MATRIX GROUPS

DANIEL FELTEY

Abstract. This paper will be a brief introduction to various groups of matrices and their algebraic, analytic, and topological properties. We consider curves in matrix groups and use them to define the dimension of a matrix group as an important invariant. We then define the matrix exponential and use it to further investigate curves in matrix groups, and to prove our main topological result that every matrix group is a manifold. We follow [1] in this paper.

Contents

1. Preliminaries: Groups, Rings, Fields, and Skew-Fields.

Before we can begin the study of matrix groups we need to define the terms that we make extensive use of in this paper. To begin we start with the definition of a group.

Definition 1.1. A *group* is a set G together with a binary operation, $*$, that is a map $* : G \times G \to G$ with $*(x, y)$ denoted $x*y$ that satisfies the following properties:

(1) $$\forall\ a, b, c \in G,\ a * (b * c) = (a * b) * c.$$

(2) $$\exists\ e \in G \text{ such that } \forall\ g \in G\ g * e = e * g = g.$$

(3) $$\forall\ g \in G,\ \exists\ g^{-1} \in G \text{ such that } g * g^{-1} = g^{-1} * g = e.$$

So a group is a pair $(G, *)$, that is, a set together with a binary operation satisfying the above properties. Usually we will refer to a group by referring only to the set G.

Definition 1.2. A group G is called *abelian* if for all x and y in G we have that $x * y = y * x$.

Date: AUGUST 22, 2008.

Definition 1.3. A subset H of G is said to be a *subgroup* of G, denoted $H \leqslant G$ if the following properties hold:

$$\text{if } g, h \in H \textit{ then } g * h \in H \tag{1}$$

$$\text{if } h \in H \textit{ then } h^{-1} \in H \tag{2}$$

Beyond these definitions this paper will assume standard results from basic group theory such as *Lagrange's Theorem.*

We now move on to the standard definitions of a *ring* and a *field.*

Definition 1.4. A *ring* is a set, R, together with two binary operations $+$ and $*$ that satisfy the following relations:

(1) $(R,\ +)$ is an abelian group.

(2) $\exists\ 1 \in R$ such that $\forall r \in R\ \ 1 * r = r * 1 = r$

(3) $\forall\ a, b, c \in R,\ \ a * (b * c) = (a * b) * c.$

(4) $\forall\ a, b, c \in R,\ \ a * (b + c) = a * b + a * c$ and $(a + b) * c = a * c + b * c$

When considering rings we call the operation, $+$, addition and the operation, $*$, multiplication. The additive identity is usually denoted by 0, and we usually add the requirement that $1 \neq 0$. A ring is called *commutative* if the operation $*$ is commutative (i.e. $x * y = y * x$ for all x and y).

Definition 1.5. A *field* is a ring, F, satisfying the following:

$$1\ \neq\ 0. \tag{1}$$

$$(F \setminus \{0\}, *) \text{ is an abelian group.} \tag{2}$$

In this paper we mainly consider the familiar fields $\mathbb{R}$, and $\mathbb{C}$ the fields of real and complex numbers respectively. We will, however, also deal with one other mathematical object that differs only slightly yet fundamentally from these two systems, that is the set of Hamiltonian Quaterions denoted $\mathbb{H}$. The quaternions form what is called a skew-field which is just a field which is not required to have the property $x * y = y * x$, in other words the group of non-zero elements is no longer required to be abelian.

Definition 1.6. The *Hamiltonian Quaternions* $\mathbb{H}$ is the set of numbers of the form $a + bi + cj + dk$ under the following conditions:

$$a, b, c, d \in \mathbb{R} \tag{1}$$

$$i^2 = j^2 = k^2 = ijk = -1 \tag{2}$$

From these relations one can derive all the multiplicative properties of the quaternions and show that they indeed form a skew-field.

2. Conjugates and Inner Products

Recall from the basic theory of complex numbers that in $\mathbb{C}$ there is a notion of the complex conjugate, that is for a complex number, $z = a + bi$ we define its complex conjugate by $\overline{z} = \overline{a + bi} = a - bi$. Clearly for a real number r, $\overline{r} = r$. We now want to extend this to the quaternions as follows.

