The Banach-Tarski Paradox

Paul Zeitz

December 13, 2014

Contents

1 Introduction 1
  1.1 Definitions .............................................. 2

2 Removing Points with Rigid-Body Motions 3

3 The Free Group 6
  3.1 Paradoxical Decomposition of the Free Group ........ 6
  3.2 A Geometric Model of the Free Group ................. 8

4 The Axiom of Choice 9
  4.1 A “Non-measurable” set .............................. 9

5 Step-by-step BTP 11
  5.1 Remove the Center ...................................... 11
  5.2 Move to the Surface ................................... 11
  5.3 Partition the Sphere with Free Group Orbits ....... 11
  5.4 Create a Weird “Orbit Representative Assembly” ... 13
  5.5 The Paradoxical Decomposition .................... 13

6 Why Did it Work? 14

1 Introduction

The Banach-Tarski Paradox (BTP) says, informally, that you can cut a sphere into finitely many pieces, and then reassemble the pieces via “rigid motions” to form two perfect copies of the original sphere. Clearly this is a paradox, for anyone with intuition about conservation of mass or volume.

In the notes that follow, we will demonstrate BTP, without skipping any details, proving everything except for one small detail, which we leave as an exercise. In Section 2, we do
some baby versions of BTP, by using rotations to remove points from circles and spheres. In Section 3, we perform BTP on an abstract group (the free group). In Section 4, we discuss the Axiom of Choice with a warm-up example on the circle. Finally, in Section 5, we put all the parts together, and show how we can transfer the BTP on the free group into the real thing, BTP on a sphere.

These notes were produced for my Spring 2009 Math 235 class, but were really written up because several of my colleagues wanted to see a short-but-complete treatment of BTP that was reasonably easy to read. I am indebted to Jonathan King, who discussed BTP as a special two-lecture bonus topic when he taught graduate analysis at U.C. Berkeley in the early 1990s. I sat in on those lectures and took careful notes, which I carried around for years. It made sense to flesh them out into something useful to people other than me. I also thank Stan Wagon for his excellent-but-not-easy-to-read book, The Banach-Tarski Paradox. It is the best, and practically the only book on the subject in English for non-specialists. But it is so complete that it is not easy to find what you want unless you already know a fair amount. So here’s my attempt to create a “one stop” BTP exposition. I hope you enjoy it.

1.1 Definitions

But first, we need some formal definitions, so that we know exactly what we are talking about.

**Definition 1.1 (Disjoint Union)** The symbol $\sqcup$ denotes disjoint union. Thus $A = B \sqcup C$ means that $A$ is the union of $B$ and $C$ and that $B \cap C = \emptyset$.

**Definition 1.2 (Rigid Motion)** A rigid motion is a translation, rotation, or composition of the two. In any given space (for example, the plane $\mathbb{R}^2$ or 3-dimensional space $\mathbb{R}^3$), the rigid motions form a group under composition. If $A$ is a set and $g$ is a rigid motion, then $g(A)$ denotes the image of $A$ under $g$. For example, suppose that $A$ is a square in the plane and $g$ is a rotation about the origin of 10 degrees counterclockwise. Then $g(A)$ is the rotated square.

**Definition 1.3 (Rigid-Motion Equivalence)** Let $A, B$ be sets (in the plane or 3-dimensional space). We say that there is a rigid-body equivalence between $A$ and $B$, written $A \sim B$, if $A$ can be partitioned into finitely many sets

$$A = A_1 \sqcup A_2 \sqcup \cdots \sqcup A_n,$$

and there are $n$ rigid motions $g_1, g_2, \cdots, g_n$ such that

$$g_1(A_1) \sqcup g_2(A_2) \sqcup \cdots \sqcup g_n(A_n) = B.$$

Informally, this says that $A$ can be cut into $n$ disjoint parts, and these parts can then be reassembled via rigid motions to form $B$.

