The Banach-Tarski Paradox

Madeline Tremblay

May 2017

I would like to give an infinite (one of the infinities that is greater than most other infinities) amount of thanks to my thesis advisor, Dr. Reed Solomon, without whose unrelenting patience and seemingly bottomless knowledge this paper would have been little more than a very long series of question marks.

1 Introduction

The Banach-Tarski Paradox is a famous theorem about the equivalence of sets. In this paper, using Karl Stromberg's version of the proof in [1] as a guide, I will begin by stating the Banach-Tarski Theorem and then proceed to prove it.

Theorem 1.1. *Banach-Tarski Theorem: If X and Y are bounded subsets of $\mathbb{R}^3$ having nonempty interiors, then there exist a natural number n and partitions $\{X_j : 1 \leq j \leq n\}$ and $\{Y_j : 1 \leq j \leq n\}$ of X and Y respectively (into n pieces each) such that X_j is congruent to Y_j for all j.*

In summary, this says that, if we have bounded sets X and Y in $\mathbb{R}^3$ such that X and Y each contains a closed ball then we can break X into some finite number of pieces and reassemble these pieces using only rotations and shifts such that we can form Y. The curious part of this theorem is that there is no condition stating that X and Y need to be the same size, or even resemble one another at all.

One implication of this theorem is that, if we had two spheres, one having a radius of 1 and the other having a radius of 2, we could partition each into a finite number n pieces and have each piece in the sphere of radius 1 be congruent to a piece in the sphere of radius 2. While this theorem may offend our common sense, the steps used to prove it should not offend our understanding of mathematical rules.

I will begin by defining some terms to be used throughout the remainder of this paper.

Definition 1.2. For $x = (x_1, x_2, x_3)$ in $\mathbb{R}^3$, the **norm** of x is the number $|x| = (x_1^2 + x_2^2 + x_3^2)^{1/2}$.

Definition 1.3. The set $\{x \in \mathbb{R}^3 : |x - a| \leq r\}$ is the **closed ball** of radius $r > 0$ centered at $a \in \mathbb{R}^3$.

Definition 1.4. An **orthogonal matrix** is a square matrix with real entries whose transpose is also its inverse. That is, a square matrix is orthogonal when the product of it and its transpose is the identity matrix.

Definition 1.5. We will use the usual definition of an **identity matrix**. That is, the identity matrix will be denoted by ι and will be the matrix:

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

Definition 1.6. A subset $X \in \mathbb{R}^3$ is **bounded** if it is contained in a closed ball.

Definition 1.7. A subset $X \in \mathbb{R}^3$ has a **nonvoid interior** if it contains a closed ball.

Definition 1.8. A **rotation** is a 3x3 orthogonal matrix ρ whose determinant equals one. Notice that this ρ is also a mapping of $\mathbb{R}^3$ onto $\mathbb{R}^3$. We write $\rho(x)$ to represent the vector obtained by multiplying ρ by the column vector x:

$$\rho(x) = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

It may at first seem odd to define a rotation in this way. However, given that the usual definition of a rotation is a transformation that preserves distances while moving all the points in a set about some axis of rotation, we will see that our definition of a rotation actually satisfies both of these characteristics. In Theorem 2.1 we will prove both that ρ preserves distances and that every possible ρ has an axis in $\mathbb{R}^3$ which passes through the origin.

Another interesting property of rotations is that they form a group. Recall that a set of matrices is considered a group if the following conditions are met:

1. Closure: Given any two elements in the set, X and Y, their product XY also falls in the set.
2. Identity: The identity element I is contained in the set.
3. Inverse: Every matrix in the set is invertible and its inverse lies in the set. That is, for all X in the set, X^{-1} is in the set and $X \cdot X^{-1} = I$.

Theorem 1.9. *The set of rotations ρ forms a group. That is, the set of 3x3 orthogonal matrices whose determinant equals one satisfies the conditions of closure, contains an identity element, and contains an inverse that corresponds to each element.*

Before we begin the proof, I will state three facts from linear algebra to be used in the proof. Let α and β be two matrices.

(1) $\det(\alpha \cdot \beta) = \det(\alpha) \cdot \det(\beta)$

(2) If α and β are invertible, then $\alpha \cdot \beta$ is invertible and $(\alpha \cdot \beta)^{-1} = \beta^{-1} \cdot \alpha^{-1}$

(3) $(\alpha \cdot \beta)^T = \beta^T \cdot \alpha^T$

Having stated these facts, I will now begin the proof of Theorem 1.9.

Proof. We need to prove that rotations satisfy each condition of being a group.

Closure: We need to prove that the product of any two rotations is still a rotation. Let α and β be rotations. Using fact (1) from above and the fact that the determinant of any rotation is 1, we can say:

$$\det(\alpha \cdot \beta) = \det(\alpha) \cdot \det(\beta) = 1 \cdot 1 = 1$$

Thus, the determinant of the product of two rotations is 1. Then, using facts (2) and (3) from above and the fact that the inverse of a rotation is the same as the transpose of that rotation, we can say:

$$(\alpha \cdot \beta)^{-1} = \beta^{-1} \cdot \alpha^{-1} = \beta^T \cdot \alpha^T = (\alpha \cdot \beta)^T$$

Thus, the inverse of a product of two rotations is the same as the transpose of the product of those two rotations. This proves that the product of any two rotations is still a rotation and that the set of rotations is closed.

Identity: We need to prove that the identity matrix is a rotation.

$$\det(\iota) = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1(1+0) + 0(0+0) + 0(0+0) = 1$$

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

This proves that the identity matrix satisfies the conditions of a rotation.

Inverse: We need to prove that the inverse of every rotation is still a rotation. By our definition of a rotation, we know that $\rho^T = \rho^{-1}$. We need to prove, given that ρ is a rotation, $\det(\rho^{-1}) = 1$ and $(\rho^{-1})^{-1} = (\rho^{-1})^T$. It's clear that, using the fact that $\rho^{-1} = \rho^T$, $\det(\rho^{-1}) = 1$. Let α be any rotation:

$$\det(\alpha) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) = 1$$

$$\det(\alpha^{-1}) = \det(\alpha^T) = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22})$$

$$\det(\alpha) = \det(\alpha^{-1}) = 1.$$

We know that $\rho^{-1} = \rho^T$. By taking the transpose of both sides, we see that this implies that $(\rho^{-1})^T = (\rho^T)^T$. When we take the transpose of a matrix twice we get back the original matrix, so this is the same as saying $(\rho^{-1})^T = \rho$. Finally, we notice that $(\rho^{-1})^{-1} = \rho$ which implies $(\rho^{-1})^{-1} = (\rho^{-1})^T$. □

Definition 1.10. A **rigid motion** (also called a **Euclidean transformation**) is a mapping r of $\mathbb{R}^3$ onto $\mathbb{R}^3$ having the form $r(x) = \rho(x) + a$ for $x \in \mathbb{R}^3$ where ρ is a fixed rotation and $a \in \mathbb{R}^3$ is fixed. That is, a rigid motion is any transformation including a rotation and a shift.

Like rotations, the set of rigid motions forms a group. This fact will be useful later on in the paper so I will prove this property now.

Theorem 1.11. *The set of rigid motions $r(x) = \rho(x) + a$ where ρ is a rotation and a is a point in $\mathbb{R}^3$ forms a group.*

Proof. As with rotations, we need to prove that the set of rigid motions is closed, contains the identity element, and contains an inverse for every element. We will begin with closure.

Closure: Let $r(x) = \rho(x) + a$ and $s(x) = \tau(x) + b$ be two rigid motions where ρ and τ are rotations and a and b are fixed points in $\mathbb{R}^3$. We now need to prove that $(r \circ s)(x)$ is a rigid motion:

$$(r \circ s)(x) = r(s(x)) = \rho(\tau(x) + b) + a = \rho(\tau(x)) + \rho(b) + a$$

Knowing that the group of rotations is closed, it is clear that $\rho(\tau(x))$ is a rotation. Then, notice that the rotation of a fixed point results in a fixed point, so that $\rho(b)$ remains some fixed point in $\mathbb{R}^3$. Thus, $(r \circ s)(x)$ is still comprised of a rotation and a fixed point in $\mathbb{R}^3$ and so is still a rigid motion. Therefore, the set of rigid motions satisfies the conditions for closure.

Identity: Let us again use $r(x) = \rho(x) + a$ as our standard form of a rigid motion. We can easily construct the identity element by setting $\rho = \iota$ and $a = 0$. Then, given any matrix x, when we apply this $r(x)$ function we get:

$$r(x) = \rho(x) + a = \iota(x) + 0 = x$$

Inverse: Let $r(x) = \rho(x) + a$ be a rigid motion where ρ is a rotation and a is a fixed point in $\mathbb{R}^3$. The inverse of $r(x)$ is $r^{-1}(x) = \rho^{-1}(x) - \rho^{-1}(a)$ since:

$$(r \circ r^{-1})(x) = r(r^{-1}(x)) = r(\rho^{-1}(x) - \rho^{-1}(a)) = \rho(\rho^{-1}(x) - \rho^{-1}(a)) + a = x - a + a = x$$

Now we need to prove that $r^{-1}(x) = \rho^{-1}(x) - \rho^{-1}(a)$ is satisfies the conditions of being a rigid motion. Since we saw in the proof of Theorem 1.9 that the inverse of a rotation is still a rotation, we know that ρ^{-1} is a rotation. Thus, using the same logic as we used in proving closure, we can see that $\rho^{-1}(x)$ is a rotation and $\rho^{-1}(a)$ is a fixed point in $\mathbb{R}^3$. Thus, $r^{-1}(x)$ satisfies the conditions of being a rigid motion, proving that the inverse of a rigid motion is still a rigid motion.

Thus, rigid motions satisfy all three characteristics necessary to be a group. □

Definition 1.12. Two subsets X and Y are said to be **congruent**, written $X \cong Y$, if there exists some rigid motion r for which $r(X) = Y$. That is, X is congruent to Y if there is a way to transform X into Y using only rotations and shifts.

Definition 1.13. A **partition** of a set of X is a family of sets whose union is X and any two members of which are disjoint. That is, $\{X_j : 1 \leq j \leq n\}$ is a partition of X if both of the following are true:

1. $X = X_1 \cup X_2 \cup \cdots \cup X_n$
2. $X_i \cap X_j = \emptyset$ if $i \neq j$

Note that some or all of the sets X_j may be empty.

2 Rotations

Now that I have established some basic definitions to be used throughout this paper, I will prove some characteristics of rotations. Rotations will be an important component of proving the Banach-Tarski Paradox.

Theorem 2.1. *All rotations ρ have the following properties:*

(1) The image of ρ of any line is still a line. That is, $\rho(b + tc) = \rho(b) + t\rho(c)$ for all $b, c \in \mathbb{R}^3$ and $t \in \mathbb{R}$.

(2) Inner products are preserved by ρ. That is, if $x, x' \in \mathbb{R}^3, \rho(x) = y$, and $\rho(x') = y'$, then

$$\sum_{i=1}^{3} y_i y_i' = \sum_{j=1}^{3} x_j x_j'$$

(3) Distances are preserved by ρ. That is, for any $x \in \mathbb{R}^3, |\rho(x)| = |x|$.

(4) If $\rho \neq \iota$, then the set $A = \{x \in \mathbb{R}^3 : \rho(x) = x\}$ is a line through the origin. That is, there is a p in $\mathbb{R}^3$ such that $A = \{tp : t \in \mathbb{R}\}$ and $|p| = 1$. We call A the axis of ρ.

(5) If q is any point in $\mathbb{R}^3$ that has the characteristics of p as described in (4), then $q = p$ or $q = -p$. We call p and $-p$ the poles of ρ.

It is this theorem that shows that our definition of a rotation aligns with the traditional definition of a rotation. Specifically, we prove first in (1) that rotations preserve the shape of lines, then strengthen this in (2) to prove that inner products are preserved by rotations, and then further strengthen this in (3) to prove that distances are preserved by rotations. In (4) we prove that each rotation has an associated axis of rotation and then in (5) we prove that this axis is in fact unique–that each rotation has in fact only one associated axis of rotation.

Proof. (1) This claim follows almost instantly from the rules of linear algebra. Since applying ρ to a line is nothing more than using matrix multiplication, the math follows the same rules. We begin by applying ρ to the standard form of a line, represented by $b + tc$ where b, c are vectors in $\mathbb{R}^3$ and T is a scalar in $\mathbb{R}$:

$$\rho(b + tc)$$

Because matrix multiplication is distributive, we can distribute ρ across the parentheses:

$$\rho(b) + \rho(tc)$$

Lastly, we can pull t out of the parentheses because it is a scalar:

$$\rho(b) + t\rho(c)$$

(2) Let $x, x' \in \mathbb{R}^3$, $\rho(x) = y$, and $\rho(x') = y'$. Since ρ is orthogonal, we know that $\rho^T = \rho^{-1}$ and thus that the product of ρ and its transpose is equal to the identity matrix. We also know that the columns of ρ are the same as the rows of ρ^T.

Using this fact, notice that $\sum_{i=1}^{3} \rho_{ij}\rho_{ik} = \iota_{jk}$. That is, summing the product of the j^{th} column of ρ and the k^{th} column of ρ (also the k^{th} row of ρ^T) across all three rows of ρ gives us the entry in the identity matrix that occurs at the intersection of the j^{th} row and the k^{th} column. Because the identity matrix has entries of 1 at all entries that occur at the same numbered row as column and entries of 0 everywhere else, it is clear that $\iota_{jk} = 1$ whenever $j = k$ and $\iota_{jk} = 0$ whenever $j \neq k$.

We use this observation to complete the following calculation:

$$\sum_{i=1}^{3} y_i y_i' = \sum_{i=1}^{3} \left(\sum_{j=1}^{3} \rho_{ij} x_j \right) \left(\sum_{k=1}^{3} \rho_{ik} x_k' \right) = \sum_{i=1}^{3} \left(\sum_{j=1}^{3} \left(\sum_{k=1}^{3} \rho_{ij}\rho_{ik} x_j x_k' \right) \right)$$

$$= \sum_{j=1}^{3} \left(\sum_{k=1}^{3} \iota_{jk} x_j x_k' \right) = \sum_{j=1}^{3} x_j x_j'$$

This proves that inner products are preserved by ρ.