Definition 2.1. For a quaternion $q \in \mathbb{H}$ we define the *conjugate* of $q = a+bi+cj+dk$ to be $\overline{q} = \overline{a + bi + cj + dk} = a - bi - cj - dk$

This notion of a conjugate shares most of its properties with that of the complex conjugate, particularly that: $q\overline{q} = \overline{q}q = a^2 + b^2 + c^2 + d^2$ when $q = a + bi + cj + dk$. Since the skew-field $\mathbb{H}$ is not commutative, however, it no longer holds true that $\overline{(q_1 q_2)} = (\overline{q_1})(\overline{q_2})$. There is a new rule, however, that is analogous to the one which holds only in the cases of the more familiar fields $\mathbb{R}$ and $\mathbb{C}$, and that rule is $\overline{(q_1 q_2)} = (\overline{q_2})(\overline{q_1})$.

We now introduce an inner-product in order to construct the orthogonal groups which will be our first examples of matrix groups.

Definition 2.2. For $k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ define the *inner-product* of two vectors $\mathbf{a} = (a_1, a_2, ..., a_n),$ *and* $\mathbf{b} = (b_1, b_2, ..., b_n)$ in the vector space k^n by

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{m=1}^{n} a_m \overline{b_m}$$

From this definition follow several properties that the inner-product satisfies which we state without proof as our first theorem.

Theorem 2.3. *For all* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in k^n$ *and* $\lambda \in k$ *the inner-product satisfies the following*

$$\langle \mathbf{a}, \mathbf{b} + \mathbf{c} \rangle = \langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{a}, \mathbf{c} \rangle \tag{1}$$

$$\langle \lambda \mathbf{a}, \mathbf{b} \rangle = \lambda \langle \mathbf{a}, \mathbf{b} \rangle \tag{2}$$

$$\overline{\langle \mathbf{a}, \mathbf{b} \rangle} = \langle \mathbf{b}, \mathbf{a} \rangle \tag{3}$$

Definition 2.4. We say that two vectors $\mathbf{a}, \mathbf{b} \in k^n$ are *orthogonal* if $\langle \mathbf{a}, \mathbf{b} \rangle = 0$.

Definition 2.5. A *basis* B for k^n is said to be *orthonormal* if for all $\mathbf{x} \in B$ $\langle \mathbf{x}, \mathbf{x} \rangle = 1$, and for all $\mathbf{x}, \mathbf{y} \in B$, $\mathbf{x} \neq \mathbf{y}$ we have that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

We finally have enough machinery to move on to the main goal of this paper, to introduce matrix groups, and we will make our constructions relying heavily on the properties of the inner-product we have defined.

3. Orthogonal Matrices

Recall that when considering vector spaces and linear maps between them that we can construct a matrix for every linear map from a vector space to itself. We call the set of all $n \times n$ matrices with coefficients in k, $M_n(k)$, these are the linear maps from k^n to k^n. We now get our first examples of matrix groups: the general linear group, and the orthogonal groups.

Definition 3.1. The *General Linear Group,* $GL_n(k)$ is the set of all matrices in $A \in M_n k$ satisfying $det(A) \neq 0$.

Note that this is in fact a group because the determinant function is multiplicative, only matrices with non-zero determinant have inverses, matrix multiplication is associative, and there is clearly an identity, namely the identity matrix possessing ones along its diagonal and zeros everywhere else.

Definition 3.2. The *Orthogonal Group* over the "field" k is the set

$$\mathscr{O}_n(k) = \{X \in GL_n(k) | \langle \mathbf{a}X, \mathbf{b}X \rangle = \langle \mathbf{a}, \mathbf{b} \rangle \text{ for all } \mathbf{a}, \mathbf{b} \in k^n\}$$

if $k = \mathbb{R}$ we denote this group by $O(n)$ and call it the *Orthogonal Group.*
if $k = \mathbb{C}$ we denote this group by $U(n)$ and call it the *Unitary Group.*
if $k = \mathbb{H}$ we denote this group by $Sp(n)$ and call it the *Sympletic Group.*

These are our first examples of matrix groups, and it is not difficult to show that these are, in fact, groups.

4. Tangent Spaces and Lie Algebras

Now that we are familiar with several examples of matrix groups we wish to investigate other, not so obvious, properties and invariants of these groups that will allow us to tell "different" groups apart. First we define a curve in a matrix group.

