COMPACT LIE GROUPS

NICHOLAS ROUSE

Abstract. The first half of the paper presents the basic definitions and results necessary for investigating Lie groups. The primary examples come from the matrix groups. The second half deals with representation theory of groups, particularly compact groups. The end result is the Peter-Weyl theorem.

Contents

1. Smooth Manifolds and Maps

Manifolds are spaces that, in a certain technical sense defined below, look like Euclidean space.

Definition 1.1. A **topological manifold of dimension n** is a topological space X satisfying:

(1) X is second-countable. That is, for the space's topology T, there exists a countable base, which is a countable collection of open sets $\{B_\alpha\}$ in X such that every set in T is the union of a subcollection of $\{B_\alpha\}$.
(2) X is a Hausdorff space. That is, for every pair of points $x, y \in X$, there exist open subsets $U, V \subseteq X$ such that $x \in U$, $y \in V$, and $U \cap V = \varnothing$.
(3) X is locally homeomorphic to open subsets of $\mathbb{R}^n$. That is, for every point $x \in X$, there exists a neighborhood $U \subseteq X$ of x such that there exists a continuous bijection with a continuous inverse (i.e. a homeomorphism) from U to an open subset of $\mathbb{R}^n$.

The first two conditions preclude pathological spaces that happen to have a homeomorphism into Euclidean space, and force a topological manifold to share more properties with Euclidean space. Indeed, Euclidean space itself is Hausdorff and second-countable. The third condition is the most important. It relates the

Date: August 29, 2014.

manifold to Euclidean space in a way that is amenable to the techniques of multivariable calculus. Of course, merely have a function that maps into Euclidean space is insufficient. We wish to impose additional structure (namely a notion of smoothness or infinite differentiability) on the homeomorphisms and hence the manifold, but first we need some more machinery to describe the homeomorphisms.

As a note on dimensionality, the definition as written requires a well-defined dimension. However, the definition leaves open the possibility that a space could be an n-dimensional and an m-dimensional manifold. A bit of sophisticated machinery that we will not develop shows this case to be impossible in the general case. However, the fact follows much more easily in the smooth case, as we shall show in the next section.

Definition 1.2. A **chart** on a manifold M is a open set $U \subseteq M$ together with a homeomorphism φ from U to an open subset of $\mathbb{R}^n$. We can write the chart as an ordered pair (U, φ). In this context the open set U is called a **coordinate domain**, and the homeomorphism ϕ is called a **coordinate map**.

Definition 1.3. An **atlas** for a manifold M is a collection of charts such that every point in M is in the coordinate domain for some chart.

Atlases are a structure to describe the entire manifold in terms of charts. It is a slight but common abuse of terminology to speak of a manifold's being covered by charts even though technically the coordinate domains cover the manifold. We are now in a position to discuss smooth structure on manifolds, but first we recall smoothness in the Euclidean context.

Definition 1.4. A **diffeomorphism** is a homeomorphism φ from U, an open subset of $\mathbb{R}^n$, to V, an open subset of $\mathbb{R}^n$ such that all orders of partial derivatives of φ and φ^{-1} are continuous. More generally, $f : U \to V$, for U, V open subsets of Euclidean space, is called **smooth** if all orders of partial derivatives are continuous.

It's important to note that this definition depends explicitly on Euclidean space. There is no obvious way to generalize partial derivatives to maps between arbitrary topological spaces such as manifolds, and we shall make no attempt to. Instead, we make our smoothness condition that compositions of coordinate maps and their inverses are smooth whenever they happen to map to and from Euclidean space.

Definition 1.5. A **smooth atlas** is an atlas such that for any two charts (U, φ), (V, ψ), $\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V)$ is a diffeomorphism. Two such charts are said to be **smoothly compatible**.

There is a technical detail that smooth atlases are not necessarily unique. There could be many smoothly compatible coordinate maps in different atlases. This consideration leads us to define a maximal atlas.

Definition 1.6. A **maximal atlas** on a manifold M is an atlas that is not a proper subset of another atlas on M. That is, any chart that is smoothly compatible with all the charts in a maximal atlas is in fact an element of that atlas.

We now have the necessary machinery to define a smooth structure on a manifold, and hence, a smooth manifold.

Definition 1.7. A **smooth manifold** is a manifold M equipped with a maximal smooth atlas $\mathcal{A}$ on M. We can write a smooth manifold as an ordered pair $(M, \mathcal{A})$. $\mathcal{A}$ is called a **smooth structure on M**.

Note, however, that finding a single smooth atlas suffices to show that a manifold is smooth because such a smooth atlas is necessarily contained in some maximal smooth atlas. Now we can generalize the notion of smooth functions between Euclidean spaces to those between manifolds.

Definition 1.8. A map f from a smooth manifold M to a smooth manifold N is called a **smooth map between manifolds** if for every point $p \in M$, there exist charts $(U,\phi),(V,\psi)$ such that $p \in U$, $f(U) \subseteq V$ (i.e the image of U under f is contained in V), and the map $\psi \circ f \circ \phi^{-1} : \phi(U) \to \psi(V)$ is smooth. We denote the set of all smooth, real-valued functions with domain M, a manifold, by $C^\infty(M)$.

2. Tangents, Differentials, and Submersions

Having generalized notions of smoothness, we want to talk about derivatives of maps between smooth manifolds. In the most basic sense, we think of a derivative as linear approximations. The problem, of course, is that manifolds lack the vector space structure to define linear maps. We remedy this problem by introducing the tangent space to a manifold.

