University of Nevada, Reno

Low-Dimensional Quaternionic Matrix Groups

A thesis submitted in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Mathematics and the Honors Program

by

Casey J. Machen

Dr. Valentin Deaconu, Thesis Advisor

May, 2011

UNIVERSITY
OF NEVADA
RENO

THE HONORS PROGRAM

We recommend that the thesis
prepared under our supervision by

Casey Machen

entitled

Low-Dimensional Quaternionic Matrix Groups

be accepted in partial fulfillment of the
requirements for the degree of

BACHELOR OF SCIENCE, MATHEMATICS

____
Valentin Deaconu, Ph.D., Thesis Mentor

____
Tamara Valentine, Ph.D., Director, **Honors Program**

May, 2011

Abstract

We focus on several properties of the Lie groups $Sp(n)$ and $SL_n(\mathbb{H})$. We discuss their Lie algebras, the exponential map from the Lie algebras to the groups, as well as when this map is surjective. Since quaternionic multiplication is not commutative, the process of calculating the exponential of a matrix in $Sp(n)$ or $SL_n(\mathbb{H})$ is more involved than the process of calculating the exponential of a matrix over the real or complex numbers. We develop processes by which this calculation may be reduced to a simpler problem, and provide an example to illustrate this. Additionally, we discuss properties of these groups such as centers, maximal tori, normalizers of the maximal tori, Weyl groups, and Clifford Algebras.

Acknowledgments

This project was made possible by a research grant from the Undergraduate Research Office and the Honors Program at the University of Nevada, Reno. I am also indebted to my mentor, Dr. Valentin Deaconu, for his expertise and support throughout the process of writing this paper.

Contents

1. Introduction

Matrix exponentiation is important in differential equations; in particular, in solving systems of linear ordinary and partial differential equations. Second, and more relevant to this project, the matrix exponential provides a map from the Lie algebra of a matrix group into the matrix group itself. The underlying theme of this project is to determine methods of calculating the exponential of a matrix with quaternion entries. To motivate the ideas, we provide details about the structure of the Lie groups $Sp(n)$ and $SL_n(\mathbb{H})$, as well as discuss some important theory.

The exponential maps of real and complex matrices are well documented (see, for example, any of [1, 2, 3, 5]); yet, we have encountered very little to no work documenting the exponential of a matrix with quaternion entries. The most prominent reason for a lack of research in this area is that, unlike the real and complex numbers, the quaternions are not commutative. Due to this, the notions of eigenvalues and of eigenvectors, with which we are familiar from linear algebra over the complex field, cannot be calculated in the usual manner. Since both eigenvalues and eigenvectors play a crucial role in calculating the exponential of a matrix, we must take a different approach when we consider matrices over the quaternions.

Our approach is to first transform a matrix over the quaternions into a complex matrix. The purpose for this is to be able to use the well-known techniques of diagonalization, SN Decomposition, Jordan forms, etc. to calculate the exponential of the matrix. When we are finished, we transform the complex matrix back into a matrix over the quaternions. One of our goals is to prove that this transformation is possible. However, by our transformation, the complex matrix representing the original quaternionic matrix has dimension twice as large; therefore, it is only practical to consider low-dimensional cases, although the theory works for any dimension.

Unfortunately, we encounter obstacles even when we consider low-dimensional cases. If we begin with a 3×3 matrix over the quaternions, our transformed matrix is a 6×6 complex matrix. The problem with this is the fact that eigenvalues are computed from the characteristic polynomial, whose degree is 6 when our original quaternionic matrix is a 3×3 matrix. Results from Galois theory emphasize the fact that there is no general formula for calculating the roots of a polynomial of degree greater than four, so we are not guaranteed a consistent method of determining the eigenvalues. Thus, our methods work best when we look at 2×2 matrices over the quaternions.

Furthermore, it is difficult to find literature that mentions a formula for the exponential of a single quaternion (i.e. a 1×1 matrix over the quaternions); our search has not lead to a satisfactory proof of such a claim. Therefore, our first task is to derive and prove the formula for the general exponential form of a single quaternion.

2. Notation

We will begin our discussion by introducing the quaternions. The quaternions arose as mathematicians tried to endow $\mathbb{R}^4$ with a multiplication operation that would make it a field. For $(a, b, c, d) \in \mathbb{R}^4$, we will write $a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$, add component-wise, and multiply with the following rules:

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

and

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \cdot \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \cdot \mathbf{i} = \mathbf{j}$$

$$\mathbf{j} \cdot \mathbf{i} = -\mathbf{k}, \quad \mathbf{k} \cdot \mathbf{j} = -\mathbf{i}, \quad \mathbf{i} \cdot \mathbf{k} = -\mathbf{j}.$$

$\mathbb{R}^4$ with the above operations forms a skew-field (which is weaker than a field, since this multiplication operation is not commutative) denoted $\mathbb{H}$, called the quaternions. From the operations shown above it is easy to see that an element of $\mathbb{H}$ commutes with all other quaternions if and only if it is real. As a matter of terminology, we will write $\mathbb{K}$ to refer to any of $\mathbb{R}, \mathbb{C}, \mathbb{H}$ when it is unambiguous.

We denote the conjugate of a quaternion $q = a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$ as $\overline{q} = a-b\mathbf{i}-c\mathbf{j}-d\mathbf{k}$. Notice that $q\overline{q} = \overline{q}q = a^2 + b^2 + c^2 + d^2$. We also define an inner product on $\mathbb{K}^n$ by

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle \;= x_1\overline{y_1} + \cdots + x_n\overline{y_n}.$$

We will write the norm of some $X \in \mathbb{K}^n$ as $|X| = \sqrt{\langle X, X \rangle}$. For $X \in \mathbb{R}^n$ or $X \in \mathbb{C}^n$, it is clear that $\langle X, X \rangle$ is a positive real number, so $|X|$ is positive. For $X \in \mathbb{H}^n$, the same conclusion follows from the above discussion where we showed $q\overline{q} = |q|^2$. Furthermore, for $q_1, q_2 \in \mathbb{H}$, it is not hard to see that $\overline{q_1 \cdot q_2} = \overline{q_2} \cdot \overline{q_1}$.

As our discussion involves square matrices, we will write $M_n(\mathbb{K})$ to mean $n \times n$ matrices with entries in $\mathbb{K}$. Let $A \in M_n(\mathbb{K})$. We will introduce the notation A^*, called the adjoint, to refer to the conjugate transpose, $\bar{A}^T$, of the matrix A, which is achieved by taking the complex (or quaternionic) conjugate of each entry in A

and then taking the transpose of the resulting matrix. It is important to notice that $(AB)^* = B^* A^*$.

For $X \in \mathbb{K}^n$ a row vector, the function $R_A : \mathbb{K}^n \to \mathbb{K}^n$ given by $R_A(X) = X \cdot A$, is a linear function over $\mathbb{K}$. We will work with R_A since the similarly defined function $L_A(X) = A \cdot X$ is not linear over $\mathbb{H}$.

3. The Injective Homomorphism Ψ_n

It is clear that we can write a complex number as $a + b\mathbf{i}$ for $a, b \in \mathbb{R}$, which leads to a bijection between $\mathbb{C}^n$ and $\mathbb{R}^{2n}$ denoted f_n:

$$f_n(a_1 + b_1\mathbf{i}, \ldots, a_n + b_n\mathbf{i}) = (a_1, b_1, \ldots, a_n, b_n).$$

This allows us to associate a complex $n \times n$ matrix with a real $2n \times 2n$ matrix through the function $\Phi_n : M_n(\mathbb{C}) \to M_{2n}(\mathbb{R})$ which is defined in terms of Φ_1:

$$\Phi_1(a + b\mathbf{i}) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

For higher n, we construct 2×2 real block matrices by applying Ψ_1 to each entry of the matrix. For example,

$$\begin{pmatrix} 1 & 2\mathbf{i} \\ 3-4\mathbf{i} & 0 \end{pmatrix} \xrightarrow{\Phi_2} \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & -2 & 0 \\ 3 & -4 & 0 & 0 \\ 4 & 3 & 0 & 0 \end{pmatrix}.$$

Due to the non-commutativity of the quaternions, they are difficult to work with, and many authors define the quaternions in complex terms. The idea is similar to the discussion of turning a complex matrix into a real matrix. We can write a quaternion as $z + w\mathbf{j}$ for $z, w \in \mathbb{C}$ which leads to the bijection between $\mathbb{H}^n$ and $\mathbb{C}^{2n}$ denoted g_n:

$$g_n(z_1 + w_1\mathbf{j}, \ldots, z_n + w_n\mathbf{j}) = (z_1, w_1, \ldots, z_n, w_n).$$

The approach is to find an injective map $\Psi_n : M_n(\mathbb{H}) \to M_{2n}(\mathbb{C})$ that makes the following diagram commute for all $A \in M_n(\mathbb{H})$:

$$\begin{tikzcd} \mathbb{H}^n \arrow[r, "g_n"] \arrow[d, "R_A"'] & \mathbb{C}^{2n} \arrow[d, "R_{\Psi_n(A)}"] \\ \mathbb{H}^n \arrow[r, "g_n"] & \mathbb{C}^{2n} \end{tikzcd}$$

This map Ψ_n will be described in terms of Ψ_1:

$$\Psi_1(z + w\mathbf{j}) = \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}.$$

To determine $\Psi_n(A)$ for some $A \in M_n(\mathbb{H})$, apply Ψ_1 to each entry of A to create 2×2 blocks as above. As with Φ_n, this map is clearly injective.

Proposition 1. For $A, B \in M_n(\mathbb{H})$ and $c \in \mathbb{R}$, we have the following:

$$\Psi_n(cA) = c \cdot \Psi_n(A) \tag{1}$$

$$\Psi_n(A + B) = \Psi_n(A) + \Psi_n(B) \tag{2}$$

$$\Psi_n(A \cdot B) = \Psi_n(A) \cdot \Psi_n(B) \tag{3}$$

Proof: The proofs of the first two properties are trivial. To see the third property, consider the following commutative diagram:

$$\begin{tikzcd}
\mathbb{H}^n \arrow[r, "g_n"] \arrow[d, "R_A"'] & \mathbb{C}^{2n} \arrow[d, "R_{\Psi_n(A)}"] \\
\mathbb{H}^n \arrow[r, "g_n"] \arrow[d, "R_B"'] & \mathbb{C}^{2n} \arrow[d, "R_{\Psi_n(B)}"] \\
\mathbb{H}^n \arrow[r, "g_n"] & \mathbb{C}^{2n}
\end{tikzcd}$$

But since $R_B \circ R_A = R_{AB}$, we also have the following commutative diagram:

$$\begin{tikzcd}
\mathbb{H}^n \arrow[r, "g_n"] \arrow[d, "R_{AB}"'] & \mathbb{C}^{2n} \arrow[d, "R_{\Psi_n(AB)}"] \\
\mathbb{H}^n \arrow[r, "g_n"] & \mathbb{C}^{2n}
\end{tikzcd}$$

Since both commutative diagrams are the same, we must have $R_{\Psi_n(A) \cdot \Psi_n(B)} = R_{\Psi_n(AB)}$, which, in turn, proves property (3). $\square$

4. The Determinant Function

Consider the matrix $A = \left(\begin{smallmatrix} \mathbf{i} & \mathbf{j} \\ \mathbf{i} & \mathbf{j} \end{smallmatrix}\right)$. This matrix is not invertible since it has a nontrivial kernel, as illustrated below:

$$R_A(\left(\begin{smallmatrix} 1 & -1 \end{smallmatrix}\right)) = \left(\begin{smallmatrix} 1 & -1 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} \mathbf{i} & \mathbf{j} \\ \mathbf{i} & \mathbf{j} \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \end{smallmatrix}\right).$$

However, if we compute the determinant as for a 2×2 matrix over the complex numbers, we find it is nonzero: $\det(\left(\begin{smallmatrix} \mathbf{i} & \mathbf{j} \\ \mathbf{i} & \mathbf{j} \end{smallmatrix}\right)) = \mathbf{ij} - \mathbf{ji} = 2\mathbf{k} \neq 0$, contradicting the fact that A is not invertible. We therefore need a better way to define the determinant function for a matrix over $\mathbb{H}$. It turns out that the image under Ψ_n of an invertible matrix is an invertible matrix, since for an $n \times n$ matrix B over the quaternions

$$\begin{aligned} \text{B is invertible} \quad &\Longleftrightarrow \quad R_B : \mathbb{H}^n \to \mathbb{H}^n \text{ is bijective} \\ &\Longleftrightarrow \quad R_{\Psi_n(B)} : \mathbb{C}^{2n} \to \mathbb{C}^{2n} \text{ is bijective} \\ &\Longleftrightarrow \quad \Psi_n(B) \text{ is invertible.} \end{aligned}$$

Thus, when we talk about the determinant of a quaternionic matrix, we mean the composition $\det \circ \Psi_n$, however, we will just write det.

5. Matrix Groups

We will first mention the matrix groups we will encounter. The general linear group of dimension n is the group of $n \times n$ invertible matrices with entries in $\mathbb{K}$:

$$GL_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid \det(A) \neq 0\}.$$

The special linear group of dimension n is a subgroup of the general linear group:

$$SL_n(\mathbb{K}) = \{A \in GL_n(\mathbb{K}) \mid \det(A) = 1\}.$$

The orthogonal group of dimension n with entries in $\mathbb{K}$ is another subgroup of the general linear group:

$$\mathcal{O}_n(\mathbb{K}) = \{A \in GL_n(\mathbb{K}) \mid A^{-1} = A^*\}.$$

When $\mathbb{K} = \mathbb{R}$, we get that $A^* = A^T$, this group is called the orthogonal group and is denoted $O(n)$. When $\mathbb{K} = \mathbb{C}$, the group is called the unitary group, $U(n)$. And when $\mathbb{K} = \mathbb{H}$, we call this group the symplectic group, $Sp(n)$. An important fact about a matrix in $\mathcal{O}_n(\mathbb{K})$ is that its rows (and also columns) form an orthonormal basis of $\mathbb{K}^n$ [2, 5]. This means that for any two distinct rows (or columns) X, Y we have $\langle X, Y \rangle = 0$ and $\langle X, X \rangle = 1$.
For a matrix $A \in \mathcal{O}_n(\mathbb{K})$, we have

$$1 = \det(I) = \det(AA^*) = \det(A)\det(\overline{A}) = \det(A)\overline{\det(A)} = |\det(A)|.$$

This observation leads us to define the special orthogonal and special unitary matrix groups:

$$SO(n) = O(n) \cap SL_n(\mathbb{R}) = \{A \in O(n) \mid det(A) = 1\}.$$

$$SU(n) = U(n) \cap SL_n(\mathbb{C}) = \{A \in U(n) \mid det(A) = 1\}.$$

We will show later, using Lie algebras, that the determinant of an element in $Sp(n)$ is 1. Therefore, there is no need for a special symplectic group.

