Volume 00, Number 0, Pages 000–000
S (XX)0000-0

APPLICATIONS OF THE AXIOM OF CHOICE: FROM THE INFINITE HAT RIDDLE TO THE BANACH-TARSKI PARADOX

(Communicated by Sudhir Murthy)

Abstract. The axiom of choice is popularly employed as a "scare tactic", at the undergraduate level, to justify the strangeness of various mathematical phenomena. This expository paper for undergraduate analysis students, reviews three novel applications of the axiom of the choice: (1) the solution to the infinite hat riddle, (2) Vitali's immeasurable subset of the unit interval, and (3) Banach-Tarski's paradox, which astonishingly reassembles a ball into two identical copies. Although these proofs are not new, our initial aim was to establish connections between paradoxes (3) and (1). However, after encountering unexpected challenges, we have redirected our focus. Our revised objective is to present these theorems through illustrative analogies paying close attention to the underlying logic, particularly the axiom of choice.

1. Introduction

The axiom of choice is a powerful tool in classical reasoning, employed freely, although implicitly, in several mathematical proofs. For example, choice implies that a countable union of countable sets is countable, and that the sequential characterization of continuity aligns with the ε, δ criterion. Despite its many forms of varying degrees of strength [4] (countable choice, dependent choice, etc.), in this paper we focus on the broadest version of the axiom.

A classical example of the bizarreness of the axiom of choice is the demonstration of immeasurable sets. We assume the reader is familiar with undergraduate analysis, including an introductory survey of measure theory. As notation, we use $\mathbb{N}$ to denote the set of natural numbers (zero omitted), $\mathbb{R}$ real numbers, $\mathcal{P}(\mathbb{R})$ as the power set of $\mathbb{R}$, and work with extended, nonnegative reals $[0, +\infty]$. We aim to prove the following theorems which, historically, have motivated measure theory.

Theorem 1.1 (Vitali's Theorem)**.** *There is no set function $\mu : \mathcal{P}(\mathbb{R}) \to [0, +\infty]$ which is translation invariant, assigns intervals their usual lengths, and is σ-additive.*

Theorem 1.2 (Extended Vitali's Theorem)**.** *For all $d \in \mathbb{N}$, there is no set function $\mu : \mathcal{P}(\mathbb{R}^d) \to [0, \infty]$ which is translation invariant, assigns rectangles their usual volume, and is σ-additive.*

For any good sense of *volume*, it's clear we impose the first two conditions. We can even impose the stronger condition of invariance to rigid transformations which includes, in addition, rotational invariance. It's less clear why we impose the third condition, of (countable) infinite additivity. Does there exist a set function which works for finite additivity? Interestingly, for $\mathbb{R}^1$ and $\mathbb{R}^2$, the answer is in the affirmative [7], but for $\mathbb{R}^3$ and higher, in the negative.

Theorem 1.3 (Corollary of Banach-Tarski's Paradox)**.** *There is no set function* $\mu : \mathcal{P}(\mathbb{R}^3) \to [0, \infty]$ *which is rotation invariant, assigns rectangles their usual volume, and is finitely additive.*

Notice the difference between conditions in Theorems 1.3 and 1.2. The condition of σ-additivity has been weakened to finite additivity, whereas translation invariance has been changed to rotational invariance.

All these theorems use the axiom of choice, to form a set of representatives from equivalence classes. Our original goal was to connect the equivalence classes formed for Theorem 1.3 to those for the *infinite hat riddle.* While the author failed in this attempt, our hope is to clearly express these proofs so that the readers, by far more capable, can find connections the author missed.