Definition 2.1. Let $f,g \in C^\infty(M)$. A **derivation** at a point $p \in M$ is a linear map $v : C^\infty(M) \to \mathbb{R}$ (here regarding $C^\infty(M)$ as a vector space over $\mathbb{R}$) that satisfies the product rule

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

The set of all such derivations is called the **tangent space to M at p** and is denoted T_pM. An element of T_pM is called a **tangent vector at p**.

It is easy to verify that the tangent space at any point of a manifold is in fact a vector space. Less obvious is that such a vector space is finite-dimensional. More remarkable is that there is an easy way to find a basis for this space. The details establishing these facts aren't particularly difficult, but they are a little long (see, for example, [1, pp. 51-57]). The basis is given by partial derivatives of the chart at whatever point the tangent space is being calculated at. More formally, for an n dimensional manifold with a chart (U,φ) and $p \in U$, define $\partial/\partial x^i|_p : C^\infty(U) \to \mathbb{R}^n$ by

$$\left.\frac{\partial}{\partial x^i}\right|_p f = \frac{\partial(f \circ \varphi^{-1})}{\partial x^i}(\varphi(p)).$$

Then, the set of maps $\{\partial/\partial x^i|_p\}_{i=1}^n$ forms a basis for T_pM. With this in mind it is natural to define differentials as linear maps between the tangent spaces.

Definition 2.2. Let M and N be smooth manifolds, $p \in M$, $v \in T_pM$, and $F : M \to N$ be a smooth map. The **differential of F at p** is a map $dF_p : T_pM \to T_{F(p)}N$. Given a derivation $v \in T_pM$, we define the map $dF_p(v)$ as the derivation at $F(p)$ that maps f to $v(f \circ F)$ for $f \in C^\infty(N)$. That is

$$dF_p(v)(f) = v(f \circ F).$$

It turns out that there are actually things we can prove with all this structure. We mentioned earlier that proving that a topological space can only be one dimension of manifold requires more machinery than we develop here. However, if the space is endowed with a smooth structure, we get this fact quite easily.

Proposition 2.3. *An n-dimensional smooth manifold cannot be locally diffeomorphic to an m-dimensional smooth manifold unless $n = m$.*

Proof. The proof depends upon a version of the chain rule. Namely, $d(G \circ F)_p = dG_{F(p)} \circ dF_p$. To see this we pick arbitrary $v \in T_pM$, $f \in C^\infty(U)$ for p and U as in the definition of differential. Then we have

$$\begin{aligned} d(G \circ F)_p(v)(f) &= v((f \circ G) \circ F) \\ &= dF_p(v)(f \circ G) \end{aligned}$$

Now we point out that since $dF_p : T_pM \to T_{F(p)}M$, we have that $dF_p(v)$ is a derivation at $F(p)$. Then

$$dF_p(v)(f \circ G) = dG_{F(p)} \circ dF_p(v)(f).$$

As v and f are arbitrary, we have the desired property. We need also an easy property of the identity map on manifolds $Id_M : M \to M, p \mapsto p$.

$$\begin{aligned} d(Id_M)_p(v)(f) &= v(f \circ Id_M) \\ &= v(f) \\ &= Id_{T_pM}(v)(f) \end{aligned}$$

Now we have that the chart φ with U as its coordinate domain is a diffeomorphism by assumption, so

$$\begin{aligned} d(\varphi \circ \varphi^{-1})_p &= d\varphi_{\varphi^{-1}(p)} \circ d\varphi_p^{-1} \\ &= Id_{T_pM} \end{aligned}$$

Then the differential of φ is an invertible linear map, which is only possible when the dimension of its range and image are identical. □

We often want to say things about submanifolds (i.e. subsets of a manifold that are manifolds in their own right), and the differential provides a tool for such investigations. We now confine ourselves to manifolds that are subsets of some ambient Euclidean space. This turns out not be a restriction at all since the Whitney Embedding Theorem states that every manifold can be embedded in some sufficiently high-dimensional Euclidean space, so every result we state actually holds for all manifolds. We make no use of this fact, however, because we are only interested in submanifolds of real and complex matrix groups, which are explicitly subsets of Euclidean space. We start with an exceedingly useful theorem that we shall not prove.

Theorem 2.4 (Inverse Function Theorem)**.** *If $f : X \to Y$ is a smooth map between manifolds whose derivative df_x at the point x is an isomorphism, then f is a local diffeomorphism.*

A local diffeomorphism is one only defined on the neighborhood of some point. The next few results will be the key tool for identifying submanifolds in the coming sections.

Definition 2.5. A smooth map between manifolds $F : M \to N$ is said to be a **submersion** at a point $x \in M$ if dF_x is surjective.

Theorem 2.6 (Local Submersion Theorem)**.** *If $f : X \to Y$, for $X \subseteq \mathbb{R}^k$, $Y \subseteq \mathbb{R}^l$, and $k \geq l$, is a submersion at x, then there exist local coordinate systems around x and $f(x)$ such that $f(x_1, ..., x_k) = (x_1, ..., x_l)$, which is called the canonical submersion.*

Proof. Since both X and Y are manifolds, we have local parametrizations, φ and ψ around x and $f(x)$, respectively. For convenience, let $\varphi(0) = x$ and $\psi(0) = f(x)$. Now define $g = \phi^{-1} \circ f \circ \varphi$, so $g : U \to V$. Since $dg_0 : \mathbb{R}^k \to \mathbb{R}^l$ is surjective, we can change our basis so that it has a $l \times k$ Jacobian matrix where the upper left $l \times l$ block matrix is the $l \times l$ identity matrix and every other entry is zero. We now define $G : U \to \mathbb{R}^k, a \mapsto (g(a), a_{l+1}, ..., a_k)$. Here $a = (a_1, ..., a_k)$, so $g(a) = (g_1(a), ..., g_l(a))$. Now we have that dG_0 is just the $k \times k$ identity matrix, which is an isomorphism. Therefore, G is a local diffeomorphism at 0. Then there also exists an inverse diffeomorphism G^{-1} of an open neighborhood U' of 0 into U. If we denote the canonical submersion by h, we have $g = h \circ G$, so $h = g \circ G^{-1}$, and we have the desired local coordinates to make f locally equivalent to h, the canonical submersion. □

Definition 2.7. Let $f : X \to Y$ be a smooth map between manifolds. Then a point $y \in Y$ is a **regular value** for f if f is a submersion at every point x such that $f(x) = y$.