Note that some of the motions $g_i$ could be the identity motion, which leaves all sets unchanged.
**Definition 1.4 (Paradoxical Decomposition)** Let \( A \) be a set (in the plane or 3-dimensional space). We say that there is a **paradoxical decomposition** of \( A \) if \( A \) can be partitioned into two subsets, each of which is rigid-body equivalent to \( A \). In other words, there exist sets \( B, C \) such that
\[
A = B \sqcup C
\]
and
\[
A \sim B, \quad A \sim C.
\]
Informally, this says that \( A \) can be cut into two disjoint parts, and each of these parts can then be cut up into finitely many pieces, and these pieces can be reassembled via rigid motions to form \( A \). Thus, starting with \( A \), we end up with two identical copies of \( A \)!

## 2 Removing Points with Rigid-Body Motions

Thus, the BTP itself says that there is a paradoxical decomposition of the solid unit sphere. We will accomplish this with a complicated algorithm which requires several steps. Some of our initial steps will involve reducing the sphere into slightly smaller sets, by removing inconvenient points. Here’s how we do this.

**Lemma 2.1 (Removing a Point)** Let \( S^1 \) denote the unit circle \( \{(x,y) : x^2 + y^2 = 1\} \) in the plane. Then
\[
S^1 \sim S^1 - \{(1,0)\};
\]
in other words, the circle can be cut and reassembled via rigid motions so as to lose a single point.

**Proof:** Pick an irrational number \( \alpha \), and let \( R_\alpha \) denote counterclockwise rotation about the origin \((0,0)\) by \( \alpha \) degrees. Let \( P = (1,0) \) be the point that we will soon remove. Now form the infinite “necklace”
\[
N := \{P, R_\alpha(P), R_\alpha(2 \alpha)(P), \ldots\}.
\]
Note that \( N \) is a countably infinite set of points. It is infinite because no two rotations of \( P \) can ever coincide (for if they did, this would force \( \alpha \) to equal 360 divided by an integer). Since \( S^1 \) has **uncountably** many points, \( U := S^1 - N \) is a non-empty set. In fact, it is an uncountable set!

Now we are ready for the rigid-body equivalence. Partition \( S^1 \) into \( N \) and \( U \), and then just observe that
\[
R_\alpha(N) = \{R_\alpha(P), R_\alpha(2 \alpha)(P), R_\alpha(3 \alpha)(P), \ldots\};
\]
in other words, when we rotate the entire necklace \( N \) by \( \alpha \) degrees, we lose exactly one point: the starting point, \( P \). Hence,
\[
R_\alpha(N) = N - \{P\}.
\]
Thus we have
\[
R_\alpha(N) \sqcup U = S^1 - \{P\},
\]
and we have reassembled the circle into a circle with a point removed.

Next, we will apply this idea in several different contexts.

**Lemma 2.2 (Removing a Point from a Sphere)** A solid sphere is rigid-body equivalent to itself minus any single point (which could be on the surface or in the interior of the sphere).

**Proof:** Let $P$ be the point you wish to remove. Consider any circle, lying entirely in or on the sphere, which contains $P$. Then repeat the process of Lemma 2.2; i.e., create an “irrational rotation necklace” $N$ which consists of a countable number of points on this circle, “starting” with $P$. Then, if you rotate $N$, it will lose $P$.

Next, we improve Lemma 2.2 to remove countable sets of points.

**Lemma 2.3 (Removing Countably Many Points from a Circle)** Let $A$ be any countable subset of $S^1$. Then $S^1 \sim S^1 - A$.

**Proof:** We use the same idea as before, but this time the entire set $A$ plays the role that the single point $P$ played in Lemma 2.2. Since $A$ is countable, we can enumerate its elements:

$$A = \{a_1, a_2, a_3, \ldots\},$$

where it is understood that the $a_k$ are measured as angles in degrees, since this will uniquely determine points on the unit circle. For example, we would denote the point $(0, 1)$ by 90.

Next, consider the differences of all of these values, i.e., the set

$$D := \{a_m - a_k : m, k \in \mathbb{N}\}.$$ 

Since $A$ is countable, so is $D$, since each difference $a_m - a_k$ can be put into a one-to-one correspondence with a lattice point $(m, k)$, and the lattice points are countable (remember the spiral?).

Then, we create an even more complicated set, by taking all the differences in $D$, and dividing each of them by $1, 2, 3, \ldots$, to get

$$W := \{(a_m - a_k)/n : m, k, n \in \mathbb{N}\}.$$ 

Once again, $W$ is countable, since it can be put into a one-to-one correspondence with the three-dimensional lattice points of natural numbers, which is countable.

