6-2017

Choice of Choice: Paradoxical Results Surrounding of the Axiom of Choice

Connor Hurley
Union College - Schenectady, NY

Choice of Choice:
Paradoxical Results Surrounding of the Axiom of Choice

Connor Hurley

March 15, 2017

Abstract

When people think of mathematics they think "right or wrong," "empirically correct" or "empirically incorrect." Formalized logically valid arguments are one important step to achieving this definitive answer; however, what about the underlying assumptions to the argument? In the early 20th century mathematicians set out to formalize these assumptions, which in mathematics are known as axioms. The most common of these axiomatic systems was the Zermelo-Fraenkel axioms. The standard axioms in this system were accepted by mathematicians as obvious, and deemed by some to be sufficiently powerful to prove all the intuitive theorems already known to mathematicians. However, this system wasn't without controversy; Zermelo included the Axiom of Choice to prove his Well Ordering Theorem. This led to unintended consequences. Imagine taking a solid, three-dimensional ball and breaking it apart into certain finite pieces. Instinctively, one would agree that no matter how these pieces are rotated, when you put them back together you should have the same ball. Surprisingly the Axiom of Choice tells us this isn't the case, that there is a way to put these pieces back together and have two identical copies of the original ball. Delving further, one can start with something the size of a pea, and after specific rotations, end up with a ball the size of our sun. The Axiom of Choice also lets us conclude that there is a way to predict the future correctly at almost any point in time. However, as many an incorrect weatherman will tell you, this too goes against what we believe. So how does one reconcile his or her concept of what's true and what the Axiom of Choice tells us to be true? Do we simply take away the Axiom of Choice? As you may expect, the answer isn't quite so simple. . .

1 Introduction

1.1 Background

When one constructs a set, the idea of having a set be a member of itself should seem preposterous. Russel's Paradox is a way of formulating a set such that this is exactly the case. Because of paradoxes such as this, in the early 20^{th} century mathematicians set out to construct a unified set of principles. The first system was formed by Ernst Zermelo, and was called Zermelo set theory. This was later refined to form todays standard set of axioms, known today as Zermelo-Fraenkel set theory. The basic idea behind the formation of an axiom system is that each individual axiom should be an easily acceptable idea which allows the proof of all "obvious" results.

When Zermelo first published his set theory system he also included a three page proof which would see him earn instant fame and a professorship at a prestigious institution [3]. Zermelo's Well Ordering Theorem says that every set can be well ordered, an extraordinary result when looking at sets such as the set of all real numbers. In order to prove this Zermelo made use of something called the Axiom of Choice.[1] At first glance, the statement of the Axiom of Choice is fairly innocuous.

Definition 1.1. The Axiom of Choice is the following assertion:
For a collection of nonempty sets K_i their product set is nonempty

An easier way to say this is that for any collection of nonempty sets, there is a method to choose an element out of each set. Once again this may not seem like a powerful tool, especially in the finite. Imagine if I asked you to define a choice function on a collection of pairs of shoes. One could easily define such a function by selecting all of the left shoes. But what if I asked you to do the same with pairs of socks? Or even worse infinite collections of infinite socks [3]? It is in situations like this where the Axiom of Choice becomes controversial.

Indeed, when we are looking at a finite collection of sets there is no need for the Axiom of Choice. It is only when we move to the infinite that the Axiom of Choice's validity becomes hazy to the observer. This phenomena was well known enough to prompt the following famous quip from Jerry Bona: "The Axiom of Choice is obviously true, the well-ordering principle obviously false, and who can tell about Zorn's lemma?" [5] [2] While Bona was simply saying this to make a point, there are well-known occurrences of controversy surrounding the Axiom of Choice. When Alfred Tarski attempted to publish his result of the equivalence of the Axiom of Choice to the statement that every infinite set A has the same cardinality as $A \times A$ in the *Comptes Rendus* he was turned down by both editors, Maurice Fréchet and Henri Lebesgue. Frénchet told Tarski that the implication between two well-known results is not worthy of publication, while Lesbegue told Tarski that the implication between two false implications wasn't of interest. [4]

While Zermelo's Well Ordering Theorem seems to be incredibly counterintuitive, it is not the most paradoxical result following from the Axiom of Choice. The most famous of these results is the Banach-Tarksi Paradox which says that one can take a solid ball and break it into finitely many pieces which can then be rotated to form two copies of the original solid ball. How can we reconcile the existence of such a paradoxical result and the Axiom

[1] Zermelo's Well Ordering Theorem was later proved to be equivalent to the Axiom of Choice.

[2] Bona was obviously joking as when he said it it, the equivalence of the three results was already well-known.

of Choice, which is used to prove many of the most important theorems of set theory? Those who looked to the other axioms of ZF to prove the validity of the Axiom of Choice were destined to be disappointed. First, Kurt Gödel proved that the Axiom of Choice was consistent with the rest of ZF. Cohen followed this proving that the negation of the Axiom of Choice was also consistent with ZF. This allowed mathematicians to explore other systems created by weakening the Axiom of Choice. While these weaker systems eliminate many of the counterintuitive theorems formed by using the Axiom of Choice, they themselves are not free of their own paradoxical results.

1.2 Set Theory Preliminaries:

While this thesis is intended to be written at a very accessible level, there are still a couple of general definitions which need to be covered. First, for a set A we will denote the cardinality of A as $|A|$. Cardinality can be seen as the size of a set, and when dealing with finite sets it is a fairly straight forward concept. If I count 5 elements of a set, then a set has a cardinality of 5. When we move the the infinite however, it becomes a little more complicated. For this reason we will say that two sets A and B have the same cardinality if there exists a bijective function between them. This is allows us to compare the size of two infinitely large sets without changing our basic interpretation of the size of finite sets.

As is standard, we will denote the set of natural numbers by $\mathbb{N}$, the set of rational numbers by $\mathbb{Q}$ and the set of real numbers by $\mathbb{R}$. Without venturing too far into the discussion of Cantor's Continuum Hypothesis, we will denote the infinite cardinals as indexed by the natural numbers using $\aleph_i$, and say that $|\mathbb{N}| = \aleph_0$.

Later we will make use of of cosets of $\mathbb{Q}$ in $\mathbb{R}$.

Definition 1.2. We will define a **coset of** $\mathbb{Q}$ **in** $\mathbb{R}$ to be a set of the form

$$x + \mathbb{Q} \text{ where } x \in \mathbb{R}$$

We will denote the set of all these cosets as $\mathbb{R}/\mathbb{Q}$.

Remark 1.3. The set of cosets of $\mathbb{Q}$ in $\mathbb{R}$ yield a partition of $\mathbb{R}$. These cosets also form an equivalence relation with $x \sim y$ if and only if $x - y \in \mathbb{Q}$.

We can also see that each coset $x + \mathbb{Q}$ is dense in $\mathbb{R}$ due to the density of $\mathbb{Q}$ in $\mathbb{R}$, so every open interval (a, b) in $\mathbb{R}$ contains at least one element from $x + \mathbb{Q}$.

In order to see this let (a, b) be an interval in $\mathbb{R}$. Then for some $x \in \mathbb{R}$ let $x + \mathbb{Q}$ be a coset of $\mathbb{Q}$ in $\mathbb{R}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$ we know that every open interval $(c, d) \subseteq \mathbb{R}$ contains at least one rational number q. Thus we know that $(a - x, b - x)$ contains some rational number q'. But then $x + q' \in (a, b) \cap (x + \mathbb{Q})$, so each coset of $\mathbb{Q}$ in $\mathbb{R}$ is dense in $\mathbb{R}$.

We will also need to dispense with a definition of Lebesgue Measure.

Definition 1.4. If $X \subset \mathbb{R}$ then the **outer measure** of X, denoted $\mu^*(X)$ is given by

$$\mu^*(X) = inf \sum_{j=0}^{n} length(Kj)$$

With $\langle K_j \rangle$ a series of intervals covering X.

Definition 1.5. If $X \subseteq \mathbf{R}$ then X is **Lebesgue measurable** iff $\forall \epsilon > 0, \exists$ a closed set F and an open set G such that

1. $F \subseteq X \subseteq G$
2. $\mu^*(G - F) < \epsilon$.

In this case $\mu^*(X)$ is renamed $\mu(X)$ and is called the Lebesgue measure of X.

Remark. Properties of Lebesgue Measure
i) For a set A, if $|A|$ is countable then $\mu(A) = 0$.
ii) If $A \subseteq B$ then $\mu(A) \leq \mu(B)$.
iii) Lebesgue measure is finitely additive, that is if $A_1, A_2, ..., A_n$ are measurable and pairwise disjoint then $\Sigma_{j=1}^{n} \mu(A_j) = \mu(\uplus_{j=1}^{n} A_j)$.
iv) Lebesgue measure is countably additive, which is the same except for a countable number of A_i.

Finally, we will use $\uplus$ to denote a disjoint union of two sets. Operationally it is the same as a standard set theory union; however, $\uplus$ indicates that the two sets have an empty intersection.

These definitions will be the foundation for our various sections. While we will need additional definitions, we will examine those in the section in which they are required.

2 Hat Problems

2.1 Introduction

One of the areas with the most obvious use of the Axiom of Choice is hat problems. The basic theme of hat problems is that there are a certain number of agents, all of which have a hat assigned to them, which they can't see. Based on the information provided to them they try to guess the color of the hat on top of their head. This abstract explanation of hat problems may seem a bit confusing, so we will examine a more concrete example.

Consider the simple example of a game played with two "prisoners" (agents) and a prison guard. These two prisoners will have a hat of one of two colors placed on their head, and if one of them can guess the color of the hat on their head correctly, both can go free. Now each prisoner, if randomly guessing, has a 50% chance of guessing the correct color hat on his or her head, which leads to a 75% chance that they will go free; however, is there a better strategy with a higher level of success? In this case there is. First, one player assumes that the hat color he sees on the opposite prisoner is the same as the hat color that he is wearing, while the other prisoner assumes that the hat color he is wearing is different from the one he observes on the other player. Following this strategy, there is no designation of hat colors that would result in neither of the prisoners guessing right (interestingly there is also no scenario where both guess right).

Another example is when there are 10 prisoners all in a line so they can see all the hat colors in front of them. Again, there are two possible selections of hat colors which can be placed on a prisoner's head, either red or green. If the warden starts at the back and works his way forward asking each prisoner what they think their hat color is, is there an optimal solution? There is such an optimal solution; however, there is no strategy that guarantees the correct guess of everyone involved (the first prisoner to guess must choose randomly). The strategy is that the prisoner in the back counts the number of red hats they observe. If there is an odd number of reds he or she would guess green. Then the prisoner in front knows the plurality of reds both in what he or she can see as well as the plurality including the hat on their own head. Thus, if he or she hears the first prisoner guess green and only sees an even number of red hats, he or she would know his own hat is red so he would guess red, which would also indicate to the prisoner in front of him that he or she sees an even number of red. This strategy guaranties that at least nine players will guess right.

As you can see, there are many types of hat problems, and by extending hat problems to the infinite we can see many more possible scenarios. However, when transitioning to infinite hat problems we run into a problem: we need to use the Axiom of Choice, which also brings in any other controversial results that may come with it. We will leave the discussion of the need of the Axiom of Choice for a later chapter after exploring more of the pros and cons associated with it.

2.2 Preliminaries:

Before we get started by looking at more complicated examples of hat problems it will be helpful to establish some general-hat theory notation; however, we will first have to briefly cover some general set theory definitions.

2.2.1 Set Theory Definitions:

Definition 2.1. Let X and Y be sets. Then the **symmetric difference** of X and Y, denoted $X \triangle Y$, is equal to $(X - Y) \cup (Y - X)$. When dealing with functions f and g, we will use this notation in a similar way. In that case $f \triangle g$ =$\{x : f(x) \neq g(x)\}$.

Definition 2.2. Let A be a set. Then we say that A is **cofinite** if A^c is finite.

2.2.2 Hat Problem Specific Preliminaries

These definitions and theorems come courtesy of Professor Chris Hardin and Professor Alan Taylor. [2]

Since we will be dealing with situations much more complex than the finite examples presented at the start of this section it will be helpful to lay down some generalized terminology. In the first example we presented there were two "prisoners"; however, in our more generalized cases we will call these "prisoners" agents and we will let A denote the set of agents. Also, in our first example the prisoners had to choose between either red or green. From here on out we will let K denote the set of hat colors and we assume that $|K| \geq 2$. If K is not uniform we will let K_a denote the set of hat colors for agent a.

Next we will define a visibility graph on A. For $a \in A$, $V(a)$ is simply the set of agents of that agent a can see; thus $V(a) \subseteq$A. Next, for every game there is a number of "strategies" which the "warden" can play. We will call these "strategies" **colorings** and let $C = \{f : f \text{ is a function and} f : A \to K\}$ denote the set of colorings. Next, we will define when two colorings are equivalent to an agent a. That is $f(a) \equiv_a g(a)$ if and only if $f(b) = g(b)$ $\forall b \in V(a)$. We will denote the set of all functions $f : A \to K$ by ${}^A K$.

Next, for an individual agent a we can define a **guess function**. This guess function is the strategy that agent a will employ based on what he or she observes. We will denote this guess function G_a as a function $G_a : \to K$ where if $f \equiv_a g$, then $G_a(f) = G_a(g)$. We will say this is the correct guess if $G_a(f) = f(a)$. Using this we can define a predictor function, $P : C \to C$ where $P(f)(a) = G_a(f)$.

2.3 The Gabay-O'Connor Theorem

Consider the two prisoner example we gave at the start of this chapter. We showed how there is a strategy ensuring that at least one prisoner guesses his or her hat color correctly, but we also noted that this strategy ensured that both didn't guess correctly. This isn't a coincidence, as it can be proven that for a finite set of agents, each choosing between two hat colors, you cannot guarantee the success of more than 50%.

But what about when we move over to the infinite? With infinitely many agents can we find a way for the agents to be assured of a success rate better than 50%? It turns out that we can. Consider the case where we have an infinite number of agents and two hat colors: red or green. We will also specify that there is complete visibility, meaning that $V(a) = A - \{a\}$. With this in mind consider the following strategy. If agent a sees an infinite number of red hats he will guess red, otherwise he will guess green. In the countable case, there will always be an infinite number of agents who guess correctly, because if an infinite number of red hats can be observed by one agent, an infinite number of red hats can be observed by all agents, so every agent will guess red. If the agents don't see an infinite number of red hats they will all guess green. Since there are only a finite number of red hats there must be an infinite number of green hats, so an infinite number of agents will guess correctly.[3]

Note that while this strategy ensures that an infinite number of agents guess correctly it doesn't necessarily guarantee that there aren't an infinite number of agents that don't guess correctly. What if we want to minimize the number of agents that guess incorrectly? Such strategies actually do exist and are referred to as finite-error predictors. By using the Axiom of Choice there are multiple results which tell us that that this is possible.[4] However, the Gabay-O'Connor theorem tells us that there is a finite-error predictor for an arbitrary number of agents and hat colors. This result seems counterintuitive when considering the odds of agents guessing correctly at random. Even with a countably infinite number of hat colorings, the odds of randomly guessing the correct answer is 0! [5] Yet, somehow there is a strategy that guarantees only finitely many guess correctly.