(3) We know that the square root of a vector's inner product with itself gives us that vector's magnitude. Thus, if inner products are preserved by ρ, then certainly distances are also preserved by ρ. We can show the complete proof by simply letting $x = x'$ in the calculation used in the proof of (2):

$$|y|^2 = \sum_{i=1}^{3} y_i y_i = \sum_{j=1}^{3} x_j x_j = |x|^2$$

This proves that $|y|^2 = |x|^2$, which is equivalent to saying that $|y| = |x|$ since both $|y|$ and $|x|$ are greater than zero. This in turn is the same as saying that $|\rho(x)| = |x|$, proving that ρ preserves distances.

(4) We will use the following definitions and facts from linear algebra in this portion of the proof. Let A be a k x k square matrix. Then:

(a) An eigenvalue of A is a scalar λ such that there is a nonzero vector x with $Ax = \lambda x$.

(b) λ is an eigenvalue of A if and only if $(A - \lambda\iota)x = 0$ has a nonzero solution.

(c) $\det(A - \lambda\iota) = 0$ is called the characteristic equation of A.

We will begin by fixing ρ to be any rotation. Assume that $\rho \neq \iota$ since otherwise every point is fixed and there is nothing to prove. The characteristic polynomial of ρ, written $f(\lambda)$, is $\det(\rho - \lambda\iota)$. We will begin this proof by showing that this matrix has an eigenvalue λ such that $\lambda = 1$ because this is the same as saying that $\rho(x) = 1 \cdot x$. Since we are looking for some vector $x \neq 0$ such that $\rho(x) = x$, the eigenvector that corresponds with this eigenvalue will be the x that we want.

Because ρ is a 3 x 3 matrix, we know that $f(\lambda)$ is a cubic polynomial having real coefficients. Because cubic functions approach ∞ and $-\infty$ at opposite ends of their range and are continuous, the intermediate value theorem guarantees that the cubic function must be equal to zero for at least one point in its range and therefore must have at least one real root.

Let λ_1, λ_2, and λ_3 be the three roots of the characteristic polynomial. Some of these roots could be complex and they possibly are not unique. Let λ_1 be its largest real root. Then the factored form of this polynomial is:

$$f(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda)$$

We can see, by letting $\lambda = 0$, that:

$$(\lambda_1 - 0)(\lambda_2 - 0)(\lambda_3 - 0) = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 = f(0) = \det(\rho - (0 \cdot \iota)) = \det(\rho) = 1$$

Recall that we know that $\det(\rho) = 1$ by definition of a rotation.

If λ_k is a real root, then the equation $(\rho - \lambda_k\iota)x = 0$ has a real solution $x \neq 0$. We know that this is true because we know that for an eigenvalue λ of ρ, there exists a corresponding eigenvector x such that $(\rho - \lambda\iota)x = 0$ and $x \neq 0$. This solution x satisfies $\rho(x) = \lambda_k x$. Then, using the fact that distances are preserved by ρ, as proved in part (3) of this theorem, we can observe the following:

$$|x| = |\rho(x)| = |\lambda_k x| = |\lambda_k||x|$$

In order for it to be true that $|x| = |\lambda_k||x|$, it must be the case that $|\lambda_k| = 1$ since $|x| \neq 0$. Then either $\lambda_k = 1$ or $\lambda_k = -1$. In particular, since we chose λ_1 to be the largest real root, we know that $\lambda_1 = 1$ or $\lambda_1 = -1$. We will now split the proof into two cases, considering separately the case where λ_2 and λ_3 are real and the case where λ_2 and λ_3 are complex.

Case 1: If λ_2 is not real, then λ_3 is its complex conjugate and:

$$\lambda_1\lambda_2\lambda_3 = f(0) = \det(\rho) = 1$$

$$\lambda_1|\lambda_2|^2 = 1$$

$$\lambda_1 = 1$$

Notice that we conclude that $\lambda_1 = 1$ because $|\lambda_2|^2 > 0$ and so λ_1 cannot be -1. This proves that we have an eigenvalue equal to one in this case of the proof.

Case 2: If λ_2 is real, then so is λ_3 and so all three roots are either -1 or 1 because those are the only real numbers for which $|\lambda_k| = 1$. Since we have already defined λ_1 to be our largest real root, we can say that $\lambda_1 = 1$ and $\lambda_2 = \lambda_3$ because:

$$\lambda_1\lambda_2\lambda_3 = f(0) = \det(\rho) = 1$$

This proves that we have an eigenvalue equal to one in this case of the proof as well and, having covered both the case in which there are complex eigenvalues and the case in which there are not complex eigenvalues, we can continue with the remainder of the proof.

Since $\lambda_1 = 1$ we can take $k = 1$ in the system of equations to find a vector $x \in \mathbb{R}^3$ such that $\rho(x) = x$ and $x \neq 0$. To construct a vector p with $|p| = 1$ such that $\rho(p) = p$, we can simply let $p = \frac{1}{|x|}x$. Then, $|p| = 1$ and p is still an eigenvector for $\lambda = 1$ since it is simply a scalar multiple of x:

$$\rho(p) = \rho\left(\frac{1}{|x|}x\right) = \frac{1}{|x|}\rho(x) = \frac{1}{|x|}\lambda x = \lambda\frac{1}{|x|}x = \lambda p = p$$

This shows us that there is some vector p with a magnitude of 1 that remains unchanged by ρ.

Recall that we defined the set A to be $A = \{x \in \mathbb{R}^3 : \rho(x) = x\}$. Thus $p \in A$. Further, if we let t be some scalar, it is also true that $tp \in A$ for all $t \in \mathbb{R}$ since:

$$\rho(tp) = t\rho(p) = tp$$

Now, we need only to find that there are no other vectors in A. Assume there exists some $u \in A$ such that $u \neq tp$ for all $t \in \mathbb{R}$ and let v be a nonzero vector that is perpendicular to the plane containing p, u, and 0. Because v is perpendicular with both p and u, the following dot products are zero:

$$\sum_{j=1}^{3} v_j p_j = \sum_{j=1}^{3} v_j u_j = 0$$

We know that $\rho(p) = p$ and $\rho(u) = u$ because $p, u \in A$ and all points in A are unchanged by ρ. Since $\rho(p) = p$ and $\rho(u) = u$, part (2) of this theorem, which states that ρ preserves dot products, implies that $\rho(v)$ is also perpendicular to this plane.

Thus, as illustrated in the following equations, $\rho(v)$ is orthogonal to u and $\rho(v)$ is orthogonal to p:

$$0 = \rho(v) \cdot \rho(u) = \rho(v) \cdot u \text{ and } 0 = \rho(v) \cdot \rho(p) = \rho(v) \cdot p$$

Then, from part (3) of this theorem, we know that $|\rho(v)| = |v|$. Since we are working in $\mathbb{R}^3$ and the plane formed by p and u is has 2 dimensions, there's only one possible line through the origin that is orthogonal to this plane. We know that both v and $\rho(v)$ fall on this line since they are orthogonal to the plane formed by p and u. Because of this, we conclude that $\rho(v) = v$ or $\rho(v) = -v$.

We know that any three linearly independent vectors in $\mathbb{R}^3$ form a basis, and so we can form a basis using v, p, and u. Thus, any vector $x \in \mathbb{R}^3$ can be written as $x = \alpha p + \beta u + \gamma v$ for appropriate scalars $\alpha, \beta, \gamma \in \mathbb{R}$.

If $x = \alpha p + \beta u + \gamma v$, then by part (1) of this theorem, we know that $\rho(x) = \alpha p + \beta u + \gamma \rho(v)$. Since $\rho \neq \iota$, we cannot have $\rho(v) = v$ and therefore $\rho(v) = -v$.

Now we construct a 3x3 matrix σ using the vectors p, u, and v and will then consider some properties of this matrix as we complete the proof:

$$\sigma = \begin{pmatrix} p_1 & u_1 & v_1 \\ p_2 & u_2 & v_2 \\ p_3 & u_3 & v_3 \end{pmatrix}$$

Because p, u, and v are linearly independent, σ is an invertible matrix and therefore has a nonzero determinant. Now consider the matrix product $\rho\sigma$, recalling that ρ leaves p and u unchanged and that $\rho(v) = -v$:

$$\rho\sigma = \begin{pmatrix} p_1 & u_1 & -v_1 \\ p_2 & u_2 & -v_2 \\ p_3 & u_3 & -v_3 \end{pmatrix}$$

Now, because $\rho\sigma$ only differs from σ in that one column is multiplied by -1, $-\det(\sigma) = \det(\rho\sigma)$. Knowing this, we can complete the following calculation to reach a contradiction:

$$-\det(\sigma) = \det(\rho\sigma) = \det(\rho)\det(\sigma) = 1 \cdot \det(\sigma) = \det(\sigma)$$

The only case in which $-\det(\sigma) = \det(\sigma)$ is if $\det(\sigma) = 0$, which contradicts our assumption that p, u, and v are linearly independent. This proves that there cannot exist any such $u \neq tp$ and thus all elements in A can be represented by the form tp where $t \in \mathbb{R}$.

(5) We can demonstrate by a quick calculation that if some element q has the properties of p as described in (4), then either $q = p$ or $q = -p$. Recall that the properties ascribed to p are that if $\rho \neq \iota$, then there is a $p \in \mathbb{R}^3$ such that the set $A = \{x \in \mathbb{R}^3 : \rho(x) = x\}$ can be represented as $A = \{tp : t \in \mathbb{R}\}$ and $|p| = 1$.

Notice that if $\{tp : t \in \mathbb{R}\} = \{tq : t \in \mathbb{R}\}$ and $|q| = |p| = 1$, then $q = tp$ for some t because in order for the sets to be equal q and p must both be elements of both sets. We can then see that:

$$t^2 = t^2|p|^2 = |tp|^2 = |q|^2 = 1$$

Thus, $t = 1$ or $t = -1$, proving that $q = p$ or $q = -p$.

□

3 The Group G

Having proven several properties about rotations in general, we will now consider two specific rotations:

$$\psi = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \phi = \begin{pmatrix} -\cos\theta & 0 & \sin\theta \\ 0 & -1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$$

In the second matrix, θ is a fixed real number that we can choose later. For now, I'll show that both of the above matrices are in fact rotations.

Proposition 3.1. *The previously defined matrices, ψ and ϕ, are rotations.*

Proof. We need to show that these are 3 x 3 invertible matrices ρ such that $\rho^T = \rho^{-1}$ and $\det(\rho) = 1$.

$$\psi \cdot \psi^T = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{-1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{-\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \psi^T = \psi^{-1}$$

$$\det(\psi) = 0 \cdot \det \begin{pmatrix} \frac{-\sqrt{3}}{2} & 0 \\ \frac{-1}{2} & 0 \end{pmatrix} - 0 \cdot \det \begin{pmatrix} \frac{-1}{2} & 0 \\ \frac{\sqrt{3}}{2} & 0 \end{pmatrix} + 1 \cdot \det \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} \end{pmatrix}$$

$$= 0 - 0 + 1\left(\frac{1}{4} - \frac{-3}{4}\right) = 1$$

$$\phi \cdot \phi^T = \begin{pmatrix} -\cos\theta & 0 & \sin\theta \\ 0 & -1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} -\cos\theta & 0 & \sin\theta \\ 0 & -1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \phi^T = \phi^{-1}$$

$$\det(\phi) = -0 \cdot \det\begin{pmatrix} 0 & \sin\theta \\ 0 & \cos\theta \end{pmatrix} - 1 \cdot \det\begin{pmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} - 0 \cdot \det\begin{pmatrix} -\cos\theta & 0 \\ \sin\theta & 0 \end{pmatrix}$$
$$= -1(-\cos^2\theta - \sin^2\theta) = 1$$

This proves that ψ and ϕ are both rotations. □

Having proved this, the next step is to show that $\psi^3 = \phi^2 = \iota$, a fact that we will use further in our next proof.

Proposition 3.2. *Using our definitions for ψ and ϕ, it is true that $\psi^3 = \phi^2 = \iota$.*

Proof. Since $\phi = \phi^T = \phi^{-1}$, we have already shown that $\phi^2 = \iota$, so it will suffice to show here that $\psi^3 = \iota$:

$$\psi \cdot \psi \cdot \psi = \begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{-1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{-\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{-1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{-1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

□

Now, let G be the set of matrices that can be obtained as a finite product of the matrices ψ and ϕ. We'll now prove that G is a group under matrix multiplication. Again, in order to prove that this is a group we need to prove that the group is closed, contains the identity matrix, and contains an inverse for every element in G.

Theorem 3.3. *The set G, defined to be all matrices that can be obtained as a finite product of the matrices ψ and ϕ, is a group.*

Proof. We need to prove that the set G satisfies each condition of being a group.

Closure: This property is immediate by the definition of G. Any product created by some matrix multiplication of ψ and ϕ is, by definition, in G.

Identity: We have already seen that ι is in G when we saw that $\psi^3 = \phi^2 = \iota$.

Inverse: From group theory, we know that if the generative elements of a group have inverses, then all elements of that group have inverses. Thus, we need only to know that ϕ, ψ, and ψ^2 have inverses and we will have proved that there exists an inverse for every element in G. We have already seen that ϕ is its own inverse and that $\psi^3 = \iota$. Because matrix multiplication is associative, we can say the following:

$$\iota = \psi \cdot \psi \cdot \psi = \psi^2 \cdot \psi = \psi \cdot \psi^2$$

This shows that ψ is the inverse of ψ^2 and that ψ^2 is the inverse of ψ. Thus, all the generative elements of G have an inverse and so every element in G has an inverse. Therefore, G is a group. □

Because G is a group, every element ρ in G such that $\rho \neq \iota$, can be written in at least one way as a product of ϕ, ψ, and ψ^2. Thus, every element of G other than ι, ϕ, ψ, and ψ^2 can be expressed in at least one of the following ways:

$$\alpha = \psi^{P_1}\phi\psi^{P_2}\phi...\psi^{P_m}\phi$$

$$\beta = \phi\psi^{P_1}\phi\psi^{P_2}...\phi\psi^{P_m}$$

$$\gamma = \phi\psi^{P_1}\phi\psi^{P_2}...\phi\psi^{P_m}\phi$$

$$\delta = \psi^{P_1}\phi\psi^{P_2}\phi...\phi\psi^{P_m}$$

We call the generative elements, ι, ϕ, ψ and ψ^2, as well as the expressions α, β, γ, and δ reduced words. In the expressions above, $m \geq 1$ and each exponent P_j is either 1 or 2 (since $P_j > 2$ can be reduced). In the case of δ, $m > 1$ or else it would be the same as α. Now, I'll take a moment to remind us of a definition before stating a theorem about the elements of G.