2. Axiom of Choice

Understanding the axiom of choice requires firm grounding in the axiomatic, logical framework of mathematics, specifically the first-order, classical language of Zermelo Frankel Set Theory (ZFC). The well-formed formulas of this language (wffs) are particular finite arrangements of the usual logical alphabet. Proofs are finite sequences of wffs where each step can be justified as either an *axiom* or by the previous steps together with some *rule of inference.*

While we won't go into further detail of formal ZFC, in summary the axiom of choice is a form of infinitary reasoning whereby infinitely many arbitrary selections can be made. It parallels the axiom of induction in Peano Arithmetic (PA), which allows for chaining infinitely many implication eliminations (*modus ponens*). Given proofs of $\forall n,\ \psi(n) \implies \psi(n+1)$ and $\psi(0)$ in PA, it's possible to prove $\psi(n)$ for any meta-logically fixed number n without induction. The proof consists of n many applications of modus ponens. However, obtaining $\forall n,\ \psi(n)$ demands an additional principle — namely, the axiom of induction.

Similarly, suppose we can prove $\forall n \in \mathbb{N}\ \exists x,\ \psi(n, x)$ in ZFC. In particular, as $\exists x,\ \psi(1, x)$ is true, we can use existential elimination to choose a witness x_1 such that $\psi(1, x_1)$ is true. This process can be continued finitely, using existential elimination n times, to form a finite sequence $x_1, x_2, \ldots, x_n$ for a meta-logically fixed number n, such that $\psi(1, x_1), \psi(2, x_2), \ldots, \psi(n, x_n)$ are true. However, we cannot necessarily form an infinite sequence $\{x_n\}$ with $\psi(n, x_n)$ true for all $n \in \mathbb{N}$.

For the underlying proof can only contain finitely many steps — the axiom of choice completes this infinitary gap.

It's worth mentioning that if we can show $\forall n \in \mathbb{N}\ \exists! x,\ \psi(n, x)$, then without the axiom of choice, from the rest of the machinery of ZFC, we can form the unique sequence $\{x_n\}$ where $\psi(n, x_n)$ holds for all $n \in \mathbb{N}$. Hence the axiom of choice becomes necessary only when we lack a process for unique selection.

Quotient constructions in ZFC are universal in mathematics. They correspond to the activity of identifying points within a set X, and partitioning it into a set of equivalence classes. This terminology 'classes' is suggestive; imagine X as a school full of students. The principal doesn't memorize each student's name, but rather divides them into 'classes' or grades, grouping together students of the same $K-12$ grade level. In a well-ordered set X, a representative student could be designated from each class, a class president so to speak, by picking the first student in an enumeration of their names. But for general partitions of X, a unique selection process isn't always available.

For example, given a partition of $[0, 1]$, there isn't necessarily a mechanism for designating unique representatives for each equivalence class. Picking the minimum of a class doesn't work as a minimum may not exist; picking the infimum is not viable either, as it may not be an element of the same equivalence class. The axiom of choice comes in handy here: it can be used to establish the existence of such a set of representatives. In this paper, the axiom of choice is used in this manner.

3. Infinite Hat Riddle

We summarize the *finite* and *infinite hat problems* as explored by Hardin and Taylor [2]. First, we describe the two-hat riddle: a black or white hat is randomly placed on 100 prisoners. They are positioned in a single file so that each can see all those hats worn by those ahead, but are unable to see their own or those of the prisoners behind. The first prisoner, who sees the rest in front, must loudly announce black or white; then the second, and so on. This guessing game proceeds down the line. During this process, if a prisoner incorrectly describes their hat color, then they're silently executed. While all prisoners can hear the previous guesses, no other form of communication is allowed besides an initial strategy session. For example, they cannot use timing or pitch to convey information. Prior to this game, unaware of their seating order and hat color assignments, prisoners are given an opportunity to devise a strategy. What is the best strategy for survival, given their lives are equally precious?