Make sure that you understand the contents of $W$. For example, if the first few values in $A$ are $6, 2\sqrt{3}, e/\pi$, then $W$ contains, among other things, $6 - 2\sqrt{3}, (2\sqrt{3} - e/\pi)/653$, etc. Since $W$ is countable, and the angles of $S^1$ are the uncountable interval of real numbers between 0 and 360, there exist angles that are not in $W$ (in fact, uncountably many). So pick one such angle, and call it $\alpha$. That will be our magic rotation angle.

This angle works, because if we rotate any point in $A$ by $\alpha$, it is guaranteed not to coincide with any other point in $A$. For example, suppose a rotation moved $a_7$ into $a_{11}$. Then the rotation had to be the angle $a_{11} - a_7$. But $W$ deliberately contained all of the
differences between the \( a_k \) values, so one thing we are sure about \( \alpha \) is that it is not such a difference. So a rotation by \( \alpha \) will not take any point in \( A \) to any other point in \( A \).

Furthermore, if we rotate more than once, say 17 times, we still will not move any point in \( A \) to another point in \( A \). For example, if 17 rotations took \( a_3 \) to \( a_8 \), then \( \alpha \) would equal \((a_8 - a_3)/17\), and once again, \( \alpha \) was cleverly defined to not equal any difference of \( a_k \) values divided by any natural number.

Consequently, \( A \) and \( R_{\alpha}(A) \) are completely disjoint. You can think of \( R_{\alpha} \) as a copy of \( A \), just rotated by \( \alpha \). Likewise, \( R_{2\alpha}(A) \) is disjoint from \( A \) as well as \( R_{\alpha}(A) \), etc. Thus we can define the “super necklace”

\[
T := A \sqcup R_{\alpha}(A) \sqcup R_{2\alpha}(A) \sqcup \ldots
\]

Finally, we observe that

\[
R_{\alpha}(T) = R_{\alpha}(A) \sqcup R_{2\alpha}(A) \sqcup R_{3\alpha}(A) \ldots = T - A,
\]

so we have our construction: form the partition

\[
S^1 = T \sqcup (S^1 - T),
\]

and then

\[
R_{\alpha}(T) \sqcup (S^1 - T) = S - A.
\]

Finally, we will modify this construction to work with the three-dimensional sphere. We define \( S^2 \) to be the surface of the unit sphere; in other words,

\[
S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}.
\]

(The exponent is two, because the surface of a sphere is a 2-dimensional object, even though it is curved in space. If you are confused by this concept, take Topology next year!)

**Lemma 2.4 (Removing Countably Many Points from a Sphere)** Let \( A \) be any countable subset of \( S^2 \). Then \( S^2 \sim S^2 - A \).

**Proof:** We use exactly the same method as above, only now our angles are longitudes on the sphere, with respect to some axis that does not pass through any points of \( A \) (otherwise, the rotation would not move every point of \( A \)). The only wrinkle is guaranteeing that such an axis exists. But this is an easy consequence of the fact that \( A \) is countable, yet the surface of the sphere is uncountable.

Let \( U \) be the union of \( A \) and the set of the antipodes of the points of \( A \) (an antipode of a point \( P \) on a sphere is the point “directly opposite” \( P \); more precisely, the ray starting at \( P \) and passing through the center of the sphere hits the surface of the sphere at the antipode of \( P \)). Clearly \( U \) is countable (it is a union of two countable sets), so there are plenty of points left over on the surface of the sphere that are not in \( U \). Let \( Q \) be one such point. By the construction of \( U \), neither \( Q \) nor the antipode of \( Q \) are members of \( A \). Now we are free to use the axis joining \( Q \) with its antipode, and we proceed as we did with the circle. ■
3 The Free Group

Let \(a, a', b, b'\) be “letters.” We will form “words” by stringing finitely many of these letters together, subject to just two simplification rules:

1. \(aa' = a'a = e\), where \(e\) is the “null” word, shorthand for the word of zero length.
2. \(bb' = b'b = e\).