Main Theorem 2.3. ***The Gabay-O'Connor Theorem***

Let the set A of agents be arbitrary, the set K of colors be arbitrary, and assume every agent sees all but finitely many of the other hats. Then there exists a predictor ensuring that all but finitely many agents guess correctly.

In order to prove this we will define an equivalence relation on the set of guesses. First let f and g be in ${}^AK$. Then we will say that if $f \triangle g$ is finite then $f \approx g$. Since $f \triangle f = 0$, clearly this relationship is reflexive. We can also clearly see that $f \triangle g = g \triangle f$ giving us transitivity. Finally for a finite m and n, if $|f \triangle g| = n$ and $|g \triangle h| = m$, we know that $|f \triangle h| \leq m + n$, so we have an equivalence relation on ${}^AK$.

Now by the Axiom of Choice we know that we can choose an element from each equivalence class. Formally put, we have a choice function Φ over these equivalence classes. This means that for each $f \in^A K$, $\Phi(f) \approx f$, and if $f \approx g$, then this choice function will select the same distinct element of their equivalence class, so $\Phi(f) = \Phi(g)$.

Let's consider what this tells us. Since each agent can see all but finitely many hats he or she can tell which equivalence class he or she is in. If he or she couldn't, what he or she

[3] In fact, in this second case cofinitely many agents guess correctly.

[4] See Hardin and Taylor 2013

[5] Can easily be proven by using the sequence $\langle f(n) \rangle = \frac{1}{n} \to 0$, with n=$|K|$. In fact, for any "reasonably definable" probability question on ω the probability is either 0 or 1. This property is called *ergodicity*.

observed would differ from the true coloring by an infinite number of agents, which can't be the case. Thus if we define the guess function to be our choice function Φ, all agents will guess based on the same equivalence class. To formalize this let's let h be our hat coloring. Then each agent a will be able to tell that he or she is in $[h]$, so $G_a(h) = \Phi(h)(a)$. Thus, each agent's guess will be the same, and more importantly, only differ from h at finitely many points, as desired.

The proof of the Gabay O'Connor Theorem clearly uses the Axiom of Choice. There are proofs that the Gabay-O'Connor theorem is independent of the weaker forms of the Axiom of Choice such as the Axiom of Dependent Choice. Perhaps more interestingly, the equivalence of the Gabay-O'Connor Theorem and Axiom of Choice is still an open question.[6]

2.4 Lenstra's Theorem

Let's look at an extension of the Gabay-O'Connor Theorem. We can take the Gabay-O'Connor Theorem and use it to create a finite error predictor **P**. There is a theorem stating that we can't guarantee more than 50% of agents guess correctly, we know that the Gabay-O'Connor is, in some sense, a maximal theorem. [2] Suppose this wasn't the case and there was a finite error predictor so that for some $n \in \mathbb{N}$, we could guarantee that fewer than n agents would guess incorrectly. Then if there were $2n+1$ agents, this would require that $n+1$ agents guess correctly, a contradiction. But if we extend the Gabay-O'Connor Theorem without guaranteeing a certain number of correct guesses we can arrive at Lenstra's Theorem, as stated below.

Theorem 2.4. ***Lenstra's Theorem:***

Consider the situation where the set of agents is arbitrary, and $|K| = 2$. Let there be full visibility. Then there exists a strategy such that either every agent will guess correctly or every agent will guess incorrectly.

Proof. Let h be a hat coloring. First we will use the Gabay-O'Connor Theorem to generate a predictor **P**. As we said before, for each agent a, changing a finite number of agents hat colors doesn't change the an agents guess, so every agent in A will be able to tell $G_a(h)$. Further, they know that only a finite number of these guesses will be incorrect. Let D be the set of agents that guess incorrectly, and let the set of agents a observes to guess incorrectly as D_a. Using this we will define a new predictor, Z by $Z_a(h) = G_a(h)$ if and only if $|D_a|$ is even.

Lets consider why this works. First, we know by the Gabay-O'Connor Theorem that D is finite so it has either even or odd cardinality. Suppose D is even. Then if a guessed correctly she can see all incorrect guesses, so she will observe an even number of incorrect guesses and will keep her guess the same. If a guessed incorrectly then $|D_a| = |D| - 1$, an odd number. Then agent a will change her guess, and guess correctly. If every agent follows this process simultaneously then every agent will guess correctly.

Next, suppose D is odd. Then for all a who guessed correctly $|D_a|$ will be odd, thus they will change their guess and all guess incorrectly. Next, all agents who guessed incorrectly will see an even number of incorrect guesses and will keep their incorrect guess. This means that every agent will guess incorrectly, as desired. $\square$

[6]For a full discussion of these results see Hardin and Taylor 2013

2.5 Weak Gabay-O'Connor Theorem

While Lenstra's Theorem is interesting, it is a lot less powerful than the Gabay-O'Connor Theorem. We will return to the Gabay-O'Connor Theorem, but let's consider the opposite. What if we want to have a strategy that assures us that all but finitely many agents will guess *incorrectly*? This result was given by Professor Alan Taylor of Union College in a handout, which he called the Weak Gabay-O'Connor Theorem. At a glance it would seem that the Gabay-O'Connor Theorem and the Weak Gabay-O'Connor Theorem would be essentially philosophically equivalent.

Theorem 2.5. ***Weak Gabay-O'Connor Theorem***
Regardless of what the set A of agents is and what the collections K_a are (provided $|K_a| \geq 2$), there is a strategy ensuring that for every assignment of colored hats, only finitely many agents guess correctly.

We can easily see how this can be proved using the Gabay-O'Connor Theorem. First, assume that for every agent $a \in A$, K_a has at least two elements. Next, for a given coloring we will use the Gabay-O'Connor theorem to get a our finite error strategy. Let's denote the Gabay-O'Connor guess for each agent a as q_a and denote the agents who guess correctly S. Then the set of agents who guessed incorrectly is $A - S$. Next, for each K_a, let's subtract q_a to give us a new set $K'_a = K_a - \{q_a\}$ for each agent a. Then for each a we will choose a new hat color from K'_a and let $\mathbf{P}$ denote the agents who guess correctly using this strategy.. Note that each agent only has one correct choice, but for all agents in S we have removed the ability for them to chose correctly. Thus $P \subseteq A - S$, so we know that P is finite, as desired.

So we can see that as a theorem, the Weak Gabay-O'Connor theorem is relatively innocuous. What is unexpected is that it can be used to prove the Axiom of Choice.

Theorem 2.6. *The Axiom of Choice follows from the Weak Gabay-O'Connor theorem.*

Proof. As before we will let A be the arbitrary set of agents, and K_a be the set of coloring for each agent $a \in A$. Next, we need a collection of non-empty sets on which to define a choice function. We will assume that the K_a are pairwise disjoint and nonempty so that they form our collection of disjoint sets. Next we will take an arbitrary hat color d, which isn't in any of the sets K_a. Let's add this d to each of our sets of colorings to form $K'_a = K_a \uplus \{d\}$.

Now let's look at the coloring that assigns d to each of the agents. We know that we can apply our Weak Gabay-O'Connor theorem to get a strategy that is correct only a finite amount of the time. Let's denote the set of agents that guess correctly S and the set of agents that guess incorrectly S^*. By the Weak Gabay-O'Connor theorem we know that S^* is finite.

Let's look at what exactly we've accomplished with S. By definition, for each a in S, we have chosen a color, that much is obvious; however, each of these color choices is wrong. This means that each color choice isn't d, meaning it comes from K_a. Thus for all but a finite number of agents in S^* we have defined a choice function.

Now let's look at S^*, the finite set of agents who chose correctly. For each a in S^* the Weak Gabay-O'Connor strategy chose d. Then for each of these agents let's examine our original set of colorings, which we will write $K^*_a = K'_a - \{d\}$. Since there are finitely many of these we can choose[7] a c from each K^*_a. Since these K^*_a are necessarily disjoint, we can

[7]While the Axiom of Choice is needed to prove the existence of a choice function for arbitrary sets, the existence of a choice function for finitely many sets is inherent within the Zermelo-Fraenkel system.

combine this finite choice function with our Weak Gabay-O'Connor strategy over S to give our desired choice function. $\square$

This result seems counterintuitive; how can the equivalence of the Gabay-O'Connor and the Axiom of Choice be an open question, while we can prove the Axiom of Choice from the Weak Gabay-O'Connor? We can attribute this to our proof of the Weak Gabay-O'Connor theorem. Although not explicitly noted, we use the Axiom of Choice when we switch each agents choice to a member of K'_a, so in this context it is not surprising that the Weak Gabay-O'Connor theorem is equivalent to the Axiom of Choice.

3 Non-Measurable Sets

3.1 Introduction

As previously discussed the idea of non-measurable sets is central to the proof of the Banach-Tarski Paradox. Without this there would be a very clear contradiction, as we would be able to double the measure of our original set by applying a rigid motion. In order to see how this is possible it will be instructive to examine the construction of a non-measurable set. While this is far from the only non-measurable set, the proof laid out below is perhaps the most famous example of a non-measurable set. This specific proof comes courtesy of Professor Belk of Bard College. [2]

Definition 3.1. The function f is the **canonical surjection** for an equivalence relation R if f is the surjective mapping $f : X \to X/R$ by $f(x) = [x]$.

Proposition 3.2. $|\mathbb{R}| = |\mathbb{Q} \times \mathbb{R}/\mathbb{Q}| = |\mathbb{R}/\mathbb{Q}|$.

Proof. (AC) First, we can examine all the equivalence classes of $\mathbb{R}/\mathbb{Q}$, which we spoke about above. Using the Axiom of Choice we can construct a set S consisting of one element from each of these equivalence classes. Since these equivalence classes partition $\mathbb{R}$ we know that $\forall r \in \mathbb{R}$ $\exists$ exactly one $s \in S$ such that $r \sim s$. Because of this we know that $r - s \in \mathbb{Q}$ so $\exists q \in \mathbb{Q}$ s.t. $r - s = q$. Thus, $r = q + s$, with a unique $s \in S$ and $q \in \mathbb{Q}$. Using this we can represent every $r \in \mathbb{R}$ uniquely as a $q + s$ for $s \in S$ and $q \in \mathbb{Q}$. Thus for $r = q + s$ we can define a function $f : \mathbb{R} \to \mathbb{Q} \times \mathbb{R}/\mathbb{Q}$, by $f(r) = (q, [s])$. This is clearly a bijection, so we can conclude $|\mathbb{R}| = |\mathbb{Q} \times \mathbb{R}/\mathbb{Q}|$.

Next, we know that $|\mathbb{Q} \times \mathbb{R}/\mathbb{Q}| = |\mathbb{Q}| \cdot |\mathbb{R}/\mathbb{Q}|$. Since these are both of infinite cardinality, we know that $|\mathbb{Q}| \cdot |\mathbb{R}/\mathbb{Q}| = |\mathbb{Q}|$ or $|\mathbb{R}/\mathbb{Q}|$, whichever is greater. Since $|\mathbb{R}| = |\mathbb{Q} \times \mathbb{R}/\mathbb{Q}|$, and $|\mathbb{Q}| < |\mathbb{R}|$, we can conclude that $|\mathbb{R}| = |\mathbb{R}/\mathbb{Q}|$, as desired. $\square$

Proof Using Vitali Sets

Central to the proof of the existence of non-measurable sets is the use of the Axiom of Choice. Since this thesis will contain proofs in both ZFC and ZF+DC, we will explicitly note when the Axiom of Choice is being used.

From the proposition above, we know there exists a bijection between $\mathbb{R}/\mathbb{Q}$ and $\mathbb{R}$. We will let $g : \mathbb{R}/\mathbb{Q} \leftrightarrow \mathbb{R}$ be this bijection. Next, we will let $p : \mathbb{R} \to \mathbb{R}/\mathbb{Q}$ be the canonical surjection of the equivalence relation defined above. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f = g \circ p$.

We can see that $\forall y \in \mathbb{R}$ the preimage of y is a coset of $\mathbb{Q}$ in $\mathbb{R}$. This is because the preimage of y under g is a coset in $\mathbb{R}/\mathbb{Q}$. But then the preimage of the canonical surjection has to be the whole equivalence class. Thus, due to the density of these cosets, we can see that any open interval (a,b) in $\mathbb{R}$ contains a point in $f^{-1}(y)$. This will lead us to our next proposition.

Proposition 3.3. *There exists a function $f : \mathbb{R} \to \mathbb{R}$ such that the image of every open interval (a,b) is all of $\mathbb{R}$.*

Proof. Consider the function f as defined above. Let $x \in \mathbb{R}$ and let (a,b) be an open interval in $\mathbb{R}$. Since every open interval contains a point in $f^{-1}([x])$, clearly $x \in f((a,b))$. □

Next we will define a Vitali set, first described by Giuseppe Vitali in 1905 in order to produce a non-measurable set. The construction of these Vitali sets explicitly uses the Axiom of Choice.

Definition 3.4. (AC)
Let $\mathbb{V} \subseteq [0,1]$. V is called a **Vitali Set** if V contains a single point from each coset of $\mathbb{Q}$ in $\mathbb{R}$.

Remark. Recall that we previously showed the density of each coset in $\mathbb{R}$. We then know that for each coset $(x+\mathbb{Q}) \cap [0,1] \neq \emptyset$. Then to construct our Vitali set we will use the Axiom of Choice to select one element from each of these nonempty intersections.

This leads us to our primary theorem.

Main Theorem 3.5. *Vitali sets are not Lebesgue measurable.*

Before we get into the proof of the theorem, we will need two lemmas.

Lemma 3.6. *Let $V \subseteq [0,1]$ be a Vitali set. Then:*
i)$\{q+V | q \in \mathbb{Q}\}$ are pairwise disjoint and
ii)$\mathbb{R} = \biguplus_{q\in\mathbb{Q}}(q+V)$.

Proof. i) Let $q, q' \in \mathbb{Q}$ such that $q \neq q'$ and let $s \in (q+V)$ and $s \in (q'+V)$. Then there exists $v, v' \in V$ such that $s = q+v$ and $s = q'+v'$. Then $s+(-v) = q$ and $s+(-v') = q'$, so $s+v$ and $s+v'$ are both in $\mathbb{Q}$ and as such are in the same coset of $\mathbb{Q}$ in $\mathbb{R}$. Hence, $v = v'$ which implies that $q = q'$, so $(q+V) = (q'+V)$, as desired.
ii) Let $x \in \mathbb{R}$. Then $x \in [x]$ and we know that $\exists v \in V$ such that $v \in [x]$. Thus $\exists q \in \mathbb{Q}$ such that $x+q = v$. Clearly $x = v+(-q)$, so $x \in q+V$ as desired. □

Lemma 3.7. *Let $V \subseteq [0,1]$ be a Vitali set, $C = \mathbb{Q} \cap [-1,1]$ and $U = \uplus_{q\in C}(q+V)$. Then $[0,1] \subseteq U \subset [-1,-2]$.*

Proof. Let $x \in [0,1]$. Then, by Lemma 1 we know there exists a $q \in \mathbb{Q}$ and a $v \in V$ such that $x = q+v$. Then $x-v = q$ and we know that $v, x \in [0,1]$, so $|x-v| \leq 1$ so $q \in [0,1]$ so $[0,1] \subseteq U$.