Definition 3.4. A **transcendental number** is a real number that is not the solution of any single-variable polynomial equation whose coefficients are all rational numbers.

Theorem 3.5. *If cos θ is a transcendental number, then each element of G other than ι has exacty one expression as a reduced word in the letters ϕ, ψ, and ψ^2.*

Proof. If no reduced word equals the identity, then no reduced word equals any other reduced word. Thus, we can complete this proof by proving only that no reduced word equals the identity. We have to do this for each case of reduced words.

Case 1: $\alpha = \psi^{P_1}\phi\psi^{P_2}\phi...\psi^{P_m}\phi$

Let $\sigma = \psi\phi$ or $\psi^2\phi$ such that $\alpha = \sigma_m\sigma_{m-1}...\sigma_2\sigma_1$. Thus:

$$\sigma = \psi\phi = \begin{pmatrix} \frac{1}{2}\cos\theta & \frac{\sqrt{3}}{2} & \frac{-1}{2}\sin\theta \\ \frac{-\sqrt{3}}{2}\cos\theta & \frac{1}{2} & \frac{\sqrt{3}}{2}\sin\theta \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \text{ or } \sigma = \psi^2\phi = \begin{pmatrix} \frac{1}{2}\cos\theta & \frac{-\sqrt{3}}{2} & \frac{-1}{2}\sin\theta \\ \frac{\sqrt{3}}{2}\cos\theta & \frac{1}{2} & \frac{-\sqrt{3}}{2}\sin\theta \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$$

It is clear right away that $\sigma \neq \iota$ because for either value of σ the a_{22} entry is $\frac{1}{2}$ whereas in ι this entry is 1. Thus, $\sigma \neq \iota$. Thus we consider the case where $\alpha = \sigma_m\sigma_{m-1}...\sigma_2\sigma_1$ and $m > 1$.

We want to prove that if $K = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, then $\sigma_m\sigma_{m-1}...\sigma_1(K) = \begin{pmatrix} \sin\theta P_{m-1}(\cos\theta) \\ \sqrt{3}\sin\theta Q_{m-1}(\cos\theta) \\ R_m(\cos\theta) \end{pmatrix}$ where P_{m-1}, Q_{m-1}, and R_m are the following recursively defined polynomials with rational coefficients and their subscripts denote their degrees:

$$P_0(x) = \frac{-1}{2},\ Q_0(x) = \pm\frac{1}{2},\ \text{and } R_1(x) = x$$

$$P_m(x) = \frac{1}{2}xP_{m-1}(x) \pm \frac{3}{2}Q_{m-1}(x) - \frac{1}{2}R_m(x)$$

$$Q_m(x) = \mp\frac{1}{2}xP_{m-1}(x) + \frac{1}{2}Q_{m-1}(x) \pm \frac{1}{2}R_m(x)$$

$$R_{m+1}(x) = (1-x^2)P_{m-1}(x) + xR_m(x)$$

Above, notice that the $\pm$ and $\mp$ signs indicate a dependence on whether $\sigma_m = \psi\phi$ or $\sigma_m = \psi^2\phi$.

The next step is to confirm that $\begin{pmatrix} \sin\theta P_{m-1}(\cos\theta) \\ \sqrt{3}\sin\theta Q_{m-1}(\cos\theta) \\ R_m(\cos\theta) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ and therefore that $\sigma_m\sigma_{m-1}...\sigma_1$ cannot be the identity no matter how many or which combination of σ matrices we are using.

Base Case: We begin by proving that the above is true when $m = 1$:

$$\sigma_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{-1}{2}\sin\theta \\ \pm\frac{1}{2}\sqrt{3}\sin\theta \\ \cos\theta \end{pmatrix} = \begin{pmatrix} \sin\theta P_0(\cos\theta) \\ \sqrt{3}\sin\theta Q_0(\cos\theta) \\ R_1(\cos\theta) \end{pmatrix}$$

This shows that the base case holds.

Induction Hypothesis: We assume that the following formula holds for m and want to prove that it also holds for $m+1$:

$$\sigma_m\sigma_{m-1}...\sigma_1(K) = \begin{pmatrix} \sin\theta P_{m-1}(\cos\theta) \\ \sqrt{3}\sin\theta Q_{m-1}(\cos\theta) \\ R_m(\cos\theta) \end{pmatrix}$$

$$\sigma_{m+1}\cdot\sigma_m\sigma_{m-1}...\sigma_1(K) = \sigma_{m+1}\cdot\begin{pmatrix} \sin\theta P_{m-1}(\cos\theta) \\ \sqrt{3}\sin\theta Q_{m-1}(\cos\theta) \\ R_m(\cos\theta) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2}\cos\theta & \frac{\pm\sqrt{3}}{2} & \frac{-1}{2}\sin\theta \\ \frac{\mp\sqrt{3}}{2}\cos\theta & \frac{1}{2} & \pm\frac{\sqrt{3}}{2}\sin\theta \\ \sin\theta & 0 & \cos\theta \end{pmatrix}\begin{pmatrix} \sin\theta P_{m-1}(\cos\theta) \\ \sqrt{3}\sin\theta Q_{m-1}(\cos\theta) \\ R_m(\cos\theta) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2}\cos\theta\sin\theta P_{m-1}(\cos\theta) \pm \sin\theta Q_{m-1}(\cos\theta) - \frac{1}{2}\sin\theta R_m(\cos\theta) \\ \mp\frac{\sqrt{3}}{2}\cos\theta\sin\theta P_{m-1}(\cos\theta) + \frac{\sqrt{3}}{2}\sin\theta Q_{m-1}(\cos\theta) \pm \frac{\sqrt{3}}{2}\sin\theta R_m(\cos\theta) \\ \sin^2\theta P_{m-1}(\cos\theta) + \cos\theta R_m(\cos\theta) \end{pmatrix}$$

$$= \begin{pmatrix} \sin\theta\left(\frac{1}{2}\cos\theta P_{m-1}(\cos\theta) \pm Q_{m-1}(\cos\theta) - \frac{1}{2}R_m(\cos\theta)\right) \\ \sqrt{3}\sin\theta\left(\mp\frac{1}{2}\cos\theta P_{m-1}(\cos\theta) + \frac{1}{2}Q_{m-1}(\cos\theta) \pm \frac{1}{2}R_m(\cos\theta)\right) \\ 1 - \cos^2\theta P_{m-1}(\cos\theta) + \cos\theta R_m(\cos\theta) \end{pmatrix}$$

$$= \begin{pmatrix} \sin\theta P_m(\cos\theta) \\ \sqrt{3}\sin\theta Q_m(\cos\theta) \\ R_{m+1}(\cos\theta) \end{pmatrix}$$

This proves through induction that $\sigma_m\sigma_{m-1}...\sigma_1(K) = \begin{pmatrix} \sin\theta P_{m-1}(\cos\theta) \\ \sqrt{3}\sin\theta Q_{m-1}(\cos\theta) \\ R_m(\cos\theta) \end{pmatrix}$ and all that is left is to show that there is no case in which $\begin{pmatrix} \sin\theta P_{m-1}(\cos\theta) \\ \sqrt{3}\sin\theta Q_{m-1}(\cos\theta) \\ R_m(\cos\theta) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. This is straightforward: since we set $\cos\theta$ to be a transcendental number, it cannot be the root of any polynomial with rational coefficients. Thus we will consider the following calculations to demonstrate that it is impossible for $R_m(\cos\theta)$ to equal 1:

$$R_m(\cos\theta) = 1$$

$$R_m(\cos\theta) - 1 = 0$$

If R_m is a polynomial with rational coefficients, then $R_m - 1$ is also a polynomial with rational coefficients. By the definition of a transcendental number, $\cos\theta$ cannot be a root of this equation and the above equations cannot be true. Thus, $\begin{pmatrix} \sin\theta P_{m-1}(\cos\theta) \\ \sqrt{3}\sin\theta Q_{m-1}(\cos\theta) \\ R_m(\cos\theta) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. This proves that there is no combination of σ values that can possibly equal ι.

We now need to consider the cases of β, γ, and δ. Notice that we could rewrite α as $\alpha = \phi\beta\phi$:

$$\phi\beta\phi = \phi\phi\psi^{P_1}\phi\psi^{P_2}...\phi\psi^{P_m}\phi = \phi^2\psi^{P_1}\phi\psi^{P_2}...\phi\psi^{P_m}\phi = \psi^{P_1}\phi\psi^{P_2}...\phi\psi^{P_m}\phi = \alpha$$

Then, assume that $\beta = \iota$ in the following short series of calculations and find a contradiction based on the fact that, as we just proved, $\alpha \neq \iota$:

$$\alpha = \phi\beta\phi = \phi\iota\phi = \phi^2 = \iota$$

.

Thus, $\beta \neq \iota$. We will now consider the case of δ. Assume that $\delta = \iota$ and that m is the smallest number for which this is true and we will seek a contradiction. Recall that we defined δ to be $\delta = \psi^{P_1}\phi\psi^{P_2}\phi...\phi\psi^{P_m}$ for $m > 1$. Thus we already know that $m > 1$. We will now consider two different cases of δ:

Case 1: Suppose that $P_1 = P_m$. Then either both are equal to 1 or both are equal to 2. Then $\psi^{P_1+P_m} = \psi^2$ or ψ^4. Now notice that $\psi^4 = \psi^3 \cdot \psi = \iota \cdot \psi = \psi$. Thus, $\psi^{P_1+P_m} = \psi^2$ or ψ. Now, using the assumption that $\delta = \iota$, consider the following equivalence:

$$\iota = \psi^{-P_1}\iota\psi^{P_1} = \psi^{-P_1}\delta\psi^{P_1} = \psi^{-P_1} \cdot \psi^{P_1}\phi\psi^{P_2}\phi...\phi\psi^{P_m} \cdot \psi^{P_1} = \phi\psi^{P_2}\phi...\phi\psi^{P_m+P_1} = \beta$$

This tells us that if $\delta = \iota$ then $\beta = \iota$, which is a contradiction. Therefore we have proven that $\delta \neq \iota$ in this case. We now consider the case where $P_1 \neq P_m$.

Case 2: Suppose that $P_1 \neq P_m$. Then one must equal 1 while the other equals 2. Therefore, $P_1 + P_m = 3$.

In the case where $m > 3$, we consider the following equivalence continuing to use the assumption that $\delta = \iota$:

$$\iota = \psi^{-P_1}\iota\psi^{P_1} = \phi\psi^{-P_1}\iota\psi^{P_1}\phi = \phi\psi^{-P_1}\delta\psi^{P_1}\phi$$

$$= \phi\psi^{-P_1} \cdot \psi^{P_1}\phi\psi^{P_2}\phi...\phi\psi^{P_m} \cdot \psi^{P_1}\phi = \phi\phi\psi^{P_2}\phi...\psi^{P_{m-1}}\phi\phi = \psi^{P_2}\phi...\phi\psi^{P_{m-1}}$$

Notice that in the above calculations we were able to add in (and later cancel out) ψ^{P_1}, ψ^{P_m}, and two ϕ's because ψ^{P_1} and ψ^{P_m} are ψ and ψ^2 and are therefore each others inverses and because ϕ is its own inverse. The final form that we constructed was again of the form δ and contradicted our definition of m which stated that m was the smallest value for which $\delta = \iota$. Thus we have reached a contradiction and we know that $m \not> 3$. Therefore, it only remains to rule out the cases in which $m = 2$ and $m = 3$.

When $m = 2$, we have:

$$\iota = \psi^{P_2}\iota\psi^{P_1} = \psi^{P_2}\delta\psi^{P_1} = \psi^{P_2}\psi^{P_1}\phi\psi^{P_2}\psi^{P_1} = \phi$$

We again used the fact that ψ^{P_1} and ψ^{P_2} must be inverses because we set $P_1 \neq P_m$ and in this case P_2 is P_m. Thus, one of these must be ψ while the other is ψ^2. Since we know that $\phi \neq \iota$, this is a contradiction and we know that $m \neq 2$.

When $m = 3$, we have:

$$\iota = \psi^{P_3}\iota\psi^{P_1} = \phi\psi^{P_3}\iota\psi^{P_1}\phi = \phi\psi^{P_3}\delta\psi^{P_1}\phi = \phi\psi^{P_3}\psi^{P_1}\phi\psi^{P_2}\phi\psi^{P_3}\psi^{P_1}\phi = \phi\phi\psi^{P_2}\phi\phi = \psi^{P_2}$$

We once again used the facts that ϕ is its own inverse and that ψ^{P_1} and ψ^{P_m} must be inverses where $m = 3$ in this case. We have arrived at a contradiction because it implies that either $\psi = \iota$ or $\psi^2 = \iota$ which we know to be untrue. Thus it is impossible that $\delta = \iota$.

Finally, the only case to consider is that of γ. However, by noticing that we can rewrite δ as $\phi\gamma\phi$ we can see easily that if $\gamma = \iota$ then:

$$\delta = \phi\gamma\phi = \phi\iota\phi = \phi^2 = \iota$$

This is again a contradiction because we just proved that $\delta \neq \iota$. Therefore, when cos θ is a transcendental number no reduced word in G is equal to ι and therefore no reduced word in G is equal to any other reduced word in G. This concludes the proof. □

In order to use the results of this proof, we choose and fix any θ such that cos θ is a transcendental number. We will now clarify some terminology before moving on to our next theorem.