Definition 4.1. If V is a finite dimensional vector space then *a curve in* V is a continuous function $\gamma : (-\epsilon, \epsilon) \to V$ where $(-\epsilon, \epsilon) \subseteq \mathbb{R}$.

We say that such a curve is differentiable at a point if it is differentiable in the traditional sense of the limit definition, and merely note that if the derivative at a point c does exist then it will be a vector in V. We denote the derivative at c by $\gamma'(c)$.

Recall from calculus that if we choose a basis for the vector space V we can represent a curve in V by an ordered pair of its coordinate functions. We want to consider matrix groups so note that each of $M_n(\mathbb{R}),\ M_n(\mathbb{C}),\ M_n(\mathbb{H})$ can be considered as subspaces of a real vector space of dimension $n^2, 2n^2,$ and $4n^2$ respectively.

Definition 4.2. A *curve in a matrix group* G is a curve in $M_n(k)$ such that for all $u \in (-\epsilon, \epsilon),\ \gamma(u) \in G$.

Since G is a group we can form products of elements in G and stay in the group. Now considering curves in groups we can consider products of curves and get new curves that stay in the matrix group.

Theorem 4.3. *If γ, σ are both curves in a matrix group G then so is the product curve $\gamma\sigma$. If both γ, and σ are differentiable at c then so is $\gamma\sigma$ and*

$$(\gamma\sigma)'(c) = \gamma(c)\sigma'(c) + \gamma'(c)\sigma(c) \tag{4.4}$$

We omit the proof of this theorem because it is not difficult and is just an exercise in symbol manipulation as it is just repeated use of the product rule for real scalar functions.

Definition 4.5. If G is a matrix group the *tangent space* to G at the identity is the set $\mathbf{T_G} = \{\gamma'(0) | \gamma : (-\epsilon, \epsilon) \mapsto G$ is a differentiable curve in G with $\gamma(0) = I\}$

Theorem 4.6. $\mathbf{T_G}$ *is a subspace of $M_n(k)$ as a real vector space.*

Proof. We prove the theorem in two parts.

- First we show that $\mathbf{T_G}$ is closed under addition.
  Suppose that $\gamma'(0)$, and $\sigma'(0)$ are in $\mathbf{T_G}$ and consider the curve $(\gamma\sigma)(t)$, then $(\gamma\sigma)(0) = \gamma(0)\sigma(0) = II = I$. Then since γ and σ are both differentiable at 0 so is $\gamma\sigma$ and so $(\gamma\sigma)'(0)$ is in $\mathbf{T_G}$. But $(\gamma\sigma)'(0) = \gamma(0)\sigma'(0) + \gamma'(0)\sigma(0) = I\sigma'(0) + \gamma'(0)I = \sigma'(0) + \gamma'(0)$. So $\mathbf{T_G}$ is closed under addition.

- We now show that $\mathbf{T_G}$ is closed under scalar multiplication.

  If $\gamma'(0)$ is in $\mathbf{T_G}$ and r is in $\mathbb{R}$ consider the curve $\sigma(t) = \gamma(rt)$. Again it is clear that $\sigma(0) = \gamma(0) = I$ and γ is differentiable at 0 hence so is σ and $\sigma'(0)$ belongs to $\mathbf{T_G}$. But $\sigma'(0) = r\gamma'(0)$ so $\mathbf{T_G}$ is closed under scalar multiplication.
- Thus $\mathbf{T_G}$ is a subspace of $M_n(k)$ and since the latter in finite dimensional so is $\mathbf{T_G}$.

□

Definition 4.7. The *dimension of a matrix group* G is the dimension of the tangent space $\mathbf{T_G}$, and is denoted by $dim(G)$.

Example 4.8. $\mathbf{T_{GL_n(\mathbb{R})}} = M_n(\mathbb{R})$, that is the tangent space of $GL_n(\mathbb{R})$ is $M_n(\mathbb{R})$ and from this it clearly follows that $dim(GL_n(\mathbb{R})) = n^2$.