It is clear that these words form a group with respect to concatenation, as long as we include the zero-length word \(e\). Certainly \(e\) is the identity element, and \(a', b'\) are respectively the inverse of \(a\) and \(b\) (so we could write them as \(a^{-1}\) and \(b^{-1}\) if we wish). Clearly, every word has an inverse. For example, the inverse of \(aaaba'\) is \(abab'a'a'a'\).

It is understood that the words in this group are always reduced according to the grammar rules. We can use exponents to simplify notation, but otherwise, each word has a unique “spelling.” For example,

\[
(a^5b^{-1}a^3)(a^{-1}b^2)(b^{-3}a^2) = a^5b^{-1}a^2b^{-1}a^2.
\]

This group is called the Free Group on Two Generators, written as \(F[a, b]\) or just \(F\) for short. The idea is that \(a\) and \(b\) have infinite order and are “independent” of one another; there is no finite non-trivial word that equals \(e\).

Note that \(F\) has \(countably\) many elements. (WHY?)

We can picture \(F\) graphically, with \(a\) and \(b\) “axes.” A word is represented by a path starting at \(e\), moving right, left, up, down depending on whether the next letter is \(a, a', b, b'\), respectively. Figure 1 on p. 7 shows all words of length 4 or less. Note that in reality, the picture continues “fractally.”

3.1 Paradoxical Decomposition of the Free Group

The reason that the free group is important to us is because it has a paradoxical decomposition, as long as we relax the idea of “rigid-body motion” to include free group elements. We begin by partitioning \(F\) into four sets that roughly correspond to the four “leaves” of the group (notice how the group looks a little like a four-leaf clover). (See Figure 2 on page 8.)

\[
F = U \sqcup D \sqcup L \sqcup R,
\]

where

- \(U\), the “upper” tree, is the set of all words that start with \(b\). If we use the notation \([b\ldots]\) for all words beginning with \(b\), we can write \(U = [b\ldots]\).
- \(D\) is the set of words starting with \(b'\). In other words, \(D = [b'\ldots]\).
- \(L\) consists of the left tree, plus all the words that contain only the letter \(a\), plus \(e\). In other words,

\[
L = [a'\ldots] \cup \{a^k : k = 0, 1, 2, 3, \ldots\}.
\]
Figure 1: The Free Group. This figure shows all free group words with four or fewer letters. Note that the actual free group has an infinite number of elements, with the “trees” growing fractally.

- $R$ is everything left! More precisely, $R$ consists of all words that begin with $a$ that contain at least one other letter.

(Remember that words are reduced; we are not allowed to, say, put $a'$ right after $a$ for the spelling of a reduced word. If we could do so, then all words, for example, would begin with $a'$)

Now let’s apply free group “motions” to these sets. Consider $b'(U)$, defined to be $\{b'u : u \in U\}$. Every word $u \in U$ starts with $b^k$, where $k$ is positive, followed by a letter that is either $a, a'$, or $e$ (in which case the word ends). Thus $b'u$ will start with $b^{k-1}$, followed by $a, a'$, or $e$. If $k = 1$, then $b'u$ will be an arbitrary word starting with $a$ or $a'$. If $k > 1$, then $b'u$ will be a word starting with $b$. In other words, $b'(U)$ will consist of all of $U$, plus all the words starting with $a$ or $a'$; in other words, three of the the four trees of $F$! The only thing missing is $D$. Thus,

$$b'(U) \sqcup D = F.$$ 

Likewise, you should verify that

$$a'(R) \sqcup L = F.$$
In other words, we were able to partition $F$ into two sets ($U \sqcup D$ and $R \sqcup L$), and each of these sets was “free-group-motion” equivalent to $F$. So we have achieved a paradoxical decomposition, at least with respect to free group elements treated as “motions.”

### 3.2 A Geometric Model of the Free Group

You may be unimpressed by this, because the free group elements are just abstract words, and not geometric motions. But we will soon transfer this paradoxical decomposition onto a geometric arena.

First we will “put” the free group onto the unit spherical surface $S^2$. We will use the same letters as before, but give them different definitions. Imagine that the sphere is centered at the origin $(0,0,0)$. Let $a$ denote counterclockwise rotation about the $x$-axis by the angle $\theta$, and let $b$ denote counterclockwise rotation about the $z$-axis by the same angle $\theta$, where $\theta$ is defined by $\cos \theta = 1/3$.