If an element $\rho \in G$ is expressed as a reduced word written $\rho = \sigma_1\sigma_2...\sigma_n$ where σ_i is either ϕ, ψ, or ψ^2 (which we just proved to be a unique expression), we call n the length of ρ. We can express the length of a word by writing $\ell(\rho) = n$. For the identity, ι, we write $\ell(\iota) = 0$. We also say that σ_1 is the first letter of ρ or that ρ begins with σ_1. We will use the usual definition of a partition as stated below.

Definition 3.6. A **partition** of a set X is a pairwise disjoint family of subsets of X whose union is X.

We will now prove a theorem about a partition of G and the elements of G that lie within those partitions.

Theorem 3.7. *There exists a partition $\{G_1, G_2, G_3\}$ of G into three nonempty subsets such that for each ρ in G we have:*

(i) $\rho \in G_1 \iff \phi\rho \in G_2 \cup G_3$

(ii) $\rho \in G_1 \iff \psi\rho \in G_2$

(iii) $\rho \in G_1 \iff \psi^2\rho \in G_3$

In essence, this theorem states that there exists some three-part partition of G for which if we multiply a "word" by one "letter," it will result in a specific change in the partitioning subset that contains the element. Before we begin the proof, it is worth noting that if ρ begins with ϕ, then $\phi\rho$ starts with ψ since $\phi^2 = \iota$.

Proof. We begin by assigning our generative elements to appropriate groups. First, we somewhat arbitrarily set $\iota \in G_1$. By the desired rules of our partition, this forces $\phi \in G_2 \cup G_3$ and so we set $\phi \in G_2$. Similarly, ψ must be in G_2 and ψ^2 must be in G_3. Thus, we have already assigned all elements of lengths 0 and 1 to groups in the following way:

$$\iota \in G_1, \phi \in G_2, \psi \in G_2, \psi^2 \in G_3 \tag{1}$$

Having assigned all elements of length 0 and length 1, we now consider elements with some length $n+1$. Suppose that each element $\sigma \in G$ such that $\ell(\sigma) \leq n$ has already been assigned to one of G_1, G_2, or G_3. We now assign all elements with length $n+1$.

If $\ell(\sigma) = n$ and σ begins with ψ or ψ^2, assign the elements to groups as follows:

$$\phi\sigma \in G_2 \text{ if } \sigma \in G_1 \tag{2}$$

$$\phi\sigma \in G_1 \text{ if } \sigma \in G_2 \cup G_3 \tag{3}$$

If $\ell(\sigma) = n$ and σ begins with ϕ, let $j = 1, 2, 3$ and let $G_4 = G_1$ and $G_5 = G_2$. Assign the elements to groups as follows:

$$\psi\sigma \in G_{j+1} \text{ if } \sigma \in G_j \tag{4}$$

$$\psi^2\sigma \in G_{j+2} \text{ if } \sigma \in G_j \tag{5}$$

Since we have now assigned every element to one and only one bucket, our partition is now formed. The assignment of any element of length n can be calculated in n steps. Now we want to check inductively that the following rules, as stated in the theorem, hold:

(i) $\rho \in G_1 \iff \phi\rho \in G_2 \cup G_3$

(ii) $\rho \in G_1 \iff \psi\rho \in G_2$

(iii) $\rho \in G_1 \iff \psi^2\rho \in G_3$

Base Case: We need to show that the above three rules hold for all of our generative elements. The case of ι is fairly simple because we chose sets in which to place the generative elements so as to not violate our rules. I will show this below:

(i) $\iota \in G_1$ and $\phi \in G_2$ so the statement $\iota \in G_1 \iff \phi \in G_2 \cup G_3$ holds.

(ii) $\iota \in G_1$ and $\psi \in G_2$ so the statement $\iota \in G_1 \iff \psi \in G_2$ holds.

(iii) $\iota \in G_1$ and $\psi^2 \in G_3$ so the statement $\iota \in G_1 \iff \psi^2 \in G_3$ holds.

Next I will show this for ϕ. Notice that since $\phi \notin G_1$, we will show the negation of each side of the equivalence:

(i) $\phi \in G_2$ and therefore $\phi \notin G_1$. $\phi\phi = \iota$ and since $\iota \in G_1$, we can say that $\phi\phi \notin G_2 \cup G_3$. Thus, the statement $\phi \notin G_1 \iff \phi\phi \notin G_2 \cup G_3$ holds.

(ii) $\phi \in G_2$ and therefore $\phi \notin G_1$. Using rule (4), we know that since $\phi \in G_2$, then $\psi\phi \in G_3$ and so $\psi\phi \notin G_2$. Thus, the statement $\phi \notin G_1 \iff \psi\phi \notin G_2$ is true.

(iii) $\phi \in G_2$ and therefore $\phi \notin G_1$. Using rule (5), we know that since $\phi \in G_2$, then $\psi^2\phi \in G_1$ and so $\psi^2\phi \notin G_3$. Therefore the statement $\phi \notin G_1 \iff \psi^2\phi \notin G_3$ holds.

Having shown that (i), (ii), and (iii) hold for ϕ we will now consider ψ. Again, since $\psi \notin G_1$ we will show the negation of each side of the equivalence:

(i) $\psi \in G_2$ and therefore $\psi \notin G_1$. Using rule (3), we see that since $\psi \in G_2$, then $\phi\psi \in G_1$ and so $\phi\psi \notin G_2 \cup G_3$. Thus, the statement $\psi \notin G_1 \iff \phi\psi \notin G_2 \cup G_3$ holds.

(ii) $\psi \in G_2$ and therefore $\psi \notin G_1$. We chose that $\psi\psi = \psi^2 \in G_3$ and so $\psi^2 \notin G_2$. Thus, the statement $\psi \notin G_1 \iff \psi\psi \notin G_2$ holds.

(iii) $\psi \in G_2$ and therefore $\psi \notin G_1$. Since $\psi^2\psi = \psi^3 = \iota$ and $\iota \in G_1$, we know that $\psi^2\psi \notin G_3$. Thus, the statement $\psi \notin G_1 \iff \psi^2\psi \notin G_3$ holds.

Having shown that (i), (ii), and (iii) hold for ψ we will now consider ψ^2. Once again, since $\psi \notin G_1$ we will show the negation of each side of the equivalence:

(i) $\psi^2 \in G_3$ and therefore $\psi^2 \notin G_1$. Using rule (3), we know that $\phi\psi^2 \in G_1$ since $\psi^2 \in G_3$. This tells us that $\phi\psi^2 \notin G_2 \cup G_3$. Thus, the statement $\psi^2 \notin G_1 \iff \phi\psi^2 \notin G_2 \cup G_3$ holds.

(ii) $\psi^2 \in G_3$ and therefore $\psi^2 \notin G_1$. Since $\psi\psi^2 = \psi^3 = \iota$ and we chose that $\iota \in G_1$, we know that $\psi\psi^2 \notin G_2$. Thus, the statement $\psi^2 \notin G_1 \iff \psi\psi^2 \notin G_2$ holds.

(iii) $\psi^2 \in G_3$ so $\psi^2 \notin G_1$. Since $\psi^2\psi^2 = \psi$ and we chose that $\psi \in G_2$, we know that $\psi^2\psi^2 \notin G_3$. Therefore, the statement $\psi^2 \notin G_1 \iff \psi^2\psi^2 \notin G_3$ holds.

This demonstrates that each case of the base case holds and so we can proceed to the induction hypothesis.

Induction Hypothesis: Assume that $n > 1$ is some integer and that our three conditions hold for all $\rho \in G$ where $\ell(\rho) < n$. Now, fix $\rho \in G$ with $\ell(\rho) = n$. I'll now consider the different cases in which ρ is comprised of different combinations of elements. Note that we will continue to use (1), (2), (3), (4), and (5) as our rules for inductively assigning elements.

Case 1: Suppose that ρ begins with ϕ. Then, with $\sigma = \rho$, we consider (4):

$$\psi\sigma \in G_{j+1} \text{ if } \sigma \in G_j$$

The above statement tells us the following:

$$\text{If } \rho \in G_1, \text{ then } \psi\rho \in G_2$$

$$\text{If } \rho \in G_2, \text{ then } \psi\rho \in G_3$$

$$\text{If } \rho \in G_3, \text{ then } \psi\rho \in G_1$$

This shows us that the only case for which $\rho \in G_1$ is when $\psi\rho \in G_2$. Thus, this shows that (ii), $\rho \in G_1 \iff \psi\rho \in G_2$, holds. Now, with $\sigma = \rho$, we will consider (5):

$$\psi^2\sigma \in G_{j+2} \text{ if } \sigma \in G_j$$

The above statement tells us the following:

$$\text{If } \rho \in G_1, \text{ then } \psi^2\rho \in G_3$$

$$\text{If } \rho \in G_2, \text{ then } \psi^2\rho \in G_1$$

$$\text{If } \rho \in G_3, \text{ then } \psi^2\rho \in G_2$$

This shows us that the only case for which $\rho \in G_1$ is when $\psi^2\rho \in G_3$, showing that (iii), $\rho \in G_1 \iff \psi^2\rho \in G_3$, holds.

Since the first element of ρ is ϕ and $\phi^2 = \iota$, we know that $\phi\rho$ has length $n-1$. Thus, we can use our induction hypothesis, which stated that all elements of length less than n adhere to conditions (i), (ii), and (iii). In this case, we will use (i) and the fact that $\phi(\phi\rho) = \rho$ to make the following statement:

$$\phi\rho \in G_1 \iff \phi(\phi\rho) \in G_2 \cup G_3 \iff \rho \in G_2 \cup G_3$$

Using this piece of our induction hypothesis, we show the following equivalence:

$$\rho \notin G_1 \iff \rho \in G_2 \cup G_3 \iff \phi(\phi\rho) = \rho \in G_2 \cup G_3 \iff \phi\rho \in G_1 \iff \phi\rho \notin G_2 \cup G_3$$

This shows that $\rho \notin G_1 \iff \phi\rho \notin G_2 \cup G_3$, proving by the contrapositive that (i) holds for ρ and thus completing the first case.

Case 2: Suppose that ρ begins with ψ. Now we will use statements (2) and (3):

$$\phi\sigma \in G_2 \text{ if } \sigma \in G_1$$

$$\phi\sigma \in G_1 \text{ if } \sigma \in G_2 \cup G_3$$

Letting $\rho = \sigma$, the above statements tells us the following:

$$\text{If } \rho \in G_1 \text{ then } \phi\rho \in G_2$$

$$\text{If } \rho \in G_2 \cup G_3 \text{ then } \phi\rho \in G_1$$

This shows us that $\rho \in G_1 \iff \phi\rho \in G_2$, proving (i) in this case. Now let $\rho = \psi\sigma$ where σ begins with ϕ so that $\psi\rho = \psi^2\sigma$. This σ then has length $n-1$ since ρ has exactly one element more than σ and so we can use our induction hypothesis that all elements with length less than n adhere to (i), (ii), and (iii). We will now use equations (4) and (5):

$$\psi\sigma \in G_{j+1} \text{ if } \sigma \in G_j$$
$$\psi^2\sigma \in G_{j+2} \text{ if } \sigma \in G_j$$

Using these rules as well as the fact that we have set $\psi\rho = \psi^2\sigma$ we make the following observation:

$$\psi\rho = \psi^2\sigma \in G_2 \iff \sigma \in G_3 \iff \rho = \psi\sigma \in G_1 \iff \psi^2\rho = \sigma \in G_3$$

This shows that $\rho \in G_1 \iff \psi\rho \in G_2$ and $\rho \in G_1 \iff \psi^2\rho \in G_3$, proving (ii) and (iii) for this case of ρ.

Case 3: Suppose that ρ begins with ψ^2. Now we will use statements (2) and (3):

$$\phi\sigma \in G_2 \text{ if } \sigma \in G_1$$

$$\phi\sigma \in G_1 \text{ if } \sigma \in G_2 \cup G_3$$

Letting $\rho = \sigma$, the above statements tells us the following:

$$\text{If } \rho \in G_1 \text{ then } \phi\rho \in G_2$$

$$\text{If } \rho \in G_2 \cup G_3 \text{ then } \phi\rho \in G_1$$

This shows that $\rho \in G_1 \iff \phi\rho \in G_2 \cup G_3$, proving that (i) holds in this case. We now let $\psi\rho = \sigma$, noticing that $\psi\rho$ starts with ψ^3 and that $\psi^3 = \iota$ and thus σ starts with ϕ and has length $n-1$. Again, we use equations (4) and (5) to get:

$$\psi\rho = \sigma \in G_2 \iff \rho = \psi^2\sigma \in G_1 \iff \sigma \in G_2 \iff \psi^2\rho = \psi\sigma \in G_3$$

$$\rho \in G_1 \iff \psi\rho \in G_2 \text{ and } \rho \in G_1 \iff \psi^2\rho \in G_3$$

This shows that $\rho \in G_1 \iff \psi\rho \in G_2$ and $\rho \in G_1 \iff \psi^2\rho \in G_3$, proving (ii) and (iii) hold in this final case and completing the proof. □

4 The Unit Sphere

The following proof uses the relatively abstract concepts that we have proved so far and applies them to the unit sphere. This process chiefly involves relabeling and demonstrating how the rotations we have seen can be applied to more tangible sets. This more concrete representation will be useful later on.

Theorem 4.1. *There exists a partition $\{P, S_1, S_2, S_3\}$ of the unit sphere $S = \{x \in \mathbb{R}^3 : |x|^2 = x_1^2 + x_2^2 + x_3^2 = 1\}$ into four subsets such that:*

(i) P *is countable.*

(ii) $\phi(S_1) = S_2 \cup S_3$

(iii) $\psi(S_1) = S_2$

(iv) $\psi^2(S_1) = S_3$

Proof. Let $P = \{p \in S : \rho(p) = p \text{ for some } \rho \in G \text{ with } \rho \neq \iota\}$. This defines the set P to be all fixed points under rotation that is not the identity, that is all points that map to themselves under some rotation $\rho \in G$ where $\rho \neq \iota$.

Because G is defined to be the set of all matrices that can be obtained as a product of a finite number of matrix factors ϕ and ψ, G is countable. Part (4) of Theorem 2.1 tell us that for any rotation ρ, there exists some $p \in \mathbb{R}^3$ having length 1 such that p is unchanged by ρ. Part (5) of Theorem 2.1 tells us that only two points, p and $-p$, have the characteristics described in part (4). Notice that when we think about all elements of length 1, we are exactly describing the surface of the unit sphere.

