Special Orthogonal Groups and Rotations

Christopher Triola

Submitted in partial fulfillment of the requirements for Honors in Mathematics at the University of Mary Washington

Fredericksburg, Virginia

April 2009

This thesis by **Christopher Triola** is accepted in its present form as satisfying the thesis requirement for Honors in Mathematics.

DATE	APPROVED
____________	______________________________
	Randall D. Helmstutler, Ph.D. thesis advisor
____________	______________________________
	Manning G. Collier, Ph.D. committee member
____________	______________________________
	J. Larry Lehman, Ph.D. committee member

Contents

Abstract

The rotational geometry of n-dimensional Euclidean space is characterized by the special orthogonal groups $SO(n)$. Elements of these groups are length-preserving linear transformations whose matrix representations possess determinant $+1$. In this paper we investigate some of the group properties of $SO(n)$. We also use linear algebra to study algebraic and geometric properties of $SO(2)$, $SO(3)$, and $SO(4)$. Through this work we find quantifiable differences between rotations in $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{R}^4$. Our focus then shifts to alternative representations of these groups, with emphasis on the ring $\mathbb{H}$ of quaternions.

1 Introduction to Rotation Groups

When we think of rotations many examples may come to mind, e.g. the wheels on a car or the Earth spinning as it orbits the Sun. These are all examples of rotations in three dimensions. This set contains all two-dimensional rotations, and as we will later show, every rotation in three dimensions can be reduced to something like a two-dimensional rotation in the right basis. What of higher dimensions? It turns out that the set of rotations on $\mathbb{R}^n$ and the operation of composition on these rotations constitute a group. It is from this perspective that we will explore rotations in two, three, and four dimensions.

Notice that one of the most fundamental properties of rotations is that they do not change distances. For instance, rotate a cube about a single axis and the distances between any two points on or within the cube will remain the same as before the rotation. This property makes rotations on $\mathbb{R}^n$ a subset of the isometries on $\mathbb{R}^n$. In fact, the set of rotations on $\mathbb{R}^n$ is a normal subgroup of the group of linear isometries on $\mathbb{R}^n$. Naturally, we will start our journey with a discussion of isometries.

1.1 Linear Isometries

Definition 1.1. An **isometry** of $\mathbb{R}^n$ is a function $f : \mathbb{R}^n \to \mathbb{R}^n$ such that, for any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have $|f(\mathbf{x}) - f(\mathbf{y})| = |\mathbf{x} - \mathbf{y}|$. That is, f preserves distances between points in $\mathbb{R}^n$.

Lemma 1.2. *Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is a function that fixes the origin. If f is an isometry then f preserves lengths of all vectors in $\mathbb{R}^n$. The converse holds if f is a linear transformation.*

Proof. Assume f is an isometry. Then it preserves distances between vectors, that is, $|f(\mathbf{x})-f(\mathbf{y})| = |\mathbf{x} - \mathbf{y}|$. Consider the distance between a vector $f(\mathbf{x})$ and the origin. Since $f(\mathbf{0}) = \mathbf{0}$, we have

$$|f(\mathbf{x})| = |f(\mathbf{x}) - f(\mathbf{0})| = |\mathbf{x} - \mathbf{0}| = |\mathbf{x}|.$$

Hence f preserves lengths.

Suppose that f preserves lengths. Given vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, if f is linear we have

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{y})| &= |f(\mathbf{x} - \mathbf{y})| \\ &= |\mathbf{x} - \mathbf{y}|. \end{aligned}$$

Thus f is an isometry. □

The origin-fixing isometries on $\mathbb{R}^n$ are called the *linear isometries* on $\mathbb{R}^n$. This lemma will be very important when we prove that linear isometries on $\mathbb{R}^n$ are in fact linear transformations on $\mathbb{R}^n$. The next lemma is important for closure purposes.

Lemma 1.3. *The composition of two isometries is an isometry. Furthermore, if they are both linear isometries, then so is the composite.*

Proof. Let f and g be two isometries on $\mathbb{R}^n$. We want to show that $f \circ g$ is another isometry on $\mathbb{R}^n$. For all $\mathbf{x}$,$\mathbf{y}$ in $\mathbb{R}^n$, we have $|f(g(\mathbf{x})) - f(g(\mathbf{y}))| = |g(\mathbf{x}) - g(\mathbf{y})| = |\mathbf{x} - \mathbf{y}|$. Hence $f \circ g$ is an isometry.

Furthermore, suppose f and g are both linear isometries. Then we know that $f(g(\mathbf{0})) = f(\mathbf{0}) = \mathbf{0}$. So if f and g are both linear, then so is $f \circ g$. □

Lemma 1.4. *Suppose the function $f : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry that moves the origin. Then the function $g : \mathbb{R}^n \to \mathbb{R}^n$ given by $g(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{0})$ is a linear isometry.*

Proof. Consider two vectors $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^n$. By definition of g,

$$|g(\mathbf{x}) - g(\mathbf{y})| = |f(\mathbf{x}) - f(\mathbf{0}) - f(\mathbf{y}) + f(\mathbf{0})| = |f(\mathbf{x}) - f(\mathbf{y})| = |\mathbf{x} - \mathbf{y}|.$$

Hence g is an isometry on $\mathbb{R}^n$.

Now, note that $g(\mathbf{0}) = f(\mathbf{0}) - f(\mathbf{0}) = \mathbf{0}$. So g fixes the origin and is hence a linear isometry. □

This allows us to construct a linear isometry $g : \mathbb{R}^n \to \mathbb{R}^n$ given any non-linear isometry $f : \mathbb{R}^n \to \mathbb{R}^n$ by the formula $g(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{0})$. Now that we only need to consider linear isometries, we will show that all of these isometries are linear transformations. Before we proceed with the proof, we must first introduce the orthogonal groups $O(n)$.

1.2 Orthogonal Groups

Consider the following subset of $n \times n$ matrices with real entries:

$$O(n) = \{A \in GL_n \mid A^{-1} = A^T\}.$$

This set is known as the *orthogonal group* of $n \times n$ matrices.

Theorem 1.5. *The set $O(n)$ is a group under matrix multiplication.*

Proof. We know that $O(n)$ possesses an identity element I. It is clear that since $A^T = A^{-1}$ every element of $O(n)$ possesses an inverse. It is also clear that matrix multiplication is by its very nature associative, hence $O(n)$ is associative under matrix multiplication. To show that $O(n)$ is closed, consider two arbitrary elements $A, B \in O(n)$ and note the following:

$$\begin{aligned}(AB)(AB)^T &= ABB^TA^T \\ &= ABB^TA^T \\ &= AA^T \\ &= I.\end{aligned}$$

Hence $(AB)^T = (AB)^{-1}$, which makes AB another element of $O(n)$. So, $O(n)$ is closed under matrix multiplication. Thus we have that it is a group. □

Note. Since $A^TA = I$, if we consider the columns of A to be vectors, then they must be orthonormal vectors.

An alternative definition of $O(n)$ is as the set of $n \times n$ matrices that preserve inner products on $\mathbb{R}^n$. Because we did not use this definition we must prove this using our definition of $O(n)$.

Lemma 1.6. *Let A be an element of $O(n)$. The transformation associated with A preserves dot products.*

Proof. If we consider vectors in $\mathbb{R}^n$ to be column matrices then we can define the dot product of $\mathbf{u}$ with $\mathbf{v}$ in $\mathbb{R}^n$ as

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}.$$

Consider the dot product of $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ after the transformation A:

$$\begin{aligned} A\mathbf{u} \cdot A\mathbf{v} &= (A\mathbf{u})^T (A\mathbf{v}) \\ &= \mathbf{u}^T A^T A \mathbf{v} \\ &= \mathbf{u}^T \mathbf{v} \\ &= \mathbf{u} \cdot \mathbf{v}. \end{aligned}$$

Hence, elements of $O(n)$ preserve dot products. □

Lemma 1.7. *If $A \in O(n)$ then A is a linear isometry.*

Proof. Let A be an element of $O(n)$. Since A preserves dot products, this means it must also preserve lengths in $\mathbb{R}^n$, since the length of a vector $\mathbf{v} \in \mathbb{R}^n$ may be defined as $|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. Furthermore, it is clear that the origin is fixed since $A\mathbf{0} = \mathbf{0}$. Thus, by Lemma 1.2, A is a linear isometry. □

So, we have shown that $O(n)$ is at least a subset of the set of linear isometries. Now, we will show containment in the other direction.