Next, let $y \in U$. Then there exists a $q \in C$ and $v \in V$ such that $y = q+v$ since $-1 \leq q \leq 1$ and $0 \leq v \leq 1$, clearly $-1 \leq y \leq 2$ so $y \in [-1,2]$, as desired. □

With these lemmas in hand we will return to the proof of our Theorem.

Proof. Let $V \subseteq [0,1]$ be Vitali, and suppose, for contradiction that V is measurable. Define $C = \mathbb{Q} \cap [-1,1]$ and $U = \biguplus_{q \in C}(q+V)$. Then U is the countable union of measurable sets, thus measurable. Also, since $[0,1] \subseteq U \subset [-1,-2]$, we know $m([0,1]) \leq m(U) \leq m([-1,-2])$, so $1 \leq m(U) \leq 3$. But since U is the countable union of Vitali sets, we know that if $m(V) = 0$ then $m(U) = 0$ and if $m(V) > 0$ then $m(U) = \infty$, a contradiction. □

3.2 Application of Vector Spaces to Non-Measurable Sets

This section also comes courtesy of Professor Belk. [2]

We may also make use of this theorem to prove the existence of additional non-measurable sets, this time using $\mathbb{R}$ as a vector space over $\mathbb{Q}$. We will do this by defining scalar multiplication as

$$\mathbb{Q} \times \mathbb{R} \to \mathbb{R}.$$

Clearly, since $\mathbb{Q} \subseteq \mathbb{R}$ and $\mathbb{R}$ has the properties of closure, associativity, and commutativity under multiplication and addition, we can see that $\mathbb{R}$ forms a vector space over $\mathbb{Q}$.

Having established that $\mathbb{R}$ forms a vector space we will use a well-known theorem as the start to our next proof.

Theorem 3.8. *Every vector space has a basis. Further, for every vector space V and every linearly independent set $S \subseteq V$, the set S can be expanded to be a basis of V.*

In order to prove this we will use Zorn's Lemma, which we will take as true for now. As previously noted, Zorn's Lemma is equivalent to the Axiom of Choice. It is instrumental in the proof of many of the most important theorems we have today, such as the Hahn-Banach Theorem and Tychonoff's Theorem.

Lemma 3.9. ***Zorn's Lemma:***

Suppose a partially ordered set P has the property that every chain has an upper bound in P. Then P contains at least one maximal element.

In the context of vector spaces we can look at the set of linearly independent sets well ordered by subset inclusion. We will formally define this notion below.

Proof. Let P be the set of linearly independent sets, for a vector space V, partially ordered by subset inclusion. That is, $P = \{I \subseteq V : I \text{ is linearly independent}\}$ and for $I_1, I_2 \in P$ $I_1 \leq I_2$ iff $I_1 \subseteq I_2$. Using this we can see a natural chain in P using our partial ordering, i.e. for $I_1 \subseteq I_2 \subseteq I_3 \subseteq ...$ $I_1 \leq I_2 \leq I_3 \leq ....$ The union of this upward chain forms its upper bound from our definition of P. As such, Zorn's Lemma tells us that there exists some maximal element $B \in P$. Let's suppose that B doesn't form a basis for V. Once again, by the construction of P we know that B is linearly independent, so we can conclude that B doesn't span V, thus there is some element $x \in V$ such that x isn't in the span of B. But then $B \cup \{x\}$ is also linearly independent. Further, we know $B \subset (B \cup \{x\}$, so $B < B + \{x\}$, a contradiction to Zorn's Lemma; therefore, we can conclude that B is a basis for V. □

This theorem allows us to conclude that the vector space of $\mathbb{R}$ over $\mathbb{Q}$ has a basis. We will call this basis H. Next we will consider the size of this basis.

Proposition 3.10. *Let H be a basis for the vector space $\mathbb{R}$ over $\mathbb{Q}$. Then H is uncountable.*

Proof. First since it spans $\mathbb{R}$ we know that its span must be uncountable. Assume that $|H| = \omega$. Then since $|\mathbb{Q}| = \omega$ we know that the set of finite linear combinations of H is countable, a contradiction. $\square$

Using Zorn's Lemma we can also "build" a basis up from $\{1\}$. Let's specify H so that it is a basis containing 1. Since H is uncountable $H - \{1\}$ is also uncountable. Let the set of all finite linear combinations of $H - \{1\}$ be denoted S. Then we know that S is uncountable. Next, assume that S isn't closed under addition. Then $\exists s_1, s_2 \in S$ such that $s_1 + s_2 = q$ for some $q^* \in \mathbb{Q}$. If $q^* = 0$, this is trivial, so we will take $q^* \neq 0$. But since s_1 and s_2 are in S then $\exists b_i, ..., b_j$ and $b_n, ..., b_m \in B - \{1\}$ [not necessarily distinct] such that for $q_i, ..., q_j$ and $q_n, ..., q_m \in \mathbb{Q}$ $q_i b_i + ... + q_j b_j + q_n b_n + ... + q_m b_m = q$. But then $q^* = 1 \cdot q^* = q_i b_i + ... + q_j b_j + q_n b_n + ... + q_m b_m$, which is a contradiction to the linear independence of B; therefore, S is closed under addition.

Next we will need to define a certain type of subgroup

Definition 3.11. Let M be a subgroup of $\mathbf{R}$. Then we say that M is **complementary** $\mathbb{Q}$ if M contains exactly one element from each coset of $\mathbb{Q}$ in $\mathbb{R}$.

This is equivalent to saying that if M is complementary to $\mathbb{R}$, then $\forall r \in \mathbb{R}$ $\exists$ a unique $m \in M$ and $q \in \mathbb{Q}$ such that $r = m + q$. This leads us to the next proposition.

Proposition 3.12. *There exists a subgroup M of $\mathbb{R}$ that is complementary to $\mathbb{Q}$.*

Proof. Lets consider the set S which we previously defined as the set of all finite linear combinations of $B - \{1\}$, where B is a basis for $\mathbb{R}$ over $\mathbb{Q}$ containing 1. We also previously established that this set is closed under addition.

Suppose, for contradiction's sake, that S isn't complementary to $\mathbb{Q}$. Then either there is a coset of $\mathbb{Q}$ in $\mathbb{R}$ such that it has an empty intersection with S or has at least two common elements of S. Since 0 is equal to the trivial linear combination of elements of $B - \{1\}$, we know that S a non-empty intersection with $\mathbb{Q}$. Next assume that there exists some coset of $\mathbb{Q}$ in $\mathbb{R}$ that has an empty intersection with S. Then, there exists an $x \in \mathbb{R} - \mathbb{Q}$ such that $\forall q \in \mathbb{Q}$, $x + q \notin S$. But since B is a basis we know that $\exists q^* \in \mathbb{Q}$ such that $x = q^* \cdot 1 + s^*$ for some $s^* \in S$. But then $x - (-q)^* = s^*$, which is a contradiction , so we know that every coset of $\mathbb{Q}$ in $\mathbb{R}$ has a nonempty intersection with S.

Next, suppose that there is a coset, $x + \mathbb{Q}$ of $\mathbb{Q}$ in $\mathbb{R}$ such that there are two elements, s_1, s_2, in S and $x + \mathbb{Q}$. Then $\exists q_1, q_2 \in \mathbb{Q}$ such that $s_1 = x + q_1$ and $s_2 = x + q_2$. But then $s_1 - s_2 = s_1 + (-s_2) = q_1 - q_2$, which is a contradiction to the closure of S. Thus we know that S is complementary to $\mathbb{Q}$, as desired. $\square$

Using this in addition to our first theorem we will prove that any subset of $\mathbb{R}$ which is complementary to $\mathbb{Q}$ is non-measurable.

Main Theorem 3.13. *Let P be a subgroup of $\mathbb{R}$ that is complementary to $\mathbb{Q}$. Then P is not Lebesgue measurable.*

Proof. First let $P \subseteq \mathbb{R}$ be complementary to $\mathbb{Q}$. Next assume for contradiction's sake that P is measurable. Then consider

$$U = \uplus_{n \in \mathbb{Q}} (n + S).$$

Since $\mathbb{Z}$ is countable and S is measurable, U is measurable.

Claim. $U \cup [0,1)$ is a Vitali set.

Proof of Claim. We will first look at how this applies to a single coset of $\mathbb{Q}$ in $\mathbb{R}$. Let $x+\mathbb{Q}$ be a coset of $\mathbb{Q}$ in $\mathbb{R}$. Then it intersects P at a unique point s. Then

$$(x+\mathbb{Q}) \cap (n+P) = \{n+s\}$$

for each $n \in \mathbb{N}$. This is equivalent to

$$(x+\mathbb{Q}) \cap U = s+\mathbb{Z}.$$

But this clearly intersects [0,1) at a single point, so each coset intersects $U \cap [0,1)$ at a single point, which proves that $U \cap [0,1)$ is a Vitali set.

Thus since $U \cap [0,1)$ is a Vitali set, clearly U isn't measurable and as such neither is P. □

4 The Banach-Tarksi Paradox

4.1 The Circle Trick

The Banach-Tarski Paradox is perhaps one of the most-well known results in mathematics due to its paradoxical nature. How could rigid motions result in a doubling of volume? To the average observer the gut reaction is to assume there must be something wrong with the proof. Since the structure of the proof has been examined many times and has been determined to be valid, the critics of the Banach-Tarski Paradox choose to focus on its use of the Axiom of Choice. In this section I will be presenting a proof of the "Circle Trick" as shown by Tom Weston [7]. The interesting aspect of this trick is that it is analogous to the Banach-Tarski Paradox in $\mathbb{R}^2$; however, it doesn't require the use of the Axiom of Choice.

Preliminaries

We will set up a few preliminaries that will be used later in the trick. We will define $\mathbf{S}^1$ as the unit circle in $\mathbb{R}^2$, which can be formally defined as $\mathbf{S}^1 = \{(x,y) \in \mathbb{R}^2 | d((x,y),(0,0)) = 1\}$ with d denoting the usual metric in $\mathbb{R}^2$. We will let the line segment (0,1) along the x axis be denoted l and we will denote $\rho(l)$ as the counterclockwise rotation of l by 1/10 of a radian. Further, let $\rho^i(l)=\rho(l) \cdot \rho(l) \cdot \rho(l) \cdot ...\rho(l)$ i times.

Remark 4.1. Since π is irrational we can see that $\rho^n(l) \neq l \; \forall n > 0$. Thus, for all $i \neq j$ $\rho^i(l) \neq \rho^j(l)$, so $\rho^{n-1}(l) \neq \rho^{-1}(l)$.

Next let $C = \uplus_{n=0}^{\infty} \rho^n(l)$, or the collection of all line segments formed by the rotation ρ. By our previous note we can see that C is an infinite collection of line segments originating from the origin and radiating outward.

The Circle Trick

We now have everything we need to see the Circle Trick. Lets start by considering $\mathbf{S}^1 \uplus C$. This can be seen as a bike wheel with an infinite number of spokes, none of which overlap. For a visual representation examine Figure 1 above.

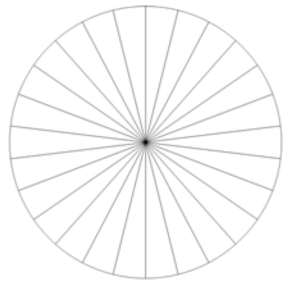

Figure 1: Think of a bike wheel like this but with an infinite number of spokes.

From our note above we can see that $\rho^{-1}(l)$ doesn't intersect $\mathbf{S}^1 \uplus C$; however, since $l \in C$ we can see that $\rho^{(-1)}(l) \in \rho^{-1}C$. Thus, via a simple mutual inclusion proof we can see that $\mathbf{S}^1 \uplus C \uplus \rho^{-1}(l) = \mathbf{S}^1 \uplus \rho^{-1}C$. Let $x \in C \uplus \rho^{-1}(l)$. If $x \in C$, then $\rho(x) \in C$, so $x \in \rho^{-1}C$. Similarly since $l \in C$, if $x \in \rho^{-1}(l)$ then $x \in \rho^{-1}C$. The other direction of this is similarly clear: if $x \in \rho^{-1}C$ then $\rho(x) \in C$, so $\rho^{-1}(\rho(x)) \in C \uplus \rho^{-1}(l)$. This would seem strange as $\mathbf{S}^1 \uplus C \uplus \rho^{-1}(l)$ has one more "spoke" than $\mathbf{S}^1 \uplus \rho^{-1}C$. Thus, we have added a an extra spoke but the two sets are equal.

The analogousness of this trick to the Banach-Tarski Paradox should be clear. By a simple rotation of a set we are adding another "spoke" to the wheel. When we move to $\mathbb{R}^3$ we need the additional power that the Axiom of Choice gives us.

4.2 Introduction to Free Groups

Before we get to our proof of the Banach-Tarski Paradox we have to cover a few preliminaries. The first of these is a brief introduction to free groups. Free groups can be thought of as the group-theoretic analogue of a basis in a vector space. This leads us to our definition of a free group.

Definition 4.2. If X is a subset of a group F, then F is a **free group** with **basis** X if, for every group G and every function $f : X \to$ G, there exists a unique homomorphism: ϕ : F$\to$G extending f.

Another more informal way to think of free groups would be as follows:

Let S be a set. Then the **free group** F_s over S consists of the set of all "expressions" formed from the elements of S. In this way we call the elements of S the **generators** of F_s. For example, if our set $S = \{a, b\}$, we would say that a and b were the generators of F_s and that the elements of F_s consisted of all of the different sequences from $\{a, b, a^{-1}, b^{-1}\}$ in which aa^{-1}, $a^{-1}a$, bb^{-1} or $b^{-1}b$ never occur.(i.e. aba^2b^{-1}, etc). We then call the identity element the **empty word**.

Let F_s be the free group over S. Then we say that F_s is of rank n if $|S| = n$.

When dealing with a free group F_s it is customary to refer to the elements of F_s as words. We can say every word has an inverse and that for a word $w = x_1x_2x_3$, the inverse is $w^{-1} = x_3^{-1}x_2^{-1}x_1^{-1}$. We will say that a word w is **reduced** if none of the generators of w are adjacent to their inverse. We can also note that the inverse of a reduced word is reduced. From this point forward we will only be dealing with reduced words unless otherwise stated.