Since these groups are all subsets of some Euclidean space, we endow them with the usual Euclidean topology. It is a well-known fact that these groups are actually *differentiable manifolds*, which essentially means that every point in the group has a neighborhood which is mapped by a smooth homeomorphism (whose inverse is also smooth) to an open ball in Euclidean space. Furthermore, both the group operation of matrix multiplication and the inverse map are smooth. A differentiable manifold with a smooth group operation and inverse map is called a *Lie group*; of course from this discussion, these matrix groups are Lie groups.

6. Topological Properties

Definition 6.1. A *matrix group* is a closed subgroup of $GL_n(\mathbb{K})$.

Notice that $GL_n(\mathbb{K})$ is open in $M_n(\mathbb{K})$, since $M_n(\mathbb{K}) \setminus GL_n(\mathbb{K}) = \det^{-1}(\{0\})$ is closed in $M_n(\mathbb{K})$ by continuity of the determinant function.

Observation 1. All the groups we mentioned above are matrix groups.

Proof: $SL_n(\mathbb{K}) = \det^{-1}(\{1\})$ which is closed in $GL_n(\mathbb{K})$ since the determinant function is continuous. To see $\mathcal{O}_n(\mathbb{K})$ is closed, let $A \in GL_n(\mathbb{K})$ and define a continuous function $\gamma : GL_n(\mathbb{K}) \to GL_n(\mathbb{K})$ by $\gamma(A) = AA^*$. Then $\mathcal{O}_n(\mathbb{K}) = \gamma^{-1}(I)$, which is closed in $GL_n(\mathbb{K})$ since the single element I is closed in $GL_n(\mathbb{K})$. We already defined $SO(n) = O(n) \cap SL_n(\mathbb{R})$ and $SU(n) = U(n) \cap SL_n(\mathbb{C})$, and since intersections of closed sets are themselves closed, we are done. $\square$

The matrix groups $\mathcal{O}_n(\mathbb{K}), SO(n),$ and $SU(n)$ are all compact since they are closed (shown in Observation 1) and bounded (their rows, and also columns, have norm 1). However, $GL_n(\mathbb{K})$ is not compact for any n because for any $m \in \mathbb{R}$, there exists an $A \in GL_n(\mathbb{K})$ with $\det(A) = m$ which implies that $GL_n(\mathbb{K})$ is not bounded. Furthermore, $SL_n(\mathbb{K})$ is not bounded for $n > 1$ since for any $r \in \mathbb{R}$, $diag(r, 1/r, 1, \ldots, 1) \in SL_n(\mathbb{K})$, showing the group is not compact for any $n > 1$. We will show below that $SL_1(\mathbb{H})$ is actually isomorphic to $Sp(1)$, thus showing $SL_1(\mathbb{H})$ is compact. By noticing $SL_1(\mathbb{R}) = SL_1(\mathbb{C}) = \{1\}$, these results will show that $SL_n(\mathbb{K})$ is compact only for $n = 1$.

Observation 2. $Sp(1) \cong SL_1(\mathbb{H})$.

Proof: Let $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in SL_1(\mathbb{H})$. By definition,

$$\begin{aligned} 1 &= \det(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \\ &= \det(\Psi_1(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})) \end{aligned}$$

$$= \det\left(\begin{pmatrix} a+b\mathbf{i} & c+d\mathbf{i} \\ -c+d\mathbf{i} & a-b\mathbf{i} \end{pmatrix}\right)$$

$$= a^2+b^2+c^2+d^2.$$

Thus,

$$SL_1(\mathbb{H}) = \{a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k} \in \mathbb{H} \mid a^2+b^2+c^2+d^2 = 1\} = S^3.$$

Any element $q \in Sp(1)$ has the property that $qq^* = 1$, but for $q = a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$, $q^* = \overline{q}$. So $q\overline{q} = |q|^2 = a^2+b^2+c^2+d^2$, which gives

$$Sp(1) = \{a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k} \in \mathbb{H} \mid a^2+b^2+c^2+d^2 = 1\} = S^3.$$

Since these equal sets have the same group operation and the same topology as well, we conclude $Sp(1) \cong SL_1(\mathbb{H})$. □

We will now discuss another important topological property: path-connectedness. As an example, we will show that $O(n)$ is not path-connected. We claim that there is no path from $A = diag(-1,1,\dots,1) \in O(n)$ to the identity, $I \in O(n)$. If there were a continuous path $\gamma : [0,1] \to O(n)$ with $\gamma(0) = A$ and $\gamma(1) = I$, then the composition $\det(\gamma([0,1]))$ would be path-connected since det is a continuous function. But for an element in $O(n)$, the determinant is either -1 or 1 and the set $\{1,-1\}$ is not path-connected. Since $\det(A) = -1$ and $\det(I) = 1$, we have a contradiction.

By noticing that $\det : GL_n(\mathbb{R}) \to (-\infty,0)\cup(0,\infty)$, a similar argument as above will show that there does not exist a continuous path from the identity to the matrix $A = diag(-1,1,\dots,1)$. This proves that $GL_n(\mathbb{R})$ is not path-connected for any n.

Hall [3] and Stillwell [4] provide proofs using only information covered up to this point that $SO(n)$, $SU(n)$, $U(n)$, $Sp(n)$, $SL_n(\mathbb{R})$, $SL_n(\mathbb{C})$, and $GL_n(\mathbb{C})$ are all path-connected. The proofs easily extend to show $SL_n(\mathbb{H})$ and $GL_n(\mathbb{H})$ are also path-connected.

7. The Exponential Map

We now define the exponential of a matrix. For any $A \in M_n(\mathbb{K})$,

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k.$$

In fact, this series converges for any square matrix A [5].

Properties of the Exponential Map

Let $X, Y \in M_n(\mathbb{K})$.

(i) If $XY = YX$, then $e^{X+Y} = e^X e^Y$.

(ii) If Y^{-1} exists, then $e^{YXY^{-1}} = Ye^X Y^{-1}$.

(iii) $(e^X)^* = e^{X^*}$.

(iv) The continuous function $\alpha : \mathbb{R} \to M_n(\mathbb{K})$ given by $\alpha(t) = e^{Xt}$ is differentiable and has derivative $\alpha'(t) = X \cdot \alpha(t)$.

(v) For $A \in M_n(\mathbb{K}),\ e^A \in GL_n(\mathbb{K})$.

(vi) If $A + A^* = 0$, then $e^A \in \mathcal{O}_n(\mathbb{K})$.

Proof: (i) We direct the reader to any of [1, 3, 5]. We will provide a counterexample in the case $XY \neq YX$ at the end of this section.

(ii) Suppose Y^{-1} exists. Then

$$\begin{aligned}
e^{YXY^{-1}} &= I + YXY^{-1} + \frac{1}{2!}(YXY^{-1})^2 + \frac{1}{3!}(YXY^{-1})^3 + \cdots \\
&= I + YXY^{-1} + \frac{1}{2!}YX^2Y^{-1} + \frac{1}{3!}YX^3Y^{-1} + \cdots \\
&= Y(I + X + \frac{1}{2!}X + \frac{1}{3!}X + \cdots)Y^{-1} = Ye^XY^{-1}.
\end{aligned}$$

(iii) We prove this directly:

$$\begin{aligned}
(e^X)^* &= (I + X + \frac{1}{2!}X^2 + \frac{1}{3!}X^3 + \cdots)^* \\
&= I^* + X^* + \frac{1}{2!}(X^2)^* + \frac{1}{3!}(X^3)^* + \cdots \\
&= I + X^* + \frac{1}{2!}(X^*)^2 + \frac{1}{3!}(X^*)^3 + \cdots = e^{X^*}.
\end{aligned}$$

The last line of the proof relies on the fact that $(X^n)^* = (X^*)^n$ which follows from the fact that for any two square matrices, $(AB)^* = B^*A^*$ as mentioned in the introduction.

(iv) To see this, notice that

$$\begin{aligned}
\alpha(t) &= I + Xt + \frac{1}{2!}X^2t^2 + \frac{1}{3!}X^3t^3 + \cdots \\
\text{which implies} \quad \alpha'(t) &= X + X^2t + \frac{1}{2!}X^3t^2 + \cdots \\
&= X(I + Xt + \frac{1}{2!}X^2t^2 + \frac{1}{3!}X^3t^3 + \cdots) = X \cdot \alpha(t).
\end{aligned}$$

The term-by-term differentiation is justified because each entry of the matrix $\alpha(t)$ is a power series in t.

(v) Since A and $-A$ commute, we will use property (i)

$$I = e^0 = e^{A-A} = e^A e^{-A} \implies (e^A)^{-1} = e^{-A} \implies e^A \in GL_n(\mathbb{K})$$

(vi) We know that $A + A^* = 0$, so $A^* = -A$. But A commutes with $-A$, so A and A^* commute. Then by property (i),

$$e^A e^{A^*} = e^{A+A^*} = e^0 = I.$$

But by property (iii), $e^{A^*} = (e^A)^*$, so $e^A(e^A)^* = I \implies e^A \in \mathcal{O}_n(\mathbb{K})$. $\square$

Proposition 2. For any $A \in M_n(\mathbb{H})$, $e^{\Psi_n(A)} = \Psi_n(e^A)$.

Proof: We will rely on the fact that Ψ_n is a continuous ring homomorphism.

$$\begin{aligned}
e^{\Psi_n(A)} &= I + \Psi_n(A) + \frac{\Psi_n(A)^2}{2!} + \frac{\Psi_n(A)^3}{3!} + \ldots \\
&= \sum_{k=0}^{\infty} \frac{\Psi_n(A)^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{\Psi_n(A^k)}{k!} \qquad \text{from property (3) of Proposition 1}
\end{aligned}$$

$$= \Psi_n\left(\sum_{k=0}^{\infty} \frac{A^k}{k!}\right) \quad \text{from properties (1) and (2) of Proposition 1}$$

$$= \Psi_n(e^A). \qquad \square$$

Since Ψ_n is an injective map, this result states that for a matrix $A \in M_n(\mathbb{H})$,

$$e^A = \Psi_n^{-1}(e^{\Psi_n(A)}).$$

In particular, in order to find the exponential of a matrix over the quaternions, we will transform it into a complex matrix, calculate its exponential, and bring that matrix back into the quaternions. This is helpful because it allows us to use the well-known techniques of diagonalization, SN Decomposition, Jordan forms, etc. defined for complex matrices. The drawback is that Ψ_n is a map from $M_n(\mathbb{H})$ into $M_{2n}(\mathbb{C})$, so the dimension of the matrix we are dealing with is doubled. For this reason we will only do computations with the low-dimensional quaternionic matrix groups.

Theorem 1. Let $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ with $v = \sqrt{b^2 + c^2 + d^2} \neq 0$, then

$$e^q = e^a(\cos v + \frac{b\mathbf{i} + c\mathbf{j} + d\mathbf{k}}{v} \sin v).$$

Proof: We can write $b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = b\mathbf{i} + (c + d\mathbf{i})\mathbf{j}$, so that

$$M = \Psi_1(b\mathbf{i} + (c + d\mathbf{i})\mathbf{j}) = \begin{pmatrix} b\mathbf{i} & c + d\mathbf{i} \\ -c + d\mathbf{i} & -b\mathbf{i} \end{pmatrix}.$$

This matrix has characteristic polynomial $\lambda^2 + v^2 = 0$ from which we have eigenvalues $\lambda = v\mathbf{i}$ and $\lambda = -v\mathbf{i}$ with corresponding eigenvectors

$$\begin{pmatrix} c + d\mathbf{i} \\ (v - b)\mathbf{i} \end{pmatrix} \text{ and } \begin{pmatrix} c + d\mathbf{i} \\ -(b + v)\mathbf{i} \end{pmatrix}, \text{ respectively.}$$

We may now diagonalize our matrix M in the standard way:

$$\begin{pmatrix} b\mathbf{i} & c+d\mathbf{i} \\ -c+d\mathbf{i} & -b\mathbf{i} \end{pmatrix} = \frac{1}{D}\begin{pmatrix} c+d\mathbf{i} & c+d\mathbf{i} \\ (v-b)\mathbf{i} & -(b+v)\mathbf{i} \end{pmatrix}\begin{pmatrix} v\mathbf{i} & 0 \\ 0 & -v\mathbf{i} \end{pmatrix}\begin{pmatrix} -(b+v)\mathbf{i} & -c-d\mathbf{i} \\ (b-v)\mathbf{i} & c+d\mathbf{i} \end{pmatrix}$$

Where $D = (c+d\mathbf{i})(-2v\mathbf{i})$ is the determinant of the matrix formed by the eigenvectors of M. We know from property (ii) of the exponential map that $e^{ABA^{-1}} = Ae^BA^{-1}$. Therefore, to calculate the exponential of M, we simply need to calculate the exponential of $diag(v\mathbf{i}, -v\mathbf{i})$ and multiply by the matrix formed by the eigenvalues on the left and on the right by the inverse of the matrix formed by the eigenvectors. But $e^{diag(v\mathbf{i},-v\mathbf{i})}$ is simply $diag(e^{v\mathbf{i}}, e^{-v\mathbf{i}})$, so we see that

$$e^M = \frac{1}{D}\begin{pmatrix} c+d\mathbf{i} & c+d\mathbf{i} \\ (v-b)\mathbf{i} & -(b+v)\mathbf{i} \end{pmatrix}\begin{pmatrix} e^{v\mathbf{i}} & 0 \\ 0 & e^{-v\mathbf{i}} \end{pmatrix}\begin{pmatrix} -(b+v)\mathbf{i} & -c-d\mathbf{i} \\ (b-v)\mathbf{i} & c+d\mathbf{i} \end{pmatrix}.$$