Proposition 1.8. *Every linear isometry is a linear transformation whose matrix is in $O(n)$.*

Proof. If we can show that for every origin-fixing isometry $f : \mathbb{R}^n \to \mathbb{R}^n$ there exists an $n \times n$ matrix A such that $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$, then f must be a linear transformation. We will now construct such a matrix.

We begin by letting the i^{th} column of the matrix be given by the vector $f(e_i)$, where e_i is the i^{th} standard basis vector for $\mathbb{R}^n$. Since f preserves dot products, the columns of A are orthonormal and thus $A \in O(n)$. Now we will show that $f = A$ by showing that $g = A^{-1} \circ f$ is the identity.

First, it is clear that g is an isometry and that $g(\mathbf{0}) = \mathbf{0}$ so that g preserves length and dot products. Also, we can see that $g(e_i) = e_i$ for all $i \in \{1, 2, \ldots, n\}$. Hence for a vector $\mathbf{x} \in \mathbb{R}^n$ we have the following (as from [4]):

$$g(\mathbf{x}) = \sum_{i=1}^{n} [g(\mathbf{x}) \cdot e_i] e_i = \sum_{i=1}^{n} [g(\mathbf{x}) \cdot g(e_i)] e_i = \sum_{i=1}^{n} [\mathbf{x} \cdot e_i] e_i = \mathbf{x}.$$

Thus $f = A$ and f is a linear transformation in $O(n)$. □

Recall that, in general, $\det(A) = \det(A^T)$ and $\det(AB) = \det(A)\det(B)$. So, for $A \in O(n)$ we find that the square of its determinant is

$$\begin{aligned} \det(A)^2 &= \det(A)\det(A^T) \\ &= \det(AA^T) \\ &= \det(I) \\ &= 1. \end{aligned}$$

Hence all orthogonal matrices must have a determinant of ± 1.

Note. The set of elements in $O(n)$ with determinant $+1$ is the set of all proper *rotations* on $\mathbb{R}^n$. As we will now prove, this set is a subgroup of $O(n)$ and it is called the *special orthogonal group*, denoted $SO(n)$.

Theorem 1.9. *The subset $SO(n) = \{A \in O(n) \mid \det(A) = 1\}$ is a subgroup of $O(n)$.*

Proof. It is clear that the identity is in $SO(n)$. Also, since $\det(A) = \det(A^T)$, every element of $SO(n)$ inherits its inverse from $O(n)$. We need only check closure. Consider the two elements $A, B \in SO(n)$. Since $\det A = \det B = 1$ we know that $\det(AB) = 1$. Hence we have shown that $SO(n)$ is closed and thus is a subgroup of $O(n)$. □

Theorem 1.10. *The group $SO(n)$ is a normal subgroup of $O(n)$. Furthermore, $O(n)/SO(n) \cong \mathbb{Z}_2$.*

Proof. Let $\{\pm 1\}$ denote the group of these elements under multiplication. Define the mapping $f : O(n) \to \{\pm 1\}$ by $f(A) = \det(A)$. Clearly f is surjective. Also, it is obvious from the properties of determinants that this is a homomorphism. Furthermore, the kernel of this homomorphism is $SO(n)$ since these are the only elements of $O(n)$ with determinant $+1$. Therefore, by the First Isomorphism Theorem we have that $O(n)/SO(n) \cong \mathbb{Z}_2$. □

1.3 Group Properties of $SO(2)$

In this section we will discuss several properties of the group $SO(2)$. We know that this is the group of rotations in the plane. We want a representation for general elements of this group to gain a better understanding of its properties.

Let us start by choosing an orthonormal basis for $\mathbb{R}^2$, $\beta = \{(1,0),(0,1)\}$. Consider the rotation by an arbitrary angle θ, denoted R_θ. We can see that the vector $(1,0)$ will transform in the following manner:

$$R_\theta(1,0) = (\cos\theta, \sin\theta).$$

This is basic trigonometry. Now we will rotate the other basis vector $(0,1)$ to find that

$$R_\theta(0,1) = (-\sin\theta, \cos\theta).$$

Now we can construct a matrix representation of R_θ with respect to β by using the transformed vectors as columns in the matrix

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

Note that $\det A = \cos^2\theta + \sin^2\theta = +1$ and that

$$\begin{aligned} AA^T &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \cos\theta\sin\theta \\ \sin\theta\cos\theta - \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Hence, ordinary rotations in the plane are indeed in $SO(2)$, as we expected.

Now, for completeness, we will show the converse. The columns of any arbitrary element of $SO(2)$ must constitute an orthonormal basis. Therefore, we can think of the columns of an element of $SO(2)$ as two orthogonal vectors on the unit circle centered at the origin in the xy-plane. It is not hard to see that the vectors $\mathbf{u} = (\cos\theta, \sin\theta)$ parameterize the unit circle centered at the origin. There are only two vectors on the unit circle that are orthogonal to this vector and they are $\mathbf{v}_1 = (-\sin\theta, \cos\theta)$ and $\mathbf{v}_2 = (\sin\theta, -\cos\theta)$. We construct a matrix using the vectors $\mathbf{u}$ and $\mathbf{v}_1$ as the columns:

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

Notice that this is the same matrix that we constructed before so we already know it lies in $SO(2)$. Now, we will examine the matrix constructed using $\mathbf{u}$ and $\mathbf{v}_2$:

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

See that the determinant of this matrix is $-\cos^2\theta - \sin^2\theta = -1$. Hence, this does not belong to $SO(2)$ and we have shown that all elements of $SO(2)$ are of the form

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

Is $SO(2)$ abelian? We will explore this question using the representation we have just derived. Consider two elements of $SO(2)$:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \; B = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

Now we check commutativity of A and B:

$$\begin{aligned} AB &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta\cos\phi \\ \sin\theta\cos\phi + \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\cos\phi \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}. \end{aligned}$$

If we swap these elements, we get

$$\begin{aligned} BA &= \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}. \end{aligned}$$

We have just shown that for any $A, B \in SO(2)$, $AB = BA$, hence $SO(2)$ is abelian. This should not be too surprising to anyone who has experience with merry-go-rounds, wheels, or arc length. Now we will consider the question of commutativity of rotations in higher dimensions.

1.4 Group Properties of $SO(n)$, $n \geq 3$

We were able to construct simple representations for elements of $SO(2)$. We would like simple representations of $SO(3)$ and $SO(4)$, but as we will discover, the determinants get to be rather complicated which makes it harder to check that a matrix is in $SO(n)$ for larger n. However, we may use our intuition about rotations and matrices to make our work easier. Since we know what elements of $SO(2)$ look like, let us take one and try to insert it in a 3×3 matrix. Consider the matrices

$$A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Notice that $A^T A = B^T B = I$ and that $\det A = \det B = 1$, hence these matrices exist in $SO(3)$. Furthermore, notice that they possess block elements of $SO(2)$. This is our trick for inserting matrices from $SO(2)$ into higher dimensions and this will be used frequently when we must construct elements of $SO(n)$ just to give examples of properties.

With these two elements of $SO(3)$ let us consider the question: is $SO(3)$ abelian? Consider the product AB:

$$AB = \begin{pmatrix} \cos\theta & -\sin\theta\cos\theta & \sin^2\theta \\ \sin\theta & \cos^2\theta & -\sin\theta\cos\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Now consider the product BA:

$$BA = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta\cos\theta & \cos^2\theta & -\sin\theta \\ \sin^2\theta & \sin\theta\cos\theta & \cos\theta \end{pmatrix}.$$

Note that in general $AB \neq BA$ and hence $SO(3)$ is not abelian. Furthermore, since elements of $SO(3)$ may be embedded in $SO(n)$ for $n > 3$, we know that $SO(2)$ is the only special orthogonal group that is abelian. However, it is worth noting that there do exist commutative subgroups of $SO(n)$ for all n. An example of a commutative subgroup in $SO(3)$ is the set of matrices of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Matrices of this form are commutative since these are rotations in the xy-plane and we already showed that $SO(2)$ is commutative.

We have established some basic properties of the special orthogonal groups and can use a parameterization of $SO(2)$ to manufacture some elements of $SO(n)$. We will now proceed to delve into the properties of $SO(n)$ as operations on $\mathbb{R}^n$ by investigating their eigenvalue structures.