Because we know that every element in G is a rotation, if $\rho_1 \in G$ is some rotation in G and $\rho_1 \neq \iota$, then for all elements $p \in S$, the set $P_1 = \{p \in S : \rho_1(p) = p\}$ contains only two elements. Then notice that P is the union of all such P_n, each of which corresponds to one unique $\rho_n \in G$. Since G is countable, we therefore know that P is also countable and (i) holds.

For each $x \in S \setminus P$, let $G(x) = \{\rho(x) : \rho \in G\}$. Notice that each such $G(x)$ is a subset of $S \setminus P$, since if $\rho(x)$ were in P then x would have had to have been in P by the definition of P.

Notice that $x \in G(x)$ since $\iota \in G$ and $x = \iota(x)$. We can also notice, by conducting the following calculation, that any two sets $G(x)$ and $G(y)$ are either disjoint or identical. That is, if some element t is in both $G(x)$ and $G(y)$, then every element in $G(x)$ is in $G(y)$ and vice versa. If some element t is in $G(x)$ but is not in $G(y)$, then no element in $G(x)$ is in $G(y)$ and vice versa. I will now show that this is true:

Suppose $t \in G(x) \cap G(y)$ for some $x, y \in S \setminus P$. Then $\rho(x) = t = \sigma(y)$ for some rotations $\rho, \sigma \in G$. Also suppose that z is some element in $G(x)$. Then, $z = \tau(x)$ for some rotation $\tau \in G$. Then, since $\rho(x) = t$, we can see that $x = \rho^{-1}(t)$ and thus that:

$$z = \tau(x) = \tau\rho^{-1}(t) = \tau\rho^{-1}\sigma(y)$$

We can see that $\tau\rho^{-1}\sigma(y) \in G(y)$ and thus that $z \in G(y)$. Since z was any arbitrary element in $G(x)$, this proves that every element in $G(x)$ is in $G(y)$. By repeating above and switching the roles of x and y, we can see that $G(y) \subset G(x)$ and thus that $G(x) = G(y)$.

Having proved that if any element is in $G(x) \cap G(y)$ then $G(x) = G(y)$, it follows that if some element in $G(x)$ is not in $G(y)$, then there cannot be any element in $G(x)$ that is in $G(y)$ and thus that $G(x)$ and $G(y)$ are disjoint.

This proves that the family of sets $\mathcal{F} = \{G(x) : x \in S \setminus P\}$ is a partition of $S \setminus P$. It remains only to prove that this partition is equivalent to one formed by the desired sets S_1, S_2, and S_3 as stated in the theorem. Begin this next portion of the proof by choosing exactly one point from each member of $\mathcal{F}$ and denote the set of points so chosen by C. The set C has the following properties:

(a) $C \subset S \setminus P$

(b) $c_1 \neq c_2$ in $C \Rightarrow G(c_1) \cap G(c_2) = \emptyset$

(c) $x \in S \setminus P \Rightarrow x \in G(c)$ for some $c \in C$

We know that (a) is true because the set C is comprised only of elements from $\mathcal{F}$ and every point in $\mathcal{F}$ is in $G(x)$ for some $x \in S \setminus P$, and as we've already stated each such $G(x)$ is a subset of $S \setminus P$.

Because we only chose one c from each disjoint set in $\mathcal{F}$, there are no two distinct points in C that fall in the same set $G(x)$, showing (b).

Because G is a group, if we have $s \in G(x)$ and $t \in G(x)$, then $s = \rho(x)$ and $t = \sigma(x)$ for some $\rho, \sigma \in G$. Thus, $\rho^{-1}(s) = x$, and $t = \sigma\rho^{-1}(s)$. This shows that we can represent any element in $G(x)$ as a function of any other element in that same set $G(x)$. This shows that (c) is true.

We now define $S_j = G_j(C) = \{\rho(c) : \rho \in G_j, c \in C\}$ for $j = 1, 2, 3$ where G_1, G_2, and G_3 are as in Theorem 3.7. Using the facts that $C \subset S \setminus P$ and $G(x) \subset S \setminus P$ if $x \in S \setminus P$, we see that $S_j \subset S \setminus P$ for each j.

Recall that $G = G_1 \cup G_2 \cup G_3$ and that, by (c), if $x \in S \setminus P$, then $x \in G(c)$ for some $c \in C$. Therefore, $S \setminus P = S_1 \cup S_2 \cup S_3$.

If $j \neq i$ for $j, i \in \{1, 2, 3\}$, then $S_j \cap S_i = \emptyset$. Otherwise, for $x \in S_j \cap S_i$, we have $x \in \rho(c_1) = \sigma(c_2)$ for some $c_1, c_2 \in C, \rho \in G_j$, and $\sigma \in G_i$. If this is true, then (b) tells us that $c_1 = c_2 = c$ and then $\sigma^{-1}\rho(c) = c$. Since we chose all c such that $c \notin P$, this means that $\sigma^{-1}\rho = \iota$ and thus that $\rho = \sigma$. But then, the element $\rho = \sigma$ was in both G_j and G_i, violating the fact that G_1, G_2, and G_3 partition G and thus that $G_j \cap G_i = \emptyset$ for $i \neq j$.

Thus, we have now determined that $\{S_1, S_2, S_3\}$ partitions $S \setminus P$ and therefore that $\{P, S_1, S_2, S_3\}$ partitions S. We now use (i)-(iii) of Theorem 3.7 to write the following:

$$\phi(S_1) = \{\phi\rho(c) : c \in G_1, c \in C\} = \{\tau(c) : \tau \in G_2 \cup G_3, c \in C\} = S_2 \cup S_3$$

$$\psi\{S_1\} = \{\psi\rho(c) : \rho \in G_1, c \in C\} = \{\tau(c) : \tau \in G_2, c \in C\} = S_2$$

$$\psi^2(S_1) = \{\psi^2\rho(c) : \rho \in G_1, c \in C\} = \{\tau \in G_3, c \in C\} = S_3$$

This shows that the remainder of this theorem holds and concludes the proof. □

Theorem 4.2. *If P is any countable subset of S, then there exists a countable set Q and a rotation ω such that $P \subset Q \subset S$ and $\omega(Q) = Q \setminus P$.*

In essence, this theorem tells us that, given any countable subset P of S, we can rotate that subset such that none of the resulting points fall in our original set P. We will rely heavily on the assumption of the countability of P throughout this proof.

Proof. Since P is countable, there can only be countably many vectors of the form $v = \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix}$ in S for which v or $-v$ is in P. Therefore, we have uncountably many vectors $v \in S$ of this form for which $v, -v \notin P$ and we select any such v.

Now let $u = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\sigma = \begin{pmatrix} v_1 & v_2 & 0 \\ -v_2 & v_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The following shows that σ satisfies the conditions of being a rotation (i.e. that its determinant is 1 and it is a 3 x 3 orthogonal matrix):

$$\det \sigma = 0(0 + 0) - 0(v_1 - 0) + 1(v_1^2 + v_2^2) = 1 \text{ (since } v \in S, |v| = v_1^2 + v_2^2 + 0^2 = 1)$$

$$\sigma\sigma^T = \begin{pmatrix} v_1 & v_2 & 0 \\ -v_2 & v_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 & -v_2 & 0 \\ v_2 & v_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

By a short calculation, we can see that $\sigma(v) = u$ and that $\sigma(-v) = -u$:

$$\sigma(v) = \begin{pmatrix} v_1 & v_2 & 0 \\ -v_2 & v_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1^2 + v_2^2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = u$$

$$\sigma(-v) = \begin{pmatrix} v_1 & v_2 & 0 \\ -v_2 & v_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -v_1 \\ -v_2 \\ 0 \end{pmatrix} = \begin{pmatrix} -v_1^2 - v_2^2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = -u$$

Since $\det(\sigma) \neq 0$, we know that σ is invertible and is therefore both one-to-one and onto. This means that there is only one element in S that maps to u and only one element that maps to $-u$. Thus, since $\sigma(v) = u$ and $\sigma(-v) = -u$ and $v, -v$ were chosen such that neither is in P, we know that there is no point p in P such that $\sigma(p) = u$ or $-u$. Thus, the set $\sigma(P)$ contains neither u nor $-u$.

Now, for real numbers t, consider the rotations:

$$\tau_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix}$$

We perform the following calculations to show that τ_t is indeed a rotation for all real t:

$$\det \tau_t = 1(\cos^2(t) + \sin^2(t)) - 0(0+0) + 0(0+0) = 1 \cdot 1 = 1$$

$$\tau_t \tau_t^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Now, notice that τ_t leaves u fixed for all t:

$$\tau_t(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

It is worth pausing here to consider geometrically what we are looking at. Recall that S is the surface of the unit sphere, or the set of all points having a distance of 1 from the origin. Then, u and $-u$ are the points where the unit sphere intersects the x-axis. We know that τ_t is a rotation that leaves u and $-u$ unchanged and since u and $-u$ are on the x-axis, we can conclude that τ_t is a rotation about the x-axis.

Let us fix t so that we can see that τ_t is a counterclockwise rotation. For simplicity, let $t = \frac{\pi}{2}$ and let us consider how this effects the vector $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$:

$$\tau_{\frac{\pi}{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ 0 & \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Thinking about where $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ fall on the unit sphere, we can see that $\tau_{\frac{\pi}{2}}$ results in a counterclockwise rotation of degree $\frac{\pi}{2}$ around the x-axis of the unit sphere. We can generalize this finding to say that τ_t is a counterclockwise rotation of degree t around the x-axis of the unit sphere.

Having considered the impact of τ_t on our set S, we will continue with this proof. The next step is to show that there are only countably many t such that:

$$\sigma(P) \cap \bigcup_{n=1}^{\infty} \tau_t^n \sigma(P) \neq \emptyset$$

We will begin by supposing that $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ are two vectors such that $x, y \in \sigma(P)$. It is worth noting first that, because $x, y \in \sigma(P)$, $x, y \neq u, -u$ because $u, -u \notin \sigma(P)$. This tells us that $-1 < x_1, y_1 < 1$ and $x_2, y_2 \neq 0$ or $x_3, y_3 \neq 0$ – otherwise x and y would not in fact be distinct from u and $-u$. Thus, $x_2^2 + x_3^2 > 0$ and $y_2^2 + y_3^2 > 0$.

Suppose first that $x_1 \neq y_1$ and notice that τ_t does not affect the first entry of any vector to which it is applied:

$$\tau_t \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 + 0 + 0 \\ 0 + a_2\cos t - a_3\sin t \\ 0 + a_2\sin t + a_3\cos t \end{pmatrix}$$

$$= \begin{pmatrix} a_1 \\ a_2\cos t - a_3\sin t \\ a_2\sin t + a_3\cos t \end{pmatrix}$$

Having noticed this, we can infer that when $x_1 \neq y_1$, there is no $t \in [0, 2\pi)$ such that $\tau_t^n(x) = y$ for any $n \geq 1$.

The slightly more complicated instance is when $x_1 = y_1$, so we will now suppose that this is the case. We will see that in this case, there are exactly n values of $t \in [0, 2\pi)$ such

that $\tau_t^n(x) = y$. First, it is important to consider further the τ function composed with itself:

$$\tau_t^n(a) = \overbrace{\tau_t(\tau_t(...(\tau_t}^{\text{n times}}(a))))$$

Written like this, it is easier to see that the process here is to apply the innermost rotation, and then the second-innermost rotation and so on, compounding the rotations as we go. It turns out that it is the same to apply the rotation τ_t n times as it is to apply the rotation τ_{nt} only one time. In summary:

$$\tau_t^n(a) = \tau_{nt}(a)$$

Now, suppose $\alpha \in [0, 2\pi)$ is the counterclockwise angle from x to y such that $\tau_\alpha(x) = y$. Then suppose $t \in [0, 2\pi)$ and that $\tau_t^n(x) = y$. Then we know that $\tau_{tn}(x) = y$. This further tells us that nt is some multiple of 2π more than α. That is, for some nonnegative integer m with $\alpha, t \in [0, 2\pi)$:

$$nt = \alpha + 2\pi m$$
$$t = \frac{\alpha}{n} + \frac{2\pi m}{n}$$

We first use the domain of α to establish more information about the value of m. As we go through these calculations, bear in mind that the end goal is to find a value of m that satisfies the condition of $0 \leq t < 2\pi$. We begin with the fact that we know α to be between 0 and 2π:

$$0 \leq \alpha < 2\pi$$
$$0 \leq \frac{\alpha}{n} < \frac{2\pi}{n}$$
$$0 \leq \frac{\alpha}{n} + \frac{2\pi m}{n} < \frac{2\pi}{n} + \frac{2\pi m}{n}$$

Notice here that the middle piece of this inequality is an expression equal to t that we saw above: $t = \frac{\alpha}{n} + \frac{2\pi m}{n}$. Thus, since we want a value of m for which $0 \leq t < 2\pi$, we want the righthand side of the equation to be less than or equal to 2π:

$$\frac{2\pi}{n} + \frac{2\pi m}{n} \leq 2\pi$$
$$\frac{1}{n} + \frac{m}{n} \leq 1$$
$$1 + m \leq n$$
$$m \leq n - 1$$

Since we defined m to be a nonnegative integer, this tells us that $m = 0, 1, ..., n-2, n-1$. Thus we have n different choices for the value of m, subsequently giving us n possible values

of t. We can obtain these values of t by plugging our values of m into the equation describing the relationship between m and t:

$$t = \frac{\alpha}{n} + \frac{2\pi m}{n}$$

$$t = \frac{\alpha}{n}, \frac{\alpha + 2\pi}{n}, \frac{\alpha + 4\pi}{n}, ..., \frac{\alpha + (n-1)(2\pi)}{n}$$