Theorem 2.8 (Preimage Theorem)**.** *If $f : X \to Y$ is a smooth map between manifolds, and y is a regular value of f, then the preimage $f^{-1}(y)$ is a submanifold of X, and* $\dim f^{-1}(y) = \dim X - \dim Y$.

Proof. Let $x \in f^{-1}(y)$. We use the local submersion theorem to choose local coordinates around x and y such that $f(x_1, ..., x_k) = (x_1, ..., x_l)$ and y corresponds to $(0, ..., 0)$ (i.e. the parametrization around y takes it to the origin). Then, locally, $f^{-1}(y) = (0, ..., 0, x_{l+1}, ..., x_k)$. Therefore, there are local coordinates for $f^{-1}(y)$ as a submanifold of $\mathbb{R}^{k-l}$. □

We now turn our attention to the smooth manifolds we're interested in, Lie groups. These manifolds also come with a group structure, which we will define in the next section. We are going to focus on a particular but important class of Lie groups, the matrix groups.

3. Lie Groups

We begin with elementary definitions about groups.

Definition 3.1. A **group** is a set G equipped with a binary operation $\cdot : G \times G \to G$ (we write the operation as $g \cdot h$ or gh rather than $\cdot(g, h)$ for $g, h \in G$) satisfying

1. $\cdot$ is associative.
2. There exists an identity element $e \in G$ such that for all $g \in G$, $e \cdot g = g \cdot e = g$.
3. For every $g \in G$, there exists an inverse element $g^{-1} \in G$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$.

We are often interested in talking about smaller groups contained in a larger group. We make this notion precise with a subgroup.

Definition 3.2. A subset H is a group G is **subgroup** of G if H is nonempty and closed under products and inverses.

A subgroup is of course a group in its own right. In later sections we will seek to describe complicated groups in terms of simpler ones. We do this with homomorphisms.

Definition 3.3. Let G and H be groups. A **group homomorphism** is a function $\varphi : G \to H$ such that for any all $g, h \in G$, $\varphi(gh) = \varphi(g)\varphi(h)$.

Of course, the juxtaposition may imply different binary operations in the different groups.

Definition 3.4. A **Lie group** G is a group that is also a smooth manifold where the group binary operation $\cdot : G \times G \to G, g, h \mapsto g \cdot h$ and the inversion map $i : G \to G, g \mapsto g^{-1}$ are both smooth maps between manifolds.

In general, Lie groups can be quite complicated, but we focus on groups of matrices, which are more tractable. We shall eventually see that for compact Lie groups, this is no restriction at all.

Definition 3.5. $M(n, \mathbb{R})$ is the set of all $n \times n$ matrices with real entries. Similarly, $M(n, \mathbb{C})$ is the set of all $n \times n$ matrices with complex entries.

Note that $M(n, \mathbb{R})$ and $M(n, \mathbb{C})$ are nothing more than $\mathbb{R}^{n^2}$ and $\mathbb{R}^{2n^2}$, respectively, so they are therefore trivially manifolds.

Definition 3.6. $GL(n, \mathbb{R})$ is the **general linear group of degree n** over $\mathbb{R}$ and consists of all invertible $n \times n$ matrices with real entries. $GL(n, \mathbb{C})$ is defined similarly. More generally, we write $GL(V)$ for linear transformations for V a finite-dimensional real or complex vector space.

Proposition 3.7. *$GL(n, \mathbb{R})$ is a Lie group of dimension n^2 with matrix multiplication as the group operation.*

Proof. We immediately see that $GL(n, \mathbb{R})$ is Hausdorff and second-countable since it's just a subspace of $\mathbb{R}^{n^2}$. Moreover, simple verification shows that the group axioms are satisfied. Now we observe that the determinant is just a smooth map, $\det : M(n, \mathbb{R}) \to \mathbb{R}$. Then $\det^{-1}(\{\mathbb{R}\backslash 0\}) = GL(n, \mathbb{R})$, and $GL(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$, so the homeomorphism into Euclidean space (and in fact diffeomorphism) is just the inclusion map from the general linear group into Euclidean space. Finally, multiplication and inversion are both smooth maps. In particular, multiplication is just a polynomial map in each entry, so it is smooth. Inversion is given by Cramer's rule, which is ultimately a rational map. Since $GL(n, \mathbb{R})$ is open (so we don't have issues taking derivatives), and the denominator is just the determinant, so we conclude that the inversion map is smooth. □

Definition 3.8. $SL(n, \mathbb{R})$ is the **special linear group of degree n** over $\mathbb{R}$ and consists of all $n \times n$ matrices with determinant 1. $SL(n, \mathbb{C})$ is defined similarly.

Some of the most developed theory of Lie groups exists for compact Lie groups. For completeness we recall the definition of compactness and the Heine-Borel theorem, though the proof, while not hard, does not concern us.

Definition 3.9. For a topological space X, a collection of open sets $\{U_\alpha\}_{\alpha \in A}$ (where A is just some indexing set) is an **open cover** for $E \subseteq X$ if $E = \cup_{\alpha \in A} U_\alpha$. E is **compact** if every open cover has a finite subcover. That is for every open cover $\{U_\alpha\}_{\alpha \in A}$ there is some indexing set $B \subseteq A$ such that B is finite, and $\{U_\alpha\}_{\alpha \in B}$ is an open cover for E.