2 Eigenvalue Structure

2.1 First Results

The objective of this section is to understand the action of elements of $SO(n)$ on $\mathbb{R}^n$. This will give us a more mathematical intuition for what a rotation actually is. But before we can investigate the eigenvalues of elements of $SO(n)$ we must first understand the tools we will use in our investigation. In this section, we will show that the following properties hold for each $A \in SO(n)$ and the roots of its characteristic polynomial:

- $\det A = 1 = \lambda_1 \lambda_2 \lambda_3 \cdots \lambda_n$, where each λ_i is a root, counted with multiplicity
- if λ is a root, so is $\overline{\lambda}$

- if λ_i is real then $|\lambda_i| = 1$.

Recall that we define *eigenvalues* for an $n \times n$ matrix A to be all $\lambda \in \mathbb{R}$ for which there exists a non-zero vector $\mathbf{r}$ such that $A\mathbf{r} = \lambda\mathbf{r}$. The vectors $\mathbf{r}$ for which this is true are called the *eigenvectors* of A. We call the set of all such eigenvectors for a given eigenvalue λ, along with the zero vector, an *eigenspace* of A. Before we can begin our investigation of how eigenvalues relate to our discussion of rotations we must establish the properties listed above.

Theorem 2.1. *If A is an $n \times n$ matrix, the determinant of A is equal to the product of the n roots (counted with multiplicity) of the characteristic polynomial of A.*

Proof. Let A be an $n \times n$ matrix with characteristic polynomial

$$p(\lambda) = \det[\lambda I - A] = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

By the fundamental theorem of algebra, we have

$$p(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

where λ_i is the i^{th} root of the polynomial. Since $p(\lambda) = \det[\lambda I - A]$, $p(0) = \det[-A] = (-1)^n \det A = a_0$. Hence $a_0 = \det A$ for even n and $a_0 = -\det A$ for odd n. Since $p(0) = a_0$ we have that

$$\begin{aligned} a_0 &= (0 - \lambda_1)(0 - \lambda_2)\cdots(0 - \lambda_n) \\ &= (-\lambda_1)(-\lambda_2)\cdots(-\lambda_n) \\ &= (-1)^n \lambda_1\lambda_2\cdots\lambda_n. \end{aligned}$$

Hence, regardless of the value of n, $\det A = \lambda_1\lambda_2\cdots\lambda_n$. □

We now turn our attention to the second property, that if λ is a root of the characteristic polynomial of A, so is its conjugate. Note that the characteristic polynomial for a given element $A \in SO(n)$ is a polynomial with real coefficients and hence complex roots of the characteristic polynomial of A come in conjugate pairs. Thus, if λ is a root of the characteristic polynomial of A then $\overline{\lambda}$ is also a root, though not distinct in the case of real roots.

We now move on to the third property. Recall that any transformation $A \in SO(n)$ is a length-preserving isometry, so for all $\mathbf{r} \in \mathbb{R}^n$ we have $|A\mathbf{r}| = |\mathbf{r}|$. It follows that if λ is a real eigenvalue with eigenvector $\mathbf{r}$ then

$$|A\mathbf{r}| = |\lambda\mathbf{r}| = |\mathbf{r}|$$

which implies that $|\lambda| = 1$. Since all real roots of the characteristic polynomial must have absolute value of 1 and all roots must multiply to 1, the complex roots must all multiply to 1.

2.2 $SO(2)$

In this section we will find and discuss all the possible roots of the characteristic polynomials of elements of $SO(2)$. Because we are in $SO(2)$ the characteristic polynomial of each element will have two roots.

Consider an arbitrary element $A \in SO(2)$. We know that $\det A = 1 = \lambda_1\lambda_2$. From the previous section we know that $|\lambda_i| = 1$ when λ_i is real. It should be clear that the only possible roots are those appearing in the following table.

Roots	Rotation Angle
$1, 1$	0
$-1, -1$	π
$\omega, \overline{\omega}$	$\theta \neq 0, \pi$

We are fortunate enough to have a simple representation for this group so we will utilize it here. Let us find the roots of the characteristic polynomial of

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in SO(2).$$

Setting the characteristic polynomial equal to 0 gives $\lambda^2 - 2\lambda\cos\theta + 1 = 0$. By the quadratic formula, we have

$$\lambda = \frac{2\cos\theta \pm \sqrt{4\cos^2\theta - 4}}{2} = \frac{2\cos\theta \pm 2\sqrt{\cos^2\theta - 1}}{2} = \cos\theta \pm \sqrt{-\sin^2\theta}.$$

Hence the roots of the characteristic polynomial for an element of $SO(2)$ all come in the form

$$\lambda = \cos\theta \pm i\sin\theta.$$

We can see that the only real value for λ occurs when $\sin\theta = 0$, which will happen only when $\theta = 0, \pi$. Thus, we have shown that all of the aforementioned cases do in fact arise from the following scenarios:

- if $\theta = 0$ then the roots are $1, 1$
- if $\theta = \pi$ then the roots are $-1, -1$
- if $\theta \neq 0, \pi$ then the roots are complex.

Note that the formula $\cos\theta \pm i\sin\theta$ gives the roots for the characteristic polynomial associated with a rotation by an angle θ. We could just as easily use this process in reverse and calculate the angle of rotation of a transformation based on the roots of its characteristic polynomial.

Now we will calculate the eigenspaces for the real eigenvalues that we found and explain them geometrically. We will accomplish this by noting that when $\theta = 0$ the rotation matrix is

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence the eigenvalue is 1 with multiplicity 2. Since all vectors $\mathbf{r} \in \mathbb{R}^2$ will be solutions to $I\mathbf{r} = \mathbf{r}$, all of $\mathbb{R}^2$ comprises the eigenspace of this matrix. The dimension of this eigenspace is thus 2.

For $\theta = \pi$, we note that the matrix representation would be

$$-I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Note that the eigenvalue is -1 with multiplicity 2, and that this is a rotation by π. Therefore, once again, all vectors $\mathbf{r} \in \mathbb{R}^2$ are solutions to $-I\mathbf{r} = -\mathbf{r}$. Hence the eigenspace for this case is again all of $\mathbb{R}^2$.

For complex roots we do not consider the eigenspaces since this would take us outside of the real numbers. Hence, we have now described the eigenvalue structure for all elements of $SO(2)$. It was what our intuition would tell us: only the identity preserves the whole plane, the negative of the identity swaps all vectors about the origin, and except for these two cases, no rotation of the plane maps a vector to a scalar multiple of itself. But we noticed the interesting fact that the roots of the characteristic polynomial come in the form $\cos\theta \pm i\sin\theta$ where θ is the angle of rotation, giving a geometric link even to complex roots. Now let us explore higher dimensions.

2.3 $SO(3)$

Consider a matrix $A \in SO(3)$. By our previously established facts and theorems, the following table displays all possible roots for the characteristic polynomial of A.

Roots	Rotation Angle
$1, 1, 1$	0
$1, -1, -1$	π
$1, \omega, \overline{\omega}$	$\theta \neq 0, \pi$

Note. When discussing rotations in $\mathbb{R}^2$ an angle was enough to describe a rotation. Now that we are dealing with $\mathbb{R}^3$, we must specify an angle and a plane. Because we are restricted to $\mathbb{R}^3$, this plane may be specified by a vector orthogonal to the plane. This vector may be thought of as what we traditionally call the axis of rotation.

Case 1. The only example of the first case is the identity. This example clearly preserves all of 3-space and hence its eigenspace has dimension 3.

Case 2. For this case we shall consider the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

We can see that once again its transpose is itself and that $AA^T = I$. We can also see that the determinant is $+1$, hence $A \in SO(3)$. Because this is diagonal, we can see that the eigenvalues are $1, -1, -1$. Now we will compute the eigenspaces for A using the following augmented matrix for $\lambda = 1$

$$\begin{pmatrix} 1-1 & 0 & 0 & 0 \\ 0 & 1+1 & 0 & 0 \\ 0 & 0 & 1+1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

and the following for $\lambda = -1$:

$$\begin{pmatrix} -1-1 & 0 & 0 & 0 \\ 0 & -1+1 & 0 & 0 \\ 0 & 0 & -1+1 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

From this we can tell that for $\lambda = 1$ the eigenspace is generated by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and for $\lambda = -1$ the generators are

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Hence for $\lambda = 1$ we have a one-dimensional eigenspace. That is, all eigenvectors lie along an axis. For $\lambda = -1$ we see that the eigenvectors make up a plane. Notice however that for $\lambda = 1$ the vectors are invariant while for $\lambda = -1$ the vectors are flipped about the origin. This tells us that for this rotation, an axis stays the same and the yz-plane is rotated by π.