Thus, we have shown that for any positive integer n, there are exactly n values of t for which the following holds for two vectors $x, y \in \sigma(P)$:

$$\tau_t^n(x) = y$$

By nature of the fact that n is any positive integer, we have countably many choices for n. Then, for each n, we then had finitely many values of t. Thus, for every n, there are still only countably many values of t for which the above equation holds. That means that there are uncountably many values for which it fails, telling us that there are uncountably many t such that the following holds:

$$\sigma(P) \cap \bigcup_{n=1}^{\infty} \tau_t^n \sigma(P) = \emptyset \tag{6}$$

We now fix any $t \in \mathbb{R}$ for which the above equation holds and write $\tau = \tau_t$. We now define the following terms:

$$\omega = \sigma^{-1}\tau\sigma$$

$$Q = P \cup \bigcup_{n=1}^{\infty} \omega^n(P)$$

In the remainder of the proof, I will show that we have chosen ω and Q such that for our fixed countable set P we have $P \subset Q \subset S$ and $\omega(Q) = Q \setminus P$. It is first helpful to notice that we defined Q so that $P \subset Q$. Furthermore, we know that ω is a rotation since it is composed of the rotations σ^{-1}, σ, and τ. Because the set of rotations is a group, $\omega^n(P) \subset S$. Thus, $P \cup \bigcup_{n=1}^{\infty} \omega^n(P) \subset S$, which implies that $Q \subset S$. Thus, $P \subset Q \subset S$ and it remains only to show that $\omega(Q) = Q \setminus P$. We continue the proof by manipulating our definition of ω:

$$\omega = \sigma^{-1}\tau\sigma$$

$$\overbrace{\omega \cdot \omega \cdot ... \cdot \omega}^{\text{n times}} = \overbrace{\sigma^{-1}\tau\sigma \cdot \sigma^{-1}\tau\sigma \cdot ... \cdot \sigma^{-1}\tau\sigma}^{\text{n times}}$$

$$\omega^n = \sigma^{-1}\tau^n\sigma$$

$$\sigma\omega^n = \tau^n\sigma$$

Using the above equality, we can rewrite equation (6) to say:

$$\sigma(P) \cap \bigcup_{n=1}^{\infty} \sigma\omega^n(P) = \emptyset$$

Since σ is a rotation, it is distributive and we can pull it out to the front of the entire equality:

$$\sigma(P \cap \bigcup_{n=1}^{\infty} \omega^n(P)) = \emptyset$$

Since all rotations are one-to-one and onto, we can infer that the following must then also be true:

$$P \cap \bigcup_{n=1}^{\infty} \omega^n(P) = \emptyset \tag{7}$$

Now consider the following:

$$\omega(Q) = \omega(P \cup \bigcup_{n=1}^{\infty} \omega^n(P)) = \omega(P) \cup \omega(\bigcup_{n=1}^{\infty} \omega^n(P)) = \omega(P) \cup \bigcup_{n=1}^{\infty} \omega^{n+1}(P) = \bigcup_{n=1}^{\infty} \omega^n(P)$$

Since P is countable and Q is the union of P and a countable number of countable sets $\omega^n(P)$ for $n = 1, 2, ..., \infty$, Q is also countable. Additionally, it's clear that $P \subset Q$ since we specifically defined Q to include the set P. Finally, we can see that since $\omega(Q) = \bigcup_{n+1}^{\infty} \omega^n(P)$ and $Q = P \cup \bigcup_{n=1}^{\infty} \omega^n(P)$, it's clear that $\omega(Q) \subset Q$. Having verified these conditions that were stated in the theorem, there remains only one point left to make.

Using the conclusion that $\omega(Q) = \bigcup_{n=1}^{\infty} \omega^n(P)$, we can insert $\omega(Q)$ into equation (7):

$$P \cap \omega(Q) = \emptyset$$

Thus, we have found an ω and a countable set Q which must contain our fixed countable set P, such that when we apply ω to Q we will never get a result that lands back in the set P no matter how many times we continue to apply ω. This concludes the proof. □

In our next proof, we will use all we have proved so far about rotations and subsets of S to partition S into ten pieces in such a way that we can rotate six of the the partition pieces to again partition S and also rotate the other four partition pieces to again partition S. This is remarkable because intuitively we would expect that once we partition S, we would need to rotate all ten partition pieces in order to re-partition S. Here I will prove that this is not the case.

Theorem 4.3. *There exists a partition $\{T_j : 1 \leq j \leq 10\}$ of the unit sphere S into ten (disjoint) subsets and a corresponding set $\{\rho_j : 1 \leq j \leq 10\}$ of rotations such that $\{\rho_j(T_j) : 1 \leq j \leq 6\}$ is a partition of S into six subsets and $\{\rho_j(T_j) : 7 \leq j \leq 10\}$ is a partition of S into four subsets. Specifically, we can take T_7, T_8, and T_9 to all be rotates of S_1 (as defined in Theorem 4.1) and take T_1, T_2, T_3, and T_{10} to all be countable.*

Proof. We continue to use the same definitions for P, S_1, S_2, S_3, ϕ and ψ. We additionally define the following terms:

$$U_1 = \phi(S_2) \qquad U_2 = \psi\phi(S_2) \qquad U_3 = \psi^2\phi(S_2)$$

$$V_1 = \phi(S_3) \qquad V_2 = \psi\phi(S_3) \qquad V_3 = \psi^2\phi(S_3)$$

Using the definitions of S_1, S_2, and S_3 as defined in Theorem 4.1, notice by the following calculations that $\{U_j, V_j\}$ is a partition of S_j for $j = 1, 2, 3$. First we see that $\{U_1, V_1\}$ partition S_1, using the fact that $\phi^2 = \iota$:

$$\phi(S_1) = S_2 \cup S_3$$

$$S_1 = \phi(S_2 \cup S_3)$$

$$S_1 = \underbrace{\phi(S_2)}_{U_1} \cup \underbrace{\phi(S_3)}_{V_1}$$

We continue with the above calculation to see that $\{U_2, V_2\}$ partitions S_2:

$$S_1 = \phi(S_2) \cup \phi(S_3)$$

$$\psi(S_1) = \psi(\phi(S_2) \cup \phi(S_3))$$

$$S_2 = \underbrace{\psi\phi(S_2)}_{U_2} \cup \underbrace{\psi\phi(S_3)}_{V_2}$$

We again begin with the same equivalence with which we began the above calculation, this time to show that $\{U_3, V_3\}$ partitions S_3:

$$S_1 = \phi(S_2) \cup \phi(S_3)$$

$$\psi^2(S_1) = \psi^2(\phi(S_2) \cup \phi(S_3))$$

$$S_3 = \underbrace{\psi^2\phi(S_2)}_{U_3} \cup \underbrace{\psi^2\phi(S_3)}_{V_3}$$

Since a rotation is a bijection, if we rotate two disjoint sets, the image of those sets will also be disjoint. Thus, since S_2 and S_3 are disjoint, then $\phi(S_2)$ and $\phi(S_3)$ are also disjoint. Therefore, U_1 and V_1 are disjoint. Using the same logic, U_2 and V_2 are disjoint, as are U_3 and V_3. Finally, we can see that all six of these subsets are in fact disjoint since $U_j, V_j \subset S_j$ for $j = 1, 2, 3$ and we know that S_1, S_2, and S_3 are disjoint.

Now, we define the following sets and rotations:

$$T_7 = U_1 \qquad T_8 = U_2 \qquad T_9 = U_3 \qquad T_{10} = P$$

$$\rho_7 = \psi^2\phi \qquad \rho_8 = \phi\psi^2 \qquad \rho_9 = \psi\phi\psi \qquad \rho_{10} = \iota$$

Notice that $\rho_{10}(T_{10}) = \iota(P) = P$. Additionally, we can see that $\rho_7(T_7) = S_1$, using the fact that $\psi(S_1) = S_2$ and $\psi^3 = \iota$ imply that $S_1 = \psi^2(S_2)$:

$$\rho_7(T_7) = \psi^2\phi(\phi(S_2)) = \psi^2(S_2) = S_1$$

Similarly, $\rho_8(T_8) = S_2$:

$$(\phi\psi^2)\psi\phi(S_2) = \phi^2(S_2) = S_2$$

In the same way, using the facts that $\psi(S_1) = S_2$ and $\psi^2(S_1) = S_3$, we can also see that $\rho_9(T_9) = S_3$:

$$\psi\phi\psi(\psi^2\phi(S_2)) = \psi\phi^2(S_2) = \psi(S_2) = \psi(\psi(S_1)) = \psi^2(S_1) = S_3$$

Thus, we've shown that $\rho_7(S_7), \rho_8(S_8), \rho_9(S_9)$, and $\rho_{10}(S_{10})$ are equivalent to S_1, S_2, S_3, and P, respectively. Since we have previously shown that S_1, S_2, S_3, and P partition S, we thus know that $\rho_7(S_7), \rho_8(S_8), \rho_9(S_9)$, and $\rho_{10}(S_{10})$ also partition S.

We will now turn our attention to the remaining portions of the partition of S and see that these can again be rotated to partition S. We've already seen that $U_1 = T_7, V_1, U_2 = T_8, V_2, U_3 = T_9$, and V_3 partition $S \setminus P$ and we know that $T_{10} = P$. Thus, $S \setminus (T_7 \cup T_8 \cup T_9 \cup T_{10}) = V_1 \cup V_2 \cup V_3$. We will now divide the sets V_j for $j = 1, 2, 3$ into six pieces. Let Q and ω retain the same definition as in the previous theorem:

$$Q = P \cup \bigcup_{n=1}^{\infty} \omega^n(P)$$

$$\omega = \sigma^{-1}\tau\sigma$$

Recall that Q is a countable set containing a countable set P such that when we apply ω to Q, the resulting points are exactly $Q \setminus P$. We now define the remainder of our T sets:

$$T_1 = \rho_8(S_1 \cap Q) \qquad T_2 = \rho_9(S_2 \cap Q) \qquad T_3 = \rho_7(S_3 \cap Q)$$

$$T_4 = \rho_8(S_1 \setminus Q) \qquad T_5 = \rho_9(S_2 \setminus Q) \qquad T_6 = \rho_7(S_3 \setminus Q)$$

We know that the sets T_1, T_2, and T_3 must be countable because Q was countable. Now, notice the following:

$$\rho_8(S_1) = \phi\psi^2(S_1) = \phi(S_3) = V_1$$

$$\rho_9(S_2) = \psi\phi\psi(S_2) = \psi\phi(S_3) = V_2$$

$$\rho_7(S_3) = \psi^2\phi(S_3) = V_3$$

Because of the way we defined T_j for $j = 1, 2, 3, 4, 5, 6$, when we look at the above equalities it is clear that T_1 and T_4 partition V_1, T_2 and T_5 partition V_2, and T_3 and T_6 partition V_3. Since we know that $\{U_j, V_j, P : j = 1, 2, 3\}$ partitions S and we have now

seen that $\{T_j : 1 \leq j \leq 10\}$ is equivalent to the set $\{U_j, V_j, P : j = 1, 2, 3\}$, we can say that $\{T_j : 1 \leq j \leq 10\}$ partitions S.

The remainder of this proof consists of defining the remaining ρ_j such that when we apply each ρ_j to each set T_j for $1 \leq j \leq 6$, the resulting sets again partition S. We first look at the terms T_j and ρ_j for $j = 4, 5, 6$ and see that $\{\rho_j(T_j) : j = 4, 5, 6\} = S \setminus Q$. I'll begin by defining the rotations ρ_j for $j = 4, 5, 6$:

$$\rho_4 = \rho_8^{-1} \qquad \rho_5 = \rho_9^{-1} \qquad \rho_6 = \rho_7^{-1}$$

We now apply each ρ_j for $j = 4, 5, 6$ to its corresponding set T_j:

$$\rho_4(T_4) = \rho_8^{-1}\rho_8(S_1 \setminus Q) = S_1 \setminus Q$$

$$\rho_5(T_5) = \rho_9^{-1}\rho_9(S_2 \setminus Q) = S_2 \setminus Q$$

$$\rho_6(T_6) = \rho_7^{-1}\rho_7(S_3 \setminus Q) = S_3 \setminus Q$$

Recalling that $P \subset Q$, the above equations clearly show that $\{\rho_j(T_j) : j = 4, 5, 6\}$ partitions $S \setminus Q$. We now define ρ_j for $j = 1, 2, 3$ and will seek to show that when we apply each of these ρ_j to its corresponding T_j we obtain a partition of Q:

$$\rho_1 = \omega^{-1}\rho_4 \qquad \rho_2 = \omega^{-1}\rho_5 \qquad \rho_3 = \omega^{-1}\rho_6$$

We now apply each ρ_j for $j = 1, 2, 3$ to its corresponding set T_j:

$$\rho_1(T_1) = \omega^{-1}\rho_4\rho_8(S_1 \cap Q) = \omega^{-1}\rho_8^{-1}\rho_8(S_1 \cap Q) = \omega^{-1}(S_1 \cap Q)$$

$$\rho_2(T_2) = \omega^{-1}\rho_5\rho_9(S_2 \cap Q) = \omega^{-1}\rho_9^{-1}\rho_9(S_2 \cap Q) = \omega^{-1}(S_2 \cap Q)$$

$$\rho_3(T_3) = \omega^{-1}\rho_6\rho_7(S_3 \cap Q) = \omega^{-1}\rho_7^{-1}\rho_7(S_3 \cap Q) = \omega^{-1}(S_3 \cap Q)$$

It is immediately apparent that $\rho_1(T_1), \rho_2(T_2)$, and $\rho_3(T_3)$ are all disjoint since they are rotations of three disjoint sets $S_1 \cap Q, S_2 \cap Q$, and $S_3 \cap Q$. However, some further work is required to show that their union is in fact Q:

$$\rho_1(T_1) \cup \rho_2(T_2) \cup \rho_3(T_3) = \omega^{-1}(S_1 \cap Q) \cup \omega^{-1}(S_2 \cap Q) \cup \omega^{-1}(S_3 \cap Q)$$

$$= \omega^{-1}((S_1 \cap Q) \cup (S_2 \cap Q) \cup (S_3 \cap Q)) = \omega^{-1}((S_1 \cup S_2 \cup S_3) \cap Q) = \omega^{-1}(Q \setminus P) = Q$$

In the above calculation, we know first that $(S_1 \cup S_2 \cup S_3) \cap Q = Q \setminus P$ because we know $S_1 \cup S_2 \cup S_3 = S \setminus P$. Then, we can see that $\omega^{-1}(Q \setminus P) = Q$ by doing a short calculation on our definition of ω:

$$\omega(Q) = Q \setminus P$$

$$\omega^{-1}\omega(Q) = \omega^{-1}(Q \setminus P)$$

$$Q = \omega^{-1}(Q \setminus P)$$

Thus, we have shown that $\{\rho_j(T_j) : j = 1, 2, 3\}$ partitions Q. Earlier, we showed that $\{\rho_j(T_j) : j = 4, 5, 6\}$ partitioned $S \setminus Q$. Thus, we can conclude that $\{\rho_j(T_j) : 1 \leq j \leq 6\}$ partitions S.