From this we get

$$\begin{aligned} e^M &= \begin{pmatrix} \frac{e^{v\mathbf{i}}+e^{-v\mathbf{i}}}{2} + \frac{b}{v}(\frac{e^{v\mathbf{i}}-e^{-v\mathbf{i}}}{2}) & \frac{e^{v\mathbf{i}}-e^{-v\mathbf{i}}}{2v\mathbf{i}}(c+d\mathbf{i}) \\ \frac{e^{v\mathbf{i}}-e^{-v\mathbf{i}}}{-2v\mathbf{i}}(c-d\mathbf{i}) & \frac{e^{v\mathbf{i}}+e^{-v\mathbf{i}}}{2} - \frac{b}{v}(\frac{e^{v\mathbf{i}}-e^{-v\mathbf{i}}}{2}) \end{pmatrix} \\ &= \begin{pmatrix} \cos v + \frac{b}{v}\mathbf{i}\sin v & \frac{c}{v}\sin v + \frac{d}{v}\mathbf{i}\sin v \\ -\frac{c}{v}\sin v + \frac{d}{v}\mathbf{i}\sin v & \cos v - \frac{b}{v}\mathbf{i}\sin v \end{pmatrix} \\ &= \Psi_1(\cos v + \frac{b}{v}\mathbf{i}\sin v + \frac{c}{v}\mathbf{j}\sin v + \frac{d}{v}\mathbf{k}\sin v) \\ &= \Psi_1(\cos v + \frac{b\mathbf{i}+c\mathbf{j}+d\mathbf{k}}{v}\sin v). \end{aligned}$$

Since $e^{a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}} = e^a e^{b\mathbf{i}+c\mathbf{j}+d\mathbf{k}}$ from the discussion in the introduction that the real numbers commute with the quaternions and from property (i) of the exponential map, we conclude that

$$e^q = e^{a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}} = e^a(\cos v + \frac{b\mathbf{i}+c\mathbf{j}+d\mathbf{k}}{v}\sin v), \text{ for } \sqrt{b^2+c^2+d^2} \neq 0.$$

Notice that when $v = \sqrt{b^2+c^2+d^2} = 0$, we must have $b = c = d = 0$. Since $\lim_{v\to 0} \frac{\sin v}{v} = 1$, we have $e^q = e^a$. □

At this point we are ready to show that $e^{X+Y} \neq e^X e^Y$ for some $X, Y \in M_n(\mathbb{K})$ (see property (i) of the exponential map).

Counterexample: Consider $\mathbf{i}, \mathbf{j} \in \mathbb{H}$. We already know that $\mathbf{i} \cdot \mathbf{j} \neq \mathbf{j} \cdot \mathbf{i}$. Using the previous theorem,

$$e^{\mathbf{i}+\mathbf{j}} = \cos\sqrt{2} + \tfrac{\mathbf{i}+\mathbf{j}}{\sqrt{2}}\sin\sqrt{2}.$$

However, $e^{\mathbf{i}}e^{\mathbf{j}} = (\cos 1 + \mathbf{i}\sin 1)(\cos 1 + \mathbf{j}\sin 1) = \cos^2 1 + (\mathbf{i}+\mathbf{j})\cos 1 \sin 1 + \mathbf{k}\sin^2 1$. Clearly these two values are not equal, in particular, the former doesn't even have a $\mathbf{k}$ term. □

8. Lie Algebras

Let S be a submanifold of $\mathbb{R}^n$. For some point $x \in S$, the tangent space to S at x is the set of initial velocity vectors of differentiable paths in S through x:

$$T_xS = \{\gamma'(0) \mid \gamma : (-\epsilon, \epsilon) \to S \text{ is differentiable with } \gamma(0) = x\}.$$

Since our matrix groups are subsets of Euclidean space, we give the following definition:

Definition 8.1. The *Lie algebra* of a matrix group G is the tangent space at the identity along with the extra bracket structure explained below. We denote it $\mathfrak{g} = T_IG$. When the group is specified, the Lie algebra of the group will be denoted by the same letters used to represent the group, except lower-case. The Lie algebra of a matrix group has the following important properties:

Properties of the Lie Algebra

(1) The Lie algebra $\mathfrak{g}$ of a matrix group G is a real vector space.

(2) The Lie algebra is closed under the Lie bracket operation $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ given by $[A, B] = AB - BA$.

Proof: For both proofs, let $A, B \in \mathfrak{g}$. Then, A, B are initial velocity vectors of differentiable paths in G through the identity. Call these paths $\alpha(t)$ and $\beta(t)$, respectively, so that $A = \alpha'(0)$ and $B = \beta'(0)$, where $\alpha(0) = I = \beta(0)$.

To prove (1), for some fixed scalars $\lambda, \mu \in \mathbb{R}$, let $\gamma(t) = \alpha(\lambda t) \cdot \beta(\mu t)$. Then $\gamma(t)$ is a differentiable path in G with $\gamma(0) = I$, so $\gamma'(0) \in \mathfrak{g}$ by definition of the tangent space. But by the product rule,

$$\gamma'(0) = \lambda\ \alpha'(0) \cdot \beta(0) + \mu\ \alpha(0) \cdot \beta'(0) = \lambda A + \mu B \in \mathfrak{g}.$$

Thus, $\mathfrak{g}$ is a real vector space.

To prove (2), fix $s \in \mathbb{R}$. Then the path $\gamma_s(t) = \alpha(s) \cdot \beta(t) \cdot \alpha^{-1}(s)$ is differentiable in G and $\gamma_s(0) = I$, so $\gamma_s'(0) \in \mathfrak{g}$. But $\gamma_s'(0) = \alpha(s) \cdot B \cdot \alpha^{-1}(s)$ is a differentiable path in $\mathfrak{g}$ (since α is differentiable), whose derivative is in $\mathfrak{g}$ because $\mathfrak{g}$ is closed under limits. If we let s vary and set $\delta(s) = \gamma_s'(0)$, then

$$\delta'(0) = \alpha'(0) \cdot B \cdot \alpha^{-1}(0) + \alpha(0) \cdot B \cdot (\alpha^{-1})'(0) = AB - BA.$$

This follows from the fact that $(\alpha^{-1})'(0)$ is $-A$ because α^{-1} traverses α in the opposite direction. So the Lie algebra is closed under the bracket operation. □

We mentioned earlier that these matrix groups are manifolds. We are now in a position to talk about their dimensions:

Definition 8.2. We define the *dimension* of a matrix group G to be the dimension of its Lie algebra $\mathfrak{g}$ as a real vector space.

Just as we used $\mathcal{O}_n(\mathbb{K})$ to denote $O(n), U(n)$, and $Sp(n)$, we will write $o(n)$ to mean $so(n), u(n)$, and $sp(n)$, the Lie algebras of $SO(n), U(n)$, and $Sp(n)$, respectively. Notice that $SO(n)$ is the connected component of $O(n)$ which contains the identity, so the Lie algebra of $O(n)$ is the same as the Lie algebra of $SO(n)$.

Proposition 3. We will list the Lie algebras of our matrix groups. When we list multiple dimensions for the same group, they are in order from $\mathbb{R}, \mathbb{C}, \mathbb{H}$.

Lie algebra	Description	Dimension
$gl_n(\mathbb{K})$	$M_n(\mathbb{K})$	$n^2, 2n^2, 4n^2$
$sl_n(\mathbb{R} \text{ or } \mathbb{C})$	$\{A \in M_n(\mathbb{R} \text{ or } \mathbb{C}) \mid trace(A) = 0\}$	$n-1, 2(n-1)$
$o(n)$	$\{A \in M_n(\mathbb{R}) \mid A + A^* = 0\}$	$\frac{n^2-n}{2}, n^2, 2n^2+n$
$su(n)$	$\{A \in u(n) \mid trace(A) = 0\}$	$n^2 - 1$

Proof: We will show the proof for $o(n)$, see any of [1, 2, 4, 5] for the remaining proofs. Let $\gamma(t)$ be a differentiable path in $\mathcal{O}_n(\mathbb{K})$ with $\gamma(0) = I$. Then $\gamma(t) \cdot \gamma(t)^* = I$ for all t. Differentiating both sides and applying the product rule yields:

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} \gamma(t) \cdot \gamma(t)^* &= \gamma'(0) + \gamma'(0)^* \\ &= 0 = \left.\frac{d}{dt}\right|_{t=0} I. \end{aligned}$$

Therefore, we have $o(n) \subseteq \{M \in M_n(\mathbb{R}) \mid M + M^* = 0\}$.
Conversely, let $A \in \{M \in M_n(\mathbb{R}) \mid M + M^* = 0\}$. Then by properties (iv) and (vi) of the exponential map, e^{At} is a differentiable path in $\mathcal{O}_n(\mathbb{K})$ through the identity, whose initial velocity vector, A, is in $o(n)$. Thus, $\{M \in M_n(\mathbb{R}) \mid M + M^* = 0\} \subseteq o(n)$, and we are done. Determining the dimensions for $o(n)$ boils down to finding the entries of the matrix that are independent of the others. First, notice that there cannot be any real elements along the diagonal, and second, that we may choose any entries either above or below the diagonal. Counting these up gives our dimensions listed above. $\square$

Note: It is not necessarily true that the Lie algebra of a matrix group is a vector space over $\mathbb{C}$ or $\mathbb{H}$ (see Property (1) of the Lie algebra). For example, $u(1) = \{x\mathbf{i} \mid x \in \mathbb{R}\}$, and clearly $x\mathbf{i} \cdot \mathbf{i} = -x \notin u(1)$. Furthermore, the Lie algebra of a matrix group is not closed under multiplication (although it is closed under the Lie bracket, see Property (2) of the Lie algebra). For example, the matrices $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ are in the Lie algebra $so(2)$ of $SO(2)$, but their product, I, is not.

For a matrix A, we will introduce the notation $A[i, j]$ to denote the matrix formed by removing row i and column j from the original matrix A. For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} [2, 2] = \begin{pmatrix} 1 & 3 \end{pmatrix}.$$

Theorem 2. Let $\gamma(t) : (-\epsilon, \epsilon) \to M_n(\mathbb{H})$ be a differentiable function with $\gamma(0) = I$, then

$$\left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)) = trace(\Psi_n(\gamma'(0))).$$

Proof: Recall that for $A \in M_n(\mathbb{H})$, $\det(A)$ means $\det(\Psi_n(A))$. So,

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)) &= \left.\frac{d}{dt}\right|_{t=0} \det(\Psi_n(\gamma(t))) \\
&= \left.\frac{d}{dt}\right|_{t=0} \sum_{m=1}^{2n} (-1)^{m+1} \cdot \Psi_n(\gamma(t))_{1m} \cdot \det(\Psi_n(\gamma(t))[1, m]) \\
&= \sum_{m=1}^{2n} (-1)^{m+1} \left.\frac{d}{dt}\right|_{t=0} \Psi_n(\gamma(t))_{1m} \cdot \det(\Psi_n(\gamma(t))[1, m]).
\end{aligned}$$

This construction of the determinant uses the $A[i, j]$ notation introduced above. In the last line, we simply distribute the derivative through the summation. Applying the product rule and the easy observation that $\frac{d}{dt}\Psi_n(\gamma(t)) = \Psi_n(\gamma'(t))$ gives us:

$$\begin{aligned}
= \sum_{m=1}^{2n} (-1)^{m+1} \Big(&\Psi_n(\gamma'(0))_{1m} \cdot \det(\Psi_n(\gamma(0))[1, m]) \\
&+ \Psi_n(\gamma(0))_{1m} \cdot \left.\frac{d}{dt}\right|_{t=0} \det(\Psi_n(\gamma(t))[1, m]) \Big).
\end{aligned}$$

In this summation, all terms except for the first will be zero. In the first line of this summation, this is because

$$\det(\Psi_n(\gamma(0))[1, m]) = \begin{cases} 1, & \text{m=1} \\ 0, & \text{otherwise.} \end{cases}$$

In the second line of the summation, $\Psi_n(\gamma(0)) = \Psi_n(I_{n\times n}) = I_{2n\times 2n}$, which shows that

$$\Psi_n(\gamma(0))_{1m} = \begin{cases} 1, & \text{m=1} \\ 0, & \text{otherwise.} \end{cases}$$

These two facts give us

$$\left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)) = \Psi_n(\gamma'(0))_{11} + \left.\frac{d}{dt}\right|_{t=0} \det(\Psi_n(\gamma(t))[1,1]).$$

This argument may be reused to show

$$\left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)[1,1]) = \Psi_n(\gamma'(0))_{22} + \left.\frac{d}{dt}\right|_{t=0} \det(\Psi_n(\gamma(t))[1,1][2,2]).$$

This process will terminate after $2n$ steps, giving

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)) &= \Psi_n(\gamma'(0))_{11} + \Psi_n(\gamma'(0))_{22} + \cdots + \Psi_n(\gamma'(0))_{2n2n} \\ &= trace(\Psi_n(\gamma'(0))). \qquad \square\end{aligned}$$

Theorem 3. For $A \in M_n(\mathbb{H})$ and $t \in \mathbb{R}$, we have

$$\det(e^{At}) = e^{trace(\Psi_n(A))\cdot t}.$$

Proof: In the last line of this proof, we rely on Theorem 2 and property (iv) of the exponential map: that e^{At} is differentiable, with derivative Ae^{At}.