Case 3. We will now deal with the most general case of $\lambda = 1, \omega, \overline{\omega}$. Consider the matrix

$$A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that $AA^T = I$ so $A^T = A^{-1}$ and that $\det(A) = 1$, so $A \in SO(3)$. Now we will solve for its eigenvalues. Setting $\det(\lambda I - A) = 0$ gives $\lambda^3 - \lambda^2 + \lambda - 1 = 0$. Since we know from before that 1 must be an eigenvalue, we divide this polynomial by $\lambda - 1$ to obtain

$$\lambda^2 + 1 = 0$$

which has solutions

$$\lambda = \pm i.$$

Hence we have a rotation whose characteristic polynomial has roots $\lambda = 1, i, -i$.

Now we will find the eigenspace for $\lambda = 1$. We start with the following augmented matrix

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1-1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

which row reduces to

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, the eigenspace is generated by

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

So the entire z-axis is preserved but the rest of 3-space is not. If we inspect what this does to a vector in the xy-plane, e.g. $(1, 1, 0)$, we find the transformed vector is $(-1, 1, 0)$. Using the dot product we find that the cosine of the angle θ between this vector and its transformed counterpart is equal to 0. This tells us that this is a rotation by $\pm\frac{\pi}{2}$. This corresponds to $\sin\theta = \pm 1$.

Recall that in $SO(2)$ we could find all roots using the formula $\cos\theta \pm i \sin\theta$, where θ is the angle of rotation. Recall also that the complex roots for this rotation were $\pm i$. These complex roots are exactly what we would expect if the same formula applied to roots in $SO(3)$. We will need more evidence to support this claim before we attempt a proof.

Though we have given examples of each possible combination of roots for $SO(3)$, we will now demonstrate the eigenvalue properties of $SO(3)$ using a more complex example.

Example. Consider the matrix

$$A = \begin{pmatrix} -\frac{5}{8} & -\frac{3\sqrt{3}}{8} & \frac{\sqrt{3}}{4} \\ \frac{3\sqrt{3}}{8} & -\frac{1}{8} & \frac{3}{4} \\ -\frac{\sqrt{3}}{4} & \frac{3}{4} & \frac{1}{2} \end{pmatrix}.$$

Note that $AA^T = I$ and that $\det A = +1$ so that A is in $SO(3)$. Now let us compute its eigenvalues. We will solve the usual equation for determining eigenvalues:

$$\det \begin{pmatrix} \lambda + \frac{5}{8} & \frac{3\sqrt{3}}{8} & -\frac{\sqrt{3}}{4} \\ -\frac{3\sqrt{3}}{8} & \lambda + \frac{1}{8} & -\frac{3}{4} \\ \frac{\sqrt{3}}{4} & -\frac{3}{4} & \lambda - \frac{1}{2} \end{pmatrix} = \lambda^3 + \frac{1}{4}\lambda^2 - \frac{1}{4}\lambda - 1 = 0.$$

As we have shown, $+1$ must be an eigenvalue so we will divide this polynomial by $\lambda - 1$ to obtain

$$\lambda^2 + \frac{5}{4}\lambda + 1 = 0.$$

We will now use the quadratic formula to solve for the remaining roots. We find that

$$\lambda = \frac{-5 \pm \sqrt{-39}}{8} = \frac{-5 \pm i\sqrt{39}}{8}.$$

Hence the characteristic polynomial of A has two complex roots that are conjugates of each other. Now we will compute the eigenspace for $\lambda = 1$ using the augmented matrix

$$\begin{pmatrix} 1 + \frac{5}{8} & \frac{3\sqrt{3}}{8} & -\frac{\sqrt{3}}{4} & 0 \\ -\frac{3\sqrt{3}}{8} & 1 + \frac{1}{8} & -\frac{3}{4} & 0 \\ \frac{\sqrt{3}}{4} & -\frac{3}{4} & 1 - \frac{1}{2} & 0 \end{pmatrix} = \begin{pmatrix} \frac{13}{8} & \frac{3\sqrt{3}}{8} & -\frac{\sqrt{3}}{4} & 0 \\ -\frac{3\sqrt{3}}{8} & \frac{9}{8} & -\frac{3}{4} & 0 \\ \frac{\sqrt{3}}{4} & -\frac{3}{4} & \frac{1}{2} & 0 \end{pmatrix}$$

which row reduces to

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{3}{2} & 1 & 0 \end{pmatrix}.$$

Hence, the eigenspace is generated by

$$\begin{pmatrix} 0 \\ 1 \\ \frac{3}{2} \end{pmatrix}.$$

Note that this transformation still leaves a line invariant.

Leonhard Euler also noticed this fact and proved his famous theorem about three-dimensional rotations that, put in our language, is:

Euler's Rotation Theorem. *If A is an element of $SO(3)$ where $A \neq I$ then A has a one-dimensional eigenspace.*

This is not to say that A does not also have a two-dimensional eigenspace but that A must have a one-dimensional eigenspace. It is this eigenspace that is known as the axis of rotation. The work we have done in this section should be enough to convince us that this is true. However, as we will see in $SO(4)$, this way of viewing rotations does not generalize at all. In fact, having an axis of rotation is a special case of how we may define planes in 3-space that only applies to $\mathbb{R}^3$. It is for this reason that we have used the term "axis" sparingly.

Before we proceed with $SO(4)$, we noted earlier that there might be some geometric significance to the complex roots but needed another example to motivate a proof. We would like to see what angle of rotation corresponds to the matrix example above. The generator of the eigenspace, $(0, 1, \frac{3}{2})$, can be seen as the axis of rotation, so we will see what this matrix does to an element of the orthogonal complement of this eigenspace. We can see from the dot product that $\mathbf{v} = (0, -\frac{3}{2}, 1)$

is orthogonal to the axis of rotation. Also note that after being transformed this vector becomes $\mathbf{v}' = (\frac{13\sqrt{3}}{16}, \frac{15}{16}, -\frac{5}{8})$. Now we will determine the cosine of the angle between these two vectors:

$$\mathbf{v} \cdot \mathbf{v}' = \frac{13}{4} \cos\theta = -\frac{65}{32} \ .$$

Solving for the cosine term we see that $\cos\theta = -\frac{5}{8}$, which is the real part of the complex roots. This tells us that even a rotation about a more general axis still follows this tendency. In a later section we will prove that this same geometric significance exists for all roots of characteristic polynomials of matrices in $SO(3)$.

2.4 $SO(4)$

To begin our discussion of $SO(4)$ we will mimic the procedure we used to discuss $SO(3)$. We begin by noting the possible eigenvalues and other complex roots of the characteristic polynomials. We proceed by giving examples of each case and discussing their geometric significance.

Note. In $\mathbb{R}^2$ there is only one plane, so we only needed to specify an angle. In $\mathbb{R}^3$ we needed to specify a plane and an angle. In $\mathbb{R}^4$ we must specify two planes and an angle for each plane. The reasoning behind this will become apparent in this section.

Roots	Rotation Angles
$1, 1; 1, 1$	$0; 0$
$-1, -1; -1, -1$	$\pi; \pi$
$1, 1; -1, -1$	$0; \pi$
$1, 1; \omega, \overline{\omega}$	$0; \theta$
$-1, -1; \omega, \overline{\omega}$	$\pi; \theta$
$\omega, \overline{\omega}; z, \overline{z}$	$\theta; \phi$

Notice that there are twice the number of cases for $SO(4)$ roots than for $SO(3)$ or $SO(2)$. This comes from the fact that we may rotate in, say, the x_1x_2-plane and then rotate in the x_3x_4-plane without disturbing our previous rotation. As we will see, $\mathbb{R}^4$ is an interesting place.

Case 1. As usual, the identity handles the first case and it is clear that the eigenspace for $I \in SO(4)$ is all of $\mathbb{R}^4$.

Case 2. It is simple to construct an example by considering $-I$. Since $-I\mathbf{x} = -\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^4$, we know that the eigenspace is once again all of $\mathbb{R}^4$. In this case, each vector in $\mathbb{R}^4$ is swapped about the origin.