Let $\mathfrak{F}(x,y)$ denote the free group generated by the set $\{x,y\}$. Next, we can define the set $\mathfrak{W}(x)$ as the set of all reduced words that begin with x. Using this notation we can see that:

$$\mathfrak{F}(x,y) = \mathfrak{W}(x) \uplus \mathfrak{W}(x^{-1}) \uplus \mathfrak{W}(y) \uplus \mathfrak{W}(y^{-1}) \uplus \{1\}.$$

This free group decomposition will be how we break up our ball in the proof of the Banach-Tarski Paradox. We will now prove two important facts about free groups which are vital to our proof of the Banach-Tarski Paradox.

Let's consider what happens when we multiply $\mathfrak{W}(x)$ by x^{-1}. This is multiplying the set of all words beginning with x by x^{-1}. Thus the first term of every word will cancel with x^{-1}.

Proposition 4.3.

$$x^{-1}\mathfrak{W}(x) = \mathfrak{W}(x) \uplus \mathfrak{W}(y) \uplus \mathfrak{W}(y^{-1}) \uplus \{1\}$$

Proof. ($\subseteq$) Let $x^{-1}w \in x^{-1}\mathfrak{W}(x)$. Since w is a word we know that $x^{-1}w \in \mathfrak{W}(x) \uplus \mathfrak{W}(x^{-1}) \uplus \mathfrak{W}(y) \uplus \mathfrak{W}(y^{-1}) \uplus \{1\}$. Suppose, for contradiction's sake, that $x^{-1}w \in \mathfrak{W}(x^{-1})$. Then $w = xx^{-1}...$, which is a contradiction since $w \in \mathfrak{W}(x)$ and as such is reduced. Thus we know $x^{-1}\mathfrak{W}(x) \subseteq \mathfrak{W}(x) \uplus \mathfrak{W}(y) \uplus \mathfrak{W}(y^{-1}) \uplus \{1\}$.
($\supseteq$)Now let $v \in \mathfrak{W}(x) \uplus \mathfrak{W}(y) \uplus \mathfrak{W}(y^{-1}) \uplus \{1\}$. Without loss of generality let $v \in \mathfrak{W}(y)$. Then we know that $xv \in \mathfrak{W}(x)$, and as such $y = x^{-1}xv \in x^{-1}\mathfrak{W}(x)$, as desired. $\square$

Since there is nothing special about the group $\mathfrak{W}(x)$ when compared to $\mathfrak{W}(x^{-1})$, $\mathfrak{W}(y)$, and $\mathfrak{W}(y^{-1})$, we can extend this proposition to hold for $\mathfrak{W}(y)$, $\mathfrak{W}(x^{-1})$ and $\mathfrak{W}(y^{-1})$. Of particular importance is the result $y^{-1}\mathfrak{W}(y) = \mathfrak{W}(x) \uplus \mathfrak{W}(y) \uplus \mathfrak{W}(x^{-1}) \uplus \{1\}$, as we will see later.

This proposition leads us to the main result which we will use in proving the Banach-Tarski Paradox, which is given by the following corollary.

Corollary 4.4.

$$\mathfrak{F}(x,y) = x^{-1}\mathfrak{W}(x) \uplus \mathfrak{W}(x^{-1})$$

and

$$\mathfrak{F}(x,y) = x^{-1}\mathfrak{W}(y) \uplus \mathfrak{W}(y^{-1}).$$

Proof. This proof is clear when looking at $x^{-1}\mathfrak{W}(x) = \mathfrak{W}(x) \uplus \mathfrak{W}(y) \uplus \mathfrak{W}(y^{-1}) \uplus \{1\}$ and $\mathfrak{F}(x,y) = \mathfrak{W}(x) \uplus \mathfrak{W}(x^{-1}) \uplus \mathfrak{W}(y) \uplus \mathfrak{W}(y^{-1}) \uplus \{1\}$. Then $\mathfrak{F}(x,y) - x^{-1}\mathfrak{W}(x) = \mathfrak{W}(x^{-1})$. Thus $\mathfrak{F}(x,y) = x^{-1}\mathfrak{W}(x) \uplus \mathfrak{W}(x^{-1})$, as desired. $\square$

How you can look at this result is that we have broken the free group generated by $\{x,y\}$ into 5 pieces and reassembled them in such a way that yields two copies of the original free group; the parallel to the Banach-Tarski Paradox should be clear. This, in combination with our final proposition in this section gives us all we need to know about free groups.

Proposition 4.5. $\mathfrak{F}(x,y)$ *is countable.*

Proof. The proof of this relies on the fact that $\mathfrak{F}(x,y)$ consists only of *finite* words that can be formed from the set $\{x,y\}$.

First let $\mathfrak{W}_n$ denote the set of all elements in $\mathfrak{W}(x,y)$ of length n. Since $\mathfrak{W}(x,y)$ is of rank two, we know that for each n, $\mathfrak{W}_n$ can have at most 4^n elements. Since $\sum_{n=1}^{\infty} \mathfrak{W}_n = \mathfrak{W}(x,y)$, we know that $\mathfrak{W}(x,y)$ is formed from a countable union of finite sets, and is thus countable. $\square$

4.3 A Free Group in SO_3

Before we get to the "full" Banach-Tarksi paradox on the ball in $\mathbb{R}^3$ we will examine the analogous result for the hollow sphere $\mathbf{S}^2$. This is a natural stepping stone as we will see later.

As we said before we will make use of Corollary 3.3 from the previous section.

In order to use this corollary we must first identify a rank 2 free group in SO_3, or the group of rotations on $\mathbf{S}^2$. For the purpose of this paper we will identify these rotations using their 3×3 matrix rotation, using the standard basis for $\mathbb{R}^3$, or the matrix

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

Thus, to denote a counter-clockwise rotation of $\frac{\pi}{2}$ around the x-axis we would write:

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

So if we represent the result of a given rotation in SO_3 by $\rho = \mathbf{A}$ on a point $p = (x, y, z) \in \mathbb{S}^2$, we would write

$$\rho(p) = \mathbf{A} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

With this notation out of the way we can proceed to our proof.

4.3.1 A Free Group of Rank 2 in $\mathbf{SO_3}$

As we previously stated we need to find a free group in SO_3 of rank 2. It is sufficient to find two independent rotations in SO_3. Let's consider the free group $\mathbf{F}(x, y)$. In order for this free group to be of rank 2 x and y need to be independent. If this weren't the case then y could be represented as some power of x, and thus all of the words we could form with x and y could be formed by using only x meaning, that $\mathbf{F}(x, y)$ was actually of rank 1.

Now let's examine two rotation. Let's let κ be a counterclockwise rotation by $cos^{-1}(\frac{1}{3})$ around the x-axis and ψ be a counter clockwise rotation by $cos^{-1}(\frac{1}{3})$ around the z axis. We can represent κ and ψ by the following matrices:

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

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

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

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

Taking the second representation above we would assert that a word in SO_3 of length n has the representation $\frac{1}{3^n}(a\sqrt{2}, b, c\sqrt{2})$ for some integers a, b, and c. We will do this by induction on the length of n. We will take $n = 0$ as our base step and look at the effect on (0,1,0), or when ρ is the identity rotation. If $n = 0$, then $\rho(0,1,0) = (0,1,0)$, as desired. Thus we may assume that this result holds for $n-1$. We now wish to prove that it holds for n.

In order to prove that this result holds for n we will examine the outcome for each rotation. First let ρ' be our rotation of length $n-1$. Then we know $\rho'(0,1,0) = (a\sqrt{2}, b, c\sqrt{2})$. In order to get to our word of length one we know that we need to multiply ρ' by $\kappa, \kappa^{-1}, \psi$ or ψ^{-1}, thus we will have four natural cases.

$$\kappa\rho'(0,1,0) = \frac{1}{3^{n-1}}\kappa(a\sqrt{2}, b, c\sqrt{2})$$

$$= \frac{1}{3^n}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & -2\sqrt{2} \\ 0 & 2\sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} a\sqrt{2} \\ b \\ c\sqrt{2} \end{pmatrix}$$

$$= \frac{1}{3^n}(3a\sqrt{2}, b-4c, (2b+c)\sqrt{2})$$

$$\kappa^{-1}\rho'(0,1,0) = \frac{1}{3^{n-1}}\kappa^{-1}(a\sqrt{2}, b, c\sqrt{2})$$

$$= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 2\sqrt{2} \\ 0 & -2\sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} a\sqrt{2} \\ b \\ c\sqrt{2} \end{pmatrix}$$

$$= \frac{1}{3^n}(3a\sqrt{2}, b+4c, (-2b+c)\sqrt{2})$$

$$\psi\rho'(0,1,0) = \frac{1}{3^{n-1}}\kappa(a\sqrt{2}, b, c\sqrt{2})$$

$$= \frac{1}{3^n}\begin{pmatrix} 1 & 2\sqrt{2} & 0 \\ -2\sqrt{2} & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} a\sqrt{2} \\ b \\ c\sqrt{2} \end{pmatrix}$$

$$= \frac{1}{3^n}((a-2b)\sqrt{2}, 4a+b, 3c\sqrt{2})$$

$$\psi^{-1}\rho'(0,1,0) = \frac{1}{3^{n-1}}\kappa^{-1}(a\sqrt{2}, b, c\sqrt{2})$$

$$= \frac{1}{3^n}\begin{pmatrix} 1 & -2\sqrt{2} & 0 \\ 2\sqrt{2} & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} a\sqrt{2} \\ b \\ c\sqrt{2} \end{pmatrix}$$

$$= \frac{1}{3^n}((a+2b)\sqrt{2}, -4a+b, 3c\sqrt{2}).$$

All of these are clearly of the desired form. Thus we have completed our proof by induction. By taking this result we can improve the efficiency of our check of the independence of our two rotations. Since κ and ψ are on two separate axes we simply check that no iteration of either of them is equal to the identity. In other words, $\forall n \in \mathbb{Z}$ $\kappa^n(0,1,0) \neq (0,1,0)$ and $\psi^n(0,1,0) \neq (0,1,0)$. If this were the case, then for $p \in S^2$, based on our previous result this simply means that for $\rho \in SO_3$, if $\rho(p) = (a,b,c)$ then $a = c = 0$ and $b = 3^n$. Since we are only checking the non-identity elements of SO_3, we can say that $n \geq 1$, which would mean that $a \equiv b \equiv c \equiv 0 \bmod 3$. We must show that $\forall n \in \mathbb{Z}$, this is not the case.

First let's define a function $N : SO_3 \to SO_3$ by $N(\rho) = N(a,b,c) = (a,b,c) \bmod 3$. Then we know from the above that

$$N(\kappa\rho) = (3a, b - 4c, 2b + c) = (0, b - c, c - b)$$

$$N(\kappa^{-1}\rho = (3a, b + 4c, -2b + c) = (0, b + c, b + c)$$

$$N(\psi\rho) = (a - 2b, 4a + b, 3c) = (a + b, a + b, 0)$$

$$N(\psi^{-1}\rho = (a + 2b, b - 4a, 3c) = (a - b, b - a, 0)$$

Taking this one step further we can see that

$$N(\kappa^2\rho) = (0, c - b, b - c).$$

$$N(\kappa^{-2})\rho = (0, -b - c, -b - c).$$

$$N(\psi^2\rho) = (-a - b, -a - b, 0).$$

$$N(\psi^{-2}\rho) = (b - a, a - b, 0).$$

And

$$N(\kappa^3\rho) = (0, b - c, c - b).$$

$$N(\kappa^{-3}\rho) = (0, b + c, b + c).$$

$$N(\psi^3\rho) = (a + b, a + b, 0).$$

$$N(\psi^{-3}\rho) = (a - b, b - a, 0),$$

which we can see is a repetition of powers of 1, giving us the following rule:

$$N(\kappa^n\rho) = \begin{cases} (0, b - c, c - b) \text{ when n is odd} \\ (0, c - b, b - c) \text{ when n is even} \end{cases}$$

$$N(\kappa^{-n}\rho) = \begin{cases} (0, b + c, b + c) \text{ when n is odd} \\ (0, -b - c, -b - c) \text{ when n is even} \end{cases}$$

$$N(\psi^n\rho) = \begin{cases} (a + b, a + b, 0) \text{ when n is odd} \\ (-a - b, -a - b, 0) \text{ when n is even} \end{cases}$$

$$N(\psi^{-n}\rho) = \begin{cases} (a - b, b - a, 0) \text{ when n is odd} \\ (b - a, a - b, 0) \text{ when n is even} \end{cases}$$

Given these rules we now need to test them. Let's return to the more concrete example of (a,b,c)=(0,1,0). Since any word $\rho \in SO_3$ is formed by alternating nonzero powers of κ

and ψ, we can form a "tree" of possible moves. Since ρ has to start with either κ or ψ, there are two possible starts to our tree. Then for some $n_1 \in \mathbb{Z}$ we can say that

$$N(\kappa^{n_1}(0,1,0)) \in \{(0,1,2),(0,2,1),(0,1,1),(0,2,2)\}$$

based on whether n_1 is positive and odd, positive and even, negative and odd, or negative and even, respectively.

This; however, is where things could become burdensome. Let's look at $N(\psi^{n_2}\kappa^{n_1})$. This gives us four possibilities for κ^{n_1} as well as for ψn_1. Mercifully there turns out to be many repetitions, leaving us with the following possibilities:

$$N(\psi^{n_2}\kappa^{n_1}(0,1,0)) \in \{(1,1,0),(2,2,0),(2,1,0),(1,2,0)\}.$$

We are even more fortunate with our second iteration of κ as the solution set is

$$N(\kappa^{n_3}\psi^{n_2}\kappa^{n_1}(0,1,0)) \in \{(0,1,2),(0,2,1),(0,1,1),(0,2,2)\}.$$

This is a repetition, so we can see without any great effort that this cycle will continue and never yield (0,1,0) as desired.

Now let's examine what happens if we start with ψ. Well, the possible solution set is

$$N(\psi^{n_1}(0,1,0)) \in \{(1,1,0),(2,2,0),(2,1,0),(1,2,0)\}, \text{so}$$

$$N(\kappa^{n_2}\psi^{n_1}(0,1,0)) \in \{(0,1,2),(0,2,1),(0,1,1),(0,2,2)\}, \text{and}$$

$$N(\psi^{n_3}\kappa^{n_2}\psi^{n_1}(0,1,0)) \in \{(1,1,0),(2,2,0),(2,1,0),(1,2,0)\}.$$

Thus, we again have a cycle, meaning that κ and ψ are independent and form a free group, as desired.