$$\begin{aligned}\frac{d}{dt}\det(e^{At}) &= \lim_{h\to 0} \frac{\det(e^{A(t+h)}) - \det(e^{At})}{h} \\ &= \lim_{h\to 0} \frac{\det(e^{At}e^{Ah}) - \det(e^{At})}{h} \\ &= \left(\lim_{h\to 0} \frac{\det(e^{Ah}) - 1}{h}\right) \cdot \det(e^{At}) \\ &= \left(\left.\frac{d}{dy}\right|_{y=0} \det(e^{Ay})\right) \cdot \det(e^{At}) \\ &= trace(\Psi_n(A)) \cdot \det(e^{At}).\end{aligned}$$

Since $\det(e^{At})\big|_{t=0} = 1$, the unique solution to this differential equation is

$$\det(e^{At}) = e^{trace(\Psi_n(A))\cdot t}. \qquad \square$$

Theorem 4. The Lie algebra $sl_n(\mathbb{H})$ of $SL_n(\mathbb{H})$ is

$$\{A \in M_n(\mathbb{H}) \mid trace(\Psi_n(A)) = 0\}.$$

Proof: Let $\gamma(t)$ be a differentiable path in $SL_n(\mathbb{H})$ with $\gamma(0) = I$, so $\det(\gamma(t)) = 1$ for all t. Then by Theorem 2,

$$\left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)) = \left.\frac{d}{dt}\right|_{t=0} 1 = 0 = trace(\Psi_n(\gamma'(0))).$$

Thus, $sl_n(\mathbb{H}) \subseteq \{M \in M_n(\mathbb{H}) \mid trace(\Psi_n(M)) = 0\}$.
Conversely, let $A \in M_n(\mathbb{H})$ with $trace(\Psi_n(A)) = 0$. Then by property (iv) of the exponential map, $\alpha(t) = e^{At}$ is a differentiable path with $\alpha(0) = I$ and $\alpha'(0) = A$. In fact, $\alpha(t)$ is in $SL_n(\mathbb{H})$ because

$$\begin{aligned} \det(e^{At}) &= e^{trace(\Psi_n(A))\cdot t} && \text{by Theorem 3} \\ &= e^0 = 1 && \text{since } trace(\Psi_n(A)) = 0. \end{aligned}$$

Therefore, $\{M \in M_n(\mathbb{H}) \mid trace(\Psi_n(M)) = 0\} \subseteq sl_n(\mathbb{H})$. The double inclusion proves the theorem. $\square$

We will now provide a characterization of $sl_n(\mathbb{H})$ that will be easier to work with since it does not involve the map Ψ_n. Let $q_{ij} \in \mathbb{H}$ denote the i, j-entry of some matrix $A \in M_n(\mathbb{H})$. We can write $q_{ij} = z_{ij} + w_{ij}\mathbf{j}$ for some $z, w \in \mathbb{C}$. Notice that

$$\begin{aligned} trace(\Psi_n(A)) &= z_{11} + \overline{z_{11}} + z_{22} + \overline{z_{22}} + \cdots + z_{nn} + \overline{z_{nn}} \\ &= 2 \cdot \mathrm{Re}(\sum_{k=1}^{n} z_{kk}) \\ &= 2 \cdot \mathrm{Re}(\sum_{k=1}^{n} q_{kk}) = 2 \cdot \mathrm{Re}(trace(A)). \end{aligned}$$

Therefore, $trace(\Psi_n(A)) = 0$ if and only if $\text{Re}(trace(A)) = 0$, so we can characterize the Lie algebra of $SL_n(\mathbb{H})$ as $sl_n(\mathbb{H}) = \{A \in M_n(\mathbb{H}) \mid \text{Re}(trace(A)) = 0\}$.

Now we can determine that the dimension of $SL_n(\mathbb{H})$ is $4n^2 - 1$. This is because every entry of a matrix in $sl_n(\mathbb{H})$ has the form $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, except for the last diagonal entry, whose real part must be the negative of the real parts of the preceding diagonal entries. Thus we have $4n^2 - 1$ choices for entries.

9. Ideals and Normal Subgroups

Definition 9.1. For a matrix group G with Lie algebra $\mathfrak{g}$, a vector subspace $\mathfrak{h} \subset \mathfrak{g}$ is called an *ideal* of $\mathfrak{g}$ if, for all $A \in \mathfrak{h}$ and $B \in \mathfrak{g}$, $[A, B] \in \mathfrak{h}$.

Theorem 5. If H is a normal subgroup of a matrix group G (whose Lie algebras are $\mathfrak{h}$ and $\mathfrak{g}$, respectively), then $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

Proof: To see this, notice that for any paths $\alpha(t) \in G$ and $\beta(s) \in H$, with $\alpha(0) = \beta(0) = I$ and $\alpha'(0) = A$ and $\beta'(0) = B$, we have $\alpha(t)\beta(s)\alpha^{-1}(t) \in H$ since H is a normal subgroup of G. Then

$$\left.\frac{d}{ds}\right|_{s=o} \alpha(t)\beta(s)\alpha^{-1}(t) = \alpha(t)\beta'(0)\alpha^{-1}(t) = \alpha(t)B\alpha^{-1}(t)$$

is a smooth path in $\mathfrak{h}$, whose derivative with respect to t is in $\mathfrak{h}$:

$$\left.\frac{d}{dt}\right|_{t=o} \alpha(t)B\alpha^{-1}(t) = \alpha'(0)B\alpha^{-1}(0) + \alpha(0)B(\alpha'(0))^{-1} = [A, B].$$

So $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. □

It turns out that $sp(n)$ has no nontrivial ideals [4]; however, $gl_n(\mathbb{H})$ does, as we will now show.

Proposition 4. $sl_n(\mathbb{H})$ is an ideal in $gl_n(\mathbb{H})$.

Proof: The determinant function, $\det : GL_n(\mathbb{H}) \to \mathbb{C}^*$ is a group homomorphism since $\det(AB) = \det(\Psi_n(AB)) = \det(\Psi_n(A)\Psi_n(B)) = \det(A)\det(B)$. Therefore, since $SL_n(\mathbb{H})$ is the kernel of the determinant function, it is normal in $GL_n(\mathbb{H})$. By Theorem 5, $sl_n(\mathbb{H})$ is an ideal in $gl_n(\mathbb{H})$. □

Proposition 5. The set $\mathcal{S} = \{x \cdot I \mid x \in \mathbb{R}\}$ is an ideal in $gl_n(\mathbb{H})$.

Proof: We will prove directly using the definition of an ideal. The set $\mathcal{S}$ is a vector subspace of $gl_n(\mathbb{H})$, since for any $X, Y \in \mathcal{S}$ and $r \in \mathbb{R}$, $r(X + Y) = rX + rY \in \mathcal{S}$. To see $\mathcal{S}$ is an ideal in $gl_n(\mathbb{H})$, let $A \in gl_n(\mathbb{H})$ and $X \in \mathcal{S}$ with $X = xI$ for some $x \in \mathbb{R}$. Then

$$\begin{aligned}[X, A] &= X \cdot A - A \cdot X = xI \cdot A - A \cdot xI \\ &= xA - xA = 0 \in \mathcal{S}.\end{aligned}$$

This last line relies on the fact that every real number commutes with any quaternion, as discussed in the introduction.

Proposition 6. Every element of $M_n(\mathbb{H})$ can be written uniquely as a sum of an element in $\mathcal{S}$ and an element in $sl_n(\mathbb{H})$.

Proof: Recall that $sl_n(\mathbb{H})$ has two equivalent definitions:

$$\{A \in M_n(\mathbb{H}) \mid trace(\Psi_n(A)) = 0\} \quad \text{and} \quad \{A \in M_n(\mathbb{H}) \mid \operatorname{Re}(trace(A)) = 0\}.$$

For this proof, we will use the latter of the two. Let $A \in M_n(\mathbb{H})$ and denote the i, j-entry of A by a_{ij}. Then

$$\operatorname{Re}(trace(A)) = \operatorname{Re}\left(\sum_{k=1}^{n}(a_{kk})\right) = \sum_{k=1}^{n} \operatorname{Re}(a_{kk}).$$

If we set $x = \frac{1}{n}\sum_{k=1}^{n} \operatorname{Re}(a_{kk})$ and $\beta_{ii} = \operatorname{Re}(a_{ii}) - x$, we see that $x, \beta_{ii} \in \mathbb{R}$ and that

$$\begin{aligned}\sum_{k=1}^{n} \beta_{kk} &= \sum_{k=1}^{n}(\operatorname{Re}(a_{kk}) - x) \\ &= \sum_{k=1}^{n}\left(\operatorname{Re}(a_{kk}) - \frac{1}{n}\sum_{m=1}^{n} \operatorname{Re}(a_{mm})\right) \\ &= \sum_{k=1}^{n} \operatorname{Re}(a_{kk}) - n \cdot \frac{1}{n}\sum_{k=1}^{n} \operatorname{Re}(a_{kk}) = 0\end{aligned}$$

Let B be the matrix formed by replacing each $\operatorname{Re}(a_{ii})$ entry of A by β_{ii} while leaving the other elements of A untouched. Then $B \in sl_n(\mathbb{H})$ since $\operatorname{Re}(trace(B)) = 0$. Furthermore, $A = B + xI$ and this decomposition is unique. $\square$

Corollary 1. Let $A \in M_n(\mathbb{H})$, then $e^A = e^x e^B$, where x and B are as in the previous proof.

Proof: Since $x \in \mathbb{R}$, x commutes with all quaternions. Furthermore, $xI \cdot B = xB$ and $B \cdot xI = xB \cdot I = xB$, so xI and B commute. By property (i) of the exponential map,

$$e^A = e^{xI+B} = e^{xI}e^B.$$

Notice that

$$\begin{aligned} e^{xI} &= I + xI + \frac{1}{2!}(xI)^2 + \frac{1}{3!}(xI)^3 + \cdots \\ &= I + xI + \frac{1}{2!}x^2 I + \frac{1}{3!}x^3 I + \cdots \\ &= e^x I. \end{aligned}$$

This gives $e^A = e^{xI}e^B = e^x I \cdot e^B = e^x \cdot e^B$. $\square$

The question of determining the exponential of any given matrix over the quaternions becomes a slightly simpler question of determining the exponential of the associated matrix over the quaternions whose real part of the trace equal to zero.

10. Maximal Tori

Definition 10.1. An *n-dimensional torus* in $GL_n(\mathbb{C})$, denoted $\mathbb{T}^n$, is a Cartesian product of n circle groups

$$\mathbb{T}^n = \{diag(e^{\mathbf{i}\theta_1}, \dots, e^{\mathbf{i}\theta_n}) \mid \theta_m \in [0, 2\pi)\}.$$

A torus is called maximal if, whenever it is contained in another torus, the tori are equal. Notice that, since the circle group is abelian, so is the n-dimensional torus. This observation will be important in determining a standard maximal torus for our matrix groups.

Proposition 7. We have the following standard maximal tori of the matrix groups we have discussed.

$$T = \{diag(\left(\begin{smallmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{smallmatrix}\right), \dots, \left(\begin{smallmatrix} \cos\theta_n & \sin\theta_n \\ -\sin\theta_n & \cos\theta_n \end{smallmatrix}\right)) \mid \theta_m \in [0, 2\pi)\} \subset SO(2n)$$

$$T = \{diag(\left(\begin{smallmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{smallmatrix}\right), \dots, \left(\begin{smallmatrix} \cos\theta_n & \sin\theta_n \\ -\sin\theta_n & \cos\theta_n \end{smallmatrix}\right), 1) \mid \theta_m \in [0, 2\pi)\} \subset SO(2n+1)$$

$$T = \{diag(e^{\mathbf{i}\theta_1}, \dots, e^{\mathbf{i}\theta_n}) \mid \theta_m \in [0, 2\pi), \textstyle\sum_{i=1}^n \theta_i = 0\} \subset SU(n)$$

$$T = \{diag(e^{\mathbf{i}\theta_1}, \dots, e^{\mathbf{i}\theta_n}) \mid \theta_m \in [0, 2\pi)\} \subset U(n)$$

$$T = \{diag(e^{\mathbf{i}\theta_1}, \dots, e^{\mathbf{i}\theta_n}) \mid \theta_m \in [0, 2\pi)\} \subset Sp(n)$$

Although we talk about the standard maximal torus, we recognize that there are other maximal tori. For example, if T is a maximal torus in a matrix group G, then so is gTg^{-1} for any $g \in G$. This is because, the conjugation map Ad_g is an automorphism of G, and thus gTg^{-1} is isomorphic to T, meaning that if gTg^{-1} were not maximal, then T would not be either.

Since we are most interested in the symplectic groups, we will prove that the standard maximal torus listed above is a maximal torus in $Sp(n)$.

Proof: Let $\mathbf{e}_i$ represent the row vector in $\mathbb{H}^n$ with a 1 in the i^{th} position and zeros elsewhere. The torus T we wrote above is clearly a torus and is in $Sp(n)$ since $T \cdot T^* = I$. Suppose this torus were contained in another maximal torus, T'. Choose an $A \in T'$, then since $T \subset T'$ and T' is abelian, A must commute with every element of the torus T we are claiming is maximal. In particular, A must commute with $diag(-1, 1, \ldots, 1)$. By writing $\mathbf{e}_1 A = a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + \cdots + a_{1n}\mathbf{e}_n$, where a_{ij} represents the i, j entry of A, we have

$$\mathbf{e}_1 \cdot A \cdot diag(-1, 1, \ldots, 1) = -a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + \cdots + a_{1n}\mathbf{e}_n$$

$$\mathbf{e}_1 \cdot diag(-1, 1, \ldots, 1) \cdot A = -a_{11}\mathbf{e}_1 - a_{12}\mathbf{e}_2 - \cdots - a_{1n}\mathbf{e}_n.$$

The equality of these implies that $a_{1j} = 0$ for $j \in \{2, 3, \ldots, n\}$. Similarly, A must commute with $diag(1, -1, 1, \ldots, 1)$, which implies that $a_{2j} = 0$ for $j \in \{1, 3, 4, \ldots, n\}$. Continuing this way yields that $A = diag(a_{11}, a_{22}, \ldots, a_{nn})$, where $a_{kk} \in \mathbb{H}$. But since A must commute with $diag(\mathbf{i}, \ldots, \mathbf{i})$, each a_{kk} must be of the form $e^{\mathbf{i}\theta_k}$ since $\mathbf{j}$ and $\mathbf{k}$ do not commute with $\mathbf{i}$. Thus, our torus above is maximal. $\square$

Properties of Maximal Tori

(i) Every maximal torus in a compact matrix group is a maximal abelian subgroup.