Proof. To show this we first note that the function $det : M_n(\mathbb{R}) \to \mathbb{R}$ that sends a matrix to its determinant is continuous since it is a polynomial in the entries of the matrix. Continuity of the determinant allows us to find an δ-ball around I in $M_n(\mathbf{R})$ such that if a matrix A is in this ball then $det(A) \neq 0$. Then for any matrix B in $M_n(\mathbf{R})$ we can define the curve $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}$ by $\gamma(t) = I + tB$. Then this is a differentiable curve through I and that has B as its derivative at 0, ie a curve satisfying $\gamma(0) = I$, and $\gamma'(0) = B$. Now it just remains to be seen that γ is a curve in $GL_n(\mathbb{R})$, and this will be true if ϵ is chosen carefully based on the determinant of B such that $I + tB$ will be in the ball around I and $det(I + tB) \neq 0$. Thus $\mathbf{T_{GL_n(\mathbb{R})}} = M_n(\mathbb{R})$ and $dim(GL_n(\mathbb{R})) = n^2$ □

We now show that the dimension of a matrix group is invariant under isomorphism of matrix groups.

Since all of our matrix groups live inside some larger vector space we can speak of continuity when considering functions from one matrix group to another. More importantly we can consider continuous homomorphisms, and from now on all homomorphisms will be assumed to be continuous. Then if we have two matrix groups G and H and a homomorphism $\phi : G \to H$ then every curve,γ, in G will give rise to a curve in H simply by the composition $\phi \circ \gamma$.

Definition 4.9. A homomorphism $\phi : G \to H$ is *smooth* if for every differentiable curve γ in G the curve $\phi \circ \gamma$ is differentiable.

Then every tangent vector in the tangent space of G will give us a corresponding vector in the tangent space of H if there is a smooth homomorphism between G and H. So if $\gamma'(0)$ is in $\mathbf{T_G}$ we define $d_\phi(\gamma'(0)) = (\phi \circ \gamma)'(0)$ to be the tangent vector in $\mathbf{T_H}$ that we get as a result of the homomorphism ϕ. This gives us a map from $\mathbf{T_G}$ to $\mathbf{T_H}$.

Definition 4.10. If $\phi : G \to H$ is a smooth homomorphism we call the map $d_\phi : \mathbf{T_G} \to \mathbf{T_H}$ *the differential of* ϕ.

Theorem 4.11. *If ϕ is a smooth homomorphism then $d_\phi : \mathbf{T_G} \to \mathbf{T_H}$ is a linear map.*

Proof. It is particularly easy to get lost in notation when proving this result so we will take it in two steps and show that d_ϕ preserves scalar multiplication and vector addition separately.

- d_ϕ preserves scalar multiplication. Consider $d_\phi(a\gamma'(0)) = (\phi \circ \sigma)'(0)$ by definition where $\sigma(t) = \gamma(at)$ since σ is the function that corresponds to the tangent vector $a\gamma'(0)$. But $(\phi \circ \sigma)'(0) = a(\phi \circ \gamma)'(0) = ad_\phi(a\gamma'(0))$, and so d_ϕ does in fact preserve scalar multiplication.
- d_ϕ preserves vector addition. Consider $d_\phi(\gamma'(0) + \sigma'(0)) = (\phi \circ \gamma\sigma)'(0)$ since the product curve $\gamma\sigma$ is the curve corresponding to the tangent vector $\gamma'(0) + \sigma'(0)$. But $(\phi \circ \gamma\sigma)'(0) = (\phi \circ \gamma)'(0) + (\phi \circ \sigma)'(0) = d_\phi(\gamma'(0)) + d_\phi(\sigma'(0))$. So d_ϕ does in fact preserve vector addition.

□

Theorem 4.12. *If $\phi : G \to H$ and $\psi : H \to K$ are smooth homomorphisms then so is $\psi \circ \phi$ and $d_{\psi\circ\phi} = d_\psi \circ d_\phi$.*

Proof. Again we show the two parts of the theorem separately.

- It is obvious that $\psi \circ \phi$ is smooth since the composition of smooth functions is smooth.
- To show $d_{\psi\circ\phi} = d_\psi \circ d_\phi$ consider $d_{\psi\circ\phi}(\gamma'(0)) = (\psi \circ \phi \circ \gamma)'(0) = (\psi \circ (\phi \circ \gamma))'(0) = d_\psi(\phi \circ \gamma)'(0) = d_\psi \circ d_\phi(\gamma'(0))$.

□

Now by corollary we get the following fundamental result which we have been building towards.