Theorem 3.10 (Heine-Borel Theorem). *A subset of Euclidean space, $E \subset \mathbb{R}^n$ is compact if and only it is closed and bounded.*

For now we are concerned with subsets of Euclidean space, so we use Heine-Borel because identifying closed and bounded sets is easier than working with covers. However, in section 6 we will use the cover definition.

Definition 3.11. $O(n)$ is the **orthogonal group of degree n** and is the set of all $n \times n$ matrices with real entries whose inverse is its transpose. That is $O(n) = \{A \in M(n, \mathbb{R}) : A^T A = AA^T = I\}$. An element of $O(n)$ is called an **orthogonal matrix**.

Proposition 3.12. *$O(n)$ is a compact Lie group of dimension $n(n-1)/2$.*

Proof. We wish to use the preimage theorem (Theorem 2.8). First we define $S(n)$ to be the space of all real symmetric matrices. Clearly, $S(n)$ can be identified (i.e. there exists a diffeomorphism) with $\mathbb{R}^{n(n+1)/2}$. Now we define the map $f : M(n, \mathbb{R}) \to S(n), A \mapsto AA^T$. Note that AA^T is in fact symmetric and that $O(n) = f^{-1}(I)$. We now wish to compute the differential. Here, we're actually just in Euclidean space, so we can use a difference quotient.

$$\begin{aligned} df_A(B) &= \lim_{h \to 0} \frac{f(A + hB) - f(A)}{h} \\ &= \lim_{h \to 0} \frac{AA^T + hBA^T + hAB^T + h^2 BB^T - AA^T}{h} \\ &= BA^T + AB^T \end{aligned}$$

As we observed before, the tangent spaces to Euclidean spaces are identical to the Euclidean spaces themselves, so we need to show $df_A : M(n, \mathbb{R}) \to S(n), B \mapsto BA^T + AB^T$ is surjective for all $A \in O(n)$. That is, we want to show that the identity is a regular value. This means that for every symmetric matrix C, there is some matrix B such that $df_A(B) = C$. Our computation of the derivative is suggestive of how to find the desired matrix. Since $(BA^T)^T = AB^T$, we observe that for a symmetric matrix C, we actually only need $BA^T = \frac{1}{2}C$ because if we have this, we immediately have $\frac{1}{2}C = \frac{1}{2}C^T = (BA^T)^T$. We can then solve for B to obtain $B = \frac{1}{2}CA$ (using orthogonality of A). Indeed,

$$df_A(\frac{1}{2}CA) = (\frac{1}{2}CA)A^T + A(\frac{1}{2}CA)^T = \frac{1}{2}CAA^T + \frac{1}{2}AA^T C = \frac{1}{2}(C + C^T) = C.$$

So I is a regular value of f, so $O(n)$ is a manifold, and the dimension is $\dim M(n, \mathbb{R}) - \dim S(n) = n^2 - n(n+1)/2 = n(n-1)/2$.

To see compactness, we first point out that $\det^{-1}(\{1\} \cup \{-1\}) = O(n)$ for the continuous map $\det : M(n, \mathbb{R}) \to \mathbb{R}$, so $O(n)$ is closed. To see bounded, we use the fact that a matrix is orthogonal if and only if its row vectors form an orthonormal basis for $\mathbb{R}^n$. This fact can be verified with a direct, though somewhat laborious, calculation. Since the row vectors form an orthonormal basis, we have that $|a_{ij}|^2 \leq 1$ for every matrix entry a_{ij}. We then have that $O(n)$ is bounded, and therefore, compact. □

Definition 3.13. $SO(n)$ is the **special orthogonal group of degree n** and is the set of orthogonal matrices with determinant 1. That is $SO(n) = \{A \in O(n) : \det A = 1\}$.

Proposition 3.14. *$SO(n)$ is a compact Lie group.*

Proof. Consider the continuous function, $\det : M(n, \mathbb{R}) \to \mathbb{R}$. $\det^{-1}(\{1\}) = SL(n, \mathbb{R})$. In particular, the real special linear group is closed because it is the preimage of a closed set under a continuous function. We already saw that $O(n)$ is compact, and hence closed set. Then $SL(n, \mathbb{R}) \cap O(n) = SO(n)$ is closed. In fact, it is a closed subset of the compact set $O(n)$, so $SO(n)$ is compact.

$SO(n)$ is also an open subset of $O(n)$. To see this, we recall that $\det : O(n) \to \{-1, +1\}$ is continuous. Then $\{+1\}$ is open as a subset of $\{-1, +1\}$. Then the preimage of $+1$ under the determinant is open and is just $SO(n)$. It follows that $SO(n)$ is a smooth submanifold, and the group properties are immediate, so $SO(n)$ is a compact Lie group. □

Definition 3.15. $U(n)$ is the **unitary group of degree n** and is the set of all $n \times n$ matrices with complex entries whose inverses are their conjugate transpose (also called the Hermitian conjugate and obtained by taking the ordinary transpose and then taking the complex conjugate of every entry). In set-builder notation, $U(n) = \{A \in M(n, \mathbb{C}) : A^*A = AA^* = I\}$. An element of $U(n)$ is called a **unitary matrix**.

Proposition 3.16. *$U(n)$ is a compact Lie group of dimension $n^2 - 1$.*

Proof. The argument is exactly the same as that for $O(n)$ with conjugate transpose in place of transpose. The only appreciable difference is that the space of matrices that are equal to their conjugate transpose (such matrices are called Hermitian matrices) has dimension $n^2 + 1$, and the space of all matrices with complex entries has dimension $2n^2$, so the dimension of $U(n)$ is $n^2 - 1$. □

$SU(n)$ is the **special unitary group of degree n** and consists of all unitary matrices with determinant 1. That is, $SU(n) = \{A \in U(n) : \det A = 1\}$.