To summarize this proof, we have shown that we are able to partition S into ten sets which we labelled $T_1, T_2, ..., T_{10}$, two disjoint subsets of which could be rotated to each partition S. We first rotated four of these sets, T_7, T_8, T_9, and T_{10} to partition S. We were additionally able to rotate the remaining six sets $T_1, T_2, ..., T_6$ to partition S a second time. □

Having proved this about the unit sphere, we now seek to prove a similar theorem about the unit ball.

5 The Unit Ball

We will now begin working with the unit ball, written B, which is the set of all points with a distance from the origin of less than or equal to one. Recall that a rigid motion is a mapping r from $\mathbb{R}^3$ onto $\mathbb{R}^3$ having the form $r(x) = \rho(x) + a$ for $x \in \mathbb{R}^3$ where ρ is a fixed rotation and $a \in \mathbb{R}^3$ is fixed.

Theorem 5.1. *There exists a partition $\{B_k : 1 \leq k \leq 40\}$ of the closed unit ball B into forty subsets and a corresponding set $\{r_k : 1 \leq k \leq 40\}$ of rigid motions such that $\{r_k(B_k) : 1 \leq k \leq 24\}$ partitions B into twenty-four subsets and $\{r_k(B_k) : 25 \leq k \leq 40\}$ partitions B into sixteen subsets.*

Proof. We will first use the results of Theorem 4.2, which state that if P is any countable subset of S,then there exists a countable set Q and a rotation ω such that $P \subset Q \subset S$ and $\omega(Q) = Q \setminus P$. We apply this theorem to the case where P is the set of a single element u where $u = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in S$ to obtain a countable set Q such that $\{u\} \subset Q \subset S$ and a rotation ρ_0 such that $\rho_0(Q) = Q \setminus \{u\}$.

Next, let $N_1 = \{\frac{1}{2}(q-u) : q \in Q\}$. We complete the following calculation, letting $q \in Q$ to show that $N_1 \subseteq B$:

$$|\frac{1}{2}(q-u)| = \frac{1}{2}|q-u| \leq \frac{1}{2}(|q| + |u|) = \frac{1}{2}(1+1) = 1$$

This tells us that every element in N_1 has a length of less than or equal to one, thus satisfying our definition for B.

Additionally define the rigid motion r_0 by $r_0(x) = \rho_0(x + \frac{1}{2}u) - \frac{1}{2}u$. I'll show here that this does in fact satisfy our definition of a rigid motion:

$$r_0(x) = \rho_0(x + \frac{1}{2}u) - \frac{1}{2}u = \rho_0(x) + \rho_0(\frac{1}{2}u) - \frac{1}{2}u = \underbrace{\rho_0(x)}_{\text{rotation}} + \underbrace{\frac{1}{2}\rho_0(u) - \frac{1}{2}u}_{\text{constant vector}}$$

We can see that the vector 0 is in N_1, since $\{u\} \in Q$ and $\frac{1}{2}(u-u) = 0$. We complete the following calculation to see that $r_0(N_1) = N_1 \setminus \{0\}$:

$$r_0\left(\frac{1}{2}(q-u)\right) = \rho_0\left(\frac{1}{2}q - \frac{1}{2}u + \frac{1}{2}u\right) - \frac{1}{2}u = \rho_0\left(\frac{1}{2}q\right) - \frac{1}{2}u = \frac{1}{2}\rho_0(q) - \frac{1}{2}u$$

Since we defined ρ_0 to be such that $\rho_0(Q) = Q \setminus \{u\}$, we know that $\rho_0(q) \neq u$ and thus that $\frac{1}{2}\rho_0(q) - \frac{1}{2}u \neq 0$. Thus, $r_0(N_1) = N_1 \setminus \{0\}$.

We now define $N_2 = B \setminus N_1, s_1 = r_0, s_2 = \iota, M_1 = s_1(N_1)$, and $M_2 = s_2(N_2)$. We can see that N_1 and N_2 partition B. The following calculation allows us to see that M_1 and M_2 partition $B \setminus \{0\}$:

$$\{M_1, M_2\} = \{r_0(N_1), \iota(N_2)\} = \{N_1 \setminus \{0\}, N_2\}$$

Since $\{N_1, N_2\}$ partitions $B, \{N_1 \setminus \{0\}, N_2\}$ partitions $B \setminus \{0\}$ and thus M_1 and M_2 partition $B \setminus \{0\}$. We now define $S' = \{y \in \mathbb{R}^3 : 0 < |y| \leq 1\}$ such that S' is the unit ball without the origin. Further, we define $T'_j = \{tx : x \in T_j, 0 < t \leq 1\}$. Since T_j for $1 \leq j \leq 10$ partitions S, it follows that T'_j for $1 \leq j \leq 10$ partitions S'. We state this more formally by using our definitions of S' and T'_j to restate Theorem 4.3 as follows:

There exists a partition $\{T'_j : 1 \leq j \leq 10\}$ of S' into ten (disjoint) subsets and a corresponding set $\{\rho_j : 1 \leq j \leq 10\}$ of rotations such that $\{\rho_j(T'_j) : 1 \leq j \leq 6\}$ is a partition of S' into six subsets and $\{\rho_j(T'_j) : 7 \leq j \leq 10\}$ is a partition of S' into four subsets.

The proof of this statement is identical to the proof of Theorem 4.3 except each S is replaced by S' and each T_j is replaced by T'_j. Notice that S' is in fact the same set as $B \setminus \{0\}$. Also notice that, for each $j = 1, 2, ..., 10$, the family $\{T'_j \cap \rho_j^{-1}(M_i) : i = 1, 2\}$ partitions T'_j. This happens because M_1 and M_2 partition S' and thus $\rho_j^{-1}(M_1)$ and $\rho_j^{-1}(M_2)$ also partition S' since ρ is invertible and thus one-to-one with $\rho_j(0) = 0$. Therefore:

$$T'_j = (T'_j \cap \rho_j^{-1}(M_1)) \cup (T'_j \cap \rho_j^{-1}(M_2)) \text{ and } \rho_j^{-1}(M_1) \cap \rho_j^{-1}(M_2) = \emptyset$$

We also need to notice that for each $j = 1, ..., 10$, $\{M_n \cap T'_j \cap \rho_j^{-1}(M_i) : n = 1, 2\}$ partitions $T'_j \cap \rho_j^{-1}(M_i)$ for $i = 1, 2$. This happens because M_1 and M_2 partition S' and since $T'_j \cap \rho_j^{-1}(M_i)$ for $i = 1, 2$ and $j = 1, ..., 10$ are subsets of S', $\{M_n \cap T'_j \cap \rho_j^{-1}(M_i) : n = 1, 2\}$ clearly partitions $T'_j \cap \rho_j^{-1}(M_i)$ for $i = 1, 2$ and $j = 1, ..., 10$. Thus, $\{M_n \cap T'_j \cap \rho_j^{-1}(M_i) : 1 \leq n \leq 2, 1 \leq i \leq 2, 1 \leq j \leq 10\}$ partitions S' into forty subsets.

Recall now that $M_1 = s_1(N_1), M_2 = s_2(N_2)$, and that N_1 and N_2 partition B while M_1 and M_2 partition $B \setminus \{0\} = S'$. Thus, by taking s_n^{-1} of our forty-set partition of S', we will have a forty-set partition of B:

$$B_{nij} = s_n^{-1}\left(M_n \cap T'_j \cap \rho_j^{-1}(M_i)\right) \text{ for } 1 \leq n \leq 2, 1 \leq i \leq 2, 1 \leq j \leq 10$$

Meanwhile, for each fixed j the four sets $\rho_j s_n(B_{nij}) = M_i \cap \rho_j(M_n \cap T'_j)$ for $n = 1, 2$ and $i = 1, 2$ form a partition of $\rho_j(T'_j)$:

$$\rho_j s_n(B_{nij}) = \rho_j s_n s_n^{-1}\left(M_n \cap T'_j \cap \rho_j^{-1}(M_i)\right) = \rho_j(M_n) \cap \rho_j(T'_j) \cap M_i \tag{8}$$

We have already established that $\{M_1, M_2\}$ and $\{\rho_j(M_1), \rho_j(M_2)\}$ each partition S'. We also already know that M_1 and M_2 are disjoint. It follows that $\rho_j(M_1)$ and $\rho_j(M_2)$ are also disjoint because ρ_j is a bijection and therefore when its input is a partition, its output is also a partition. Therefore, for each fixed j, $\rho_j(M_n) \cap \rho_j(T'_j) \cap M_i = \rho_j s_n(B_{nij})$ for $n = 1, 2$ and $i = 1, 2$ partitions $\rho_j(T'_j)$.

Using our adapted version of Theorem 4.3, we know that $\{\rho_j(T'_j) : 1 \leq j \leq 6\}$ partitions S' and that $\{\rho_j(T'_j) : 7 \leq j \leq 10\}$ also partitions S'. Because for each fixed j the four sets $\rho_j s_n(B_{nij}) = M_i \cap \rho_j(M_n \cap T'_j)$ for $n = 1, 2$ and $i = 1, 2$ form a partition of $\rho_j(T'_j)$, we can use the results of Theorem 4.3 to conclude that the following two families are also each a partition of S':

$$\{\rho_j s_n(B_{nij}) : 1 \leq n \leq 2, 1 \leq i \leq 2, 1 \leq j \leq 6\}$$

$$\{\rho_j s_n(B_{nij}) : 1 \leq n \leq 2, 1 \leq i \leq 2, 7 \leq j \leq 10\}$$

Now, let us fix i in the above families and use equation (8) to see that the following families of twelve and eight sets are each a partition of M_i:

$$\{\rho_j s_n(B_{nij}) : 1 \leq n \leq 2, 1 \leq j \leq 6\} = \{\rho_j(M_i) \cap \rho_j(T'_j) \cap M_i : 1 \leq n \leq 2, 1 \leq j \leq 6\}$$

$$\{\rho_j s_n(B_{nij}) : 1 \leq n \leq 2, 7 \leq j \leq 10\} = \{\rho_j(M_i) \cap \rho_j(T'_j) \cap M_i : 1 \leq n \leq 2, 7 \leq j \leq 10\}$$

This is true because without fixing i each family is a partition of all of S', and so when we fix i it follows immediately that we have a partition of M_i.

Then, as we did earlier, we can map these partitions of M_i to partitions of N_i by applying s_i^{-1}. In the following calculation, i is fixed:

$$\{s_i^{-1}\rho_j s_n(B_{nij}) : 1 \leq n \leq 2, 1 \leq j \leq 6\}$$

$$= \{s_i^{-1}\rho_j s_n\left(s_n^{-1}\left(M_n \cap T'_j \cap \rho_j^{-1}(M_i)\right)\right) : 1 \leq n \leq 2, 1 \leq j \leq 6\}$$

$$= \{s_i^{-1}\rho_j(M_n) \cap s_i^{-1}(T'_j) \cap s_i^{-1}(M_i) : 1 \leq n \leq 2, 1 \leq j \leq 6\}$$

$$= \{s_i^{-1}\rho_j(M_n) \cap s_i^{-1}(T'_j) \cap N_i : 1 \leq n \leq 2, 1 \leq j \leq 6\}$$

We use the fact that s_i^{-1} is a bijection to know that the above expression partitions N_i. The fact that s_i^{-1} is one-to-one ensures that the partition of M_i remains disjoint when we project it onto N_i. The fact that s_i^{-1} is onto ensure that the image of the partition covers all of N_i. Thus, we have formed a partition of N_i. We form a second partition of

N_i by repeating the above calculation except by letting $7 \leq j \leq 10$. Let us now define $r_{nij} = s_i^{-1}\rho_j s_n$ so that we can rewrite the above equation to say:

$$\{r_{nij}(B_{nij}) : 1 \leq n \leq 2, 1 \leq i \leq 2, 1 \leq j \leq 6\}$$

$$\{r_{nij}(B_{nij}) : 1 \leq n \leq 2, 1 \leq i \leq 2, 7 \leq j \leq 10\}$$

Because when we fixed i these families each partitioned N_i and because N_i for $i = 1, 2$ partitions B, now that we are allowing $i = 1, 2$, the above families of twenty-four and sixteen sets, respectively, each partition B. Finally, let us relabel our sets B_{nij} and rigid motions r_{nij} as follows:

$$r_{nij} = r_1, r_2, ..., r_{24} \text{ for } 1 \leq n \leq 2, 1 \leq i \leq 2, 1 \leq j \leq 6$$

$$r_{nij} = r_{25}, r_2, ..., r_{40} \text{ for } 1 \leq n \leq 2, 1 \leq i \leq 2, 7 \leq j \leq 10$$

$$B_{nij} = r_1, r_2, ..., r_{24} \text{ for } 1 \leq n \leq 2, 1 \leq i \leq 2, 1 \leq j \leq 6$$

$$B_{nij} = r_{25}, r_2, ..., r_{40} \text{ for } 1 \leq n \leq 2, 1 \leq i \leq 2, 7 \leq j \leq 10$$

Thus we have forty sets B_k for $k = 1, 2, ..., 40$ which partition B and forty rigid motions r_k for $k = 1, 2, ..., 40$ such that $\{r_k(B_k) : 1 \leq k \leq 24\}$ partitions B and $\{r_k(B_k) : 25 \leq k \leq 40\}$ also partitions B. This concludes our proof. □

In summary, we have found a way to partition B into forty pieces and defined forty rigid motions such that we are able to again partition B twice using only those forty pieces and rigid motions. This is remarkable because we have seemingly doubled the size of the set we are working with–without changing the contents of our set at all!