Theorem 4.13. *If $\phi : G \to H$ is a smooth isomorphism then $d_\phi : \mathbf{T_G} \to \mathbf{T_H}$ is a linear isomorphism and so $dim(G) = dim(H)$.*

Proof. $\phi \circ \phi^{-1} = id_H$ so $d_\phi \circ\ d_{\phi^{-1}} = id_{\mathbf{T_H}}$ which implies that d_ϕ is surjective. Similarly since $\phi^{-1} \circ \phi = id_G$ we have that $d_{\phi^{-1}} \circ\ d_\phi = id_{\mathbf{T_G}}$ implies that d_ϕ is injective, hence it is a bijection and thus a linear isomorphism thus $dim(G) = dim(H)$. □

5. The Matrix Exponential and One-Parameter Subgroups

Now that we are familiar with curves in a matrix group we have a way of finding the tangent space of a matrix group it is natural to ask that if we know a vector in the tangent space of a matrix group can we easily find a curve in the matrix group corresponding to this vector? To answer this question we introduce the matrix exponential. We consider only the case of real matrices in this section and note that constructions for complex and quaternion matrices are similar.

Definition 5.1. For a matrix A in $M_n(\mathbb{R})$ we define the *matrix exponential of A* to be the map $exp : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ where we denote $exp(A)$ by e^A and define

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

Making the convention that $A^0 = I$ for all real matrices.

Thus the matrix exponential is basically identical to the real exponential, and just as the real exponential converges for all real numbers so does the matrix exponential for all real matrices.

Theorem 5.2. *If A is in $M_n(\mathbb{R})$ then e^A converges.*

Proof. We will prove this result by bounding all of the entries of e^A with a clearly convergent series.

Let a_{ij} be the entry of a in row i and column j. Then set $m = \max_{1 \leq i,j \leq n} |a_{ij}|$. Now consider the largest(in absolute value) elements in each of I, A, A^2... Clearly 1 is the largest element in I, and m is the largest element in A. It is then easy to show that nm^2 is greater than or equal to the largest element in A^2 and in general $n^{k-1}m^k$ is greater than or equal to the largest element in A^k.

Now consider the series

$$1 + m + \frac{nm^2}{2} + \frac{n^2m^3}{6} + ... + \frac{n^{k-1}m^k}{k!} + ... = 1 + \sum_{k=1}^{\infty} \frac{n^{k-1}m^k}{k!}$$

It is easy to check using the ratio test that this series does in fact converge,and actually converges to

$$\frac{e^{nm} + n - 1}{n}.$$

Thus we have a convergent series which dominates all entries of e^A so e^A does converge for all real matrices. □

Theorem 5.3. *If A and B are in $M_n(\mathbb{R})$ such that $AB = BA$ then $e^{A+B} = e^A e^B$.*

Corollary 5.4. *If A is in $M_n(\mathbb{R})$ then e^A is in $GL_n(\mathbb{R})$.*

Proof. A and $-A$ commute so $e^{A+(-A)} = e^A e^{-A}$, but $e^{A+(-A)} = e^0 = I$, and this implies $det(e^A)det(e^{-A}) = det(e^A e^{-A}) = det(I) = 1$ which implies that $det(e^A) \neq 0$ hence $e^A \in GL_n(\mathbb{R})$. □

Theorem 5.5. *If A is in $M_n(\mathbb{R})$ and B is in $GL_n(\mathbb{R})$ then $e^{BAB^{-1}} = Be^AB^{-1}$.*

Proof. This follows because $(BAB^{-1})^n = BA^nB^{-1}$ and $B(A+C)B^{-1} = BAB^{-1} + BCB^{-1}$. □

Just as in the case of real numbers where we can define an inverse for the exponential function we can do a similar thing for real matrices.

Definition 5.6. For a real matrix X we define the *matrix logarithm* by the series

$$log(X) = (X - I) - \frac{(X - I)^2}{2} + \frac{(X - I)^3}{3} - ... = \sum_{k=1}^{\infty} \frac{(X - I)^k}{k}$$

And just as in the real case where this series only converges for x near 1 for real matrices this series only converges for matrices X near I.