4.4 The Hausdorff Paradox

Now that we have established that these two rotations are independent, we can examine the free group generated by κ and ψ. We will denote this group $\mathfrak{F}(\kappa,\psi)$, and attempt to apply our paradoxical free group corollary to it. This should give us our desired paradox on $\mathbf{S}^2$. Before we begin, let's start by remembering that

$$\mathfrak{F}(\kappa,\psi) = \mathfrak{W}(\kappa) \uplus \mathfrak{W}(\kappa^{-1}) \uplus \mathfrak{W}(\psi) \uplus \mathfrak{W}(\psi^{-1}) \uplus 1.$$

First, we will specify an equivalence relation on $\mathbf{S}^2$. Let p, q be in $\mathbf{S}^2$. Then we will say that p is equivalent to q if and only if there exists some rotation ρ in $\mathfrak{F}(\kappa,\psi)$ such that $\rho(p) = q$. By the structure of free groups, clearly this forms an equivalence relation. Thus for any point p in $\mathbf{S}^2$, the orbit of p is

$$\{\rho(p)|\rho \in \mathfrak{F}(\kappa,\psi)\}.$$

Since we previously established that $\mathfrak{F}(\kappa,\psi)$ is countable, we know that each of these equivalence classes must be countable. Hence, since $\mathbf{S}^2$ is uncountable, we know there must be uncountably many equivalence classes. Next, we will invoke the Axiom of Choice to take one element from each of these equivalence classes, and call this set $\mathcal{M}_0$. Clearly by this definition we can see that

$$\mathbf{S}^2 = \mathfrak{F}(\kappa,\psi)\mathcal{M}_0 = \{\rho(p)|\rho \in \mathfrak{F}(x,y)p \in \mathcal{M}_0\}.$$

By extending this using our partition of $\mathfrak{F}(\kappa,\psi)$, we get that

$$\mathbf{S}^2 = \mathfrak{F}(\kappa,\psi)\mathcal{M}_0 = \mathfrak{W}(\kappa)\mathcal{M}_0 \cup \mathfrak{W}(\kappa^{-1})\mathcal{M}_0 \cup \mathfrak{W}(\psi)\mathcal{M}_0 \cup \mathfrak{W}(\psi^{-1})\mathcal{M}_0 \cup \mathcal{M}_0.$$

Unfortunately these sets are no longer necessarily disjoint. For example, consider what happens if $(1,0,0)$ is in $\mathcal{M}_0$. Then, since κ is a rotation around the x-axis, $\kappa(1,0,0)$ would be in $\mathcal{M}_0$ and $\mathfrak{W}(\kappa)\mathcal{M}_0$. In fact, any point in $\mathcal{M}_0$ which lies on the axis of rotation for some ρ in $\mathfrak{F}(\kappa,\psi)$ would be in multiple sets of our decomposition. However, since $\mathfrak{F}(\kappa,\psi)$ is countable, we know that the number of points on these axis of rotations is countable, and we can define a countable set

$$\mathcal{D} = \{p \in \mathcal{M}_0 | p \text{ is a fixed point of some } \rho \in (\mathfrak{F}(\kappa,\psi) - \{1\})\}.$$

We can now consider our free group on $\mathbf{S}^2 - \mathcal{D}$. We must first prove closure under the group action $\mathfrak{F}(\kappa,\psi)$ on $\mathbf{S}^2 - \mathcal{D}$. In explicit terms this means that for all p in $\mathbf{S}^2 - \mathcal{D}$ and for all ρ in $\mathfrak{F}(\kappa,\psi)$, $\rho(p)$ in $\mathfrak{F}(\kappa,\psi)$.

We will proceed using a proof by contrapositive. Assume, for contradiction's sake, that there exists some p' in $\mathbf{S}^2 - \mathcal{D}$ and ρ in $\mathfrak{F}(\kappa,\psi)$ such that $\rho(p')$ isn't in $\mathbf{S}^2 - \mathcal{D}$. Then we know that$\rho(p')$ is in $\mathcal{D}$, and as such there exists some γ in $\mathfrak{F}(\kappa,\psi)$, such that $\gamma(\rho(p')) = \rho(p')$. But then $\rho^{-1}(\gamma(\rho(p))) = p$, so p is in $\mathcal{D}$ a contradiction, as desired.

Since we have established that $\mathbf{S}^2 - \mathcal{D}$ is closed under the group action of $\mathfrak{F}(\kappa,\psi)$, we will now attempt to create a choice set parallel to $\mathcal{M}_0$. Let's define $\mathcal{M}$ to be the set consisting of one element from each equivalence class of $\mathfrak{F}(\kappa,\psi)$ on $\mathbf{S}^2 - \mathcal{D}$. Then we will need to prove that this with our group action $\mathfrak{F}(\kappa,\psi)$ forms a partition of $\mathbf{S}^2 - \mathcal{D}$. Due to the construction of $\mathcal{M}$ we can see that every element of $\mathbf{S}^2 - \mathcal{D}$ is in at least one of these equivalence classes; however, we still need to prove disjointness so let ρ_1 and ρ_2 be in $\mathfrak{F}(\kappa,\psi)$ with $\rho_1(\mathcal{M}) \cap \rho_2(\mathcal{M}) \neq \emptyset$. Then there exists a p, p_1 and p_2 such that $\rho_1(p_1) = p = \rho_2(p_2)$ which means that $\rho_2^{-1}\rho_1(p_1) = p_2$. Since p_1 and p_2 are in the same equivalence class we know that $p_1 = p_2$. Thus we know that $\rho_1 = \rho_2$, as desired.

Now we can see that our partition of $\mathfrak{F}(\kappa,\psi)$ can be used to give us a partition of $\mathbf{S}^2 - \mathcal{D}$:

$$\mathbf{S}^2 - \mathcal{D} = \mathfrak{F}(\kappa,\psi)\mathcal{M} = \mathfrak{W}(\kappa)\mathcal{M} \uplus \mathfrak{W}(\kappa^{-1})\mathcal{M} \uplus \mathfrak{W}(\psi)\mathcal{M} \uplus \mathfrak{W}(\psi^{-1})\mathcal{M} \uplus \mathcal{M}. \tag{4.1}$$

Next, we will rename these groups for notation sake. This will make it easier to see the application of our free group corollary:

$$\mathcal{W}_\kappa = \mathfrak{W}(\kappa)\mathcal{M}$$

$$\mathcal{W}_{\kappa^{-1}} = \mathfrak{W}(\kappa)\mathcal{M}$$

$$\mathcal{W}_\psi = \mathfrak{W}(\psi)\mathcal{M}$$

$$\mathcal{W}_{\psi^{-1}} = \mathfrak{W}(\kappa^{-1})\mathcal{M}.$$

So we can see from (1) that

$$\mathbf{S}^2 - \mathcal{D} = \mathcal{W}_\kappa \uplus \mathcal{W}_{\kappa^{-1}} \uplus \mathcal{W}_\psi \uplus \mathcal{W}_{\psi^{-1}} \uplus \mathcal{M}, \tag{4.2}$$

to which we can apply our corollary. This gives us:

$$\mathbf{S}^2 - \mathcal{D} = \kappa^{-1}\mathcal{W}_\kappa \uplus \mathcal{W}_{\kappa^{-1}}, \text{ and}$$

$$\mathbf{S}^2 - \mathcal{D} = \psi^{-1}\mathcal{W}_\psi \uplus \mathcal{W}_{\psi^{-1}}.$$

What this says is that we can take two of the pieces of our partition of $\mathbf{S}^2 - \mathcal{D}$, namely $\mathcal{W}_\kappa$ and $\mathcal{W}_\psi$, and rotate $\mathcal{W}_\kappa$ by κ^{-1} and $\mathcal{W}_\psi$ by ψ^{-1}, and then re-add the other pieces of our partition to get two copies of $\mathbf{S}^2 - \mathcal{D}$. This is what's known as the Hausdorff Paradox. We can state this below as:

Main Theorem 4.6. ***Theorem: The Hausdorff Paradox***
There is a countable set $\mathcal{D} \subseteq \mathbf{S}^2$ such that $\mathbf{S}^2 - D$ can be divided into 5 pieces and rotated to form two copies of $\mathbf{S}^2 - D$.

Now if we look at precisely what this theorem says it isn't particularly exciting. In layman terms, it only applies to a hollow ball with infinitely many points missing. It also requires the use of some ambiguous and uncountable subsets of $\mathbf{S}^2 - \mathcal{D}$. Thus, the next step of this is to find a way that we can apply this result to all of $\mathbf{S}^2$.

4.5 The Banach-Tarski Paradox on $\mathbf{S}^2$

The Hausdorff Paradox tells us that we can break apart $\mathbf{S}^2 - \mathcal{D}$, rotate the pieces, and reassemble them into 2 copies of $\mathbf{S}^2 - \mathcal{D}$. In order to generalize this result to $\mathbf{S}^2$, we will have to do something to alleviate the problem of $\mathcal{D}$.

Let's think back to the circle trick we looked at earlier. We used a counterclockwise rotation by $\frac{1}{10}\pi$ to "create" another line segment in $\mathbb{R}^2$. But this trick can also be used in reverse! We will define our ρ and l as we did before. That is, ρ is our counterclockwise rotation by $\frac{1}{10}\pi$ around the origin, and l is the line segment (0,1). Remember that before we defined our C as $C = \uplus_{n=0}^{\infty}\rho(l)$. This time we will change this, but only slightly: let's define C^* to be $C^* = \uplus_{n=\mathbf{1}}^{\infty} \rho^n(l)$. Thus $C^* = C - \{l\}$. We can look at this as "erasing" l from our infinitely spoked wheel, but we can get l back by applying ρ^{-1} to $C^*(l)$. We will try to apply this sort of idea to with $\mathbf{S}^2 - \mathcal{D}$ to "erase" $\mathcal{D}$.

Before we do this we will give three facts about disjoint unions which we will need to make use of later:

Remark 4.7. Let $\mathcal{A}$ be a set formed from the disjoint union of its subsets, $\mathcal{A}_i$, which we can denote

$$\mathcal{A} = \uplus_i \mathcal{A}_i.$$

Let's say that we can then form each $\mathcal{A}_i$ from the disjoint union of its subsets $\mathcal{A}_{ij}$, which we can denote similarly as

$$\mathcal{A}_i = \uplus_j \mathcal{A}_{ij}.$$

Then we can write:

$$\mathcal{A} = \uplus_{i,j}\ \mathcal{A}_{ij}.$$

Remark 4.8. If $\mathcal{A} \subseteq \mathbf{S}^2$ is formed through the union of disjoint subsets $\mathcal{A}_i$, then for any rotation ρ in SO_3

$$\rho(\mathcal{A}) = \ \uplus_i\ \rho(\mathcal{A}_i).$$

Remark 4.9. If $\mathcal{A} \subseteq \mathbf{S}^2$, and is composed of disjoint subsets as above, then for $\mathcal{B} \subseteq \mathbf{S}^2$:

$$\mathcal{A} \cap \mathcal{B} = \uplus_i (\mathcal{A}_i \cap \mathcal{B}).$$

Using these three facts will make our proof of the Banach-Tarski Paradox much easier as we won't have to verify that our actions don't damage our partition.

Next we will set about finding a trick analogous to the circle trick put forward earlier. First we will chose an axis which travels through the center of $\mathbf{S}^2$, but doesn't intersect $\mathcal{D}$. Since $\mathcal{D}$ is constructed using the axes of rotation from elements of $\mathfrak{F}(\kappa, \psi)$, we know that there are countably many axes which we can't choose. Since $\mathbf{S}^2$ is uncountable, we know that there are uncountably many points in $\mathbf{S}^2$. Thus there are uncountably many axes that run through $\mathbf{S}^2$ (since there are half as many axes as there are points), and we can choose an axis composed of points not in D. We will denote this axis by E. Further, we will denote the counterclockwise rotation of θ radians around E by E_θ.

Since $\mathcal{D}$ is countable, we know that $\mathcal{D} \times \mathcal{D}$ is also countable. We can think of the set $\mathcal{D} \times \mathcal{D}$ as a set of possible ways that a point in $\mathcal{D}$ can be returned to a point in $\mathcal{D}$ by some type of rotation around an axis by an element of $(0, 2\pi)$. [Note that for a specific axis the number of possible rotations returning an element of $\mathcal{D}$ to an element of $\mathcal{D}$ will most of the time be far fewer.] This may be clearer to see as the following set. Define

$$T = \{\theta \in [0, 2\pi) | \exists p \in \mathcal{D} \text{ and } n > 0 \in \mathbb{N} \text{ s.t. } E_{n \cdot \theta}(p) \in \mathcal{D}\}.$$

Since we already established that $|\mathcal{D} \times \mathcal{D}| = \omega$, we know the size of this set is less than or equal to $n \cdot \aleph_0 = \aleph_0$. Since $[0, 2\pi)$ is uncountable, we know that we can choose a θ_0 in $[0, 2\pi)$ such that θ_0 isn't in T. We will denote E_{θ_0} as σ. Note that not only is $\sigma(\mathcal{D}) \cap \mathcal{D} = \emptyset$, we can say that for all $n \geq 1$, $\sigma^n(\mathcal{D}) \cap \mathcal{D} = \emptyset$. [Note that this is similar to what we described in our circle trick with $C^*(l)$.] We can also note that for all positive integers m, $\sigma^{n+m}(\mathcal{D}) \cap \sigma^m(\mathcal{D}) = \emptyset$, so for $m \neq n$ we have $\sigma^n(\mathcal{D}) \cap \sigma^m(\mathcal{D}) = \emptyset$.

Next we will define a set similar to C from the circle trick, but for our set D. Define

$$\Omega = \uplus_{n=0}^{\infty} \sigma^n(\mathcal{D}).$$

Since $\sigma^0(\mathcal{D}) \in \Omega$, we can see that $\mathcal{D} \subset \Omega$.

Next, clearly $\mathbf{S}^2 = \mathbf{S}^2 - \Omega \uplus \Omega$. And similarly to how we "erased" l,

$$\begin{aligned}
\sigma\Omega &= \sigma \uplus_{n=0}^{\infty} \sigma^n(\mathcal{D}) \\
&= \uplus_{n=0}^{\infty} \sigma\sigma^n(\mathcal{D}) \\
&= \uplus_{n=0}^{\infty} \sigma^{n+1}\mathcal{D} \\
&= \uplus_{n=1}^{\infty} \sigma^n\mathcal{D} \\
&= \Omega - \mathcal{D}.
\end{aligned}$$

Combining this with the obvious fact gives us

$$\mathbf{S}^2 - \mathcal{D} = (\mathbf{S} - \Omega) \uplus (\Omega - \mathcal{D}) = (\mathbf{S}^2 - \Omega) \uplus \sigma\Omega.$$

What this says is that we can take $\mathbf{S}^2$, break it into two parts, multiply the parts and be left with $\mathbf{S}^2 - \mathcal{D}$. Once we get to this point we already know how to use the Hausdorff Paradox to duplicate $\mathbf{S}^2 - \mathcal{D}$. Recognizing this, it is merely a burden of notation to apply the Hausdorff Paradox, and eventually get two copies of $\mathbf{S}^2$.