Does this rotation take place about an axis? In three dimensions we had a similar case in which a plane of vectors was rotated by π. However that had an axis because only a plane was involved and there was an extra degree of freedom which we called the axis. In this case, however, all vectors are included in the eigenspace so there is no favored axis of rotation here: the entire space is flipped. In three-space we think of rotations as occurring about one-dimensional axes, but what is really going on is that a plane is being rotated and any vector with a non-trivial projection in that plane will also be rotated. However, only the component in the direction of the projection will be rotated: the component orthogonal to the plane of rotation will remain unchanged. If we remember, this is what happened in two-space as well. In fact, it is by embedding the two-dimensional rotations into the higher dimensions that we have obtained most of the examples in this section. So, looking at it from this perspective, the first case would be rotation by 0 radians. The second would be a rotation by π of, say, the x_1x_2-plane followed by a rotation of the x_3x_4-plane by π. We can see that this acts exactly as we claim it does:

$$\begin{pmatrix} \cos\pi & -\sin\pi & 0 & 0 \\ \sin\pi & \cos\pi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\pi & -\sin\pi \\ 0 & 0 & \sin\pi & \cos\pi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$
$$= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ -x_3 \\ -x_4 \end{pmatrix} = \begin{pmatrix} -x_1 \\ -x_2 \\ -x_3 \\ -x_4 \end{pmatrix}.$$

Thinking about rotations in this way may help to alleviate the confusion that is common when trying to visualize higher dimensions.

Case 3. For this case, we can use the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

It should be clear that the eigenvalues for this matrix are $1, 1, -1, -1$. Solving for the eigenspaces, we will need to use two augmented matrices: $(I - A|0)$ and $(-I - A|0)$. In the first case we find an eigenspace generated by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

This eigenspace is simply a copy of the x_1x_2-plane. For the second augmented matrix we obtain an eigenspace generated by

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

This eigenspace is a copy of the x_3x_4-plane. However, if we once again use this matrix to see where an arbitrary vector is sent we find

$$A\mathbf{x} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ -x_3 \\ -x_4 \end{pmatrix}.$$

Hence we have a rotation by π of the x_3x_4-plane and a rotation by 0 of the x_1x_2-plane.

Case 4. We will show that the following matrix satisfies this case and then examine the implications of such a rotation:

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

It is not hard to see that $BB^T = I$ and $\det(B) = 1$ and hence $B \in SO(4)$. Now we can move on to its eigenvalues which we will find in the usual way:

$$\det(\lambda I - B) = \det \begin{pmatrix} \lambda - 1 & 0 & 0 & 0 \\ 0 & \lambda - 1 & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & -1 & \lambda \end{pmatrix} = (\lambda - 1)(\lambda - 1)(\lambda^2 + 1) = 0.$$

We see that $\lambda = 1$ appears as a solution with multiplicity 2 and we have $\lambda = \pm i$ as the other two solutions. We have thus shown that this is an example of Case 4. Now we must find the eigenspace for $\lambda = 1$ using the augmented matrix

$$\begin{pmatrix} 1-1 & 0 & 0 & 0 & 0 \\ 0 & 1-1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix}$$

which row reduces to

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Hence, the eigenspace is generated by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

which generates a copy of the x_1x_2-plane. To avoid confusion and to see what happens to the other two parameters, let us see where B maps an arbitrary vector:

$$B\mathbf{x} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ -x_4 \\ x_3 \end{pmatrix}.$$

This rotation leaves the x_1x_2-plane invariant (as we said before) and it rotates the x_3x_4-plane by $\frac{\pi}{2}$.

Case 5. For the case where $\lambda = -1, -1, \omega, \overline{\omega}$, we will use our previous example as inspiration to construct the following matrix:

$$B' = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Once again, it is easy to show that this is an element of $SO(4)$. When we compute its eigenvalues we find that they are $\lambda = -1, -1, i, -i$, so this is in fact an example of Case 5. Watching what this matrix does to an arbitrary vector, however, we see that:

$$B'\mathbf{x} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -x_1 \\ -x_2 \\ -x_4 \\ x_3 \end{pmatrix}.$$

This still rotates the x_3x_4-plane by $\frac{\pi}{2}$ but it also rotates the x_1x_2-plane by π.

Case 6. This is the case where there are no real eigenvalues. We might think that such an example would have to be a very exotic matrix. However, using what we've learned about rotations, all we need is a matrix that doesn't map any vector in $\mathbb{R}^4$ to a scalar multiple of itself. So, if we rotate the x_1x_2-plane by $\frac{\pi}{2}$ and do the same to the x_3x_4-plane, that action should satisfy this case. Let's do that using the matrix

$$C = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

We now compute the eigenvalues of this matrix:

$$\det(\lambda I - C) = \det \begin{pmatrix} \lambda & 1 & 0 & 0 \\ -1 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & -1 & \lambda \end{pmatrix} = \lambda^4 + 2\lambda^2 + 1 = 0.$$

This factors into

$$(\lambda^2 + 1)^2 = 0.$$

Hence, the eigenvalues are $\lambda = \pm i$, each with multiplicity two, which gives us an example of Case 6. For completeness, we will now verify that it transforms vectors the way we constructed it to:

$$C\mathbf{x} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \\ -x_4 \\ x_3 \end{pmatrix}.$$

As we can see this is a rotation of the x_1x_2-plane by $\frac{\pi}{2}$ and then a rotation of the x_3x_4-plane by $\frac{\pi}{2}$, just as we claimed during its construction.

2.5 $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{R}^4$

The following are important properties to note about the relationships between rotations in two dimensions and those in three and four dimensions. With these properties we will be able to simplify our work in the final section while remaining mathematically rigorous in the theorems to come.

Theorem 2.2. *Any arbitrary element of $SO(3)$ may be written as the composition of rotations in the planes generated by the three standard orthogonal basis vectors of $\mathbb{R}^3$.*

Proof. It is easy to show that the following matrices represent rotations by θ_1, θ_2, and θ_3 in the xy-, yz- and xz-planes respectively:

$$R_z = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_2 & -\sin\theta_2 \\ 0 & \sin\theta_2 & \cos\theta_2 \end{pmatrix}, R_y = \begin{pmatrix} \cos\theta_3 & 0 & -\sin\theta_3 \\ 0 & 1 & 0 \\ \sin\theta_3 & 0 & \cos\theta_3 \end{pmatrix}.$$

As we know, all rotations in $\mathbb{R}^3$—aside from the identity—have a one-dimensional eigenspace about which we rotate by an angle θ. A unit vector $\mathbf{n}$ in this eigenspace may be parameterized as

$$\mathbf{n} = (\cos\alpha\sin\beta, \sin\alpha\sin\beta, \cos\beta).$$

This is because the one-dimensional eigenspace, or axis of rotation, must intersect the unit sphere centered at the origin and the above can be shown to parameterize the unit sphere.

We may align the axis of rotation with the z-axis as follows:

$$\begin{aligned} R_y(\beta)R_z(\alpha)^T\mathbf{n} &= \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\alpha\sin\beta \\ \sin\alpha\sin\beta \\ \cos\beta \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

We can thus perform the rotation by θ in the xy-plane and then transform back. All we need to know is the angle of rotation θ and the angles α, β which specify the rotation's one-dimensional eigenspace to write any element of $SO(3)$ as

$$R_z(\alpha)R_y(\beta)^T R_z(\theta)R_y(\beta)R_z(\alpha)^T.$$

This is not unique since we could just as easily transform the eigenspace to align with either of the other two axes. Hence, as we claimed, we may write any rotation in $SO(3)$ as the product of rotations in the orthogonal planes of $\mathbb{R}^3$. □

Recall earlier we noticed a pattern indicative of a possible geometric interpretation of complex eigenvalues. Now we will state and prove this fact in the following theorem.

Theorem 2.3. *For any element $A \in SO(3)$ with rotation angle θ, the roots of the characteristic polynomial of A are 1 and $\cos\theta \pm i\sin\theta$.*

Proof. Let A be an element of $SO(3)$ with axis generated by $\mathbf{n} = (\cos\alpha\sin\beta, \sin\alpha\sin\beta, \cos\beta)$. By Theorem 2.2, we may represent A as the product

$$\begin{aligned} A &= R_z(\alpha)R_y(\beta)^T R_z(\theta)R_y(\beta)R_z(\alpha)^T \\ &= (R_z(\alpha)R_y(\beta)^T)R_z(\theta)(R_z(\alpha)R_y(\beta)^T)^T \\ &= PR_z(\theta)P^T. \end{aligned}$$

This means that the matrix A is similar to the matrix

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The characteristic polynomial of this matrix is

$$(\lambda - 1)(\lambda^2 - 2\lambda\cos\theta + 1).$$

Dividing by $\lambda - 1$ we find that the roots that are not equal to 1 are $\cos\theta \pm i\sin\theta$. Since A and $R_z(\theta)$ are similar they share the same characteristic polynomial. Hence for any rotation by θ in $\mathbb{R}^3$ the roots of the characteristic polynomial are 1, $\cos\theta + i\sin\theta$, and $\cos\theta - i\sin\theta$. □