Theorem 3.1. *There exists a strategy for the two-hat riddle which guarantees the survival of all prisoners, except the first to guess.*

Proof. Each arrangment of prisoners can be seen as a finite sequence of 0's and 1's where the elements correspond to the two possible hat colors. The prisoners learn arithmetic in $\mathbb{Z}_2$, and agree for the first prisoner to announce the sum of all the hats in front of them. While they aren't guaranteed survival, the rest are. For every prisoner after can take the immediate previous announced color, and infer their own by computing the sum of all hats they see. □

Nothing is special about the number of prisoners chosen or hat colors available. So long as there are finitely many of both, the above proof still works; prisoners compute sums in some Z_n. Interestingly, no matter the cardinality of the set of hat colors, the above proof appears to work. The prisoners just have to collectively decide on a group structure. However, this is generally only possible with the axiom of choice. Also observe this strategy fails in the case of countably infinite prisoners, where it's impossible to compute an infinite sum in a group.

The *infinite hat riddle* is similar to the two-hat problem, except there are countably infinite prisoners, an infinite possibility of hat of colors, and, to make matters worse, prisoners are no longer able to hear previous guesses. However, they're told when a guess has been made so they can infer their position in the sequence.

Theorem 3.2. *There exists a strategy for the infinite hat riddle which guarantees the survival of almost all prisoners, i.e. except for finitely many.*

Proof. Let X be the set of hat colors; each possible configuration of prisoners can be regarded as a sequence in X. Define the equivalence relation $\sim$ on the set of all sequences in X by idenftiying those which are eventually equal. Using the axiom of choice, the prisoners agree on some choice function $f : {}^{X}/{\sim} \to X$, so that $f(A) \in A$ for all $A \in {}^{X}/{\sim}$. They additionally have the function $[\{x_n\}]$ at their disposal, which sends each sequence $\{x_n\}$ to its equivalence class.

Now the prisoner agree to carry out the following instructions:

(1) Form a sequence $\{x_n\}$, in X, from the prisoners in front.
(2) Compute $f([\{x_n\}])$, i.e. the chosen representative of $[\{x_n\}]$.
(3) Guess the color given by $f([\{x_n\}])$, at the prisoner's inferred position.

Indeed when this strategy is followed, only finitely many prisoners may be executed. While they don't know the exact hat distribution for all prisoners, they know its eventual behavior and, thus, know the equivalence class the correct distribution belongs to. Eventually, all prisoners will guess correctly. □

Remarkably in this strategy, while a given prisoner cannot tell if they survive, they can infer a large enough subsequent tail of prisoners which will. The prisoners can find comfort in being able to compute the maximum number of deaths in this horrific experiment.

4. Vitali's Immeasurable Set in $\mathbb{R}$

We prove Theorems 1.1 and 1.2 following the usual strategy outlined in [5]. First, we define an equivalence relation on $\mathbb{R}$ which identifies points that are rationally separated; second, we use the axiom of choice to form a complete set of representatives in the unit interval; third, we argue the set formed is immeasurable.

Definition 4.1. The relation $\sim$ on $\mathbb{R}$ is given by $x \sim y \iff y - x \in \mathbb{Q}$.

It's clear $\sim$ is an equivalence relation on $\mathbb{R}$. For each $x \in \mathbb{R}$, we define its equivalence class as usual: $[x] := \{y \in \mathbb{R} \mid x \sim y\}$. The set of equivalence classes, denoted $\mathbb{R}/\mathbb{Q}$, forms a partition of $\mathbb{R}$. The notation is sensible: we imagine dividing $\mathbb{R}$ into uncountably disjoint *copies* of $\mathbb{Q}$. The set $\mathbb{Q}$ itself, expressible as $[0]$, is one such copy and its irrational translates form the rest. We formalize this intuition.