(ii) Let G be a compact, connected matrix group and T its standard maximal torus. Then

$$G = \bigcup_{x \in G} xTx^{-1}.$$

We mentioned earlier that conjugation of the standard maximal torus by an invertible element of the matrix group gives another maximal torus. The converse is a consequence of property (ii) above: any two maximal tori in a matrix group are conjugate by an element of the group. Property (ii) also states that any element of the matrix group is conjugate to an element in the standard maximal torus. In our proof of the maximal torus in $Sp(2)$, we noted property (i). It follows that the center of a compact matrix group is contained in the maximal torus, since every element which commutes with the entire group must commute with the maximal torus, and is therefore contained in the torus. The centers are given by [2]:

$$Z(SO(2n)) = \{I, -I\}, \quad Z(SO(2n+1)) = \{I\}, \quad Z(SU(n)) = \{\omega I \mid \omega^n = 1\},$$

$$Z(U(n)) = \{e^{\mathbf{i}\theta} I \mid \theta \in [0, 2\pi)\}, \quad Z(Sp(n)) = \{I, -I\}.$$

We will illustrate how the properties of maximal tori for compact, connected matrix groups listed above do not hold for non-compact or disconnected matrix groups, such as $GL_n(\mathbb{H})$ and $O(n)$.

Observation 3. $T = \{diag(e^{\mathbf{i}\theta_1}, \ldots, e^{\mathbf{i}\theta_n}) \mid \theta_m \in [0, 2\pi)\}$ is a maximal torus in $SL_n(\mathbb{H})$ and in $GL_n(\mathbb{H})$.

Proof: Clearly T is a torus. We must first show that this torus is contained in $SL_n(\mathbb{H})$, and therefore in $GL_n(\mathbb{H})$. This follows from

$$\begin{aligned} \det(T) &= \det(\Psi_n(diag(e^{\mathbf{i}\theta_1}, \ldots, e^{\mathbf{i}\theta_n})) \\ &= \det(diag(e^{\mathbf{i}\theta_1}, e^{-\mathbf{i}\theta_1}, \ldots, e^{\mathbf{i}\theta_n}, e^{-\mathbf{i}\theta_n})) \\ &= e^{\mathbf{i}\theta_1 - \mathbf{i}\theta_1 \cdots + \mathbf{i}\theta_n - \mathbf{i}\theta_n} = e^0 = 1. \end{aligned}$$

At this point, the same argument used in the proof of Proposition 7 will show that this torus is maximal. $\square$

When we discussed compactness of matrix groups, we showed that $GL_n(\mathbb{H})$ is not compact for any n. Since real numbers commute with all quaternions, it is easy to see that for $x \in \mathbb{R}$ and $x \neq 0$, that $A = diag(x, \ldots, x) \in Z(GL_n(\mathbb{H}))$ but when $|x| \neq 1$, A is not in the standard maximal torus in $GL_n(\mathbb{H})$. Thus, the center of $GL_n(\mathbb{H})$ is not contained in the standard maximal torus as before when we dealt with compact matrix groups. Furthermore, when we mentioned connectedness of matrix groups, we showed that $O(n)$ is not connected for any n. It is clear that $O(n)$ and $SO(n)$ have the same maximal tori, since $SO(n)$ is the connected component of the identity and tori are path-connected. However, $Z(O(n)) = \{I, -I\}$ for all n, while

$$Z(SO(n)) = \begin{cases} \{I, -I\}, & \text{for } n \text{ even} \\ \{I\}, & \text{for } n \text{ odd.} \end{cases}$$

This discussion suggests that maximal tori are only useful in compact, connected matrix groups. Therefore, we will omit further discussion of maximal tori in non-compact or in disconnected matrix groups.

We may also talk about the Lie algebra of the standard maximal torus. This notion makes sense because the identity is in the standard maximal torus, and since tori are path-connected, we may find smooth paths through the identity. Since our standard maximal torus looks like a Cartesian product of circle groups (whose tangent space to the identity is the set $\{x\mathbf{i} \mid x \in \mathbb{R}\}$), we expect its Lie algebra to look like a Cartesian product of Lie algebras of the circle group. This is in fact the case, and we have the following:

Theorem 6. Let G be a matrix group with Lie algebra $\mathfrak{g}$. Then since the maximal torus of G is contained in G, the Lie algebra of the maximal torus, denoted $\tau(\mathfrak{g})$ is contained in $\mathfrak{g}$. Thus,

$$\tau(so(2n)) = \{diag(\left(\begin{smallmatrix} 0 & \theta_1 \\ -\theta_1 & 0 \end{smallmatrix}\right), \ldots, \left(\begin{smallmatrix} 0 & \theta_n \\ -\theta_n & 0 \end{smallmatrix}\right)) \mid \theta_k \in \mathbb{R}\}$$

$$\tau(so(2n+1)) = \{diag(\left(\begin{smallmatrix} 0 & \theta_1 \\ -\theta_1 & 0 \end{smallmatrix}\right), \ldots, \left(\begin{smallmatrix} 0 & \theta_n \\ -\theta_n & 0 \end{smallmatrix}\right), 1) \mid \theta_k \in \mathbb{R}\}$$

$$\tau(su(n)) = \{diag(\theta_1\mathbf{i}, \ldots, \theta_n\mathbf{i}) \mid \theta_k \in \mathbb{R} \text{ and } \textstyle\sum_{k=1}^{n} \theta_k = 0\}$$

$$\tau(u(n)) = \{diag(\theta_1\mathbf{i}, \ldots, \theta_n\mathbf{i}) \mid \theta_k \in \mathbb{R}\}$$

$$\tau(sp(n)) = \{diag(\theta_1\mathbf{i}, \ldots, \theta_n\mathbf{i}) \mid \theta_k \in \mathbb{R}\}.$$

Proof: We will prove that $\tau(sp(n)) = \{diag(\theta_1, \ldots, \theta_n) \mid \theta_i \in \mathbb{R}\}$. Let $\mathbb{T}^n$ represent the standard maximal torus in $Sp(n)$ and let $\gamma(t)$ be a differentiable path in $\mathbb{T}^n$ with $\gamma(0) = I$. Since $\gamma(t)$ is in $\mathbb{T}^n$, it must be diagonal and each diagonal element must have norm 1. Therefore, the real part of $\gamma(t)$ reaches a maximum at 0, since $\gamma(0) = I$. Thus, the real part of $\gamma'(0)$ is zero, so $\gamma'(0) = diag(\mathbf{i}\theta_1, \ldots, \mathbf{i}\theta_n)$, so that $\tau(sp(n)) \subseteq \{diag(\theta_1\mathbf{i}, \ldots, \theta_n\mathbf{i}) \mid \theta_k \in \mathbb{R}\}$. Conversely, let $A \in \{diag(\theta_1\mathbf{i}, \ldots, \theta_n\mathbf{i}) \mid \theta_k \in \mathbb{R}\}$, then $\delta(t) = e^{At}$ is a differentiable path in $\mathbb{T}^n$ with $\delta(0) = I$ and $\delta'(0) = diag(\mathbf{i}\theta_1, \ldots, \mathbf{i}\theta_n)$, which implies $\{diag(\theta_1\mathbf{i}, \ldots, \theta_n\mathbf{i}) \mid \theta_k \in \mathbb{R}\} \subseteq \tau(sp(n))$. Thus, $\tau(sp(n)) = \{diag(\theta_1\mathbf{i}, \ldots, \theta_n\mathbf{i}) \mid \theta_k \in \mathbb{R}\}$.

The proofs for $\tau(su(n))$ and $\tau(u(n))$ are similar. The proofs for $\tau(so(2n))$ and $\tau(so(2n+1))$ use the same argument and the fact that:

$$\begin{aligned} exp(\left(\begin{smallmatrix} 0 & \theta \\ -\theta & 0 \end{smallmatrix}\right)) &= I + \left(\begin{smallmatrix} 0 & \theta \\ -\theta & 0 \end{smallmatrix}\right) + \frac{1}{2!}\left(\begin{smallmatrix} 0 & \theta \\ -\theta & 0 \end{smallmatrix}\right)^2 + \frac{1}{3!}\left(\begin{smallmatrix} 0 & \theta \\ -\theta & 0 \end{smallmatrix}\right)^3 + \frac{1}{4!}\left(\begin{smallmatrix} 0 & \theta \\ -\theta & 0 \end{smallmatrix}\right)^4 + \cdots \\ &= I + \left(\begin{smallmatrix} 0 & \theta \\ -\theta & 0 \end{smallmatrix}\right) - \frac{1}{2!}\left(\begin{smallmatrix} \theta^2 & 0 \\ 0 & \theta^2 \end{smallmatrix}\right) - \frac{1}{3!}\left(\begin{smallmatrix} 0 & \theta^3 \\ -\theta^3 & 0 \end{smallmatrix}\right) + \frac{1}{4!}\left(\begin{smallmatrix} \theta^4 & 0 \\ 0 & \theta^4 \end{smallmatrix}\right) + \cdots \\ &= \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right). \end{aligned}$$

□

It is clear from this proof that, for any of the matrix groups listed above, the exponential map is surjective from $\tau(\mathfrak{g})$ onto the standard maximal torus of the matrix group. From property (ii) of maximal tori, we know that for any $g \in G$, there exists an $x \in G$ and a T in the standard maximal torus of G, such that $g = xTx^{-1}$. Surjectivity of the exponential map implies $T = e^t$ for some $t \in \tau(\mathfrak{g})$. Then

$$g = xTx^{-1} = xe^t x^{-1} = e^{xtx^{-1}}.$$

Since $xtx^{-1} \in \mathfrak{g}$, it follows that exp: $\mathfrak{g} \to G$ is surjective (where G is one of the compact, connected groups $SO(n), SU(n), U(n), Sp(n)$ from above).

Although we mentioned path-connectedness of these four matrix groups, our discussion allows for an easy proof of this fact. By surjectivity of the exponential map onto these groups, and from property (vi) of the exponential map, it follows that every element $A \in G$, where $e^{\alpha} = A$ for $\alpha \in \mathfrak{g}$, can be connected to the identity by the path $e^{\alpha t}$.

We mentioned above that there are special orthogonal and special unitary groups, but no special symplectic group. It turns out that there is no need for a special symplectic group, since the determinant of an element in the symplectic group is 1, as we now show.

Theorem 7. The determinant of an element in $Sp(n)$ is 1.

Proof: Let $A \in Sp(n)$. Since the exponential map is surjective from $sp(n)$ onto $Sp(n)$, there exists an element $\alpha \in sp(n)$ such that $e^{\alpha} = A$. But since $\alpha \in sp(n)$, $\alpha + \alpha^* = 0$ from construction of the Lie algebra of $Sp(n)$. Among other things, this implies that there are no real entries along the diagonal, i.e. $\mathrm{Re}(trace(\alpha)) = 0$. Recall that the Lie algebra of $SL_n(\mathbb{H})$ had two equivalent definitions, one of which was $\{M \in M_n(\mathbb{H}) \mid \mathrm{Re}(trace(M)) = 0\}$. Therefore $\alpha \in sl_n(\mathbb{H})$ and so $sp(n) \subset sl_n(\mathbb{H})$. But the exponential map takes elements of $sl_n(\mathbb{H})$ into $SL_n(\mathbb{H})$, as we showed in the proof of Theorem 4. This implies that $A = e^{\alpha} \in SL_n(\mathbb{H})$, so that $\det(A) = 1$. Since A was an arbitrary element of $Sp(n)$, we conclude that $Sp(n) \subset SL_n(\mathbb{H})$. $\square$

11. Weyl Group of $Sp(n)$

Definition 11.1. The *normalizer* of a nonempty subset H of a group G is the subgroup

$$N_G(H) = \{g \in G \mid gHg^{-1} = H\}.$$

In fact, if H is a subgroup of G, then H is normal in $N_G(H)$.

Definition 11.2. Let T be a maximal torus in some compact matrix group G. We define the *Weyl group* of G as $N_G(T)/T$.

It turns out that under this definition, the Weyl group of a matrix group is well defined, for if we used another maximal torus T', then by the property that maximal tori are conjugate by an element in the matrix group, the conjugation map by this element would induce an isomorphism of these maximal tori and their normalizers, and therefore their Weyl groups.

Lemma 1. There are only two automorphisms of the circle group $\mathbb{T}^1 = S^1$.

Proof: Let $\phi : S^1 \to S^1$ be an isomorphism, then $\phi(1) = 1$. Furthermore, since $\phi(-1)^2 = \phi(1) = 1$ implies $\phi(-1) = -1$ or $\phi(-1) = 1$, injectivity of ϕ implies $\phi(-1) = -1$. Next we see that $\phi(\mathbf{i})^2 = \phi(-1) = -1$ and similarly $\phi(-\mathbf{i})^2 = -1$, so $\phi(\mathbf{i})$ can take two possible values, namely $\mathbf{i}$ or $-\mathbf{i}$, which leaves the remaining choice for $\phi(-\mathbf{i})$. We see that this gives rise to the only two automorphisms of S^1, the identity map and the conjugation map: $z \mapsto z$ and $z \mapsto \overline{z}$, respectively. $\square$

Theorem 8. The Weyl group of $Sp(1)$ is $N_{Sp(1)}(\mathbb{T}^1)/\mathbb{T}^1 \cong \mathbb{Z}_2$.