Proposition 3.17. *$SU(n)$ is a compact Lie group.*

Proof. The argument is analogous to that showing $SO(n)$ is compact. □

Showing that $SU(n)$ is in fact a Lie group is a bit harder, and we omit the details for space. The proof essentially relies on showing that 1 is a regular value of the determinant map.

4. Group Representations

In general Lie groups can be quite complex. A goal of representation theory is to describe Lie groups in terms of homomorphisms into the general linear group, a well-understood group. This approach is most fruitful in the case of compact Lie groups, and we shall specialize to them when appropriate. We concern ourselves only with finite dimensional representations.

Definition 4.1. A **group representation** of a Lie group G is a smooth homomorphism Φ of G into $GL(V)$ for V a finite-dimensional complex vector space. Such a representation is often phrased as a representation of G on V.

When we write $\Phi(g)$, we have a matrix (or, if you prefer, a linear transformation) in mind, so an expression like $\Phi(g)(v)$ for $v \in V$ is just a vector transformed by $\Phi(g)$.

Definition 4.2. Two representations of G, Φ on V and Φ' on V' are **equivalent** if there is an invertible linear map $T : V \to V'$ such that $\Phi'(g)T = T\Phi(g)$ for all $g \in G$.

Definition 4.3. An **invariant subspace** U for a representation $\Phi(g)$ is one such that for every $u \in U$, $\Phi(g)(u) \in U$ for all $g \in G$.

Obviously V itself and the zero space, 0, are both invariant subspaces. In fact, representations that only have these invariant subspaces are of great importance.

Definition 4.4. A representation Φ of a group G on a nonzero, finite dimensional, vector space V is said to be **irreducible** if the only invariant subspaces are V and 0.

We also wish to generalize our notion of unitary operators.

Definition 4.5. The **adjoint**, A^*, of a linear transformation A on a nonzero, finite dimensional, complex inner product space with respect to a Hermitian inner product $\langle \cdot, \cdot \rangle$ is the unique transformation satisfying $\langle Av, w \rangle = \langle v, A^* w \rangle$ for all $v, w \in V$. As before such a transformation is **unitary** if $AA^* = I$. Regarding $\Phi(g)$ as a linear transformation then, we say that a representation Φ of G on V is unitary with respect to a Hermitian inner product if $\Phi(g)^*\Phi(g) = I$ for all $g \in G$.

The main reason to depart from the concrete notion of conjugate transpose is that we will actually define our Hermitian inner product to make an arbitrary representation on a compact group unitary. An exceedingly useful fact is that any invariant subspace also implies the existence of another.

Proposition 4.6. *For a unitary representation G on V the orthogonal complement (denoted $U^\perp$) to an invariant subspace is also an invariant subspace.*

Proof. Recall that the orthogonal complement to a vector subspace $U \subseteq V$ equipped with an inner product $\langle \cdot, \cdot \rangle$ is defined to be $U^\perp = \{u^\perp \in V : \langle u^\perp, u \rangle = 0 \, \forall u \in U\}$. Now take $u^\perp \in U^\perp$, $u \in U$, and consider $\langle \Phi(g)u^\perp, u \rangle = \langle u^\perp, \Phi(g)^* u \rangle$. Of course we are considering whatever Hermitian product that Φ is unitary with respect to. Now since $\Phi(g)^* u \in U$, we have $\langle \Phi(g)u^\perp, u \rangle = 0$ for all $u^\perp \in U^\perp, u \in U$, so $\Phi(g)u^\perp$ must always be in the orthogonal complement, so $U^\perp$ is an invariant subspace. □

5. Haar Measure and Applications

Many of our subsequent results for compact Lie groups rely on the existence of a unique (up to a multiplicative constant we normalize to 1) bi-invariant measure on such spaces. Thus we may define Lebesgue integrals on compact Lie groups. This measure is called the Haar measure. There are a number of constructions of the Haar measure, but we take its existence for granted.

Proposition 5.1. *If Φ is a representation of G on V, a finite-dimensional, complex vector space, then there exists a Hermitian product on V such that Φ is unitary.*

Proof. Take any Hermitian product $\langle \cdot, \cdot \rangle$ on V and define $(\cdot, \cdot)$ by

$$(u, v) = \int_G \langle \Phi(x)u, \Phi(x)v \rangle \, dx.$$

The properties of a Hermitian inner product are all immediate from calculation. □

The program of representation theory involves finding all representations (up to isomorphism) of a given group G. In general this is about as hard as it sounds (and it sounds hard), but for compact Lie groups we can reduce the problem to finding the irreducible representations. The next theorem makes this notion precise.

Theorem 5.2. *Let Φ be a representation of G on V. Then Φ is the direct sum of irreducible representations. That is, there exists a finite set $\{V_i\}_{i=1}^k$ such that Φ acts irreducibly on each V_i, and $V = V_1 \oplus ... \oplus V_k$.*

Proof. First, we form a Hermitian inner product as in the previous proposition. Then Φ is a unitary representation. Find some proper invariant subspace (if there isn't one, we have the desired decomposition) $U \subset V$. As shown earlier, $U^\perp$ is also an invariant subspace. Now we have $V = U \oplus U^\perp$. Continuing this process on $U^\perp$ we must eventually (because V is finite-dimensional) obtain irreducible representations and thus the desired decomposition. □

Our next few results give some nice properties of linear maps involving irreducible representations.