Lemma 4.1. *There exists a surjection from $\mathbb{R}$ to $\mathbb{R}/\mathbb{Q}$.*

Proof. $\mathbb{R}/\mathbb{Q} = \{[x] : x \in \mathbb{R}\}$ is formed by the axiom of replacement. Thus, the canonical map $[\cdot] : \mathbb{R} \to \mathbb{R}/\mathbb{Q}$ given by $x \mapsto [x]$, is a surjection. □

Lemma 4.2. *For each $x \in \mathbb{R}$, the equivalence class $[x]$ is countably infinite.*

Proof. For $x \in \mathbb{R}$ fixed, define $f : \mathbb{Q} \to [x]$ as rational translations of x, i.e. by $f(q) := x + q$ for all q. Indeed f is a bijection. □

Lemma 4.3. $\mathbb{R}/\mathbb{Q}$ *is uncountable.*

Proof. Assume $\mathbb{R}/\mathbb{Q}$ is countable. Since $\mathbb{R} = \bigcup \mathbb{R}/\mathbb{Q}$, by Lemma 4.2, we can express an uncountable set as the countable union of countable sets. That this is a contradiction is a consequence of the axiom of choice. □

We regard equivalence classes $[x]$ as *rational combs*. In this analogy, the teeth of a fixed comb represent elements of $[x]$. Any rational displacement from one tooth will land on another, and the distance between any two teeth is, again, rational. The set $\mathbb{R}/\mathbb{Q}$ contains uncountably many such combs. No two distinct combs touch each other but each comb is identical in shape to the others, offset only by an irrational distance. Thus, one could conceptually detach these combs, each firmly rooted in $\mathbb{R}$, stack them atop one another, and find it impossible to tell them apart.

Of course, this process of pulling the combs apart is untenable. Any instrument capable of grip interacts, by the density of the reals, with infinitely many combs at once. Overcoming this requires extraordinary precision and dexterity.

Lemma 4.4. $\mathbb{R}/\mathbb{Q}$ *is equinumerous with* $\mathbb{R}$.

Proof. From the axiom of choice, select a choice function $f : \mathbb{R}/\mathbb{Q} \to \mathbb{R}$ such that $f(X) \in X$ for all $X \in \mathbb{R}/\mathbb{Q}$. Notice f is an injection since $f([x]) = f([y])$ implies $[x]$ intersects $[y]$ and, thus, are the same.

However, f is not surjective, e.g. there is no equivalence class which gets sent to $f([0]) + 1$. To remedy this, the idea is we can use $f([x])$ as base-points, for each $x \in \mathbb{R}$, and translate by rational amounts.

Define the map $F : \mathbb{R}/\mathbb{Q} \times \mathbb{Q} \to \mathbb{R}$ by $F(X, q) := f(X) + q$ for all X, q. It's easy to show F is a bijection. *(Injection)* Assume $F(X, q) = F(Y, p)$ for $X, Y \in \mathbb{R}/\mathbb{Q}$ and $q, p \in \mathbb{Q}$. Then $f(X) \sim f(Y)$ which implies $X = Y$. Hence $p = q$ as well. *(Surjection)* Fix $x \in \mathbb{R}$. Then observe $x - f([x])$ is rational as both $x, f([x]) \in [x]$ by design. Therefore $F([x], x - f([x])) = x$.

Hence, $\mathbb{R}$ is equinumerous with $\mathbb{R}/\mathbb{Q} \times \mathbb{Q}$, so it remains to show $\mathbb{R}/\mathbb{Q} \times \mathbb{Q}$ is equinumerous with $\mathbb{R}/\mathbb{Q}$. The diagram of injections

$$\mathbb{R}/\mathbb{Q} \to \mathbb{R}/\mathbb{Q} \times \mathbb{Q} \to \mathbb{R}/\mathbb{Q} \times \mathbb{R}/\mathbb{Q} \to \mathbb{R}/\mathbb{Q},$$

pastes. While the first injection is easy to construct, the other two involve the axiom of choice [3] with Lemmas 4.1 and 4.3. From Cantor-Bernstein, all pairs in this chain are in bijection; in particular, $\mathbb{R}/\mathbb{Q} \times \mathbb{Q}$ is equinumerous with $\mathbb{R}/\mathbb{Q}$. $\square$

We now define an immeasurable set. We form a set of representatives for each $X \in \mathbb{R}/\mathbb{Q}$, ensuring these values lie in $[0, 1]$.