Proof: From Proposition 7, the standard maximal torus in $Sp(1)$ is $\mathbb{T}^1 = (e^{\mathbf{i}\theta}) \cong S^1$. From the previous lemma, we want to find the elements of $Sp(1)$ which induce the two automorphisms of the circle group by conjugation. If conjugation by $q \in Sp(1)$ induces the identity automorphism, then q commutes with every element in $\mathbb{T}^1$, and must be in $\mathbb{T}^1$ (since $\mathbf{j}$ and $\mathbf{k}$ do not commute with $\mathbf{i}$). If $q \in Sp(1)$ induces the other automorphism, then

$$q(x + y\mathbf{i})q^{-1} = x - y\mathbf{i}.$$

So we seek $q \in Sp(1)$ with $q\mathbf{i}q^{-1} = -\mathbf{i}$. But $q = a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$ with $a^2+b^2+c^2+d^2 = 1$. Since $q \in Sp(1)$, $q^{-1} = \overline{q}$. So,

$$\begin{aligned} q\mathbf{i}q^{-1} &= (a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})\mathbf{i}(a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}) \\ &= (a^2 + b^2 - c^2 - d^2)\mathbf{i} + 2(ad + bc)\mathbf{j} - 2(ac - bd)\mathbf{k}. \end{aligned}$$

Setting this equal to $-\mathbf{i}$ implies $a = b = 0$. Thus $q = c\mathbf{j} + d\mathbf{k}$, so that the normalizer of $\mathbb{T}^1$ in $Sp(1)$ is

$$N_{Sp(1)}(\mathbb{T}^1) = \mathbb{T}^1 \sqcup \{c\mathbf{j} + d\mathbf{k} \mid c^2 + d^2 = 1\} \quad \text{and} \quad N_{Sp(1)}(\mathbb{T}^1)/\mathbb{T}^1 \cong \mathbb{Z}_2. \qquad \square$$

Lemma 2. There are 8 automorphisms of the torus $\mathbb{T}^2 = S^1 \times S^1$.

Proof: From Lemma 1, we know there are two automorphisms of the circle group. Similarly, the 8 automorphisms of the 2-torus are those that permute and/or conjugate the two factors. In particular,

$$(*) \quad \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \mapsto \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix}, \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \mapsto \begin{pmatrix} \overline{z}_1 & 0 \\ 0 & z_2 \end{pmatrix}, \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \mapsto \begin{pmatrix} z_1 & 0 \\ 0 & \overline{z}_2 \end{pmatrix}, \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \mapsto \begin{pmatrix} \overline{z}_1 & 0 \\ 0 & \overline{z}_2 \end{pmatrix},$$

$$\begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \mapsto \begin{pmatrix} z_2 & 0 \\ 0 & z_1 \end{pmatrix}, \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \mapsto \begin{pmatrix} \overline{z}_2 & 0 \\ 0 & z_1 \end{pmatrix}, \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \mapsto \begin{pmatrix} z_2 & 0 \\ 0 & \overline{z}_1 \end{pmatrix}, \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \mapsto \begin{pmatrix} \overline{z}_2 & 0 \\ 0 & \overline{z}_1 \end{pmatrix}.$$

$\square$

Theorem 9. The Weyl group of $Sp(2)$ is $N_{Sp(2)}(\mathbb{T}^2)/\mathbb{T}^2 \cong D_4$.

Proof: We will first show how to determine the normalizer of $\mathbb{T}^2 \subset Sp(2)$. Since there are only 8 automorphisms of the 2-torus, we will show that there are elements of $Sp(2)$ that induce these automorphisms by conjugation. In particular, we need to find $q \in Sp(2)$ such that $q\mathbb{T}^2q^{-1} = q\mathbb{T}^2q^* = \mathbb{T}^2$. Let us first consider whether or not the automorphisms exchange the two circle groups that make up $\mathbb{T}^2$, i.e. if we let Ad_q be the conjugation map by $q \in Sp(2)$, we are looking to see if

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \xrightarrow{Ad_q} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

or if

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \xrightarrow{Ad_q} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Relating this to what we found in (*) of Lemma 2, the former represents the automorphisms that do not exchange the circle groups (shown in the first line of (*)), while the latter represents those automorphisms that do (shown on the second line of (*)). Thus,

$$\begin{pmatrix} q_1 & q_2 \\ q_3 & q_4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \overline{q}_1 & \overline{q}_3 \\ \overline{q}_2 & \overline{q}_4 \end{pmatrix} = \begin{pmatrix} |q_1|^2 - |q_2|^2 & q_1\overline{q}_3 - q_2\overline{q}_4 \\ q_3\overline{q}_1 - q_4\overline{q}_2 & |q_3|^2 - |q_4|^2 \end{pmatrix}.$$

Setting this equal to $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ yields $|q_1| = |q_4| = 1$ and $|q_2| = |q_3| = 0$. On the other hand, if we set this equal to $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, it follows that $|q_1| = |q_4| = 0$ and $|q_2| = |q_3| = 1$. This follows from the fact that $|q_1|^2 + |q_2|^2 = 1$ and $|q_3|^2 + |q_4|^2 = 1$ since the rows of the matrix $\left(\begin{smallmatrix} q_1 & q_2 \\ q_3 & q_4 \end{smallmatrix}\right)$ form an orthonormal basis of $\mathbb{H}^2$. This tells us that an element in the normalizer of the standard maximal torus in $Sp(2)$ is either diagonal or has entries only in the off-diagonal positions. To determine more information about each q_n, we must consider how the above matrices act on, say, $\left(\begin{smallmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{i} \end{smallmatrix}\right)$. We have:

$$\begin{pmatrix} q_1 & 0 \\ 0 & q_4 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{i} \end{pmatrix} \cdot \begin{pmatrix} \overline{q}_1 & 0 \\ 0 & \overline{q}_4 \end{pmatrix} = \begin{pmatrix} q_1\mathbf{i}\overline{q}_1 & 0 \\ 0 & q_4\mathbf{i}\overline{q}_4 \end{pmatrix}$$

or

$$\begin{pmatrix} 0 & q_2 \\ q_3 & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{i} \end{pmatrix} \cdot \begin{pmatrix} 0 & \overline{q}_3 \\ \overline{q}_2 & 0 \end{pmatrix} = \begin{pmatrix} q_2\mathbf{i}\overline{q}_2 & 0 \\ 0 & q_3\mathbf{i}\overline{q}_3 \end{pmatrix}.$$

From this, if we write $q_n = a_n + b_n\mathbf{i} + c_n\mathbf{j} + d_n\mathbf{k}$, we get that

$$(\dagger) \qquad q_n\mathbf{i}\overline{q}_n = (a_n^2 + b_n^2 - c_n^2 - d_n^2)\mathbf{i} + 2(a_nd_n + b_nc_n)\mathbf{j} + 2(b_nd_n - a_nc_n)\mathbf{k}.$$

Therefore, if $q_n\mathbf{i}\overline{q}_n = \mathbf{i}$, we see that $c_n = d_n = 0$ in ($\dagger$) since $|q_n|^2 = 1$. So, $q_n = e^{\mathbf{i}\theta_n}$, for some $\theta_n \in [0, 2\pi)$. This is the automorphism that is the identity map on one of the S^1 factors of $\mathbb{T}^2 = S^1 \times S^1$.

However, if $q_n\mathbf{i}\overline{q}_n = -\mathbf{i}$, then in ($\dagger$), it must be that $a_n = b_n = 0$. Thus, $q_n = c_n\mathbf{j} + d_n\mathbf{k} = e^{\mathbf{i}\theta_n}\mathbf{j}$ for some θ_n. Of course, this is the automorphism that takes the complex conjugate of one of the S^1 factors in $\mathbb{T}^2$. We conclude that the normalizer of $\mathbb{T}^2$ in $Sp(2)$ is:

$$N_{Sp(2)}(\mathbb{T}^2) = \Big\{ \begin{pmatrix} e^{\mathbf{i}\theta_1} & 0 \\ 0 & e^{\mathbf{i}\theta_2} \end{pmatrix}, \begin{pmatrix} e^{\mathbf{i}\theta_1}\mathbf{j} & 0 \\ 0 & e^{\mathbf{i}\theta_2} \end{pmatrix}, \begin{pmatrix} e^{\mathbf{i}\theta_1} & 0 \\ 0 & e^{\mathbf{i}\theta_2}\mathbf{j} \end{pmatrix}, \begin{pmatrix} e^{\mathbf{i}\theta_1}\mathbf{j} & 0 \\ 0 & e^{\mathbf{i}\theta_2}\mathbf{j} \end{pmatrix},$$
$$\begin{pmatrix} 0 & e^{\mathbf{i}\theta_1} \\ e^{\mathbf{i}\theta_2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & e^{\mathbf{i}\theta_1}\mathbf{j} \\ e^{\mathbf{i}\theta_2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & e^{\mathbf{i}\theta_1} \\ e^{\mathbf{i}\theta_2}\mathbf{j} & 0 \end{pmatrix}, \begin{pmatrix} 0 & e^{\mathbf{i}\theta_1}\mathbf{j} \\ e^{\mathbf{i}\theta_2}\mathbf{j} & 0 \end{pmatrix} \;\Big|\; \theta_1, \theta_2 \in [0, 2\pi) \Big\}.$$

These elements correspond with those automorphisms of the 2-torus we found above in (*).

Therefore, the Weyl group, $N_{Sp(2)}(\mathbb{T}^2)/\mathbb{T}^2$, contains 8 elements. To further characterize the Weyl group, notice that $N_{Sp(2)}(\mathbb{T}^2)$ is not commutative:

$$\begin{pmatrix} e^{\mathbf{i}\theta_1} & 0 \\ 0 & e^{\mathbf{i}\theta_2}\mathbf{j} \end{pmatrix} \cdot \begin{pmatrix} 0 & e^{\mathbf{i}\theta_3} \\ e^{\mathbf{i}\theta_4} & 0 \end{pmatrix} = \begin{pmatrix} 0 & e^{\mathbf{i}(\theta_1+\theta_3)} \\ e^{\mathbf{i}(\theta_2-\theta_4)}\mathbf{j} & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & e^{\mathbf{i}\theta_3} \\ e^{\mathbf{i}\theta_4} & 0 \end{pmatrix} \cdot \begin{pmatrix} e^{\mathbf{i}\theta_1} & 0 \\ 0 & e^{\mathbf{i}\theta_2}\mathbf{j} \end{pmatrix} = \begin{pmatrix} 0 & e^{\mathbf{i}(\theta_2+\theta_3)}\mathbf{j} \\ e^{\mathbf{i}(\theta_1+\theta_4)} & 0 \end{pmatrix}$$

so the Weyl group cannot be commutative. Furthermore, it is not difficult to see that the only two elements of $N_{Sp(2)}(\mathbb{T}^2)$ that have order four are:

$$\begin{pmatrix} 0 & e^{\mathbf{i}\theta_1}\mathbf{j} \\ e^{\mathbf{i}\theta_2} & 0 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 & e^{\mathbf{i}\theta_1} \\ e^{\mathbf{i}\theta_2}\mathbf{j} & 0 \end{pmatrix}$$

Therefore, the Weyl group has only two elements of order 4. The remaining elements (besides the identity) have order 2. This is enough information to completely classify the Weyl group of $Sp(2)$: $N_{Sp(2)}(\mathbb{T}^2)/\mathbb{T}^2 \cong D_4$. $\square$

Lemma 3. There are $2^n n!$ automorphisms of the n-torus, $\mathbb{T}^n = S^1 \times \cdots \times S^1$.

The same reasoning applies here as it did before. Write $(z_1, z_2, \ldots, z_n)$ to represent each factor of $S^1 \times \cdots \times S^1$. Then we have the choice to send each z_i to either z_i or $\overline{z_i}$. This gives us our 2^n choices. Now, we may permute these objects (i.e. send some z_i to z_j). Of course, the number of ways to permute n objects is $n!$. Since these choices are independent, we get $2^n n!$ automorphisms.

Corollary 2. The normalizer of the maximal torus in $Sp(n)$ contains $2^n n!$ components.

Proof: This is just a generalization of the proof of the Weyl group of $Sp(2)$, we just find the elements of $Sp(n)$ that induce, by conjugation, these $2^n n!$ automorphisms of the maximal torus. In the proof of the Weyl group of $Sp(2)$, that in order to induce complex conjugation, we need to conjugate by $e^{\mathbf{i}\theta}\mathbf{j}$, see (†) in the proof of Theorem 9. If we are given an automorphism of $\mathbb{T}^n$ that takes the complex conjugate of entries $a_1, a_2, \ldots, a_k$ and permutes the entries $b_1, b_2, \ldots, b_l$, then we can construct an element of $Sp(n)$ that induces the given automorphism by conjugation in the following manner. Begin with the maximal torus $\mathbb{T}^n$ as a diagonal matrix as in Definition 8.1. For each a_i, multiply the (a_i, a_i)-entry of $\mathbb{T}^n$ on the right with $\mathbf{j}$. This will give an element of $Sp(n)$ which induces the automorphism that conjugates the entries $a_1, a_2, \ldots, a_k$. To have this matrix in $Sp(n)$ also permute the entries $b_1, b_2, \ldots, b_l$, take the resulting matrix from the first step and permute the columns $b_1, b_2, \ldots, b_l$ as they are permuted in the given automorphism of $\mathbb{T}^n$. This process will give an element of $Sp(n)$ that induces the given automorphism by conjugation. The computations of the Weyl groups of $Sp(1)$ and $Sp(2)$ illustrate this idea.

It follows that the Weyl group of $Sp(n)$ is isomorphic to the semidirect product $(\mathbb{Z}_2)^n \rtimes S_n$. $\square$

In computing the normalizers of the standard maximal torus in $Sp(n)$, we observed the following:

Observation 4. Let $\mathbb{T}^n$ be the standard maximal torus in $Sp(n)$. Then the $2^n n!$ components of $N_{Sp(n)}(\mathbb{T}^n)$ are all disjoint and isomorphic to $\mathbb{T}^n$. In other words, inside $Sp(n)$, there are $2^n n!$ disjoint $\mathbb{T}^n$.