Proposition 2.4. *Any arbitrary element of $SO(4)$ may be written as the composition of rotations in the planes generated by the four orthogonal basis vectors of $\mathbb{R}^4$.*

The proof involves changing the basis of an arbitrary element of $SO(4)$ such that the two planes of rotation are made to coincide with the x_1x_2-plane and the x_3x_4-plane. We would change the basis using products of the matrices

$$\begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} \cos\theta_2 & 0 & -\sin\theta_2 & 0 \\ 0 & 1 & 0 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} \cos\theta_3 & 0 & 0 & -\sin\theta_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sin\theta_3 & 0 & 0 & \cos\theta_3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_4 & -\sin\theta_4 & 0 \\ 0 & \sin\theta_4 & \cos\theta_4 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_5 & 0 & -\sin\theta_5 \\ 0 & 0 & 1 & 0 \\ 0 & \sin\theta_5 & 0 & \cos\theta_5 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_6 & -\sin\theta_6 \\ 0 & 0 & \sin\theta_6 & \cos\theta_6 \end{pmatrix}.$$

We would then perform our rotation using the matrices

$$\begin{pmatrix} \cos\phi_1 & -\sin\phi_1 & 0 & 0 \\ \sin\phi_1 & \cos\phi_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\phi_2 & -\sin\phi_2 \\ 0 & 0 & \sin\phi_2 & \cos\phi_2 \end{pmatrix}$$

and then change back using the appropriate transposes of the matrices above. Without an axis it is more difficult to see how to change the basis appropriately, so an actual proof will be omitted. However, this is a known fact that will be assumed later.

3 Other Representations

Recall that the reason we have been using matrices was that rotations turned out to be linear transformations and matrices are natural representations of linear transformations. However, we must not believe that these are the only possible ways of expressing rotations. After all, we can easily represent a rotation in $\mathbb{R}^3$ by using just an axis and an angle. In this section we will turn our attention to alternative representations of the special orthogonal groups. Specifically, we will consider those representations given by complex numbers and quaternions.

3.1 $\mathbb{C}$ and $SO(2)$

A complex number is a number with a real and an imaginary component, written $z = a + bi$. For all elements $z_1 = a + bi$, $z_2 = c + di \in \mathbb{C}$ we define addition as $z_1 + z_2 = (a + c) + (b + d)i$. It is simple to check that with this addition operation $\mathbb{C}$ is a two-dimensional vector space over the real numbers which makes it isomorphic to $\mathbb{R}^2$ as a vector space. If $z = x + yi$ then we define the *conjugate* to be $\overline{z} = x - yi$. Also, we have a multiplication operation defined for every pair of elements as above by $z_1 z_2 = (ac - bd) + (ad + bc)i$. With this knowledge we may define the *norm* of $z \in \mathbb{C}$ as

$$|z| = \sqrt{z\overline{z}}.$$

Recall that the norm is the length of the vector. Another useful property to note is that the norm is multiplicative, that is,

$$|z_1 z_2| = |z_1||z_2|.$$

With this knowledge we are able to express the unit circle, or one-sphere S^1, as

$$S^1 = \{z \in \mathbb{C} \ : \ |z| = 1\}.$$

Lemma 3.1. *The one-sphere is an abelian group under complex multiplication.*

Proof. Recall that the set of non-zero complex numbers is a group under multiplication. This means we only need to check that S^1 is a subgroup of $\mathbb{C} - \{0\}$. We start by checking closure. Let z_1, z_2 be elements of S^1. Then we have that

$$|z_1 z_2| = |z_1||z_2| = 1 \cdot 1 = 1.$$

Hence S^1 is closed under the multiplication operation of $\mathbb{C}$. It is clear that $|1| = 1$ and it should also be clear that this is the identity under multiplication. Hence S^1 possesses an identity. Also, every element $z \in S^1$ has an inverse in $\mathbb{C}$ such that $zz^{-1} = 1$. Then $|zz^{-1}| = |1| = 1$ and $|z| = |z^{-1}| = 1$, hence if z is in S^1 so is its inverse. Associativity is inherited from the complex numbers and we have shown that S^1 is a group under complex multiplication. Complex multiplication is commutative, so S^1 is abelian. □

Recall that the set of unit complex numbers may be written in the form $e^{i\theta} = \cos\theta + i\sin\theta$. This new form will aid us in the proof of the main theorem for this section.

Theorem 3.2. *The group $SO(2)$ is isomorphic to S^1.*

Proof. Recall that any element in $SO(2)$ may be represented as a matrix

$$A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

where θ is the angle of rotation, chosen to be in $[0, 2\pi)$. Consider the mapping $f \ : \ S^1 \to SO(2)$, where $f(e^{i\theta}) = A(\theta)$. See that

$$\begin{aligned} f(e^{i\theta})f(e^{i\phi}) &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & -(\cos\theta\sin\phi + \sin\theta\cos\phi) \\ \sin\theta\cos\phi + \cos\theta\sin\phi & \cos\theta\cos\phi - \sin\theta\sin\phi \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}. \end{aligned}$$

Now see that

$$\begin{aligned} f(e^{i\theta}e^{i\phi}) &= f(e^{i(\theta+\phi)}) \\ &= \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix} \\ &= f(e^{i\theta})f(e^{i\phi}). \end{aligned}$$

Thus f is a homomorphism. By our representation, every rotation in $SO(2)$ is of the form $A(\theta)$ for some angle θ, and in this case $f(e^{i\theta}) = A(\theta)$. Hence f is surjective. Since θ is chosen to be in $[0, 2\pi)$ it is uniquely determined, and hence f is injective. Therefore S^1 is isomorphic to $SO(2)$. □

3.2 $\mathbb{H}$ and $SO(3)$

We have shown that S^1 is a group under complex multiplication and that it is isomorphic to $SO(2)$. We would now like to find representations for higher-dimensional rotations. Since the elements of $SO(3)$ can be seen as invariant points on the unit sphere S^2 in $\mathbb{R}^3$, we turn our attention to S^2. Note that an element of $SO(3)$ is not just a point on a sphere: it contains another piece of information, the angle of rotation. Therefore, we look to the next dimension up, the three-sphere S^3. But before we can hope to find a representation using elements of the three-sphere, we must establish its multiplication rules. To begin we would like to find a set that is a four-dimensional vector space and whose non-zero elements form a group under multiplication. Thus we must introduce Hamilton's quaternions.

The set of *quaternions* $\mathbb{H}$ is the set of generalized complex numbers $q = a_0 + a_1 i + a_2 j + a_3 k$, where i, j, k are imaginary numbers satisfying the properties:

- $i^2 = j^2 = k^2 = ijk = -1$
- $ij = k$
- $jk = i$
- $ki = j$
- $ji = -k$
- $kj = -i$
- $ik = -j$.

We define the addition of quaternions similarly to the way we defined complex addition. Explicitly, for every pair of elements $q_1 = a_0 + a_1 i + a_2 j + a_3 k$, $q_2 = b_0 + b_1 i + b_2 j + b_3 k \in \mathbb{H}$ we define their addition component-wise:

$$q_1 + q_2 = (a_0 + b_0) + (a_1 + b_1)i + (a_2 + b_2)j + (a_3 + b_3)k$$

where all coefficients are real numbers. Thus we can see that addition is commutative in $\mathbb{H}$. In fact, it is not hard to see that $\mathbb{H}$ makes up a four-dimensional vector space and hence is isomorphic to $\mathbb{R}^4$ as a vector space.