Definition 4.2. Put $\mathcal{A} := \{X \cap [0, 1] : X \in \mathbb{R}/\mathbb{Q}\}$.

Lemma 4.5. *$\mathcal{A}$ is a collection of nonempty sets.*

Proof. Fix $X \in \mathbb{R}/\mathbb{Q}$ and let x witness X nonempty. Should $x \in [0, 1]$ then we are done. Otherwise, roughly, we add or subtract by the floor of $|x|$ to ensure $x \in [0, 1]$. $\square$

Definition 4.3. Let $f : \mathcal{A} \to \mathbb{R}$ be a function such that $f(X) \in X$ for all $X \in \mathcal{A}$. Define rational translations by $f(\mathcal{A}) + q := \{x + q : x \in f(\mathcal{A})\}$ for all $q \in \mathbb{Q}$.

Observe that f is well-defined by the axiom of choice and Lemma 4.5.

Lemma 4.6. *For all $q, p \in \mathbb{Q}$, either $f(\mathcal{A}) + q$ and $f(\mathcal{A}) + p$ are disjoint or equal.*

Proof. Fix $q, p \in \mathbb{Q}$ and suppose there is a common point $x \in (f(\mathcal{A}) + q) \cap (f(\mathcal{A}) + p)$. So $x = y + q$ and $x = z + p$ for some $y, z \in f(\mathcal{A})$. Observe that $x \in [y]$ and $x \in [z]$ and therefore $y \sim z$. As f picks out precisely one representative from each equivalence class, it follows $y = z$ and, hence, $p = q$. $\square$

In continuation of the analogy, each equivalence class in $\mathbb{R}/\mathbb{Q}$ can be thought of as a distinctly colored comb. The axiom of choice plays the role of infinitely precise tweezers, capable of selecting exactly one tooth from each colored comb intersecting $[0,1]$. The chosen teeth are reattached to form a single, *rainbow-colored* comb represented by $f(\mathcal{A})$. This set is the crux of the proof of Theorem 1.1. By duplicating the set infinitely many times and translating each copy by a distinct rational in $[-1,1]$, their whole union covers $[0,1]$ but is contained in $[-1,2]$. Thus, $f(\mathcal{A})$, as we'll show, is immeasurable.

Proof. (Theorem 1.1). Suppose there were such a set function μ. Our proof transforms into the demonstration that,

$$[0,1] \subset \bigcup_{q\in\mathbb{Q}\cap[-1,1]} (f(\mathcal{A})+q) \subset [-1,2].$$

For once shown, it follows from monotonicity of μ (derived from σ-additivity) and disjointness of the family $\{f(\mathcal{A})+q\}_{q\in\mathbb{Q}}$ (from Lemma 4.6), that:

$$1=\mu([0,1]) \leq \mu\left(\bigcup_{q\in\mathbb{Q}\cap[-1,1]} (f(\mathcal{A})+q)\right) = \mu(f(\mathcal{A}))\cdot\infty \leq \mu([-1,2]) = 3.$$

The situation is hopeless; no matter what value we make $\mu(f(\mathcal{A}))$, zero or otherwise, the inequality is violated.

To this end fix $x\in[0,1]$. Since $[x]\cap[0,1]\in\mathcal{A}$, the chosen representative $\alpha := f([x]\cap[0,1])$ satisfies $\alpha\sim x$. As $\alpha\in[0,1]$ it follows $x=\alpha+q$ for some rational $q\in[-1,1]$ and, hence, $x\in\bigcup_{q\in\mathbb{Q}\cap[-1,1]}(f(\mathcal{A})+q)$.

Now fix $x\in\bigcup_{q\in\mathbb{Q}\cap[-1,1]}(f(\mathcal{A})+q)$ so that $x=\alpha+q$ for some rational $q\in[-1,1]$ and $\alpha\in f(\mathcal{A})$. As $\alpha\in[0,1]$ it follows $x\in[-1,2]$. $\square$

As a corollary, the above proof can be generalized to larger spaces $\mathbb{R}^d$, by showing the Cartesian product $f(\mathcal{A})\times[0,1]^{d-1}$ is immeasurable in $[0,1]^d$.