It turns out that the rotations $a, b, a^{-1}, b^{-1}$ behave exactly like their free group analogues. In other words, every different finite reduced word containing these letters is distinct. For example, there are no relations such as $a^6 = e$ (which would be the case if, say, $\theta$
were equal to 60 degrees), or more complicated ones like $a^6 ba^{-11} b^4 = e$. In our geometric model, when we compose two rotations, we get another rotation.\footnote{The fact that the composition of two rotations on a sphere is another rotation is not obvious. It was first proven by Euler. See Tristan Needham’s Visual Complex Analysis for a good discussion of this.} So the collection of all of these finite “rotation words” is a group of rotations that is isomorphic to $F$. And it is a group of rigid-body motions!

This isomorphism is not obvious: it is a hard exercise in linear algebra to prove that every reduced word is distinct. We will not show this proof in this document. This is the only “gap” in our discussion of the BTP, though.\footnote{For a proof, and lots more information about BTP, see The Banach-Tarski Paradox, by Stan Wagon.}

Subsequently, when we refer to $F$, we mean the free group of rotations, not the free group of words. They are, after all, the same thing, essentially.

But first, we will take a detour and look at a subtle thing that happens with some infinite sets.

4 The Axiom of Choice

BTP requires the Axiom of Choice (AoC). This axiom states that given a collection of sets (possibly infinite; possibly uncountably infinite!), one can choose a representative from each set. In more formal language, if $S_\alpha$ is a collection of sets, where $\alpha$ ranges over a possibly infinite “index” set $I$, there exists a function $f$ whose domain is $I$ and whose range is the union of all $S_\alpha$, in other words, there exists

$$ f : I \to \bigcup_{\alpha \in I} S_\alpha, $$

such that for each $\alpha \in I$, $f(\alpha) \in S_\alpha$. (We use $\alpha$ in the subscript to emphasize that $\alpha$ may not just be an ordinary integer, but instead an index from an uncountable set.)

Here is a simple example: Let $I = \mathbb{N} = \{1, 2, 3 \ldots \}$ and for each $\alpha \in \mathbb{N}$, define $S_\alpha = \{\alpha n : n \in \mathbb{N}\}$. In this case, each $\alpha$ is just a positive integer, and $S_\alpha$ is the set of positive multiples of $\alpha$. Then AoC merely means that it is possible to pick a multiple of $m$ for each positive integer $m$. That’s easy: one way would be $1, 2, 3, 4, 5, \ldots$. Another is $1, 4, 9, 16, 25, \ldots$, etc. These are both explicit constructions, so we really didn’t need an axiom to get them!

Here is another example, with an uncountable index set: Let $I = \mathbb{R}$. For each $m \in \mathbb{R}$, let $S_m$ denote the set of all lines in the plane with slope $m$. Then the axiom of choice says that we can choose lines with every possible real slope. Again, we can construct this explicitly (try $y = mx$ for each $m \in \mathbb{R}$) and do not need to rely on a dumb axiom.

4.1 A “Non-measurable” set

But sometimes, we need the axiom, because we cannot form an explicit construction. Here is a much nastier example, the “Vitali decomposition of the circle.” We will use AoC to produce a near-paradox.
Consider a circle $C$ with circumference 1 unit. Designate a point on this circle by 0. Now consider the points on the circle whose distance to 0 along the circumference is a rational number. Since the rationals are countable, there are countably many such points; call them $\{r_1, r_2, r_3, \ldots\}$. This is just the set of nonnegative rational numbers less than 1.

Now, for each point $x \in C$, consider the set of points $Q_x$ which are a rational distance (along the circumference) from $x$. For example, $Q_0 = \{r_1, r_2, r_3, \ldots\}$. In fact, for any $x \in C$, $Q_x = \{x + r_i, i = 1, 2, 3, \ldots\}$, where it is understood that we perform addition “modulo one;” for example, $0.81 + 0.34 = 0.15$.