Just like in the complex numbers we have a conjugate and it is defined in a similar fashion. If $q = a_0 + a_1 i + a_2 j + a_3 k$ is a quaternion we say that the *conjugate* of q, denoted $\overline{q}$, is given by $\overline{q} = a_0 - a_1 i - a_2 j - a_3 k$. Another similarity between the quaternions and the complex numbers is that we may define a multiplication operation on $\mathbb{H}$. This quaternionic multiplication works using the regular distribution laws combined with the identities above, so that for every pair of elements $q_1 = a_0 + a_1 i + a_2 j + a_3 k$, $q_2 = b_0 + b_1 i + b_2 j + b_3 k \in \mathbb{H}$ we define $q_1 q_2$ by

$$\begin{aligned} q_1 q_2 = {} & (a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3) + (a_0 b_1 + a_1 b_0 + a_2 b_3 - a_3 b_2)i + (a_0 b_2 + a_2 b_0 + a_3 b_1 - a_1 b_3)j \\ & + (a_0 b_3 + a_3 b_0 + a_1 b_2 - a_2 b_1)k. \end{aligned}$$

Notice that this does not always commute. Now let us see what happens when $q_2 = \overline{q_1}$:

$$q_1\overline{q_1} = (a_0^2 + a_1^2 + a_2^2 + a_3^2) + (0)i + (0)j + (0)k = a_0^2 + a_1^2 + a_2^2 + a_3^2.$$

This expression happens to be the square of the norm for a vector in $\mathbb{R}^4$, hence we define the *norm* of a quaternion q to be

$$|q| = \sqrt{q\overline{q}}.$$

It is not hard to show that the following property holds for all quaternions q_1, q_2:

$$|q_1q_2| = |q_1||q_2|.$$

Now we are prepared to define the three-sphere using quaternions. Since the three-sphere S^3 is the set of all unit length vectors in $\mathbb{R}^4$ (or $\mathbb{H}$) we may define S^3 as follows:

$$S^3 = \{q \in \mathbb{H} \ : \ |q| = 1\}.$$

We can now prove the following result.

Lemma 3.3. *The three-sphere is a non-abelian group under quaternionic multiplication.*

Proof. Consider the norm of the multiplication of two elements $q_1, q_2 \in S^3$:

$$|q_1q_2| = |q_1||q_2| = 1.$$

Hence S^3 is closed under quaternionic multiplication.

Next, we can see that the set contains an identity, namely the number 1. We know this acts as the identity because of how real numbers distribute over the quaternions. Also note that for all $q \in S^3$ we have $|q|^2 = q\overline{q} = 1$, hence $\overline{q}$ is the inverse of q. It is clear that $q^{-1} \in S^3$.

Finally, S^3 inherits associativity from $\mathbb{H}$. Thus S^3 is a group under quaternionic multiplication. Furthermore, this group is not abelian, since quaternions rarely commute. □

Now that we know that this is a group we are in a position to prove the following theorem.

Theorem 3.4. *There exists a two-to-one homomorphism from S^3 onto $SO(3)$.*

Proof. To show this we must demonstrate that every quaternion of unit length may be mapped to a rotation in $\mathbb{R}^3$ and that every rotation in $\mathbb{R}^3$ has two quaternion representations.

Let $\mathbf{n} = a_1i + a_2j + a_3k$ be such that $|\mathbf{n}| = 1$. Consider a unit quaternion written in the form

$$q = \cos\phi + \mathbf{n}\sin\phi.$$

It is not hard to show that by varying both the direction of $\mathbf{n}$ and the angle ϕ we may represent any element of S^3 in this form.

We will show that we can use quaternionic multiplication to construct rotations in $\mathbb{R}^3$. Let $\mathbf{r} = (x, y, z)$ be a vector in $\mathbb{R}^3$. We could just as well write this vector as $\mathbf{r} = (0, x, y, z)$ in $\mathbb{R}^4$. Since $\mathbb{R}^4$ is isomorphic to $\mathbb{H}$ as a vector space we could also write this vector as $\mathbf{r} = xi + yj + zk$. If we multiply by a unit quaternion q then we see that $|q\mathbf{r}| = |\mathbf{r}|$, so that multiplication by q is length-preserving. However, this multiplication will in general yield a non-zero real component in $q\mathbf{r}$, which will prevent us from mapping the result back to $\mathbb{R}^3$. So we need to find a multiplication whose result does not contain a real part. We will try the following:

$$q\mathbf{r}\overline{q} = (\cos\phi + \mathbf{n}\sin\phi)\mathbf{r}(\cos\phi - \mathbf{n}\sin\phi).$$

When we expand this we see that it does not contain a real part:

$$\begin{aligned} q\mathbf{r}\overline{q} = & [\cos^2\phi x + 2\cos\phi\sin\phi(za_2 - ya_3) - \sin^2\phi(a_1(-xa_1 - ya_2 - za_3) + a_2(xa_2 - ya_1) - a_3(za_1 - xa_3))]i \\ & + [\cos^2\phi y + 2\cos\phi\sin\phi(xa_3 - za_1) - \sin^2\phi(a_2(-xa_1 - ya_2 - za_3) + a_3(ya_3 - za_2) - a_1(xa_2 - ya_1))]j \\ & + [\cos^2\phi z + 2\cos\phi\sin\phi(ya_1 - xa_2) - \sin^2\phi(a_3(-xa_1 - ya_2 - za_3) + a_1(za_1 - xa_3) - a_2(ya_3 - za_2))]k. \end{aligned}$$

This is a step in the right direction because this vector has only three components.

We will now construct the matrix representation of this linear transformation on $\mathbb{R}^3$. By evaluating $q\mathbf{r}\overline{q}$ as $\mathbf{r}$ ranges over the standard basis vectors, we see that the matrix representation of this operation is

$$A = \begin{pmatrix} \cos^2\phi + \sin^2\phi(a_1^2 - a_2^2 - a_3^2) & -2\cos\phi\sin\phi a_3 + 2\sin^2\phi a_1 a_2 & 2\cos\phi\sin\phi a_2 + 2\sin^2\phi a_1 a_3 \\ 2\cos\phi\sin\phi a_3 + 2\sin^2\phi a_1 a_2 & \cos^2\phi + \sin^2\phi(a_2^2 - a_1^2 - a_3^2) & -2\cos\phi\sin\phi a_1 + 2\sin^2\phi a_2 a_3 \\ -2\cos\phi\sin\phi a_2 + 2\sin^2\phi a_1 a_3 & 2\cos\phi\sin\phi a_1 + 2\sin^2\phi a_2 a_3 & \cos^2\phi + \sin^2\phi(a_3^2 - a_2^2 - a_1^2) \end{pmatrix}.$$

Furthermore, it can be shown that $\det A = 1$ and that $A^T A = I$. Given a unit length quaternion q, we define a transformation f_q on $\mathbb{R}^3$ by $f_q(\mathbf{r}) = q\mathbf{r}\overline{q}$. The argument above proves that $f_q \in SO(3)$.

Define a map $f : S^3 \to SO(3)$ by $f(q) = f_q$. We claim that f is a two-to-one surjective homomorphism. Consider two elements $q_1, q_2 \in S^3$. Note the following:

$$\begin{aligned} f(q_1q_2)(\mathbf{r}) &= q_1q_2\mathbf{r}\overline{q_1q_2} \\ &= q_1q_2\mathbf{r}(q_1q_2)^{-1} \\ &= q_1q_2\mathbf{r}q_2^{-1}q_1^{-1} \\ &= (f(q_1) \circ f(q_2))(\mathbf{r}). \end{aligned}$$

Thus, f is a homomorphism. Now we will show that this map is onto.

Consider the matrix associated with the quaternion $q = \cos\phi + \mathbf{n}\sin\phi$ when $\mathbf{n} = (1, 0, 0)$. We find this by setting $a_1 = 1$ and $a_2 = a_3 = 0$ in the matrix A above. We find that this matrix is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos^2\phi - \sin^2\phi & -2\cos\phi\sin\phi \\ 0 & 2\cos\phi\sin\phi & \cos^2\phi - \sin^2\phi \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\phi & -\sin 2\phi \\ 0 & \sin 2\phi & \cos 2\phi \end{pmatrix}.$$

Note that it is not a rotation by ϕ but instead a rotation by 2ϕ in the yz-plane. Notice that if $\mathbf{n}$ had coincided with the y- or z-axis we would have obtained a similar matrix, except it would represent a rotation of 2ϕ about the chosen axis. Since we can generate rotations about the standard orthogonal basis vectors using quaternions, we may invoke Theorem 2.2 to conclude that this homomorphism is onto.

To see that f is two-to-one, we will show that the kernel of f contains only two elements. It is clear that f_1 and f_{-1} are both the identity rotation. We claim that ± 1 are the only elements in the kernel. Suppose that $q \in \ker(f)$, so that f_q is the identity rotation. Then $q\mathbf{r}\overline{q} = \mathbf{r}$ for all $\mathbf{r} = xi + yj + zk$. This would imply that $q\mathbf{r} = \mathbf{r}q$ for all such $\mathbf{r}$. The only quaternions that commute with all pure quaternions are the reals, so q must be real. Since q must be unit length, we conclude that $q = \pm 1$. Hence f is two-to-one. □

While this is not an isomorphism because it is two-to-one, by the First Isomorphism Theorem we see that $S^3/\mathbb{Z}_2 \cong SO(3)$. This result is important because it allows us to prove a famous result from topology using only linear and abstract algebra. First, note that in S^3 every point (x_1, x_2, x_3, x_4) has an *antipode* $(-x_1, -x_2, -x_3, -x_4)$. This relation leads to a normal subgroup of order 2 which is isomorphic to $\mathbb{Z}_2$. If we take S^3 modulo this subgroup we obtain real projective space $\mathbb{R}P^3$. There is a famous result from topology that states $\mathbb{R}P^3 \cong SO(3)$. However, we have arrived at this result using mostly linear algebra.