Proof. (Theorem 1.2). For $d\geq 2$, the proof similarly relies on the demonstration:

$$[0,1]^d \subset \bigcup_{q\in\mathbb{Q}\cap[-1,1]} (f(\mathcal{A})+q)\times[0,1]^{d-1} \subset [-1,2]\times[0,1]^{d-1}.$$

Note that the ambiguity of the middle expression doesn't matter; both interpretations, whether it's the union of products or the product of a union, are the same. To show these inclusions, the last $d-1$ coordinates remain stationary, whereas the first coordinate uses the same proof above. $\square$

5. Banach-Tarski Paradox

We synthesize and reflect, in our own words, Daniel's formalized proof [1] of the Banach-Tarski paradox (which implies Theorem 1.3). Roughly, the paradox says that a ball can be *equidecomposed* into two identical copies. Two subsets of $\mathbb{R}^3$ are said to be *equidecomposable* if we can partition one into finitely many pieces and reassemble them, by rigid transformations, into the other. We have to use all the pieces in a non-overlapping fashion, so that the process is reversible. Thus equidecomposability is an equivalence relation.

Let S^2 be the set of all points of the unit sphere in $\mathbb{R}^3$, and D^3 of the closed unit ball. The idea is to define an equivalence relation on S^2 and to use the axiom of choice to select a set of representatives. From the given selection, one can re-partition S^2 into six parts. Then one can magically double these parts, by just rotations, and reassemble them elsewhere.

Any mechanism which partitions the surface of a sphere, necessarily partitions the volume — at least, without its center. For we can take radial lines which connect the origin to the surface, and use the partition of the surface to carve out a partition of the volume. The treatment of a center is another matter.

Definition 5.1. Let angle θ be a fixed, irrational multiple of π and let x, x^{-1}, z, z^{-1} represent rotations of $\mathbb{R}^3$ by clockwise rotation by θ of the $+x$, $-x$, $+z$, and $-z$ axes respectively. x corresponds to polar rotation by θ, and z to azimuthal.

Definition 5.2. A *path* is a finite sequence of rotations x, x^{-1}, z, and z^{-1} where there is no consecutive occurrence of xx^{-1}, $x^{-1}x$, zz^{-1}, or $z^{-1}z$. A path $\{p_i\}$ *connects* $a \in \mathbb{R}^3$ to $b \in \mathbb{R}^3$ only when $(p_n \circ \ldots p_2 \circ p_1)(a) = b$.

Paths correspond to composing finite rotations by twisting the $+x$ or $+z$ axes either clockwise or counter-clockwise by θ. Of course, given any *pseudo-path* where inverse pairs appear, such as xx^{-1} or $z^{-1}z$, then one can remove annihilating pairs to form a bonafide path described in Definition 5.2.

By convention, the empty sequence is the identify path which keeps each point, in $\mathbb{R}^3$, stationary.

Definition 5.3. Define the equivalence relation $\sim$ on S^2 by $p \sim q$ if and only if there exists a path which connects a with b. The equivalence class $[a]$ is called the *orbit* of $a \in S^2$.

It's easy to check this is an equivalence relation and that each orbit is a countable set of points.

Definition 5.4. A point $a \in S^2$ is said to be a *fixed point* of a path exactly when the path connects a to a. Let $D \subset S^2$ be the union of orbits of all fixed points for nonempty paths.

Lemma 5.1. *D is countable.*

Proof. By Euler's rotation theorem [6], any composition of rotations is again a rotation about some principal axis. Provided the rotation is not the identity map, the corresponding axis is unique. It turns out by picking θ to be an irrational multiple of π, it's impossible for identity to result from the composition along a nonempty path. Thus, every nonempty path has exactly two fixed points: the poles associated with its unique axis of rotation. Then as there are countably many paths, there are countably many fixed points, so D is a countable union of countable sets and, thus, countable. $\square$

Observe that $\sim$ is an equivalence relation on D^c, that is $S^2 - D$. Parallel to Definition 4.2 and Lemma 4.5, we can use the axiom of choice to find representatives of ${}^{(D^c)}/\sim$ which lie in the complement of D.