In the next section, we seek to further generalize this concept beyond the unit ball to any two nonempty, bounded subsets of $\mathbb{R}^3$.

6 Piecewise Congruence

I begin by defining piecewise congruence, then establish some properties related to piecewise congruence, and finally use those properties to complete our proof the Banach-Tarski Paradox in our final two theorems.

Definition 6.1. Two subsets X and Y of $\mathbb{R}^3$ are **piecewise congruent**, written $X \sim Y$ if, for some natural number n, there exists a partition $\{X_j : 1 \leq j \leq n\}$ of X into n subsets and a corresponding set $\{f_j : 1 \leq j \leq n\}$ of rigid motions such that $\{f_j(X_j) : 1 \leq j \leq n\}$ is a partition of Y.

In other words, two sets are considered piecewise congruent if we are able to take a finite partition of X and use only rigid motions to transform it into a finite partition of Y. If X is piecewise congruent to a subset of Y, we write $X \lesssim Y$.

Theorem 6.2. *For subsets X, Y, and Z of $\mathbb{R}^3$, we have:*

(1) $X \sim X$

(2) $X \sim Y \Rightarrow Y \sim X$

(3) $X \sim Y$ *and* $Y \sim Z \Rightarrow X \sim Z$

(4) $X \sim Y \Rightarrow X \lesssim Y$

(5) $X \lesssim Y$ *and* $Y \lesssim Z \Rightarrow X \lesssim Z$

(6) $X \subseteq Y \Rightarrow X \lesssim Y$

(7) $X \lesssim Y$ *and* $Y \lesssim X \Rightarrow X \sim Y$

Properties (1)-(6) are easy to check. Because of this, we will only complete the proof of property (7). The proof here is based on a well-known proof of the Schröder-Bernstein Theorem.

Proof. Suppose that $X \sim Y_0$ and $Y \sim X_0$ where $Y_0 \subset Y$ and $X_0 \subset X$. Let $\{X_j : 1 \leq j \leq n\}$ and $\{Y_i : 1 \leq i \leq m\}$ be partitions of X and Y, respectively, and let $\{f_j : 1 \leq j \leq n\}$ and $\{g_i : 1 \leq i \leq m\}$ be sets of rigid motions such that $\{f_j(X_j) : 1 \leq j \leq n\}$ is a partition of Y_0 and $\{g_i(T_i) : 1 \leq i \leq m\}$ is a partition of X_0.

First, define f on X by $f(x) = f_j(x)$ if $x \in X_j$ and define g on Y by $g(y) = g_i(y)$ if $y \in Y_i$. Then, for a set $E \subset X$, define $E' \subset X$ by the following:

$$E' = X \setminus g[Y \setminus f(E)]$$

Notice by the following calculation that if F is also a subset of X such that $E \subset F \subset X$, then $E' \subset F'$:

$$\begin{aligned} E \subset F &\Rightarrow f[E] \subset f[F] \\ &\Rightarrow Y \setminus f[E] \supset Y \setminus f[F] \\ &\Rightarrow g\left[Y \setminus f[E]\right] \supset g\left[Y \setminus f[F]\right] \\ &\Rightarrow X \setminus g[Y \setminus f[E]] \subset X \setminus g[Y \setminus f[F]] \\ &\Rightarrow E' \subset F' \end{aligned}$$

Now, let $\mathfrak{D} = \{E : E \subset X, E \subset E'\}$. Here, $\emptyset \in \mathfrak{D}$ because $\emptyset \subset X$ and $\emptyset \subset \emptyset'$ because $\emptyset$ is contained in every set. Let $D = \bigcup \mathfrak{D}$ be the union of all sets that belong to $\mathfrak{D}$.

For each $E \in \mathfrak{D}$, we know that $E \subset D \subset X$, and thus that $E' \subset D'$. Using this fact and the knowledge that we defined $\mathfrak{D}$ to contain only sets E for which $E \subset E'$, we further know that $E \subset D'$.

Since we have just determined that, for every set E in $\mathfrak{D}$, $E \subset D'$, and since D is the union of all sets belonging to $\mathfrak{D}$, we now know that $D \subset D'$. Then, since $D \subset D' \subset X$, we know that $D' \subset (D')'$.

Then, because of our definition of $\mathfrak{D}$, we know that $D' \in \mathfrak{D}$ and so $D' \subset D$ since D is the union of all sets that belong to $\mathfrak{D}$. We can conclude that $D' = D$. Using this fact as well as our earlier definition of E', we write out D' and then perform some calculations to find $X \setminus D$:

$$D = D' = X \setminus g[Y \setminus f[D]]$$
$$D = (g[Y \setminus f[D]])^C$$
$$D^C = (g[Y \setminus f[D]])^{C^C}$$
$$X \setminus D = g[Y \setminus f[D]]$$

Because we've defined g to always map to X_0, we then know that $X \setminus D \subset X_0$. Now define the following for $1 \leq j \leq n$ and $1 \leq i \leq m$:

$$A_j = D \cap X_j \qquad A_{n+i} = g_i[Y_i \setminus f[D]] \qquad h_j = f_j \qquad h_{n+i} = g_i^{-1}$$

Since $D \subset X$, the sets A_j for $1 \leq j \leq n$ partition D. Additionally, since g is a bijection from Y onto X_0 and Y_i is a partition of Y for $1 \leq i \leq m$, the sets A_{n+i} partition $X \setminus D$. Then:

$$h_j(A_j) = h_j(D \cap X_j) = f_j(D \cap X_j) = f_j(D) \cap f_j(X_j)$$

This tells us that the sets $h_j(A_j)$ for $1 \leq j \leq n$ partition $f(D)$. Furthermore:

$$h_{n+i}(A_{n+i}) = g_i^{-1}(g[Y_i \setminus f[D]]) = Y_i \setminus f[D]$$

This shows that the sets $h_{n+i}(A_{n+i})$ partition $Y \setminus f[D]$.

Thus, since D and $X \setminus D$ partition X while $f(D)$ and $Y \setminus f(D)$ partition Y, we have found $2n$ many sets, A_j and A_{n+i} for $1 \leq j \leq n$ and $1 \leq i \leq n$, which partition X and can be transformed by specific rigid motions, h_j and h_{n+i} respectively, such that their images partition Y. Therefore, $X \sim Y$ and we have completed the proof. □

Before beginning the next proof, I will state the following definitions.

Definition 6.3. A **closed ball** in $\mathbb{R}^3$ is any set of the form $A = \{x \in \mathbb{R}^3 : |x - a| \leq \epsilon\}$ where $a \in \mathbb{R}^3$ and $\epsilon > 0$ are given.

Definition 6.4. A **translate** of a set $A \subset \mathbb{R}^3$ is any set of the form $A + b = \{x + b : x \in A\}$ where $b \in \mathbb{R}^3$ is given.

Theorem 6.5. *If $A \subset \mathbb{R}^3$ is a closed ball and if $A_1, A_2, ..., A_n$ are a finite number of translates of A, then:*

$$A \sim \bigcup_{j=1}^{n} A_j$$

Proof. Suppose that A is some closed ball in $\mathbb{R}^3$ centered around the origin. That is, let A be the set $A = \{x \in \mathbb{R}^3 : |x| \leq \epsilon\}$ for some $\epsilon > 0$. Then, choose any $a \in \mathbb{R}^3$ for which $|a| > 2\epsilon$ and let $A' = A + a = \{y \in \mathbb{R}^3 : |y - a| \leq \epsilon\}$. This tells us that A' is also a closed ball in $\mathbb{R}^3$ centered around a. Because $|a| > 2\epsilon$, A and A' do not overlap.

The next step is to use the results from Theorem 5.1 to show that $A \sim (A \cup A')$. Let B_k and r_k be defined in the following way, as they were in Theorem 5.1: $\{B_k : 1 \leq k \leq 40\}$ is a partition of the closed unit ball B and $\{r_k : 1 \leq k \leq 40\}$ is a set of corresponding rigid motions such that each $\{r_k(B_k) : 1 \leq k \leq 24\}$ and $\{r_k(B_k) : 25 \leq k \leq 40\}$ partitions B.

We now adopt the following notation. For any set $D \subset \mathbb{R}^3$ and any $\delta > 0$, let $\delta D = \{\delta x : x \in D\}$. Let us now consider the set $\{\epsilon B_k : 1 \leq k \leq 40\}$. Notice that we have exactly scaled the forty-piece partition of B to now partition the set A. That is, $\epsilon B = A$ and the sets ϵB_k are disjoint for all $1 \leq k \leq 40$.

We now define the rigid motions s_k for $1 \leq k \leq 40$ as follows:

$$s_k(x) = \epsilon r_k\left(\frac{1}{\epsilon}x\right) \text{ if } 1 \leq k \leq 24$$

$$s_k(x) = \epsilon r_k\left(\frac{1}{\epsilon}x\right) + a \text{ if } 25 \leq k \leq 40$$

Here, we are taking each point in A, scaling it by $\frac{1}{\epsilon}$ so that we then have all the points in B, applying the rigid motions r_k such that the result is then two separate partitions of B, and then scaling this result by ϵ so that we then have two partitions of A. Then, for $25 \leq k \leq 40$, we are shifting this partition by a such that we then have a partition of A'. In summary, when we apply s_k to A, the result is a partition of A and a partition of A'.

Since we chose a such that A and A' would be disjoint, we have therefore created a partition of $A \cup A'$. That is, we were able to transform a 40-set partition of A into a partition of $A \cup A'$ using 40 rigid motions. Thus, $A \sim (A \cup A')$.

We now use induction to complete the proof that if $A_1, ..., A_n$ are a finite number of translates of A, then $A \sim \bigcup_{j=1}^{n} A_j$.

Base Case: We begin by proving that $A \sim \bigcup_{j=1}^{n} A_j$ when $n = 1$. We can easily show that A is piecewise congruent to a single translate of itself. Let A_1 be a translate of A. That is, A_1 is the set A shifted by some value $x \in \mathbb{R}^3$, written $A_1 = \{a + x : a \in A\}$. In this case, we are translating A to A_1 by the rigid motion $r_x = x + a$ and since rigid motions preserve partitions, applying r_x to any partition of A will result in a partition of A_1. Thus $A \sim A_1$, proving the base case.

Induction Hypothesis: We assume that $n > 1$ is such that A is piecewise congruent to the union of any $n - 1$ of its translates and let $A_1, ..., A_n$ be any n of its translates. We now seek to prove that A is then piecewise congruent to all n of these translates.

Since, according to our induction hypothesis, $A \sim (A_1 \cup ... \cup A_{n-1})$ and $A_1 \cup ... \cup A_{n-1}$ is clearly a subset of $A_1 \cup ... \cup A_n$, therefore $A \lesssim (A_1 \cup ... \cup A_n)$. Then, using our results

from Theorem 6.2, we need only to show that $(A_1 \cup ... \cup A_n) \lesssim A$ and we will have that $A \sim (A_1 \cup ... \cup A_n)$.

Notice first that since A' and A_n are both translates of A they are also translates of one another. Thus, $A_n \sim A'$.

Then, notice that $A_n \setminus (A_1 \cup ... \cup A_{n-1}) \subseteq A_n$ (they are equivalent if there is no overlap between A_n and any of the other sets $A_1, ..., A_n$). Thus, we can see that:

$$A_n \setminus (A_1 \cup ... \cup A_{n-1}) \lesssim A'$$

We then use the following fact to complete the proof:

$$A_1 \cup ... \cup A_n = (A_1 \cup ... \cup A_{n-1}) \cup (A_n \setminus (A_1 \cup ... \cup A_{n-1}))$$

Recall that $A_1 \cup ... \cup A_{n-1} \sim A$ by our induction hypothesis and that $(A_n \setminus (A_1 \cup ... \cup A_{n-1})) \lesssim A'$, as we have already established.

Because A and A' are disjoint, we now know that the following is true:

$$(A_1 \cup ... \cup A_{n-1}) \cup (A_n \setminus (A_1 \cup ... \cup A_{n-1})) \lesssim A \cup A'$$

By the first part of this proof, we know that $A \cup A' \sim A$ and so we can say the following:

$$(A_1 \cup ... \cup A_{n-1}) \cup (A_n \setminus (A_1 \cup ... \cup A_{n-1})) \lesssim A$$

Thus, we have shown both that $A_1 \cup ... \cup A_n \lesssim A$ and that $A \lesssim A_1 \cup ... \cup A_n$ and so, by Theorem 6.2, we have that $A_1 \cup ... \cup A_n \sim A$. □

Finally, we are ready to prove the final portion of the Banach-Tarski Paradox.

Theorem 6.6. *If X and Y are bounded subsets of $\mathbb{R}^3$ having non-empty interiors, then $X \sim Y$.*

Proof. Let X and Y be bounded subsets of $\mathbb{R}^3$ having non-empty interiors. Begin by choosing interior points $a \in X$ and $b \in Y$. Then choose $\epsilon > 0$ such that the following satisfies $A + a \subset X$ and $A + b \subset Y$:

$$A = \{x \in \mathbb{R}^3 : |x| \leq \epsilon\}$$

Thus, we again have that A is a closed ball having radius ϵ centered around the origin and we have chosen a and b such that shifting A by a results in a subset of X and shifting A by b results in a subset of Y.

Since X is bounded, X can be contained in the union of finitely many translates of A, which we will write as $A_1, ..., A_n$. Thus:

$$X \subset A_1 \cup ... \cup A_n$$

Using the above statement as well as the result from Theorem 6.5 that the set A is piecewise congruent to any finite number of its translates, we can make the following conclusion:

$$X \lesssim A$$

Because $A + a \subset X$ and $A \sim A + a$ since $A + a$ is simply a translate of A, we know that:

$$A \lesssim X$$

Using our results from Theorem 6.2, we therefore have that $X \sim A$.

Using the same logic, we also have that $Y \sim A$. Then, once again using Theorem 6.2, we finally have that:

$$X \sim Y$$

□

This final result concludes the proof of the Banach-Tarski Paradox.

References

[1] Karl Stromberg, "The Banach-Tarski Paradox," *American Mathematical Monthly* **86 (3)**, 1979, 151-161.