However, before we get into this we will put forward a couple of facts which will hopefully make the proof easier to follow. First, since $\mathcal{D} \subset \Omega$, we can say that

$$\mathbf{S}^2 - \Omega = (\mathbf{S}^2 - \mathcal{D}) \cap (\mathbf{S}^2 - \Omega).$$

And substituting this in below we can see that:

$$\begin{aligned}\Omega &= \sigma^{-1}\sigma(\Omega)\\ &= \sigma^{-1}(\Omega - \mathcal{D})\\ &= \sigma^{-1}[(\mathbf{S}^2 - \mathcal{D}) \cap \Omega].\end{aligned}$$

So we can break down $\mathbf{S}^2$ as

$$\begin{aligned}\mathbf{S}^2 &= (\mathbf{S}^2 - \Omega) \uplus \Omega\\ &= [(\mathbf{S}^2 - \mathcal{D}) \cap (\mathbf{S}^2 - \Omega)] \uplus \sigma^{-1}[(\mathbf{S}^2 - \mathcal{D}) \cap \Omega].\end{aligned}$$

This then gives us a natural way to break down $\mathbf{S}^2$:

$$\mathbf{S}^2 = [(\mathcal{W}_\kappa \uplus \mathcal{W}_{\kappa^{-1}} \uplus \mathcal{W}_\psi \uplus \mathcal{W}_{\psi^{-1}} \uplus \mathcal{M}) \cap (\mathbf{S}^2 - \Omega)] \uplus \sigma^{-1}[(\mathcal{W}_\kappa \uplus \mathcal{W}_{\kappa^{-1}} \uplus \mathcal{W}_\psi \uplus \mathcal{W}_{\psi^{-1}} \uplus \mathcal{M}) \cap \Omega].$$

If we take a step back we can represent $\mathbf{S}^2 - \mathcal{D}$ as

$$\mathbf{S}^2 - \mathcal{D} = [(\mathcal{W}_\kappa \uplus \mathcal{W}_{\kappa^{-1}} \uplus \mathcal{W}_\psi \uplus \mathcal{W}_{\psi^{-1}} \uplus \mathcal{M}) \cap (\mathbf{S}^2 - \Omega)] \uplus [(\mathcal{W}_\kappa \uplus \mathcal{W}_{\kappa^{-1}} \uplus \mathcal{W}_\psi \uplus \mathcal{W}_{\psi^{-1}} \uplus \mathcal{M}) \cap \Omega].$$

We can then apply our corollary to this representation.This gives us

$$\mathbf{S}^2 = [(\kappa^{-1}\mathcal{W}_\kappa \uplus \mathcal{W}_{\kappa^{-1}}) \cap \mathbf{S}^2 - \Omega] \uplus \sigma^{-1}[(\kappa^{-1}\mathcal{W}_\kappa \uplus \mathcal{W}_{\kappa^{-1}}) \cap \Omega], \text{ and}$$

$$\mathbf{S}^2 = [(\psi^{-1}\mathcal{W}_\psi \uplus \mathcal{W}_{\psi^{-1}}) \cap \mathbf{S}^2 - \Omega] \uplus \sigma^{-1}[(\psi^{-1}\mathcal{W}_\psi \uplus \mathcal{W}_{\psi^{-1}}) \cap \Omega].$$

These are helpful; however, we are starting with $\mathbf{S}^2 - \mathcal{D}$. If we want to get the full Banach-Tarski Paradox we must start with $\mathbf{S}^2$. In order to do this we will add more notation.

In the following notation, the superscript 0 indicates that we are looking at a subset of $\mathbf{S}^2 - \Omega$ while a superscript of 1 indicates that we are dealing with the more complicated subsets of Ω. With this in mind we will define:

$$\begin{aligned}\mathcal{M}^0 &= \mathcal{M} \cap (\mathbf{S}^2 - \Omega)\\ \mathcal{W}^0_\kappa &= \mathcal{W}_\kappa \cap (\mathbf{S}^2 - \Omega)\\ \mathcal{W}^0_{\kappa^{-1}} &= \mathcal{W}_{\kappa^{-1}} \cap (\mathbf{S}^2 - \Omega)\\ \mathcal{W}^0_\psi &= \mathcal{W}_\psi \cap (\mathbf{S}^2 - \Omega)\\ \mathcal{W}^0_{\psi^{-1}} &= \mathcal{W}_{\psi^{-1}} \cap (\mathbf{S}^2 - \Omega)\\ \mathcal{M}^1 &= \mathcal{M} \cap \Omega\\ \mathcal{W}^1_\kappa &= \mathcal{W}_\kappa \cap \Omega\\ \mathcal{W}^1_{\kappa^{-1}} &= \mathcal{W}_{\kappa^{-1}} \cap \Omega\\ \mathcal{W}^1_\psi &= \mathcal{W}_\psi \cap \Omega\\ \mathcal{W}^1_{\psi^{-1}} &= \mathcal{W}_{\psi^{-1}} \cap \Omega.\end{aligned}$$

We can then see that

$$\mathbf{S}^2 = \mathcal{M}^0 \uplus \mathcal{W}^0_\kappa \uplus \mathcal{W}^0_{\kappa^{-1}} \uplus \mathcal{W}^0_\psi \uplus \mathcal{W}^0_{\psi^{-1}} \uplus \sigma^{-1}(\mathcal{M}^1 \uplus \mathcal{W}^1_\kappa \uplus \mathcal{W}^1_{\kappa^{-1}} \uplus \mathcal{W}^1_\psi \uplus \mathcal{W}^1_{\psi^{-1}})$$

$$= \mathcal{M}^0 \uplus \mathcal{W}^0_\kappa \uplus \mathcal{W}^0_{\kappa^{-1}} \uplus \mathcal{W}^0_\psi \uplus \mathcal{W}^0_{\psi^{-1}} \uplus \sigma^{-1}\mathcal{M}^1 \uplus \sigma^{-1}\mathcal{W}^1_\kappa \uplus \sigma^{-1}\mathcal{W}^1_{\kappa^{-1}} \uplus \sigma^{-1}\mathcal{W}^1_\psi \uplus \sigma^{-1}\mathcal{W}^1_{\psi^{-1}}.$$

While this may seem like a lot to digest, our notation should give an adequate indication of which piece of $\mathbf{S}^2$ we are looking at. But when we apply our corollary we will have to apply a rotation by either κ^{-1} or ψ^{-1}. This would be a problem if we are rotating Ω, as Ω is only closed under rotations by σ. Thus, before we start we will rotate Ω forward by κ or ψ, before rotating it back as part of our corollary. As such we have to introduce even more notation:

$$\mathcal{W}_\kappa^{00} = \mathcal{W}_\kappa^0 \cap \kappa(\mathbf{S}^2 - \Omega) = \mathcal{W}_\kappa \cap (\mathbf{S}^2 - \Omega) \cap \kappa(\mathbf{S}^2 - \Omega)$$
$$\mathcal{W}_\kappa^{01} = \mathcal{W}_\kappa^0 \cap \kappa(\Omega) = \mathcal{W}_\kappa \cap (\mathbf{S}^2 - \Omega) \cap \kappa(\Omega)$$
$$\mathcal{W}_\kappa^{10} = \mathcal{W}_\kappa^0 \cap \kappa(\mathbf{S}^2 - \Omega) = \mathcal{W}_\kappa \cap (\Omega) \cap \kappa(\mathbf{S}^2 - \Omega)$$
$$\mathcal{W}_\kappa^{11} = \mathcal{W}_\kappa^0 \cap \kappa(\Omega) = \mathcal{W}_\kappa \cap (\Omega) \cap \kappa(\Omega)$$
$$\mathcal{W}_\psi^{00} = \mathcal{W}_\psi^0 \cap \psi(\mathbf{S}^2 - \Omega) = \mathcal{W}_\psi \cap (\mathbf{S}^2 - \Omega) \cap \psi(\mathbf{S}^2 - \Omega)$$
$$\mathcal{W}_\psi^{01} = \mathcal{W}_\psi^0 \cap \psi(\Omega) = \mathcal{W}_\psi \cap (\mathbf{S}^2 - \Omega) \cap \psi(\Omega)$$
$$\mathcal{W}_\psi^{10} = \mathcal{W}_\psi^0 \cap \psi(\mathbf{S}^2 - \Omega) = \mathcal{W}_\psi \cap (\Omega) \cap \psi(\mathbf{S}^2 - \Omega)$$
$$\mathcal{W}_\psi^{11} = \mathcal{W}_\psi^0 \cap \psi(\Omega) = \mathcal{W}_\psi \cap (\Omega) \cap \psi(\Omega).$$

Now our decomposition of $\mathbf{S}^2$ is:

$$\mathbf{S}^2 = \mathcal{M}^0 \uplus \mathcal{M}^1 \uplus \mathcal{W}_\kappa^{00} \uplus \mathcal{W}_\kappa^{01} \uplus \sigma^{-1}\mathcal{W}_\kappa^{10} \uplus \sigma^{-1}\mathcal{W}_\kappa^{11} \uplus \mathcal{W}_{\kappa^-1}^0 \uplus \sigma^{-1}\mathcal{W}_{\kappa^-1}^0 \uplus \mathcal{W}_\psi^{00} \uplus \mathcal{W}_\psi^{01} \uplus \sigma^{-1}\mathcal{W}_\psi^{10} \uplus \sigma^{-1}\mathcal{W}_\psi^{11} \uplus \mathcal{W}_{\psi^-1}^0 \uplus \sigma^{-1}\mathcal{W}_{\psi^-1}^0.$$

This finally gives us a sufficiently thorough decomposition of $\mathbf{S}^2$ to explicitly apply our corollary. This then gives us

$$\mathbf{S}^2 = \kappa^{-1}\mathcal{W}_\kappa^{00} \uplus \kappa^{-1}\mathcal{W}_\kappa^{01} \uplus \sigma^{-1}\kappa^{-1}\mathcal{W}_\kappa^{10} \uplus \sigma^{-1}\kappa^{-1}\mathcal{W}_\kappa^{11} \uplus \mathcal{W}_{\kappa^-1}^0 \uplus \sigma^{-1}\mathcal{W}_{\kappa^-1}^0, \text{ and}$$
$$\mathbf{S}^2 = \psi^{-1}\mathcal{W}_\psi^{00} \uplus \psi^{-1}\mathcal{W}_\psi^{01} \uplus \sigma^{-1}\psi^{-1}\mathcal{W}_\psi^{10} \uplus \sigma^{-1}\psi^{-1}\mathcal{W}_\psi^{11} \uplus \mathcal{W}_{\psi^-1}^0 \uplus \sigma^{-1}\mathcal{W}_{\psi^-1}^0.$$

Notice how, for reasons previously discussed, we apply our corollary to $\mathbf{S}^2 - \mathcal{D}$ before rotating by σ^{-1}.

In order to make our next proof easier we will rewrite this decomposition one more time. Let

$$\mathcal{A}_1 = \mathcal{M}^0 \ \mathcal{A}_2 = \mathcal{M}^1$$
$$\mathcal{A}_3 = \mathcal{W}_\kappa^{00} \ \mathcal{A}_4 = \sigma^{-1}\mathcal{W}_\kappa^{10}$$
$$\mathcal{A}_5 = \mathcal{W}_\kappa^{01} \ \mathcal{A}_6 = \sigma^{-1}\mathcal{W}_\kappa^{11}$$
$$\mathcal{A}_7 = \mathcal{W}_{\kappa^{-1}}^0 \ \mathcal{A}_8 = \sigma^{-1}\mathcal{W}_{\kappa^{-1}}^{01}$$
$$\mathcal{A}_9 = \mathcal{W}_\psi^{00} \ \mathcal{A}_{10} = \sigma^{-1}\mathcal{W}_\psi^{10}$$
$$\mathcal{A}_{11} = \mathcal{W}_\psi^{01} \ \mathcal{A}_{12} = \sigma^{-1}\mathcal{W}_\psi^{11}$$
$$\mathcal{A}_{13} = \mathcal{W}_{\psi^{-1}}^0 \ \mathcal{A}_{14} = \sigma^{-1}\mathcal{W}_{\psi^{-1}}^{01}.$$

This then gives us:

$$\mathbf{S}^2 = \mathcal{A}_1 \uplus \mathcal{A}_2 \uplus \mathcal{A}_3 \uplus \mathcal{A}_4 \uplus \mathcal{A}_5 \uplus \mathcal{A}_6 \uplus \mathcal{A}_7 \uplus \mathcal{A}_8 \uplus \mathcal{A}_9 \uplus \mathcal{A}_{10} \uplus \mathcal{A}_{11} \uplus \mathcal{A}_{12} \uplus \mathcal{A}_{13} \uplus \mathcal{A}_{14}, \text{ so}$$
$$\mathbf{S}^2 = \kappa^{-1}\mathcal{A}_3 \uplus \kappa^{-1}\sigma\mathcal{A}_4 \uplus \sigma^{-1}\kappa^{-1}\mathcal{A}_5 \uplus \sigma^{-1}\kappa^{-1}\sigma\mathcal{A}_6 \uplus \mathcal{A}_7 \uplus \mathcal{A}_8, \text{ and}$$
$$\mathbf{S}^2 = \psi^{-1}\mathcal{A}_9 \uplus \psi^{-1}\sigma\mathcal{A}_{10} \uplus \sigma^{-1}\psi^{-1}\mathcal{A}_{11} \uplus \sigma^{-1}\psi^{-1}\sigma\mathcal{A}_{12} \uplus \mathcal{A}_{13} \uplus \mathcal{A}_{14}.$$

Note that we rotate $\mathcal{A}_4$, $\mathcal{A}_6$, $\mathcal{A}_{10}$ and $\mathcal{A}_{12}$ by sigma first because above they are first intersected with $\mathbf{S}^2 - \Omega$. This means that we don't have to "erase" them as we would an element of $\mathcal{D}$. This then completes the Banach-Tarksi paradox for $\mathbf{S}^2$, which is explicitly stated below.

Main Theorem 4.10. ***The Banach-Tarski Paradox on*** $\mathbf{S}^2$

The sphere $\mathbf{S}^2$ *in* $\mathbb{R}^3$ *can be partitioned into a finite number of pieces which can be rotated to form two copies of* $\mathbf{S}^2$.

As we will see next, there is a very natural step to the Banach-Tarski Paradox on $\mathbf{B}^3$.

4.6 The Banach-Tarksi on $\mathbf{B}^3$

The Banach-Tarski Paradox on $\mathbf{B}^3$ is one of the most well known paradoxes in math. It is far more famous than the two proceeding theorems: the Hausdorff Paradox, and the Banach-Tarski Paradox on $\mathbf{S}^2$. Perhaps this is due to the tangible nature of $\mathbf{B}^3$. Certainly, one could argue that it gives a more paradoxical feel.