Definition 5.5. Put $\mathcal{A} := \{[a] : a \in D^c\}$.

Lemma 5.2. *$\mathcal{A}$ is a collection of nonempty sets contained in D^c.*

Proof. Consider $[a]$ for $a \notin D$; first it's clear $a \in [a]$. Next if $[a]$ intersects D at a point b, then $[a] = [b] \subset D$ which implies $a \in D$, contradicting the given. $\square$

Similarly from the axiom of choice and Lemma 5.2, we can define a choice function $f : \mathcal{A} \to D^c$.

Definition 5.6. Let $f : \mathcal{A} \to D^c$ be a function such that $f(X) \in X$ for all $X \in \mathcal{A}$.

Here are the sets we consider for decomposition.

(1) The set D of fixed points and elements in their orbits, called *poles*.
(2) The image $f(\mathcal{A})$ of representatives called *initial points*.
(3) The set $S(x)$ of points connected by a path, from an initial point, that ends with x, called *x-endpoints*.
(4) The set $S(x^{-1})$ of points connected by a path, from an initial point, that ends with x^{-1}, called *x^{-1}-endpoints*.
(5) The set $S(z)$ of points connected by a path, from an initial point, that ends with z, called *z-endpoints*.
(6) The set $S(z^{-1})$ of points connected by a path, from an initial point, that ends with z^{-1}, called *z^{-1}-endpoints*.

Lemma 5.3. *The sets described above partition S^2.*

Proof. D contains the north pole $(0, 0, 1)$; by Lemma 5.2 $f(\mathcal{A})$ is nonempty, say inhabited by $a \in D^c$; and the points $x(a)$, $x^{-1}(a)$, $z(a)$, and $z^{-1}(a)$ are elements of the corresponding set of endpoints. Thus, these sets are nonempty.

Now we show they're disjoint. By design, initial points don't lie in D, and their orbits are also disjoint from D. Now for the disjointness of $S(x)$ and $S(z)$, suppose a point is simultaneously an x-endpoint and z-endpoint. Then $(p_n \circ \cdots \circ p_1)(a) = (q_m \circ \cdots \circ q_1)(b)$ for paths $\{p_i\}, \{q_j\}$ and initial points a, b, with $p_n = x$ and $q_m = z$.

Observe $(q_1^{-1} \circ \cdots \circ q_m^{-1} \circ p_n \circ \cdots \circ p_1)(a) = b$ and the concatenation of $\{p_i\}$ followed by the opposite of $\{q_j\}$, is a legitimate path; for $q_m^{-1} p_n$ is $z^{-1}x$ which is not an inverse pair. Hence, $a \sim b$. Since $f(\mathcal{A})$ picks out precisely one representative from each class, $a = b$. But now a is a fixed point of the rotation $q_1^{-1} \circ \cdots \circ q_m^{-1} \circ p_n \circ \cdots \circ p_1$ which contradicts $a \notin D$. This same argument works for the rest of the endpoints.

Finally we prove the union of all these sets is S^2. Since D is already given, it's enough to show $S(x) \cup \cdots \cup S(z^{-1}) \cup f(\mathcal{A})$ equals D^c. Let $a \notin D$; the set $[a]$ has a representative $f([a])$. Either $f([a]) = a$, in which case a is an initial point, or the path which connects $f([a])$ with a ends in one of x, x^{-1}, z, z^{-1} and is, consequently, some endpoint. Either way, we are done. □