One may wonder, since $U(n)$ and $Sp(n)$ have the same maximal torus, if the Weyl groups are also the same. However, the answer is no. This is because the normalizers of the maximal tori in their respective groups are different, although the maximal tori are the same. In our constructions of the normalizers of the maximal tori in $Sp(n)$ (Theorems 8 and 9), we were able to show that every automorphism of the n-torus was induced by conjugation with certain elements of $Sp(n)$. This is not the case with $U(n)$, particularly because individual elements commute. For example, there are no elements of $U(1)$ that induce the complex conjugation automorphism of its maximal torus, S^1. To see this, notice that for $z \in U(1)$ and $e^{\mathbf{i}\theta} \in S^1$, $ze^{\mathbf{i}\theta}z^{-1} = zz^{-1}e^{\mathbf{i}\theta} = e^{\mathbf{i}\theta}$.

Perhaps this was not very insightful since $S^1 \cong U(1)$. We will consider $U(2)$ and its maximal torus $\mathbb{T}^2 \cong S^1 \times S^1$. Recall that there are 8 automorphisms of $\mathbb{T}^2$ shown in (*) of Lemma 2. Since we were not able to induce the complex conjugation automorphism of S^1 above, we do not expect to be able to do so in this case either. In the computation of Theorem 9, we showed that an element of $Sp(2)$ which induces an automorphism of $\mathbb{T}^2$ by conjugation must be either diagonal or have only elements in the off-diagonal positions. The same argument applies for the $U(2)$ case. But

$$\begin{pmatrix} z_1 & 0 \\ 0 & z_4 \end{pmatrix} \cdot \begin{pmatrix} e^{\mathbf{i}\theta_1} & 0 \\ 0 & e^{\mathbf{i}\theta_2} \end{pmatrix} \cdot \begin{pmatrix} \overline{z}_1 & 0 \\ 0 & \overline{z}_4 \end{pmatrix} = \begin{pmatrix} e^{\mathbf{i}\theta_1} & 0 \\ 0 & e^{\mathbf{i}\theta_2} \end{pmatrix}$$

$$\begin{pmatrix} 0 & z_2 \\ z_3 & 0 \end{pmatrix} \cdot \begin{pmatrix} e^{\mathbf{i}\theta_1} & 0 \\ 0 & e^{\mathbf{i}\theta_2} \end{pmatrix} \cdot \begin{pmatrix} 0 & \overline{z}_3 \\ \overline{z}_2 & 0 \end{pmatrix} = \begin{pmatrix} e^{\mathbf{i}\theta_1} & 0 \\ 0 & e^{\mathbf{i}\theta_2} \end{pmatrix}.$$

So we cannot induce the complex conjugation automorphisms. This argument generalizes for $U(n)$ as it did in Corollary 2 for $Sp(n)$, and we see that the number of elements in the normalizer of $U(n)$ come from only permuting the individual factors of the n-torus. Therefore, the normalizer of the maximal torus in $U(n)$ has $n!$ components and the Weyl group is isomorphic to S_n.

12. Clifford Algebras

Definition 12.1. A *Clifford Algebra*, denoted C_k, is a finite-dimensional real algebra generated by the elements $e_1, e_2, \ldots, e_k$ which satisfy the following relations:

$$(\diamond) \qquad e_m e_n = \begin{cases} -e_n e_m, & \text{if } n \neq m \\ -1, & \text{if } n = m. \end{cases}$$

We also include the element $e_0 = 1$, meaning we set $C_0 = \mathbb{R}$.

By *real algebra*, we mean a real vector space which is also an associative ring with a multiplicative identity such that $(ax)(by) = (ab)(xy)$ for all $a, b \in \mathbb{R}$ and x, y in the algebra. C_k is a real vector space of dimension 2^k, whose basis is obtained by taking 1, each e_r, and products of them.

Proposition 8. $C_1 \cong \mathbb{C},\ C_2 \cong \mathbb{H},\ C_3 \cong \mathbb{H} \oplus \mathbb{H},\ C_4 \cong M_2(\mathbb{H})$, and $C_5 \cong M_4(\mathbb{C})$.

Proof: Let $\left(\begin{smallmatrix} q_1 & 0 \\ 0 & q_2 \end{smallmatrix}\right)$ represent (q_1, q_2). To determine C_1 through C_4, we use the following identifications:

$$(\dagger) \qquad e_1 \xrightarrow{\chi} \left(\begin{smallmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{i} \end{smallmatrix}\right) \qquad e_2 \xrightarrow{\chi} \left(\begin{smallmatrix} \mathbf{j} & 0 \\ 0 & \mathbf{j} \end{smallmatrix}\right) \qquad e_3 \xrightarrow{\chi} \left(\begin{smallmatrix} \mathbf{k} & 0 \\ 0 & -\mathbf{k} \end{smallmatrix}\right) \qquad e_4 \xrightarrow{\chi} \left(\begin{smallmatrix} 0 & \mathbf{k} \\ \mathbf{k} & 0 \end{smallmatrix}\right)$$

Of course e_0 is identified with the identity matrix. It is not hard to see that the matrices satisfy the relations of the e_i's given above in $(\diamond)$, so the map is injective. In order to prove the isomorphisms we claimed above, it remains to show that the map is surjective, which is clear for C_1 and C_2. For C_3, it suffices to show that we can obtain the standard basis of $\mathbb{H} \oplus \mathbb{H}$ through linear combinations of products of the e_i's. The standard basis of $\mathbb{H} \oplus \mathbb{H}$ is $\{(1,0), (\mathbf{i},0), (\mathbf{j},0), (\mathbf{k},0), (0,1), (0,\mathbf{i}), (0,\mathbf{j}), (0,\mathbf{k})\}$. Notice that

$$-\frac{1}{2}\begin{pmatrix} \mathbf{k} & 0 \\ 0 & -\mathbf{k} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{j} & 0 \\ 0 & \mathbf{j} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{i} \end{pmatrix} = \begin{pmatrix} \mathbf{i} & 0 \\ 0 & 0 \end{pmatrix}.$$

The remaining basis elements are obtained in a similar manner.

To show $C_4 = M_2(\mathbb{H})$, we can use the result for C_3 and the observation that multiplication of the basis elements of C_3 above by $\left(\begin{smallmatrix} \mathbf{k} & 0 \\ 0 & -\mathbf{k} \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} 0 & \mathbf{k} \\ \mathbf{k} & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ will achieve the off-diagonal basis elements.

For C_5, we must use the identifications:

$$(\ddagger) \qquad e_1 \xrightarrow{\chi} \begin{pmatrix} \mathbf{i} & 0 & 0 & 0 \\ 0 & -\mathbf{i} & 0 & 0 \\ 0 & 0 & \mathbf{i} & 0 \\ 0 & 0 & 0 & -\mathbf{i} \end{pmatrix} \qquad e_2 \xrightarrow{\chi} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad e_3 \xrightarrow{\chi} \begin{pmatrix} 0 & \mathbf{i} & 0 & 0 \\ \mathbf{i} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mathbf{i} \\ 0 & 0 & -\mathbf{i} & 0 \end{pmatrix}$$
$$e_4 \xrightarrow{\chi} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad e_5 \xrightarrow{\chi} \begin{pmatrix} 0 & 0 & 0 & -\mathbf{i} \\ 0 & 0 & -\mathbf{i} & 0 \\ 0 & -\mathbf{i} & 0 & 0 \\ -\mathbf{i} & 0 & 0 & 0 \end{pmatrix}$$

It is routine to verify that the matrices satisfy the relations of the e_i's given in ($\diamond$). Since the only element of C_5 that is mapped to I under χ is 1, we have that χ is an injective homomorphism. We will now show surjectivity. As before, we only need to show that we can achieve every basis element of $M_4(\mathbb{C})$.

First, $\chi(\frac{1}{2}(1+e_1e_4e_5)) = diag(0,1,1,0)$ and $\chi(\frac{1}{2}(e_2e_3e_4e_5 - e_1e_2e_3)) = diag(0,-1,1,0)$, so by adding these and dividing by 2 we get $diag(0,0,1,0)$. Multiplication on the right and/or left by $\chi(e_4)$ will permute the 1 throughout the center positions, i.e. we can get an individual 1 into any of the positions marked with a cross:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \times & \times & 0 \\ 0 & \times & \times & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Multiplying one of these basis elements by $\chi(e_2)$ on the right or left allows us to get a 1 into each position marked with a cross:

$$\begin{pmatrix} 0 & \times & \times & 0 \\ \times & 0 & 0 & \times \\ \times & 0 & 0 & \times \\ 0 & \times & \times & 0 \end{pmatrix}$$

Multiplying one of these matrices by $\chi(e_2)$ on the right or left finishes the process of getting an individual 1 into any position with zeros elsewhere. We can get an $\mathbf{i}$ into any position as well by multiplying $diag(0,0,1,0)$ found above by $\chi(e_1)$ and repeating the process. $\square$

We will define an automorphism, α, of C_k by:

$$\alpha(e_{i_1}e_{i_2}\cdots e_{i_t}) = (-1)^t e_{i_1}e_{i_2}\cdots e_{i_t} = \begin{cases} e_{i_1}e_{i_2}\cdots e_{i_t}, & \text{if } t \text{ is even} \\ -e_{i_1}e_{i_2}\cdots e_{i_t}, & \text{if } t \text{ is odd.} \end{cases}$$

We also define a conjugation on C_k:

$$\overline{(e_{i_1}e_{i_2}\cdots e_{i_t})} = (-1)^t e_{i_t}e_{i_{t-1}}\cdots e_{i_1}.$$

Notice that this agrees with the complex/quaternionic conjugation we have already described, since

$$\overline{1} = 1, \quad \overline{e_r} = -e_r, \quad \overline{x+y} = \overline{x} + \overline{y} \text{ and } \overline{x \cdot y} = \overline{y} \cdot \overline{x} \quad \text{for } x, y \in C_k$$

By identifying $\mathbb{R}^k$ with $span\{e_1, \ldots, e_k\}$, we see that

$$S^{k-1} = \{a_1e_1 + \cdots + a_ke_k \mid a_1^2 + \cdots + a_k^2 = 1\} \subset \mathbb{R}^k \subset C_k.$$

For any $u \in S^{k-1}$, we will show $u \cdot \overline{u} = 1$. Let $u = a_1e_1 + \cdots + a_ke_k$, then

$$\begin{aligned} u \cdot \overline{u} &= (a_1e_1 + \cdots + a_ke_k) \cdot \overline{(a_1e_1 + \cdots + a_ke_k)} \\ &= (a_1e_1 + \cdots + a_ke_k) \cdot (-a_1e_1 - \cdots - a_ke_k) \\ &= a_1^2 + \cdots + a_k^2 = 1, \quad \text{since } u \in S^{k-1}. \end{aligned}$$

Thus, u is a unit in C_k. If we let C_k^* denote the group of units in C_k, it follows that $S^{k-1} \subset C_k^*$. In particular, the subgroup generated by S^{k-1}, denoted $\langle S^{k-1} \rangle$, is also contained in C_k^*.

Definition 12.2. $\mathrm{Pin}(k) = \langle S^{k-1} \rangle \subset C_k^*$.

Let $u \in \mathrm{Pin}(k)$. Then $u = u_1u_2\cdots u_m$, where $u_i \in S^{k-1}$. Thus,

$$\alpha(u) = \alpha(u_1u_2\cdots u_m)$$

$$= \alpha(u_1)\cdots\alpha(u_m) \quad \text{since } \alpha \text{ is a homomorphism.}$$

But since $u_i \in S^{k-1}$, $u_i = a_{1_i}e_1 + \cdots + a_{k_i}e_k$ where $a_{1_i}^2 + \cdots + a_{k_i}^2 = 1$. From this, it follows that $\alpha(u_i) = \alpha(a_{1_i}e_1 + \cdots + a_{k_i}e_k) = -u_i$, which implies:

$$\alpha(u) = \begin{cases} u, & \text{if } m \text{ is even} \\ -u, & \text{if } m \text{ is odd.} \end{cases}$$

Furthermore, $u^{-1} = (u_1u_2\cdots u_m)^{-1} = (-1)^m u_m\cdots u_2u_1 = \overline{u_1u_2\cdots u_m}$ which shows, as above:

$$\alpha(u^{-1}) = \begin{cases} u^{-1}, & \text{if } m \text{ is even} \\ -u^{-1}, & \text{if } m \text{ is odd.} \end{cases}$$

These observations lead us to define an even/odd grading on $\mathrm{Pin}(k)$ depending on whether or not an element $u \in \mathrm{Pin}(k)$ is a product of an even or odd number of elements of S^{k-1}:

$$\mathrm{Pin}(k)^+ = \{u \in \mathrm{Pin}(k) \mid \alpha(u) = u\}$$

$$\mathrm{Pin}(k)^- = \{u \in \mathrm{Pin}(k) \mid \alpha(u) = -u\}.$$

It is clear that $\mathrm{Pin}(k) = \mathrm{Pin}(k)^+ \cup\ \mathrm{Pin}(k)^-$ with $\mathrm{Pin}(k)^+ \cap\ \mathrm{Pin}(k)^- = \emptyset$.
For $u, v \in \mathrm{Pin}(k)$, notice that

$$\alpha(uv) = \alpha(u)\alpha(v) = \begin{cases} uv, & \text{if } u, v \in \mathrm{Pin}(k)^+ \text{ or } u, v \in \mathrm{Pin}(k)^- \\ -uv, & \text{if } u \in \mathrm{Pin}(k)^+,\ v \in \mathrm{Pin}(k)^- \text{ or vice versa.} \end{cases}$$

Definition 12.3. $\mathrm{Spin}(k) = \mathrm{Pin}(k)^+$.

The previous discussion shows that $\mathrm{Spin}(k)$ is a subgroup of $\mathrm{Pin}(k)$.

Definition 12.4. For $u \in \mathrm{Pin}(k)$ and $x \in \mathbb{R}^k$, we define the function $\rho : \mathrm{Pin}(k) \to O(k)$ by

$$\rho_u(x) = \alpha(u)x\overline{u}.$$

Theorem 10. $\rho : \mathrm{Pin}(k) \to O(k)$ is a 2-to-1 surjective homomorphism [1, 2].