3.3 $\mathbb{H} \times \mathbb{H}$ and $SO(4)$

We showed that S^1 is isomorphic to $SO(2)$. We then showed that there is a two-to-one homomorphism from S^3 onto $SO(3)$. We will now attempt to find a similar representation of $SO(4)$. Intuition might suggest we attempt to find another sphere to represent $SO(4)$. However, from our previous work with $SO(4)$, we know there are fundamental differences in the behavior of rotations in $\mathbb{R}^4$. We understand from earlier that $SO(3)$ is like $SO(2)$ except we may specify the plane in which to rotate by giving the orthogonal complement of that plane, namely the axis. This had the effect of taking away commutativity for $SO(3)$. However, elements of $SO(4)$ are not just single elements of $SO(2)$ oriented in odd directions, they are more like pairs of noncommutative elements of $SO(2)$. Elements of $SO(3)$ can be seen as noncommutative elements of $SO(2)$ so perhaps we should investigate $S^3 \times S^3$. As we will soon see, this choice is, in fact, correct.

Theorem 3.5. *There exists a two-to-one, surjective homomorphism $f : S^3 \times S^3 \to SO(4)$.*

Proof. Let $\mathbf{x} = (x_1, x_2, x_3, x_4)$ be a vector in $\mathbb{R}^4$. Since $\mathbb{H}$ is a vector space isomorphic to $\mathbb{R}^4$ we may represent this vector as the quaternion $\mathbf{x} = x_1 + x_2 i + x_3 j + x_4 k$. Given two unit length quaternions q_1, q_2, we define a transformation f_{q_1,q_2} on $\mathbb{R}^4$ by

$$f_{q_1,q_2}(\mathbf{x}) = q_1 \mathbf{x} \overline{q_2}.$$

We claim that $f_{q_1,q_2} \in SO(4)$. We know that $|q_1 \mathbf{x} \overline{q_2}| = |\mathbf{x}|$ so the transformation f_{q_1,q_2} preserves length. Also note that if we let $\mathbf{x} = \mathbf{0}$ then $q_1 \mathbf{x} \overline{q_2} = \mathbf{0}$. Thus such a transformation is also origin-preserving and hence linear. Since f_{q_1,q_2} is a length-preserving linear transformation on $\mathbb{R}^4$, we know that it is an element of $O(4)$.

We will now construct the matrix representation of a general transformation f_{q_1,q_2}. Let $q_1 = a_0 + a_1 i + a_2 j + a_3 k$ and $q_2 = b_0 - b_1 i - b_2 j - b_3 k$ be two unit quaternions. Then $f_{q_1,q_2}(\mathbf{x}) = q_1 \mathbf{x} \overline{q_2}$ where $\mathbf{x}$ is an arbitrary vector in $\mathbb{R}^4$. Notice that $|\mathbf{x}\overline{q_2}| = |\mathbf{x}|$ and that $\mathbf{0}\overline{q_2} = \mathbf{0}$, so that f_{1,q_2} alone is a length-preserving linear transformation on $\mathbb{R}^4$. Hence it is in $O(4)$ and may be represented with a matrix, specifically

$$\begin{pmatrix} b_0 & -b_1 & -b_2 & -b_3 \\ b_1 & b_0 & b_3 & -b_2 \\ b_2 & -b_3 & b_0 & b_1 \\ b_3 & b_2 & -b_1 & b_0 \end{pmatrix}.$$

Solving for the determinant we find that it is 1, so the linear transformation f_{1,q_2} is an element of $SO(4)$. Along a similar argument we can see that $f_{q_1,1}$ is a linear transformation that preserves length. Hence we may find that its matrix representation is

$$\begin{pmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{pmatrix}.$$

The determinant of this matrix is also 1, so $f_{q_1,1}$ is in $SO(4)$. Now a general transformation f_{q_1,q_2} is simply the composition of these two transformations. Hence its matrix representation is the product of the matrix representations of f_{1,q_2} and $f_{q_1,1}$. When we multiply these two matrices together we find that the matrix representation for a general transformation f_{q_1,q_2} is

$$\begin{pmatrix} a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3 & -a_0b_1 - a_1b_0 + a_2b_3 - a_3b_2 & -a_0b_2 - a_1b_3 - a_2b_0 + a_3b_1 & -a_0b_3 + a_1b_2 - a_2b_1 - a_3b_0 \\ a_1b_0 + a_0b_1 - a_3b_2 + a_2b_3 & -a_1b_1 + a_0b_0 + a_3b_3 + a_2b_2 & -a_1b_2 + a_0b_3 - a_3b_0 - a_2b_1 & -a_1b_3 - a_0b_2 - a_3b_1 + a_2b_0 \\ a_2b_0 + a_3b_1 + a_0b_2 - a_1b_3 & -a_2b_1 + a_3b_0 - a_0b_3 - a_1b_2 & -a_2b_2 + a_3b_3 + a_0b_0 + a_1b_1 & -a_2b_3 - a_3b_2 + a_0b_1 - a_1b_0 \\ a_3b_0 - a_2b_1 + a_1b_2 + a_0b_3 & -a_3b_1 - a_2b_0 - a_1b_3 + a_0b_2 & -a_3b_2 - a_2b_3 + a_1b_0 - a_0b_1 & -a_3b_3 + a_2b_2 + a_1b_1 + a_0b_0 \end{pmatrix}.$$

Since this matrix is the product of two elements of $SO(4)$ it must also be an element of $SO(4)$.

We define $f : S^3 \times S^3 \to SO(4)$ by $f(q_1, q_2) = f_{q_1,q_2}$. We claim that f is a surjective, two-to-one homomorphism. Now consider quaternions $q_1, q_2, q_3, q_4 \in S^3$. We check that f is a homomorphism:

$$\begin{aligned} f(q_1q_3, q_2q_4)(\mathbf{x}) &= q_1q_3\mathbf{x}\overline{q_2q_4} \\ &= q_1q_3\mathbf{x}(q_2q_4)^{-1} \\ &= q_1q_3\mathbf{x}q_4^{-1}q_2^{-1} \\ &= (f(q_1, q_2) \circ f(q_3, q_4))(\mathbf{x}). \end{aligned}$$

Hence f is a homomorphism, as claimed.

Now, letting $a_0 = b_0 = \cos\theta$ and $a_3 = b_3 = \sin\theta$ in the matrix above we obtain

$$\begin{pmatrix} \cos 2\theta & 0 & 0 & -\sin 2\theta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sin 2\theta & 0 & 0 & \cos 2\theta \end{pmatrix}.$$

Notice that this is a rotation by 2θ in the x_1x_4-plane. We can similarly select rotations to occur in any of the six coordinate planes of $\mathbb{R}^4$. Recall that Proposition 2.4 stated that every rotation in $SO(4)$ may be expressed as a product of the resulting six matrices. Thus, by Proposition 2.4 we may conclude that f is onto.

Furthermore, notice that $f(1, 1)$ is the identity and so is $f(-1, -1)$. Assume that $f(q_1, q_2)$ is the identity. Then it must be true that $q_1\mathbf{x}\overline{q_2} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^4$. Let $\mathbf{x} = 1$ to see that q_1 must equal q_2. Now we have $q_1\mathbf{x}\overline{q_1} = \mathbf{x}$ and so q_1 commutes with every quaternion $\mathbf{x}$. Hence q_1 is real. Since q_1 is unit length, it must be ± 1. Thus the kernel of f contains only two elements, so f is two-to-one. □

References

[1] Simon Altmann, *Rotations, Quaternions, and Double Groups*, Dover Publications, Mineola, NY, 1986.

[2] Andrew Hanson, *Visualizing Quaternions*, Morgan Kaufmann, San Francisco, CA, 2006.

[3] Ron Larson, Bruce Edwards, and David Falvo, *Elementary Linear Algebra*, 5 ed., Houghton Mifflin, Boston, MA, 2004.

[4] Kristopher Tapp, *Matrix Groups for Undergraduates*, Student Mathematical Library, vol. 29, American Mathematical Society, Providence, RI, 2005.