Theorem 5.7. *For X* near *I in $M_n(\mathbb{R})$ $log(X)$ converges.*

Proof. The proof of this theorem is basically identical to the one above, except that we must be careful about what is meant by near. So we set $Y = X - I$ and suppose that $|y_{ij}| < \epsilon$ we then proceed as before and bound the entries of $log(X)$. Proceeding in this manner it turns out that $log(X)$ converges making the requirement that $\epsilon < \frac{1}{n}$ thus giving a more concrete notion to what we mean by near. □

Theorem 5.8. *Let U be a neighborhood in $M_n(\mathbb{R})$ on which log is defined and let V be a neighborhood of 0 in $M_n(\mathbb{R})$ such that $exp(V) \subseteq U$ then*

$$\text{For } X \in U \;\; e^{log(X)} = X \tag{1}$$

$$\text{For } Y \in V \;\; log(e^{Y}) = Y \tag{2}$$

This theorem is proved by extensive symbol manipulation with the definitions for the matrix exponential and logarithm so it is omitted. The important thing to note is that this theorem says that the matrix exponential is one-to-one on a neighborhood of 0. This would seem obvious if we get too preoccupied in analogy with the real exponential since that function is one-to-one everywhere. So we make a brief aside to present an example that shows the matrix exponential is not in general one-to-one.

Example 5.9. Let $X = \begin{pmatrix} 0 & 2\pi \\ -2\pi & 0 \end{pmatrix}$ then $e^X = I$

Proof. Note that X is conjugate to the matrix $Y = \begin{pmatrix} 2\pi i & 0 \\ 0 & -2\pi i \end{pmatrix}$ since if $Z = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ then $X = ZYZ^{-1}$. Then since Y clearly maps to the identity under the matrix exponential by theorem 5.5 above X will also map to the identity. $\square$

Theorem 5.10. *If X and Y are near I such that $log(X)$, $log(Y)$, and $log(XY)$ are all defined and $log(X)$ and $log(Y)$ commute then*

$$log(XY) = log(X) + log(Y)$$

Proof. $e^{log(XY)} = XY = e^{log(X)}e^{log(Y)} = e^{log(X)+log(Y)}$ and since exp is one-to-one near 0 we have that $log(XY) = log(X) + log(Y)$. $\square$

Now we return to the question of curves in matrix groups.

Definition 5.11. A *one-parameter-subgroup,* γ, in a matrix group G is a smoth homomorphism $\gamma : \mathbb{R} \to G$.

As with curves in matrix groups we will usually only care about a one-parameter-subgroup restricted to some small interval around 0 in $\mathbb{R}$.

Example 5.12. For any matrix A in $M_n(\mathbb{R})$ the function $\gamma(u) = e^{uA}$ is a one-parameter-subgroup of $GL_n(\mathbb{R})$ with $\gamma'(0) = A$.

This example should suggest how we can find curves in a matrix group given an element of its tangent space.

Theorem 5.13. *If γ is a one-parameter-subgroup in $GL_n(k)$ then there exists a matrix A in $M_n(k)$ such that $\gamma(u) = e^{uA}$.*

Proof. Let $\sigma(u) = log(\gamma(u)$ then $\sigma(u)$ is a curve in $M_n(k)$ satisfying $\gamma(u) = e^{\sigma(u)}$.

Set $\sigma'(0) = A$, so now we want to show that $\sigma(u)$ is a line through 0 in $M_n(k)$ or equivalently that $\sigma(u) = uA$. Now fix u in $\mathbb{R}$ and consider

$$\begin{aligned}
\sigma'(u) &= \lim_{h\to 0} \frac{\sigma(u+h)-\sigma(u)}{h} \\
&= \lim_{h\to 0} \frac{log(\gamma(u+h)) - log(\gamma(u))}{h} \\
&= \lim_{h\to 0} \frac{log(\gamma(u)\gamma(h)) - log(\gamma(u))}{h} \\
&= \lim_{h\to 0} \frac{log(\gamma(u)) + log(\gamma(h)) - log(\gamma(u))}{h} \\
&= \lim_{h\to 0} \frac{log(\gamma(h))}{h} \\
&= \lim_{h\to 0} \frac{\sigma(h)}{h} \\
&= \sigma'(0)
\end{aligned}$$

This proves that $\sigma'(u) = A$ for all u since its derivative is independent of u thus $\sigma(u) = uA$ hence $\gamma(u) = e^{uA}$. □

Then in general we have the following theorem.