The outline of the proof is as follows. Since $O(n)$ is generated by reflections and ρ_u is a reflection in $\mathbb{R}^n$ through the hyperplane perpendicular to u, ρ is a homomorphism of $\mathrm{Pin}(n)$ onto $O(n)$ with $\ker \rho = \{-1, 1\}$.

Theorem 11. $\mathrm{Spin}(k) = \rho^{-1}(SO(k))$. Furthermore, $\mathrm{Spin}(k)$ is a normal subgroup of $\mathrm{Pin}(k)$ which is closed, compact, and path connected [1].

Define the map $\delta : C_{k-1} \to C_k$ by $\delta(e_r) = e_r e_k$, for $r \in \{1, 2, \ldots, k-1\}$. This definition is sufficient because we specify where the basis elements of C_{k-1} go under the map δ. Then

$$\delta(e_r)^2 = (e_r e_k)^2 = e_r e_k e_r e_k = -e_r e_r e_k e_k = -1$$

$$\delta(e_r)\delta(e_s) = (e_r e_k)(e_s e_k) = -e_k e_r e_s e_k = e_k e_s e_r e_k = -(e_s e_k)(e_r e_k) = -\delta(e_s)\delta(e_r).$$

So δ extends to an algebra homomorphism since it maintains the relations in ($\diamond$).

Recall our even/odd grading of $\mathrm{Pin}(k)$ preceding Definition 12.3, and notice that we may define the same even/odd grading on C_k. Then, $e_i e_k e_j e_k = -e_i e_j e_k e_k = e_i e_j$, so, in fact, δ gives an isomorphism of C_{k-1} onto the even graded elements of C_k.

Proposition 9. For $x \in C_k$ and $A \in M_n(\mathbb{K})$, the isomorphism χ from (†) and (‡) in Proposition 8 and the isomorphism δ satisfy the following properties:

(1) $\overline{\delta(x)} = \delta(\overline{x})$

(2) $\chi(A^*) = \overline{\chi(A)}$

Proof: To prove both of these, it is sufficient to show they hold on the basis elements.

For (1), $\overline{\delta(e_i)} = \overline{e_i e_k} = -e_i e_k = \delta(-e_i) = \delta(\overline{e_i})$.

For (2), we must check the property holds on the basis matrices, A, above in (†) and (‡). Let $\chi(A) = e_i$. It is easy to see that for each A that $A^* = -A$, in which case $\chi(A^*) = \chi(-A) = -\chi(A) = -e_i = \overline{\chi(A)}$. □

Theorem 12. We have the following isomorphisms:

$$\begin{aligned} Sp(1) &\cong \text{Spin}(3) \\ Sp(1) \times Sp(1) &\cong \text{Spin}(4) \\ Sp(2) &\cong \text{Spin}(5). \end{aligned}$$

Proof: Let ι represent the inclusion map, and consider the compositions

$$\begin{aligned} &Sp(1) \xrightarrow{\iota} \mathbb{H} \xrightarrow{\chi} C_2 \xrightarrow{\delta} C_3 \\ &Sp(1) \times Sp(1) \xrightarrow{\iota} \mathbb{H} \oplus \mathbb{H} \xrightarrow{\chi} C_3 \xrightarrow{\delta} C_4 \\ &Sp(2) \xrightarrow{\iota} M_2(\mathbb{H}) \xrightarrow{\chi} C_4 \xrightarrow{\delta} C_5. \end{aligned}$$

Every matrix $A \in Sp(1),\ Sp(1) \times Sp(1)$, or $Sp(2)$, has the property that $A \cdot A^* = I$, which gives

$$\begin{aligned} 1 = \delta(1) &= \delta(\chi(I)) \\ &= \delta(\chi(A \cdot A^*)) \\ &= \delta(\chi(A) \cdot \overline{\chi(A)}) && \text{from part (2) of Proposition 9} \\ &= \delta(\chi(A)) \cdot \delta(\overline{\chi(A)}) \\ &= \delta(\chi(A)) \cdot \overline{\delta(\chi(A))} && \text{from part (1) of Proposition 9} \end{aligned}$$

The discussion preceding Definition 12.3 showed that if $x \in \text{Spin}(k)$, then x was a linear combination of elements of $\text{Pin}(k)^+$ (the even graded elements of $\text{Pin}(k)$) and that $x \cdot \overline{x} = 1$. But since $\delta(\chi(A)) \cdot \overline{\delta(\chi(A))} = 1$, we obtain all even graded units whose inverse is their conjugate. Therefore,

$$\mathrm{Spin}(3) \subseteq (\delta \circ \chi)(Sp(1))$$

$$\mathrm{Spin}(4) \subseteq (\delta \circ \chi)(Sp(1) \times Sp(1))$$

$$\mathrm{Spin}(5) \subseteq (\delta \circ \chi)(Sp(2))$$

Since both $(\delta \circ \chi)(Sp(1))$ and Spin(3) are closed, path-connected manifolds of dimension 3, with one contained in the other, they must be equal [2]. The same result follows because both $(\delta \circ \chi)(Sp(1) \times Sp(1))$ and Spin(4) are closed, path-connected manifolds of dimension 6 and because both $(\delta \circ \chi)(Sp(2))$ and Spin(5) are closed, path-connected manifolds of dimension 10. $\square$

From Theorems 10 and 11, we know that Spin(k) is a double cover of $SO(k)$. Then the isomorphisms we just proved show that $Sp(1)$ (respectively $Sp(1) \times Sp(1)$ and $Sp(2)$) is a double cover of $SO(3)$ (respectively $SO(4)$ and $SO(5)$). Baker [1] shows that the Lie algebra of Spin(k) is isomorphic to the Lie algebra of $SO(k)$. This allows us to use the fact that smoothly isomorphic Lie groups have isomorphic Lie algebras [5], whereby we conclude that $sp(1) \cong so(3)$, $sp(1) \times sp(1) \cong so(4)$, and that $sp(2) \cong so(5)$.

Furthermore, since the fundamental group, $\pi_1(Sp(n))$, is trivial, it follows that $Sp(1)$ (respectively $Sp(1) \times Sp(1)$ and $Sp(2)$) is a universal cover of $SO(3)$ (respectively $SO(4)$ and $SO(5)$). We may say that $\pi_1(SO(3)) \cong \pi_1(SO(4)) \cong \pi_1(SO(5)) \cong \mathbb{Z}_2$ and that $SO(3)$ looks like the real-projective space $\mathbb{RP}^3$.

13. An Example

We will calculate the exponential of the following matrix:

$$A = \begin{pmatrix} \mathbf{i} & \mathbf{j} \\ \mathbf{j} & -\mathbf{i} \end{pmatrix} \in sp(2).$$

It is straightforward to find

$$\Psi_2(A) = \begin{pmatrix} \mathbf{i} & 0 & 0 & 1 \\ 0 & -\mathbf{i} & -1 & 0 \\ 0 & 1 & -\mathbf{i} & 0 \\ -1 & 0 & 0 & \mathbf{i} \end{pmatrix}.$$

Solving for the eigenvalues of the resulting complex matrix yields:

$$\det(\Psi_2(A) - \lambda \cdot I) = \begin{vmatrix} \mathbf{i}-\lambda & 0 & 0 & 1 \\ 0 & -\mathbf{i}-\lambda & -1 & 0 \\ 0 & 1 & -\mathbf{i}-\lambda & 0 \\ -1 & 0 & 0 & \mathbf{i}-\lambda \end{vmatrix}$$

$$= \lambda^2(\lambda^2 + 4) = 0.$$

$$\Longrightarrow \lambda = 0, \pm 2\mathbf{i}.$$

We may solve for the eigenvectors:

For $\lambda = 0$:

$$\Psi_2(A) \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Longrightarrow \begin{cases} x_1 = x_4 \mathbf{i} \\ x_2 = x_3 \mathbf{i} \end{cases}$$

$$\Longrightarrow \text{ the corresponding eigenvectors are } \begin{pmatrix} \mathbf{i} \\ 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ -\mathbf{i} \\ -1 \\ 0 \end{pmatrix}.$$

Similarly for the remaining eigenvalues:

For $\lambda = 2\mathbf{i}$ we get the eigenvector $\begin{pmatrix} \mathbf{i} \\ 0 \\ 0 \\ -1 \end{pmatrix}$

and for $\lambda = -2\mathbf{i}$ we get the eigenvector $\begin{pmatrix} 0 \\ -\mathbf{i} \\ 1 \\ 0 \end{pmatrix}$.

Let B be the matrix whose columns are formed from the eigenvectors of $\Psi_2(A)$. Thus,

$$B = \begin{pmatrix} \mathbf{i} & 0 & \mathbf{i} & 0 \\ 0 & -\mathbf{i} & 0 & -\mathbf{i} \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix} \quad \text{and} \quad B^{-1} = \frac{1}{2}\begin{pmatrix} -\mathbf{i} & 0 & 0 & 1 \\ 0 & \mathbf{i} & -1 & 0 \\ -\mathbf{i} & 0 & 0 & -1 \\ 0 & \mathbf{i} & 1 & 0 \end{pmatrix}.$$

Therefore, $\Psi_2(A) = B \cdot diag(0, 0, 2\mathbf{i}, -2\mathbf{i}) \cdot B^{-1}$ by diagonalization. Since Ψ_n is a ring homomorphism and B and B^{-1} are images of some quaternionic matrices under Ψ_2, we then have

$$\begin{aligned} A &= \Psi_2^{-1}(B \cdot diag(0, 0, 2\mathbf{i}, -2\mathbf{i}) \cdot B^{-1}) \\ &= \Psi_2^{-1}(B) \cdot \Psi_2^{-1}(diag(0, 0, 2\mathbf{i}, -2\mathbf{i})) \cdot \Psi_2^{-1}(B^{-1}) \\ &= \frac{1}{2}\begin{pmatrix} \mathbf{i} & \mathbf{i} \\ -\mathbf{j} & \mathbf{j} \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 2\mathbf{i} \end{pmatrix} \cdot \begin{pmatrix} -\mathbf{i} & \mathbf{j} \\ -\mathbf{i} & -\mathbf{j} \end{pmatrix}. \end{aligned} \tag{13.1}$$

By using Property (ii) of the exponential map, we can compute e^A:

$$\begin{aligned} e^A &= \exp\left(\frac{1}{2}\begin{pmatrix} \mathbf{i} & \mathbf{i} \\ -\mathbf{j} & \mathbf{j} \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 2\mathbf{i} \end{pmatrix} \cdot \begin{pmatrix} -\mathbf{i} & \mathbf{j} \\ -\mathbf{i} & -\mathbf{j} \end{pmatrix} \right) \\ &= \frac{1}{2}\begin{pmatrix} \mathbf{i} & \mathbf{i} \\ -\mathbf{j} & \mathbf{j} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & e^{2\mathbf{i}} \end{pmatrix} \cdot \begin{pmatrix} -\mathbf{i} & \mathbf{j} \\ -\mathbf{i} & -\mathbf{j} \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 + e^{2\mathbf{i}} & \mathbf{k}(1 - e^{-2\mathbf{i}}) \\ -\mathbf{k}(1 - e^{2\mathbf{i}}) & 1 + e^{-2\mathbf{i}} \end{pmatrix}. \end{aligned}$$

To verify that this matrix is in $Sp(2)$, first notice that

$$(e^A)^* = \frac{1}{2}\begin{pmatrix} 1 + e^{-2\mathbf{i}} & \mathbf{k}(1 - e^{2\mathbf{i}}) \\ -\mathbf{k}(1 - e^{-2\mathbf{i}}) & 1 + e^{2\mathbf{i}} \end{pmatrix}.$$

Then it is not difficult to check that $e^A \cdot (e^A)^* = I$.

14. Conclusions

In the previous example, if B and B^{-1} were not images of some quaternionic matrices under Ψ_2, we would have to consider

$$e^{\Psi_2(A)} = B \cdot e^{diag(0,0,2\mathbf{i},-2\mathbf{i})} \cdot B^{-1} = B \cdot diag(1,1,e^{2\mathbf{i}},e^{-2\mathbf{i}}) \cdot B^{-1}.$$

and then use Proposition 2 which says that

$$e^A = \Psi_2^{-1}(e^{\Psi_2(A)}) = \Psi_2^{-1}(B \cdot diag(1,1,e^{2\mathbf{i}},e^{-2\mathbf{i}}) \cdot B^{-1}).$$

However, since B and B^{-1} were images of some quaternionic matrices under Ψ_2, we were able to diagonalize the complex matrix $\Psi_2(A)$, and use that to diagonalize the initial matrix A, as in equation (13.1).

This example provides a possible application of the map Ψ_n to define a linear algebra over the quaternions. By transforming a quaternionic matrix into a complex matrix via Ψ_n, we may diagonalize the complex matrix (assuming it is diagonalizable) and write it in the form PDP^{-1}, where P is the matrix formed by the eigenvectors of the complex matrix. If P and P^{-1} are both images under Ψ_n of some quaternionic matrix, we then apply Ψ_n^{-1} to obtain a diagonalization of the quaternionic matrix we began with.

In terms of calculating the exponential of a matrix in $M_n(\mathbb{H})$, Corollary 1 reduces the computation to finding the exponential of a matrix in the Lie algebra $sl_n(\mathbb{H})$. Although this does reduce the computation slightly, further study could potentially reduce this problem even more.

References

[1] Baker, A. (2002). *Matrix groups: An introduction to Lie group theory.* London: Springer-Verlag.

[2] Curtis, M. (1984). *Matrix groups, Second edition.* New York: Springer-Verlag.

[3] Hall, B. (2003). *Lie groups, Lie algebras, and representations: An elementary introduction.* New York: Springer-Verlag.

[4] Stillwell, J. (2008). *Naive Lie theory.* New York: Springer Science + Business Media.

[5] Tapp, K. (2005). *Matrix groups for undergraduates.* Providence, RI: American Mathematical Society.