In order to prove our paradox on $\mathbf{B}^3$, we will rely on these-lesser known results. As a reminder, to get to our $\mathbf{S}^2$ paradox we built on the Hausdorff Paradox by finding a way to "erase" our rotational fixed points, $\mathcal{D}$. In a similar sense, we will have to "erase" a trouble point when we try and extend this paradox to $\mathbf{B}^3$. First, since we already established

$$\mathbf{S}^2 = \kappa^{-1}\mathcal{A}_3 \uplus \kappa^{-1}\sigma\mathcal{A}_4 \uplus \sigma^{-1}\kappa^{-1}\mathcal{A}_5 \uplus \sigma^{-1}\kappa^{-1}\sigma\mathcal{A}_6 \uplus \mathcal{A}_7 \uplus \mathcal{A}_8, \text{ and}$$

$$\mathbf{S}^2 = \psi^{-1}\mathcal{A}_9 \uplus \psi^{-1}\sigma\mathcal{A}_{10} \uplus \sigma^{-1}\psi^{-1}\mathcal{A}_{11} \uplus \sigma^{-1}\psi^{-1}\sigma\mathcal{A}_{12} \uplus \mathcal{A}_{13} \uplus \mathcal{A}_{14},$$

we will attempt to simplify notation by defining the following rotations:

$$\rho_3 = \kappa^{-1} \ \rho_4 = \kappa^{-1}\sigma$$

$$\rho_5 = \sigma^{-1}\kappa^{-1} \ \rho_6 = \sigma^{-1}\kappa^{-1}\sigma$$

$$\rho_7 = 1 \ \rho_8 = 1$$

$$\rho_9 = \psi^{-1} \ \rho_{10} = \psi^{-10}\sigma$$

$$\rho_{11} = \sigma^{-1}\psi^{-1} \ \rho_{12} = \sigma^{-1}\psi^{-1}\sigma$$

$$\rho_{13} = 1 \ \rho_{14} = 1.$$

This simplifies our notation so that in order to use our corollary, we simply have to match each ρ_i with its corresponding $\mathcal{A}_i$. Because of this, we can write the Banach-Tarski Paradox on $\mathbf{S}^2$ as

$$\mathbf{S}^2 = \biguplus_{i=3}^{8} \rho_i\mathcal{A}_i = \biguplus_{i=9}^{14} \rho_i\mathcal{A}_i.$$

Next we will have transition from subsets of $\mathbf{S}^2$ to subsets of $\mathbf{B}^3$. Given that we already have defined sets $\mathcal{A}_i$, it would seem to make the most sense to extend these sets inward towards the origin. In fact, if we do this there are very few complications in moving from $\mathbf{S}^2$ to $\mathbf{B}^3$.

First let us define our partition of $\mathbf{B}^3$ using the following sets. Let

$$\mathcal{B}_i \subseteq \mathbf{B}^3, \text{ for } 1 \le i \le 14, \text{ and using spherical coordinates define}$$

$$\mathcal{B}_i = \{(r, \theta, \phi) | (1, \theta, \phi) \in \mathcal{A}_i, 0 < r \le 1\}.$$

This defines each $\mathcal{B}_i$ as we described above. Make sure to notice that the origin isn't in any $\mathcal{B}_i$ as it is a fixed point of every rotation we have so far defined. Because of this we can write

$$\mathbf{B}^3 - 0 = \uplus_{i=1}^{14}\mathcal{B}_i.$$

Our new rotation also allows us to easily apply our corollary these sets and gives us

$$\mathbf{B}^3 - 0 = \uplus_{i=3}^{8} \rho_i \mathcal{B}_i = \uplus_{i=9}^{14} \rho_i \mathcal{B}_i.$$

So all we have to do to get the Banach-Tarski Paradox on $\mathbf{B}^3$ is to find a way to "erase" the origin as we previously "erased" $\mathcal{D}$. However, this is made a little more difficult as every rotation running through the origin has 0 as a fixed point. Fortunately, now that we are dealing with the solid ball we can define rotations not running through the origin without losing closure properties.

Let's do exactly this and take $p = (0, 0, \frac{1}{10})$. We will then let s be the line running through p while parallel to the y-axis. Next we will need a rotation similar to those we've used before, so let τ be a $\theta = 1$ radian counterclockwise rotation around s. Clearly this is similar to the rotation used in the circle trick. We can see that for all positive n, $\tau^n(0) \neq 0$. From this we can also conclude that for all $m \neq n$, $\tau^n(0) \neq \tau^m(0)$.

Note τ isn't unique. Hypothetically, p could be any non-zero point with $\mathrm{d}(p, 0) < 0.5$, while θ can be any rational value.

Similarly to before, lets define a set

$$\mathfrak{T} = \uplus_{n=0}^{\infty} \tau^n(0).$$

Then we can see that $\tau(\mathfrak{T}) = \mathfrak{T} - 0$, so in parallel to our use of Ω with $\mathbf{S}^2$, we can say that

$$\mathbf{B}^3 = (\mathbf{B}^3 - \mathfrak{T}) \uplus \mathfrak{T}, \text{ and}$$
$$\mathbf{B}^3 - 0 = (\mathbf{B}^3 - \mathfrak{T}) \cap \tau(\mathfrak{T}).$$

Also, similarly to how$\mathcal{D} \subset \Omega$, we can see $0 \in \mathfrak{T}$, so

$$\mathbf{B}^3 = [(\mathbf{B}^3 - 0) \cap (\mathbf{B}^3 - 0) \uplus \tau^{-1}[(\mathbf{B}^3 - 0) \cap \mathfrak{T}].$$

So we can apply our corollary to

$$\mathbf{B}^3 - 0 = [\uplus_{n=1}^{14} \rho_i \mathcal{B}_i \cap (\mathbf{B}^3 - \mathfrak{T})] \uplus [\uplus_{n=1}^{14} \rho_i \mathcal{B}_i \cap (\mathfrak{T})]$$

to get

$$\mathbf{B}^3 = [\uplus_{n=1}^{14} \mathcal{B}_i \cap (\mathbf{B}^3 - \mathfrak{T})] \uplus \tau^{-1}[\uplus_{n=1}^{14} \mathcal{B}_i \cap (\mathfrak{T})]$$
$$\mathbf{B}^3 = [\uplus_{n=3}^{8} \rho_i \mathcal{B}_i \cap (\mathbf{B}^3 - \mathfrak{T})] \uplus \tau^{-1}[\uplus_{n=3}^{8} \rho_i \mathcal{B}_i \cap (\mathfrak{T})]$$
$$\mathbf{B}^3 = [\uplus_{n=9}^{14} \rho_i \mathcal{B}_i \cap (\mathbf{B}^3 - \mathfrak{T})] \uplus \tau^{-1}[\uplus_{n=9}^{14} \rho_i \mathcal{B}_i \cap (\mathfrak{T})].$$

Once again, we started with $\mathbf{B}^3 - 0$, in order to get the full Banach-Tarski Paradox we will need to find a way to start with $\mathbf{B}^3$. Unfortunately this again means that we need to add more notation, however, it is a format similarly to what we've done before. Recall how we further decomposed our partition of $\mathbf{S}^2$ using $\mathcal{W}_\rho$ into 4 pieces. For $\mathcal{W}_\rho^{ij}$, we said that $i = 0$ denoted that we were using the intersection of $\mathcal{W}_\rho$ with $\mathbf{S}^2 - \Omega$, while $i = 1$ indicated that we were using the intersection of $\mathcal{W}_\rho$ with Ω. Similarly, $j = 0$ signaled a further intersection with $\rho^{-1}(\mathbf{S}^2 - \Omega)$, while $j = 1$ signaled an intersection with $\rho^{-1}(\Omega)$. We will use these same partitioning rules except this time on each $\mathcal{B}_i$ and ρ_i. So for $i = 1, 2$,

$$\mathcal{B}_i^0 = \mathcal{B}_i \cap (\mathbf{B}^3 - \mathfrak{T}), \text{ and } \mathcal{B}_i^1 = \mathcal{B}_i \cap (\mathfrak{T}),$$

Next, for $3 \leq i \leq 14$, define

$$\mathcal{B}_i^{00} = \mathcal{B}_i \cap (\mathbf{B}^3 - \mathfrak{T}) \cap \rho_i^{-1}(\mathbf{B}^3 - \mathfrak{T})$$
$$\mathcal{B}_i^{01} = \mathcal{B}_i \cap (\mathbf{B}^3 - \mathfrak{T}) \cap \rho_i^{-1}(\mathfrak{T})$$
$$\mathcal{B}_i^{10} = \mathcal{B}_i \cap (\mathfrak{T}) \cap \rho_i^{-1}(\mathbf{B}^3 - \mathfrak{T})$$
$$\mathcal{B}_i^{11} = \mathcal{B}_i \cap (\mathfrak{T}) \cap \rho_i^{-1}(\mathfrak{T}).$$

This leaves us with a possible 50 sets partitioning $\mathbf{B}^3$; however, many of these sets will be empty. This is inconsequential to the proof, and singling out these empty sets for elimination would be unnecessarily burdensome.

Using this partition we can represent $\mathbf{B}^3$ as

$$\begin{aligned}\mathbf{B}^3 = \mathcal{B}_1^0 \uplus \tau^{-1}\mathcal{B}_1^1 \uplus \mathcal{B}_1^0 \uplus \tau^{-1}\mathcal{B}_1^1 \uplus\ [\uplus_{n=3}^{14}\mathcal{B}_n^{00}] \uplus\ [\uplus_{n=3}^{14}\mathcal{B}_n^{01}] \uplus\ [\uplus_{n=3}^{14}\tau^{-1}\mathcal{B}_n^{10}] \uplus\ [\uplus_{n=3}^{14}\tau^{-1}\mathcal{B}_n^{00}]\\ = [\uplus_{n=3}^{8}\rho_n\mathcal{B}_n^{00}] \uplus\ [\uplus_{n=3}^{8}\rho_n\mathcal{B}_n^{01}] \uplus\ [\uplus_{n=3}^{8}\tau^{-1}\rho_n\mathcal{B}_n^{10}] \uplus\ [\uplus_{n=3}^{8}\tau^{-1}\rho_n\mathcal{B}_n^{00}]\\ = [\uplus_{n=9}^{14}\rho_n\mathcal{B}_n^{00}] \uplus\ [\uplus_{n=9}^{14}\rho_n\mathcal{B}_n^{01}] \uplus\ [\uplus_{n=9}^{14}\tau^{-1}\rho_n\mathcal{B}_n^{10}] \uplus\ [\uplus_{n=9}^{14}\tau^{-1}\rho_n\mathcal{B}_n^{00}].\end{aligned}$$

Note that we don't have to add the inverse words separately as they are accounted for in the formula for their respective ρ's. $[\rho_{7/8/13/14} = 1]$

This gives us the Banach-Tarski Paradox on $\mathbf{B}^3$. We can see that this proof isn't unique to the unit ball; we can extend this paradox by changing the radius of our $\mathcal{B}_i \subseteq \mathbf{B}^3$. Explicitly stated, this is:

Main Theorem 4.11. ***The Banach-Tarski Paradox on*** $\mathbf{B}^3$***:***
Any solid ball $\boldsymbol{B}$ *in* $\mathbb{R}^3$ *can be cut into finitely many pieces which can be rotated to form two copies of* $\boldsymbol{B}$*.*

5 The Division Paradox

5.1 Introduction

As previously noted, a weakening of the Axiom of Choice may be sufficient to overcome the Banach-Tarski Paradox; however, this doesn't address any other problems which may come about. This section addresses a consequence of an intuitive weakening of ZFC which would eliminate the Banach-Tarksi Paradox, known as the Division Paradox. First discovered by Mycielski and Sierpiński, the Division Paradox is an undesirable result dealing with the cosets of $\mathbb{Q}$ in $\mathbb{R}$, which we previously used in our proof of non-measurable sets. In this section we will rely on a proof put forward by Professor Alan Taylor of Union College and Professor Stan Wagon of Macalester College. [6]

5.2 Preliminaries

First we will dispense with some definitions.

Definition 5.1. Let X be a topological space. Let $A \subseteq X$. Then A is said to be **meager** if A can be expressed as a countable union of nowhere dense sets in X.

Definition 5.2. Let $A \subseteq \mathbb{R}$. Then A is **$\mathbb{Q}$-invariant** if $\forall q \in \mathbb{Q}\ A + q = A$.

Remark 5.3. Consider the group $(\mathbb{R}, +)$ and its subgroup of $(\mathbb{Q}, +)$. Then we can form the quotient group $(\mathbb{R}/\mathbb{Q}, \oplus)$, where $\mathbb{R}/\mathbb{Q}$ is the collection of equivalence classes with $x \sim y$ iff $x - y \in \mathbb{Q}$, and $[x] \oplus [y] = [x + y]$. Clearly each of these equivalence classes forms a $\mathbb{Q}$-invariant set.

With these definitions in mind we have to first establish a theorem having to do with Lebesgue measure.

Theorem 5.4. ***The Zero-One Rule:***
Assume that a set A is $\mathbb{Q}$-invariant. Then either A or A^c has Lebesgue measure zero.

Proof. Suppose a set A is $\mathbb{Q}$-invariant and is Lebesgue measurable. Let $B = A \cap [0, 1]$ with $\mu(B) = \alpha = 1 \cdot \alpha$. Notice that we are saying that part the portion of A in [0,1] has measure α. We want to use the fact that A is $\mathbb{Q}$-invariant to prove that for any interval $\mathbf{J}$ with rational endpoints, $\mu(A \cap \mathbf{J}) = \alpha\mu(\mathbf{J})$.

First let's let m be rational and consider $A \cap [m, m + 1]$. Since $\alpha = A \cap [0, 1]$, we know that

$$\alpha = inf\{\sum_{j=0}^{\infty} \text{length}(K_J) : \langle K_j \rangle \text{ is a sequence of closed intervals covering } (A \cap [0, 1])\}.$$

Next, since m is a rational number and A is $\mathbb{Q}$-invariant, we know that

$$A \cap [m, m + 1] = A \cap [m + 0, m + 1] = A \cap ([0, 1] + m),$$

but for any rational number q, we know that $A = A + q$ so $A = A + m$, as such

$$A \cap ([0, 1] + m) = (A + m) \cap ([0, 1] + m) = (A \cap [0, 1]) + m,$$

so the sequence $\langle K_j \rangle + m$ would cover $A \cap [m, m+1]$. Since shifting intervals doesn't change their length we can conclude that $\mu(A \cap [m, m + 1]) = \alpha$. Since we know that measure is countably additive we know that for any natural number n, $\mu(A \cap [0, n]) = \alpha \cdot n$. We can also divide this result into subintervals by any natural number $z - \{0\}$ to get $\mu(A \cap [0, \frac{n}{z}]) = \alpha \cdot \frac{n}{z}$. By shifting this result we get our result that $\mu(A \cap \mathbf{J}) = \alpha\mu(\mathbf{J})$.