Therefore, each $Q_x$ is a countable set, yet $C$ is uncountable. You should recognize that each $Q_x$ is an equivalence class with respect to the equivalence relation $\sim$ defined by $x \sim y$ if and only if $x - y$ is rational. In other words, we could write $Q_x = [x]$. (Carefully verify that $\sim$ is indeed an equivalence relation.)

Thus, $C$ can be partitioned by these equivalence classes $Q_x$. In other words, there is an index set $I$ such that $C = \bigsqcup_{\alpha \in I} Q_\alpha$.

Now use AoC to select a single point from each distinct $Q_x$. Call the set of selected points $W$. Notice the following things about $W$.

1. $W$ is uncountable. This is because each $Q_x$ is countable, so there has to be uncountably many of them, or otherwise $C$ would be a countable union of countable sets, but we proved that that is countable. So since there are uncountably many distinct $Q_x$ sets, our selection $W$ will be uncountable. In other words, the index set $I$ above is uncountable.

2. For each $r_i$, the set $W + r_i = \{w + r_i : w \in W\}$ (in other words, $W$ “translated” by a rational distance $r_i$) is disjoint from $W$. Here’s why: suppose there was a point $y$ in both sets. Then $y \in W$ and $y \in W + r_i$. That would mean that $y = w + r_i$ for some $w \in W$, which would mean that $y$ and $w$ are a rational distance apart, which means that $y \in Q_w$. But we explicitly constructed $W$ so that it has a single point from each $Q_x$.

3. In fact, each point $x \in C$ is a member of $W + r_i$ for some natural number $i$. In other words, every point in $C$ is a rational distance from a point in $W$. The reason: either $x$ was selected in the construction of $W$ (in which case $x \in W + 0$), or else $x \in Q_y$ for some $y$. Now $W$ contains a single member from $Q_y$, call it $z$. Then $x$ is a rational distance from $y$, and $z$ is a rational distance from $y$, so $x$ is a rational distance from $z$.

4. Therefore, $C$ can be expressed as the disjoint union $C = (W + r_1) \sqcup (W + r_2) \sqcup (W + r_3) \sqcup \cdots$. 
5. Here’s the near-paradoxical part: What is the “measure” (length) of $W$? If it is zero, then surely each $W + r_i$ also has zero length, since it is just a translate of $W$. Then $C$ can be expressed as a countable union of zero-length sets. Yet $C$ has length 1. If the length of $W$ is a positive value $p$, then surely $W + r_i$ also has length $p$, and we have expressed $C$ as a countable union of sets with length $p$, which would add up to an infinite length. In other words, the existence of $W$ forces us to think of $C$ as either having zero or infinite length, or else to give up the idea that a set (such as $W$) can have a defined length! In fact, such sets are called non-measurable.

A key part of BTP, as you will see, is a weird unmeasurable set, similar to $W$ above, but far more intricate, produced by selecting a single point from each of an infinite collection of sets. Since the selection is not something that can be explicitly constructed, we will need AoC.

5 Step-by-step BTP

We are now ready to do the BTP. We will take the solid unit ball $D^3$, partition it into two parts, and show that each part is rigid-body equivalent to $D^3$.

5.1 Remove the Center

Start with $D^3$, the solid unit ball. Use Lemma 2.2 to remove the center with rigid body-motions.

5.2 Move to the Surface

Now that the center is gone, we can move all of our work to the surface $S^2$ of the ball. Here’s why: Suppose we have a set $A$ lying entirely in the surface. Now consider the “cone” of points in the ball consisting of every ray emanating from the origin to each point of $A$. Every set on the surface has a corresponding cone. The only problem is that disjoint sets on the surface do not have disjoint cones—every cone includes the center of the sphere. But we removed the center!

So now, we will just work on the surface $S^2$, knowing that whatever sets we manipulate on the surface induces analogous set manipulations with “conical” subsets of $D^3$ minus its center. So if we can paradoxically decompose $S^2$, then we can paradoxically decompose $D^3$ minus its center, and thus we can paradoxically decompose $D^3$ itself.

