lambda$, there exists a unique (up to isomorphism) holomorphic line bundle $\mathscr{L}_X \to G/T$ such that (i) the action of G on G/T lifts to $\mathscr{L}_\lambda$ and (ii) T acts on the fiber of $\mathscr{L}_\lambda$ at eT by e^{λ}. So G acts on the space of holomorphic sections of $\mathscr{L}_\lambda$.

Theorem 10.5.1 (Borel-Weil). *For $\lambda \in \Lambda$ dominant,*

$$H^0(G/T, \mathscr{O}(\mathscr{L}_\lambda))$$

is the space of irreducible representations of highest weight λ and

$$H^p(G/T, \mathscr{O}(\mathscr{L}_\lambda)) = 0 \text{ for } p > 0.$$

Remark 10.5.2. The Borel-Weil-Bott theorem gives a description of $H^p(G/T, \mathscr{O}(\mathscr{L}_\lambda))$ for all $\lambda \in \Lambda$ and integers p.