Theorem 5.3 (Schur's Lemma)**.** *If Φ and Φ' are irreducible representations of G on V and V', respectively, and $L : V \to V'$ is a linear map such that $\Phi'(g)L = L\Phi(g)$ for all $g \in G$, then either L is a bijection or $L = 0$.*

Proof. Because Φ is irreducible, the kernel of L is either 0 or V. Then, the image of L must be either V' or 0. Then L, unless $L = 0$ (in the case where $\ker L = V$), must be a bijection. □

Corollary 5.4. *Suppose Φ is an irreducible representation of G on V. If there exists a linear map $L : V \to V$ with $\Phi(g)L = L\Phi(g)$ for all $g \in G$, then L is a scalar map.*

Proof. Take λ to be a eigenvalue for L. Consider $L-\lambda I$. We see that $\Phi(g)(L-\lambda I) = \Phi(g)L - \lambda\Phi(g) = L\Phi(g) - \lambda\Phi(g) = (L - \lambda I)\Phi(g)$, so $L - \lambda I$ commutes with $\Phi(g)$ for all $g \in G$. Because λ is an eigenvalue, $L-\lambda I$ is not bijective (it has determinant 0), so by Schur's Lemma, $L - \lambda I = 0$, so $L = \lambda I$, a scalar map. □

A simple consequence is that for an abelian group (i.e. $g \cdot h = h \cdot g \quad \forall g, h \in G$) the finite-dimensional, irreducible representations are all 1-dimensional. To see this, we take $L = \Phi(g_0)$, with L as in Corollary 5.7 and g_0 some element of G. Then, $\Phi(g)$ just maps into a 1-dimensional complex vector space. As an example we can consider $U(1)$, which is just the unit circle. Since $U(1)$ is abelian, the irreducible representations are maps from $U(1)$ to $GL(1, \mathbb{C})$. Of course, since Φ is continuous and $U(1)$ compact, we actually need to map into a compact subset of $GL(1, \mathbb{C})$. Indeed, it turns out that the irreducible representations map $U(1)$ to itself and the maps are given by $e^{i\theta} \mapsto e^{ik\theta}$ for $k \in \mathbb{Z}$.

We conclude this section with a very useful theorem. The proof is not particularly difficult, but it omitted for space. It can be found in [4, p.241].

Theorem 5.5 (Schur Orthogonality Relation)**.** *If Φ and Φ' are inequivalent irreducible unitary representations of G on finite-dimensional vector spaces V and V', respectively, with inner products $(\cdot, \cdot)$, then*

$$\int_G (\Phi(x)u, v)\overline{(\Phi'(x)u', v')}\, dx = 0 \quad \forall u, v \in V, \text{ and } u', v' \in V'.$$

6. The Peter-Weyl Theorem and Consequences

The goal of this section is to prove the Peter-Weyl theorem, which provides a basis for $L^2(G)$. We illustrate its application with $U(1)$, which is just the circle. The proof of the theorem relies on three lemmas. We do assume a bit of measure and L^2 theory in this section, but any standard treatment of those topics should provide the necessary results and conventions.

Lemma 6.1. *Let G be a compact group and $h \in L^2(G)$, then the function $f : G \to L^2(G), y \mapsto h(y^{-1}x)$ is continuous.*

Proof. The space of continuous functions on G is dense in $L^2(G)$, so we find a continuous function c such that $\|h - c\| < \epsilon/3$. Since G is compact, c is uniformly continuous. Then there exists an open neighborhood of the identity ("1" here) such that

$$|c(y^{-1}x) - c(y^{-1}x)| < \epsilon/3$$

for all $x \in G$ whenever $y_1^{-1}y_2 \in U$. Now we compute.

$$\begin{aligned} &||h(y_1^{-1}x) - h(y_2^{-1}x)||_{2,x} \\ \leq\ &||h(y_1^{-1}x) - c(y_1^{-1}x)||_{2,x} + ||c(y_1^{-1}x) - c(y_2^{-1}x)||_{2,x} + ||c(y_2^{-1}x) - h(y_2^{-1}x)||_{2,x} \\ =\ &2||h - c||_2 + ||c(y_1^{-1}x) - c(y_2^{-1}x)||_{2,x} \\ \leq\ &2||h - c||_2 + \sup_{x \in G} |c(y_1^{-1}x) - c(y_2^{-1}x)| < 2\epsilon/3 + \epsilon/3 = \epsilon. \end{aligned}$$

□

Lemma 6.2. *Let G be a compact group and $h \in L^2(G)$. For every $\epsilon > 0$, there are finitely many Borel sets $E_i \subseteq G$ such that $\{E_i\}$ disjointly covers G and finitely many $y_i \in G$ such that*

$$||h(y^{-1}x) - h(y_i^{-1}x|| < \epsilon \qquad \text{for all } i \text{ and for all } y \in E_i.$$

Proof. By Lemma 6.1, we can choose an open neighborhood U of 1 such that $||h(gx) - h(x)||_{2,x} < \epsilon$ for all $g \in U$. To see this, we use the terms from the proof of Lemma 6.1 and choose $y_1^{-1} = g$ and $y_2 = 1$. Now, we can take an arbitrary element of the group z_0 and have $||h(gz_0x) - h(z_0x)||_{2,x} < \epsilon$ whenever $g \in U$ by the choice (again with y_1, y_2 as in the proof of Lemma 6.1) $y_1^{-1} = gz_0$ and $y_2^{-1} = z_0$, so $y_1^{-1}y_2 = gz_0z_0^{-1} = g$. We see then that Uz_0 (i.e. the right translate of U by z_0) are open neighborhoods of z_0. Now we can cover G with the neighborhoods Uz_i since each such set is an open neighborhood of z_i. Since G is compact, we can choose a finite subcover $\{Uz_i\}_{i=1}^n$. We set $Uz_i = U_i$. We now define $F_j = U_j \setminus \cup_{i=1}^{j-1} U_i$. To finish the lemma, we set $y_i = z_i^{-1}$ (with y_i as in the statement of the lemma), then form $\{E_i\}$ as we formed $\{F_i\}$ except with y_i rather than z_i, and $\{E_i\}$ has the desired properties. □