5.3 Partition the Sphere with Free Group Orbits

Let $x$ be an arbitrary point on $S^2$. Consider the orbit of $x$ with respect to $F$, defined by

$$F(x) := \{ r(x) : r \in F \}.$$  

Remember that $F$ consists of countably many rotations, so $F(x)$ will be a set of countably many points, namely the images of $x$ under each of these rotations. Each orbit $F(x)$ is a
“copy” of $F$ in a sense. If you tried to visualize an orbit, it would appear to cover the entire surface, but it is only a countable subset of the surface.

There is one technical detail: for some points $x$, the orbit $F(x)$ is not an “exact” copy of $F$, because $x$ may be a fixed point for one or more particular rotation. For example, the “north pole” $(0, 0, 1)$ is fixed by the rotation $b$, since this rotates about the $z$-axis. In fact, the north pole is fixed by every power of $b$. Each rotation in $F$ has an axis, and hence has 2 fixed points. These fixed points, it turns out, are “bad.” So define

$$B := \{ x \in S^2 : \text{there exists } r \in F, r \neq e, \text{ such that } r(x) = x \}$$

to be this set of bad fixed points. Note that $B$ is countable. By Lemma 2.3, we can remove $B$! (More precisely, $S^2 \sim S^2 - B$).

So now, we will no longer work with $S^2$, but with its slightly smaller cousin $S^2 - B$. Removing the pesky fixed points means that for each $x \in S^2 - B$, the points in the orbit $F(x)$ are in a one-to-one correspondence with the rotations of $F$. In other words, if $r$ and $t$ are two different rotations in $F$, then $r(x)$ and $t(x)$ are different points in $S^2$. To see this, suppose that $r(x) = t(x)$. Then $r^{-1}t(x) = x$. But since $x$ is not a fixed point of any non-identity rotation, the only possibility is that $r^{-1}t = e$, which means that $r = t$. (Notice that this argument depends on removing the bad set $B$ from $S^2$.)

Now we define an orbit-equivalence relation. For points $x, y \in S^2 - B$, define $x \sim y$ whenever $x$ and $y$ lie on the same orbit. If $x \sim y$, then there exist rotations $r, t \in F$ such that $r(x) = t(y)$. Thus, $x = r^{-1}t(y)$. Since $r^{-1}t$ is just another element of $F$, we can say that $x \sim y$ means that there exists $u \in F$ such that $x = u(y)$. It is easy to verify that $\sim$ is an equivalence relation:

- For any $x \in S^2 - B$, it is certainly true that $x \sim x$, since the orbit $F(x)$ includes $x$ (because $x = e(x)$).
- If $x \sim y$, then there exists $u \in F$ such that $x = u(y)$. Thus $y = u^{-1}(x)$, which means that $y \sim x$.
- Transitivity is an exercise for you.

What are the equivalence classes? The orbits, of course! By the definition of equivalence class, if $x \in S^2 - B$, we have

$$[x] = \{ y \in S^2 - B : y \sim x \} = \{ y \in S^2 - B : \text{there exists } u \in F \text{ such that } y = ux \} = \{ ux : u \in F \} = F(x).$$

Each equivalence class is a countable set, since $F$ is countable. By the theorem that we proved in class relating equivalence classes to partitions, we conclude that

$S^2 - B$ can be partitioned into disjoint orbits.

---

3I know that $\sim$ is also used for rigid-body equivalence, but it should be clear from context which we are using. Rigid-body equivalence is for sets, but this equivalence is between points.
5.4 Create a Weird “Orbit Representative Assembly”

Now we will mimic the ideas developed in Section 4 to create a weird non-measureable set. We have shown that $S^2 - B$ can be partitioned into disjoint orbits. We may imagine a sphere divided up by rubber bands that miraculously do not overlap, but in reality the orbits are far more complex sets, more like clouds of dense points that seemingly completely cover the sphere. Notice that there are uncountably many orbits, since $S^2 - B$ is uncountable and each orbit is countable (so if there were only countably many orbits, $S^2 - B$ would be a countable union of countable sets, and hence countable.

Using the Axiom of Choice, pick a single representative from each disjoint orbit. Call this set of orbit representatives $W$. Clearly $W$ is an uncountable set. By the definition of $W$, if $x, y$ are two distinct points in $W$, then the orbits $F(x)$ and $F(y)$ must be disjoint. This must be so, because orbits are either disjoint or completely the same, and if they are the same, then $x \sim y$. But $W$ is constructed explicitly so that no two points in $W$ lie on the same orbit.