Theorem 5.14. *If A is a vector in the tangent space $\mathbf{T_G}$ of a matrix group G then there is a unique one-parameter-subgroup, γ in G satisfying $\gamma'(0) = A$.*

6. Some Topology

The main result of this section proves that every matrix group is a manifold. We assume many familiar definitions from topology, mainly (path) connectedness, compactness, and continuity. Since all of the groups we consider here live inside the vector space $M_n(k)$ we can consider them as subspaces and hence discuss their topological properties. Now in the context of topology we can finally define a matrix group exactly.

Definition 6.1. A *matrix group* is a closed subgroup of the group $GL_n(k)$.

Definition 6.2. A topological space X is an *n-manifold* if for every point x in X there is an open neighborhood O_x containing x such that O_x is homeomorphic to an open ball in $\mathbb{R}^n$.

We say that an n-manifold has dimension n, and this should suggest that a matrix group of dimension n would turn out to be an n-manifold. Before we prove this result, however, we need one short lemma.

Lemma 6.3. *If G is a matrix group then the function $L_x : G \to G$ defined by $L_x(g) = xg$ is a homeomorphism.*

Proof. Given $\epsilon > 0$ Set $\delta = \frac{\epsilon}{|x|}$ then if $|g-h| \leq \delta$ then $|L_x(g) - L_x(h)| = |xg - xh| = |x||g-h| < |x|\delta = |x|\frac{\epsilon}{|x|} = \epsilon$ and L_x is continuous. Then note that $L_{x^{-1}}$ is the inverse of L_x which is clearly continuous for the same reason L_x is. Thus L_x is a homeomorphism. □

Theorem 6.4. *A matrix group of dimension n is an n-manifold.*

Proof. We follow the proof presented in [1] for this result.

The exponential function is a continuous map from $\mathbf{T_G}$ the n-dimensional tangent space of G to G. We have shown that it is one-to-one on a neighborhood of 0, and has a continuous inverse, the matrix logarithm. Call this neighborhood of 0, O then for any x in G we can consider the map $L_x \circ exp : O \to G$. Since the exponential map will send O to a neighborhood of I the composition will map to a neighborhood of the point x. Thus since this composition results in a homeomorphism we have that G is an n-manifold. $\square$

We conclude this paper in a slightly greater generality by classifying subgroups of topological groups based on the properties of the space.

Theorem 6.5. *In a topological group G every open or closed subgroup of finite index is both open and closed.*

Proof. Let H be a subgroup of finite index in G. Then H is homeomorphic to all of its cosets. So if H is open so is the union of the cosets of H that are not equal to H, but this implies H is also closed. Since that union is finite if H were closed then that union is also closed and H is also open, so any subgroup of finite index in G is both open and closed. $\square$

Theorem 6.6. *If G is a connected group then every closed (or equivalently open) subgroup H of G has infinite index.*

Proof. By the above theorem every subgroup of finite index is both open and closed so if a subgroup H has finite index then we can partition G into H and the cosets of H not equal to H, these are both open sets, but this is impossible if G is connected, thus H cannot have finite index. $\square$

Theorem 6.7. *A connected group G cannot have an open subgroup.*

Proof. The reasoning here is identical to the above, partition G into the disjoint union of H and the union of cosets distinct from H, both sets are open implying G is disconnected, a contradiction. $\square$

Theorem 6.8. *If G is a compact group then every open subgroup H of G has finite index.*

Proof. Let H be an open subgroup of G. Then the cosets of H partition G and hence form an open cover of G, since G is compact there is a finite sub-cover, but since the cosets partition G this means that there are only finitely many cosets and hence H has finite index in G. $\square$

Now returning to matrix groups, which are also topological groups, if we consider a matrix group G that is both compact and connected we can classify all of its matrix subgroups.

Definition 6.9. A *matrix subgroup* of a matrix group G is a subgroup H of G that is also a matrix group (ie H is closed in $GL_n(k)$).

Theorem 6.10. *If G is a compact and connected group then every matrix subgroup H of G is closed and has infinite index.*

Proof. This follows from the preceding theorems. $\square$

References

[1] Morton L. Curtis. Matrix Groups. Springer-Verlag. 1979.
[2] Kristopher Tapp. Matrix Groups for Undergraduates. American Mathematical Society. 2005.