Lemma 6.3. *Let G be a compact group, $f \in L^1(G)$, and $h \in L^2(G)$. Define $F : G \to \mathbb{R}, x \mapsto \int_G f(y)h(y^{-1}x)\,dy$. Then F is the limit in $L^2(G)$ of a sequence of linear combinations of left translates of h. That is for every $\epsilon > 0$, there exists a finite set $\{c_i\}_{i=1}^k$ such that*

$$\left\| F(x) - \sum_{i=1}^{k} c_i h(y_i^{-1}x) \right\|_{2,x} < \epsilon$$

Proof. Given $\epsilon > 0$, we choose y_i and E_i as in the statement of Lemma 6.2. Choose $c_i = \int_{E_i} f(y)dy$. We can now compute.

$$\begin{aligned}
&\left\| \int_G f(y)h(y^{-1}x)\,dy - \sum_i^k c_i h(y_i^{-1}x) \right\|_{2,x} \\
&\leq \left\| \sum_{i=1}^k \int_{E_i} |f(y)||h(y^{-1}x) - h(y_i^{-1}x)|\,dy \right\|_{2,x} \\
&\leq \sum_{i=1}^k \int_{E_i} |f(y)|\|h(y^{-1}x) - h(y_i^{-1}x)\|_{2,x}\,dy \\
&\leq \sum_{i=1}^k \int_{E_i} |f(y)|\epsilon\,dy \\
&= \epsilon \int_G |f(y)|\,dy = \epsilon\|f\|_1.
\end{aligned}$$

□

We need another definition for the statement of the Peter-Weyl theorem.

Definition 6.4. A **matrix coefficient** for Φ a unitary, finite-dimensional representation of G on V with Hermitian inner product $\langle\cdot,\cdot\rangle$ is any function of the form $x \mapsto \langle\Phi(x)u, v\rangle$ for $x \in G$ and $u, v \in V$.

The term matrix coefficient springs from fact that if we consider a basis $\{e_i\}_{i=1}^n$ for V, then $\langle\Phi(x)u, v\rangle$ is just the ij entry in the matrix $\Phi(x)$. Arbitrary $u, v \in V$ are just linear combinations of basis elements, so a general matrix coefficient can be realized as a linear combination of literal matrices.

Theorem 6.5 (Peter-Weyl Theorem)**.** *If G is a compact group, then the span of all matrix coefficients for all finite-dimensional, irreducible, unitary representations of G is dense in $L^2(G)$.*

Proof. Recall that a matrix coefficient is a function of the form $(\Phi(x)u, v)$ for a unitary representation Φ and the associated Hermitian inner product $(\cdot,\cdot)$. Then if $h(x) = (\Phi(x)u, v)$, then

$$\begin{aligned}
\overline{h(x^{-1})} &= (\Phi(x)v, u) \\
h(gx) &= (\Phi(x)u, \Phi(x^{-1}v) \\
h(xg) &= \Phi(x)\Phi(g)u, v)
\end{aligned}$$

are also matrix coefficients for Φ. Let U be the closure in $L^2(G)$ of all finite-dimensional, irreducible, unitary representations. Then for a particular matrix coefficient, $h(x) \in U$, we have that $\overline{h(x^{-1})}$, $h(gx)$, and $h(xg)$ are also in U. We now argue by contradiction and suppose that $U \neq L^2(G)$. Then $U^\perp \neq \emptyset$, and $h(x) \in U^\perp$ implies $\overline{h(x^{-1})}, h(gx), h(xg) \in U^\perp$. We now wish to form a nonzero continuous function in $U^\perp$. Since $U^\perp$ is nonempty, there is some $H \in L^2(G)$ such that $H \neq 0$. Then around every neighborhood of the identity, 1, we form

$$F_N(x) = \frac{1}{|N|}\int_G I_N(y)H(y^{-1}x)\,dy.$$

Here I_N is the characteristic function for N, so $I_N(x) = 1$ if $x \in N$, $I_N(x) = 0$ if $x \notin N$, and $|N|$ is the Haar measure of the neighborhood. Because I_N and H are in $L^2(G)$, we can use the Schwarz inequality to see that F_N is continuous. Further, as N shrinks to the identity, F_N approaches H in L^2. We conclude that some F_N is not the 0 function. Now we invoke our lemma: each linear combination of left translates of H is in $U^\perp$, so F_N is in $U^\perp$. We conclude that $U^\perp$ contains a nonzero continuous function. In particular it has some nonzero value at the identity. For convenience, we can use scalar multiplication and translation to create a new continuous function, F_1, in $U^\perp$ with the property that $F_1(1)$ is real and nonzero. We now create yet another continuous function in $U^\perp$

$$F_2(x) = \int_G F_1(yxy^{-1})\, dy.$$

We also observe that $F(gxg^{-1}) = F(x)$ (since we're integrating over the entire group) for all $g \in G$, and $F_2(1) = F_1(1)$ (and is still real and nonzero). Finally we set

$$F(x) = F_2(x) + \overline{F(x^{-1})}.$$

Similar to before we have F continuous and in $U^\perp$, $F(gxg^{-1}) = F(x)$, $F(1)$ real and nonzero (equal to $2F_2(1)$ this time), and $F(x) = \overline{F(x^{-1})}$. Then, F is not the 0 function in $L^2(G)$. Now let $k(x,y) = F(x^{-1}y)$ and define an integral transform T by