Therefore, if we form the orbit of each point of $W$, we will recover all the disjoint orbits of $S^2 - B$. Thus the orbit of $W$ is all of $S^2 - B$. In other words,

$$F(W) := \{F(w) : w \in W\} = S^2 - B.$$ 

5.5 The Paradoxical Decomposition

Make sure you understand the notation: $F$ is a set of rotations, $W$ is a set of points, and $F(W)$ is the set of the images of each rotation applied to each point in $W$. So for any set $V$ of rotations, $V(W)$ would be the set formed by taking every rotation in $V$ and applying it to every point in $W$. In other words,

$$V(W) = \{r \in V : r(W)\} = \{V(w) : w \in W\}.$$ 

The reason that we are spending time on this notation is because we want to see what happens with some of the subsets of $F$ that we have already investigated. Remember that we partitioned $F$ into the four sets

$$F = U \sqcup D \sqcup L \sqcup R.$$ 

Let us now define

$$U' := U(W), D' = D(W), L' = L(W), R' = R(W).$$ 

For example, $D'$ is equal to the set of points that are images of all the rotations of $D$, applied to all the points of $W$. Since $U, D, L, R$ are disjoint, so are the sets $U', D', L', R'$. And furthermore,

$$U' \sqcup D' \sqcup L' \sqcup R' = S^2 - B.$$ 

This is so, because the left hand side is the same as $(U \sqcup D \sqcup L \sqcup R)(W)$, which is just $F(W)$. 

But recall that \( b^{-1}U \sqcup D = F \). This immediately yields (make sure you understand each step!)

\[
\begin{align*}
    b^{-1}U' \sqcup D' &= b^{-1}U(W) \sqcup D(W) \\
                      &= (b^{-1}U \sqcup D)(W) \\
                      &= F(W) \\
                      &= S^2 - B.
\end{align*}
\]

In other words, we applied \( b^{-1} \) to \( U' \) and took the union of this with \( D' \), and we got \( S^2 - B \). But \( U' \) and \( D' \) are just sets of points, and \( b^{-1} \) is a very simple rigid-body motion (a rotation about the \( z \)-axis). What we have shown is that

\[
\text{\( U' \sqcup D' \) is rigid-body equivalent to \( S^2 - B \).}
\]

And this is the real thing: not some abstract group operation, but an actual simple rotation. And since we also have \( a^{-1}L \sqcup R = F \), we can immediately realize this on the sphere and conclude that

\[
\begin{align*}
    a^{-1}L' \sqcup R' &= a^{-1}L(W) \sqcup R(W) \\
                       &= (a^{-1}L \sqcup R)(W) \\
                       &= F(W) \\
                       &= S^2 - B.
\end{align*}
\]

Thus

\[
\text{\( L' \sqcup R' \) is rigid-body equivalent to \( S^2 - B \).}
\]

Putting these together, we conclude that

\[
\text{There is a paradoxical decomposition of \( S^2 - B \).}
\]

And since \( S^2 - B \) is rigid-body equivalent to \( S^2 \), that means that there is a paradoxical decomposition of \( S^2 \). But that means that there is a paradoxical decomposition of \( D^3 \) minus the center (using conical sets). But since \( D^3 \) minus its center is rigid-body equivalent to \( D^3 \), that means, finally, that there is a paradoxical decomposition of \( D^3 \).

We have established the Banach-Tarski Paradox!

\[
\square
\]

6 Why Did it Work?

Your intuition should rebel at this point. How in the world can you partition a sphere so that the parts can be reassembled via rigid-body motions to form two spheres? It’s impossible. Yet we did it, and the rotations used were very simple ones.

But to actually make it happen, we needed to construct the weird set \( W \), which is non-measurable, and quite bizarre. You cannot actually construct it. You can only theoretically deduce that it exists (as long as you accept the Axiom of Choice). At this point, you
probably want to denounce the Axiom of Choice. But as you continue in your mathematical education, you may find that it’s just too much trouble to avoid it.

If that makes your head spin, that’s OK. These fundamental issues of mathematics are not simple. But they are worth pondering!