Next, suppose that $\alpha < 1$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$ we can say that there is a series of closed intervals covering A $\{M_i\}_{i=0}^{\infty}$, such that $\sum_{i=0}^{\infty} \text{length}(M_j) = \beta$, with $\alpha < \beta < 1$. Let $\epsilon = \alpha(1 - \beta)$. Since the length of this series has to approach zero there exists a natural number n such that $\bigcup_{i=0}^{n} M_j$ covers all of B except for a set of measure less than ϵ. Thus we have

$$\alpha = \mu(B) < \epsilon + \left(\bigcup_{i \leq n} B \cap K_i\right) \leq \epsilon + \sum_{i \leq 1} \mu(B \cap K_i) = \epsilon + \alpha \sum_{i \leq n} \mu(K_i) \leq \alpha(1 - \beta) + \alpha\beta = \alpha,$$

a contradiction so we know that $\alpha = 0$ or $\alpha = 1$, as desired. □

5.3 The Division Paradox

With this in hand we can begin to look at our proof of the Division Paradox. The proof of this will drift from our usual domain of ZFC. Without the Axiom of Choice we go in a couple different directions. The most obvious of which would be to use a weaker form of the Axiom of Choice, such as the Axiom of Dependent Choice (**DC**). In order to absolutely disallow the Banach-Tarksi paradox we will also assume that all sets are Lebesgue measurable (**LB**). This works because if all sets were Lebesgue measurable, then in our proof of the Banach-Tarksi Paradox each of our subsets of $\mathbf{B}^3$ would be measurable, so any rotation wouldn't change the measure of these sets. Let's say that $\mathbf{B}^3$ has measure γ. Then the result of rotations would result in two sets, each with measure γ, which would be a contradiction. Thus we will assume ZF + DC + LM.[8]

Main Theorem 5.5. ***The Division Paradox*** *(ZF+DC+LM)*[9]
$|\mathbb{R}| < |\mathbb{R}/\mathbb{Q}|$

This would seem incredibly counterintuitive, especially when looking at the finite. Imagine if I told said that there were 50 baskets, each filled with a different type of fruit. Then you would be able to say with certainty that between the baskets there were at least 50 total pieces of fruit. In this analogy we would say that each basket makes up an equivalence classes of fruit. If we look at the set $\mathbb{R}$ partitioned by the cosets of $\mathbb{Q}$, the Division Paradox says that there are more cosets of $\mathbb{Q}$ in $\mathbb{R}$ than there are real numbers. In the context of our fruit analogy this says that there are more types of fruit than there are total fruit. [10]

As we proved in chapter 2, when taking the Axiom of Choice we can define a bijection from $\mathbb{R}$ to $\mathbb{R}/\mathbb{Q}$. The Division Paradox is exactly that this result doesn't hold without taking the Axiom of Choice. In order to show this we will use the following theorem.

Theorem 5.6. $|\mathbb{R}| \leq |\mathbb{R}/\mathbb{Q}|$

Proof. In order to prove this we will define an injection from the set $\mathbb{R}$ to the set $\mathbb{R}/\mathbb{Q}$. The Schröder-Bernstein theorem tells us that if we can define such an injection we have our desired result. [11]

First let's look at the size of $\mathbb{N}$ and $\mathbb{Q}$. The identity mapping $i : \mathbb{N} \to \mathbb{Q}$ is clearly an injection. We can easily define a bijection between $\mathbb{N}$ and $\mathbb{Z}$ by sending 0 to 0, the odds to the set of positive integers and the even natural numbers to the negative integers. Using the Cantor Pairing function we can define a bijection between $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$. Using this we can map an injection from the rationals to $\mathbb{N} \times \mathbb{Z}$. Thus we have

$$|\mathbb{N}| \leq |\mathbb{Q}| \leq |\mathbb{N} \times \mathbb{Z}| = |\mathbb{N}|, \text{ so we know that}$$

$$|\mathbb{N}| = |\mathbb{Q}|. \tag{5.1}$$

We can then see that $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{N} \times \mathbb{Q}| = |\mathbb{Q} \times \mathbb{Q}|$.

By using equation (1.1) we can index the rational numbers in [0,1] by the natural numbers. We will denote this $\{q_i\}_{i \in \mathbb{N}}$. Next we will build a binary tree of closed intervals starting with [0,1]. We will say that each interval is contained in its parent and is disjoint from the other interval coming from the same parent. We can use induction to see that at each level of our tree, the closed intervals are pairwise disjoint. We will denote each left branch as a

[8] Interestingly for this to be consistent we also have to assume the existence of inaccessible cardinals.
[9] We can also prove this assuming that every subset of $\mathbb{R}$ has the Property of Baire, with little structural difference.
[10] There is a further analogy that takes this one set further. If looking at a professional sports this would be like saying there are more leagues than teams and more teams than players. For a further discussion of this see [6].
[11] While many proofs of the Schröder-Bernstein theorem use the Axiom of Choice, it is not necessary.

Tree Diagram

$$[0,1]$$

$$([a_0, a_1] \quad \cap \quad [b_0, b_1]) = \emptyset$$

$$\emptyset = ([a_{00}, a_{10}] \cap [a_{10}, a_{11}]) \quad ([b_{00}, b_{01}] \cap [b_{10}, b_{11}]) = \emptyset$$

and so on.

Figure 2: a and b are included as clarification, to see the true binary representation replace a with 0 and b with 1.

0 and each right branch as a 1, so each branch is uniquely denoted using a binary representation. For an example, see the diagram below, replacing a with 0 and b with 1. Using Cantor's Intersection Theorem, we can see that the intersection of each of these branches is nonempty. In each of these branches the length of the interval converges to 0 so it must converge to a single point. This representation works perfectly as we can represent each of the elements of [0,1] in binary form, so we will map each $x \in [0,1]$ to its branch with the same binary representation. We denote this branch y_x and write our function $\lambda(x) = [y_x]$. Thus for a level of our tree q_n, we want to space our intervals such that for all $m < n$, any two intervals $\mathbf{I}$ and $\mathbf{J}$, no point in $\mathbf{I}$ has a distance q_m from any point in $\mathbf{J}$. Thus for any $x, z \in [0,1]$ we know that $y_x - y_z$ is irrational so y_x and y_z are not equivalent.

So as you can imagine, how we split these intervals is the most important part of this argument. We will make the interval to the left have a length less than $\frac{1}{3}$ of that to its right. Thus for an arbitrary level of our tree diagram the lengths of each interval are

$$[\frac{1}{3}] \text{ of } \rightarrow [\frac{1}{3}] \text{ of } \rightarrow [\frac{1}{3}] \text{ of } \rightarrow [\frac{1}{3}] \text{ of } \rightarrow ...$$

So for any shift of the left interval, it doesn't intersect all of the thirds of the interval to its right, or the interval two to the right, and so on inductively. Suppose it did, and it intersects all three intervals. Let's let the measure of the initial interval to the right be α. Then if our new left interval has a length of β, we know $\beta < \frac{1}{3}\alpha$. We also know that the middle third of our right interval has a measure of $\frac{1}{3}\alpha$. Next, since these intervals are connected we know that the middle third of right interval is a subset of our shifted left interval. We can go even further and say that it is a proper subset of our left interval. Since a shift doesn't change the measure of an interval we have

$$\frac{1}{3}\alpha < \beta < \frac{1}{3}\alpha,$$

a contradiction.

Let's think back to our indexing of the rationals $\{q_i\}_{i \in \mathbb{N}}$. Since we also showed that $|\mathbb{N}| = |\mathbb{Q} \times \mathbb{Q}|$, we also know that there exists a bijection between $\mathbb{N}$ and $\mathbb{Q} \times \mathbb{Q}$, so we can index $\mathbb{Q} \times \mathbb{Q}$ using $\mathbb{N}$. We denote this indexing $\{p_i\}_{i \in \mathbb{Q}}$. Let's do this with $p_1 = (q_1, q_1)$,

$p_2 = (q_2, q_1)$, $p_3 = (q_2, q_2)$, $p_4 = (q_3, q_1)$, ... This representation then equates each interval between triangle numbers with the first coordinate.[12] Then for each level of our tree indexed by $\{p_i\}_{i\in\mathbb{N}}$ we will shrink each interval so that its length is a third of the interval to its right. We must also choose which third of our interval we want to specify. We will do this using the second coordinate of our $p_n = (q_i, q_j)$ representation. For a given interval $a_{01\ldots0}$[13] at the level $q_i, q_j)$ of our tree diagram, we will choose $a_{01\ldots01}$ as a subset of the third of $a_{01\ldots0}$ not intersected by $a_{01\ldots00} + q_j$. If there are multiple such thirds of $a_{01\ldots0}$ choose the leftmost.[14]

Tree Diagram P-Indexed

$p_1 = (q_1, q_1)$		$[0,1]$	
$p_2 = (q_2, q_1)$	$([a_0, a_1]$	$\cap$	$[b_0, b_1]) = \emptyset$
$p_3 = (q_2, q_2)$	$\emptyset = ([a_{00}, a_{10}] \cap [a_{10}, a_{11}])$		$([b_{00}, b_{01}] \cap [b_{10}, b_{11}]) = \emptyset$

and so on.

Figure 3: a and b are included as clarification, to see the true binary representation replace a with 0 and b with 1.

Thus for each level of our tree $p_n = (q_i, q_j)$, we know for all $j < i$, any two intervals **I** and **J**, no point in **I** has a distance q_j from any point in **J**, as desired.

We will simplify our tree diagram one last time. Let's now index the levels of our tree with q_n. We want to relate this to our tree diagram indexed by $p_b = (q_i, q_j)$. Lets make the level of our new tree diagram denoted q_n equal to level of our previous diagram denoted by p_k, where $k = \frac{1}{2}(n(n+1)) + 1$. Thus each level of our new tree diagram is equal to the first level of our old tree diagram with q_{n+1} in the first coordinate. We can then be sure that for each level of our tree p_n, we know that for all $m < n$, any two intervals **I** and **J**, no point in **I** has a distance q_m from any point in **J**, as desired.

Tree Diagram Q-Indexed

$q_1 = p_1$		$[0,1]$	
$q_2 = p_4$	$([a_0, a_1]$	$\cap$	$[b_0, b_1]) = \emptyset$
$q_3 = p_7$	$\emptyset = ([a_{00}, a_{10}] \cap [a_{10}, a_{11}])$		$([b_{00}, b_{01}] \cap [b_{10}, b_{11}]) = \emptyset$

and so on.

Figure 4: a and b are included as clarification, to see the true binary representation replace a with 0 and b with 1.

If we look inductively down the tree this gives us an injection from [0,1] to $\mathbb{R}/\mathbb{Q}$. We can see $f : (0,1) \Rightarrow \mathbb{R}$ given by $f(x) = (\tan(\frac{\pi}{2}(2x-1))$ is a bijection, and since $(0,1) \subseteq [0,1] \subseteq \mathbb{R}$,

[12] The $n^t h$ triangle number is given by $\frac{1}{2} \cdot n(n+1)$. We will use this later

[13] The 0's and 1's are arbitrary.

[14] We can define this choice function since at each interval gives us a finite choice, and at each level there are finite intervals.

we know there exists a bijection between [0,1] and $\mathbb{R}$. Hence, there is an injection from $\mathbb{R}$ to $\mathbb{R}/\mathbb{Q}$, so $|\mathbb{R}| \leq |\mathbb{R}/\mathbb{Q}|$ as desired. $\square$

We are now ready to proceed with our proof of the Division Paradox.

Proof. Suppose that $|\mathbb{R}/\mathbb{Q}| \leq \mathbb{R}$. Then there exists an injection from $\mathbb{R}/\mathbb{Q}$ to $\mathbb{R}$. Since $\mathbb{R}$ is linearly ordered, this injection would induce a linear ordering of $\mathbb{R}/\mathbb{Q}$. We can then define

$$A = \{x \in \mathbb{R} : [x] \leq [-x]\}.$$

Let's examine this set a more closely. Suppose x and $-x$ are in A. Then $x - (-x) \in \mathbb{Q}$, so $2x \in \mathbb{Q}$, so $x \in \mathbb{Q}$. This tells us that the $x \sim -x$ if and only if $x \in \mathbb{Q}$, so $\mathbb{Q} \subseteq A$. Since A is composed of the cosets of $\mathbb{Q}$ in $\mathbb{R}$ we can also say that A is $\mathbb{Q}$-invariant. Now let's look at elements of $A - \mathbb{Q}$. We know either $x \in A$ or $x \in \mathbb{R} - A$, so we can define a bijection $\rho : A - \mathbb{Q} \to \mathbb{R} - A$ by $\rho(x) = -x$. Further, this bijection preserves measure, as any cover for $A - \mathbb{Q}$, would cover $\mathbb{R} - A$ once negated.

Since all sets are Lebesgue measurable clearly A is measurable. By the Zero-One Rule we know that either A or $\mathbb{R} - A$ has measure zero. But since $\mathbb{Q}$ is countable, we know that A is measure zero if and only if $A - \mathbb{Q}$ is measure zero. $A - \mathbb{Q}$ is measure zero if and only if $\mathbb{R} - A$ is measure zero. But then A and A^c are both measure zero, a contradiction to the Zero-One Rule. Thus we know that there is no linear ordering of $\mathbb{R}/\mathbb{Q}$. This is again a contradiction, so we know that no injection exists from $\mathbb{R}/\mathbb{Q}$ to $\mathbb{R}$ and by the Schröder-Bernstein theorem, $|\mathbb{R}/\mathbb{Q}| \not\leq |\mathbb{R}|$, so $|\mathbb{R}| < |\mathbb{R}/\mathbb{Q}|$. $\square$

6 Conclusion

While the Banach-Tarski Paradox may be seen as undesirable to many mathematicians, many would feel the same about the Division Paradox. While weighing the magnitude of the Banach-Tarski Paradox and the Division Paradox, one may want to examine the actual role of the Axiom of Choice in their proof. As we noted at the start of this paper, the Axiom of Choice is not needed for a finite collection of finite sets. Instead, one can see the Banach-Tarski Paradox as a result of infinite sets. We would not be able to use our free group corollary without relying on the fact that $\aleph_0 - 1 = \aleph_0$. Additionally, upon further inspection this is not the only paradoxical result having to do with the infinite. Without the Axiom of Choice, Galileo was able prove that there exists an injective function from the set of integers to the set of squares. [5]

This concept of infinite sets necessarily removes the Banach-Tarski Paradox from any physical reality which we may encounter. As a result, any preconceived concept of measure which we may have is not applicable. It would then seem that the Division Paradox would be unpalatable, even when compared to the Banach-Tarski Paradox. This leads us to conclude that ZF+LM+DC is not an acceptable axiomatic system. Finally, the Axiom of Choice and its seemly paradoxical results are not so obviously false as to necessitate its exclusion from a reasonable axiomatic system. In fact, many equivalent results are intuitive enough to conclude that the exclusion of the Axiom of Choice would be unreasonably costly.

References

1. J. Belk, *Non-Measurable Sets*, http://math.bard.edu/belk/math461/NonMeasurableSets.pdf
2. C. Hardin and A. Taylor, *The Mathematics of Coordinated Inference: A Study of Generalized Hat Problems*, Springer, 2013.
3. H. Herrlich, *Axiom of Choice*, Berlin, Springer, 2006.
4. J. Mycielski, *A System of Axioms of Set Theory for the Rationalists*
5. S. Wagon, The Banach-Tarski Paradox, *Encyclopedia of Mathematics and Its Application*, Press Syndicate of the University of Cambridge, 1985.
6. S. Wagon and A. Taylor, *A Paradox Arising from the Elimination of a Paradox*, 2017.
7. T. Weston, *Banach-Tarski Paradox*