$$Tf(x) = \int_G k(x,y)f(y)\, dy = \int_G F(x^{-1}y)f(y)\, dy$$

for $f \in L^2(G)$. We now borrow some functional analysis. We clearly have that $k(x,y) = \overline{k(x,y)}$ and

$$\int_{G\times G} |k(x,y)|^2\, dx\, dy < \infty.$$

This makes T a nonzero Hilbert-Schmidt operator from $L^2(G)$ into itself, so the Hilbert-Schmidt theorem states that it has a real nonzero eigenvalue λ and that the eigenspace V_λ for λ is finite-dimensional. We now wish to show that V_λ is invariant under left translation by g. By this we mean that $(L(g)f)(x) = f(g^{-1}x)$. Moreover,

$$TL(g)f(x) = \int_G F(x^{-1}y)f(g^{-1}y)\, dy = \int_G F(x^{-1}gy)f(y)\, dy$$

$$= Tf(g^{-1}x) = \lambda f(g^{-1}x) = \lambda L(g)f(x).$$

The second inequality follows because we just replace y with gy but still integrate over the same elements. Then, by previous lemma, we have a function on G that take G into V_λ by the continuous assignment $g \mapsto L(g)f$. Then, L is a finite-dimensional representation of G on V_λ. Since we can decompose representations into irreducible representations, we have some nonempty invariant subspace W_λ. Choose $\{f_1, ..., f_n\}$ an orthonormal basis for W_λ. Then there are matrix coefficients for W_λ given in terms of its inner product. That is

$$h_{ij}(x) = (L(x)f_j, f_i) = \int_G f_j(x^{-1}y)\overline{f_i(y)}\, dy.$$

Moreover, since $h_{ij}(x)$ is a matrix coefficient, it belongs to U. However, F is $U^\perp$, so their inner product must be 0. We now compute.

$$\begin{aligned} 0 &= \int_G F(x)\overline{h_{ii}(x)}\,dx \\ &= \int_G \int_G F(x)\overline{f_i(x^{-1}y)}f_i(y)\,dy\,dx. \end{aligned}$$

Changing the order of integration (this is, of course, Fubini's theorem, an important theorem with a decidedly nontrivial proof but one whose details are of no concern to us) we obtain

$$\begin{aligned} &= \int_G \int_G F(x)\overline{f_i(x^{-1}y)}f_i(y)\,dx\,dy \\ &= \int_G \int_G F(yx^{-1})\overline{f_i(x)}f_i(y)\,dx\,dy \end{aligned}$$

We now change variables back, use the definition of T and the fact that $F(gxg^{-1}) = F(x)$ to obtain

$$\begin{aligned} &= \int_G \Big(F(x^{-1}y)f_i(y)\,dy\Big)\overline{f_i(x)}\,dx \\ &= \int_G (Tf_i(x))\overline{f_i(x)}\,dx \\ &= \lambda \int_G |f_i(x)|^2\,dx \end{aligned}$$

But this implies that $0 = \lambda \int_G |f_i(x)|^2\,dx$ for λ nonzero and all i. Since W_λ was assumed to be nonempty, we conclude that $U^\perp$ must be empty, so $U = L^2(G)$. □

Returning to our simple example of the circle identified with $U(1)$, we see that applying the Peter-Weyl theorem gives us that the matrix coefficients of the irreducible representations $\{e^{i\theta} \mapsto e^{ik\theta}\}_{k\in\mathbb{Z}}$ form an orthonormal basis for $L^2(U(1))$. Indeed, this is the standard basis for $L^2(U(1))$.

We conclude with a couple corollaries to the Peter-Weyl theorem. The first has a simple proof. The second's proof is within reach, but it is a bit long and relies on an intermediate corollary, so we omit it.

Corollary 6.6. *If $\{\Phi^{(\alpha)}\}$ (for α an index) is a maximal set of mutually inequivalent finite-dimensional representations of a compact group G (by maximal we mean not properly contained in some other such set) and $\{(d)^{1/2}\Phi^{(\alpha)}_{ij}(x)\}$ is a corresponding set of matrix coefficients, where d is the degree of $\Phi^{(\alpha)}$ and is equal to the dimension of the underlying vector space, then $\{(d)^{1/2}\Phi^{(\alpha)}_{ij}(x)\}$ is dense is an orthonormal basis for $L^2(G)$.*

Proof. Since the representations are mutually inequivalent, the Schur orthogonality relation gives orthogonality. The factor of the square root of the degree gives normality. The Peter-Weyl theorem shows that the set is an orthonormal basis for the Hilbert space $L^2(G)$. □

Corollary 6.7. *If G is a compact Lie group, then there exits a one-to-one finite-dimensional representation of G. Therefore, G is isomorphic to a closed matrix group.*

Acknowledgments. I owe a great deal of thanks to my mentor Sergei Sagatov, whose instruction, guidance, and patience broadened my mathematical horizons and made possible this paper. Among other things, Sergei provided the idea and references for the paper, taught me the necessary techniques, and provided corrections and suggestions to numerous drafts of the paper. I must also thank Peter May whose efforts continue to make possible the University of Chicago REU.

References

[1] John M. Lee. Introduction to Smooth Manifolds. Springer, New York, 2nd Edition, 2013.
[2] John M. Lee. Riemannian Manifolds: An Introduction to Curvature. Springer, New York, 1997.
[3] Victor Guillemin and Alan Pollack. Differential Topology. AMS Chelsea, Providence, 1974.
[4] Anthony W. Knapp. Lie Groups: Beyond an Introduction, Birkhäuser, New York, 2nd Edition, 2002.