The crux of the paradox comes from the observation the set of all points in $S(x^{-1})$ rotated by x, precisely equals the union $f(\mathcal{A}) \cup S(x^{-1}) \cup S(z) \cup S(z^{-1})$. Similarly, rotating $S(z^{-1})$ by z equals $f(\mathcal{A}) \cup S(x) \cup S(x^{-1}) \cup S(z^{-1})$. This procedure is magical; imagine S^2 as a spherical bouquet of flowers, each point on the sphere representing a flower, and the radial line connecting it the origin, its stem. There are six distinct varieties of flowers, corresponding to the sets described in the decomposition. We separate this spherical bouquet into six smaller ones, each comprising of flowers of the same likeness, being sure to keep its original shape in-tact. When we twirl one of these bouquets, say $S(x^{-1})$, by angle θ along the $+x$ axis, something extraordinary happens. This rotation morphs $S(x^{-1})$ into a combined bouquet of four varieties: $f(\mathcal{A})$, $S(x^{-1})$, $S(z)$, and $S(z^{-1})$. The following lemma makes this clear.

Lemma 5.4. *Rotations of $S(x^{-1})$ by x and $S(z^{-1})$ by z, satisfy:*

$$\begin{aligned} xS(x^{-1}) := \{x(a) : a \in S(x^{-1})\} &= f(\mathcal{A}) \cup S(x^{-1}) \cup S(z) \cup S(z^{-1}), \text{ and} \\ zS(z^{-1}) := \{z(a) : a \in S(z^{-1})\} &= f(\mathcal{A}) \cup S(x) \cup S(x^{-1}) \cup S(z^{-1}). \end{aligned}$$

Proof. We prove one of the equalities; the proof of the other is similar.

(⊂) Take any x^{-1}-endpoint so that it can be expressed as $(p_n \circ \cdots \circ p_1)(a)$ for an initial point a and path $\{p_i\}$ with $p_n = x^{-1}$. Applying rotation x to this point allows it to be expressed either as a, when $n = 1$, or $(p_{n-1} \circ \cdots \circ p_1)(a)$ otherwise. Note in the latter case, $p_{n-1} \neq x$ as the original path would, otherwise, contain xx^{-1}. Therefore, the resulting point belongs to $f(\mathcal{A}) \cup S(x^{-1}) \cup S(z) \cup S(z^{-1})$.

(⊃) Fix $a \in f(\mathcal{A}) \cup S(x^{-1}) \cup S(z) \cup S(z^{-1})$. No matter the case, we can express a as $x(x^{-1}(a))$, where $x^{-1}(a) \in S(x^{-1})$ since a is not an x-endpoint. □

We'd like to conclude D^c is equidecomposable with two copies of D^c by the above trick taking $f(\mathcal{A})$, $xS(x^{-1})$, and $S(x)$ to make one D^c sphere, and $zS(z^{-1})$ and $S(z)$ to make another. However we face a setback: $xS(x^{-1})$ double counts $f(\mathcal{A})$ as they overlap. This can be remedied.

Lemma 5.5. *D^c, that is S^2 without D, is equidecomposable with two identical copies of D^c.*

Proof. By equidecomposability with two *copies* of D^c, we mean D^c union a far enough translation of D^c, so that there is no overlap. To handle the complication mentioned earlier, define $G := x^{-1}f(\mathcal{A}) \cup x^{-2}f(\mathcal{A}) \cup \dots$. We partition D^c into the following five sets:

(1) $f(\mathcal{A}) \cup G$
(2) $S(x^{-1}) - G$
(3) $S(x)$
(4) $S(z)$
(5) $S(z^{-1})$.

Indeed the first three sets will be used to form one copy of D^c and the last two will form the other. To see the second, take uniform translations of $S(z)$ and $zS(z^{-1})$. By Lemma 5.4, the union of these disjoint sets forms a translation of D^c. To see the first, we show the family $\{f(\mathcal{A}) \cup G, x\left(S(x^{-1}) - G\right), S(x)\}$ partitions D^c.

It's easy to verify, from Lemma 5.4, the union of the elements in this family equals D^c. So it remains to show the family is pairwise disjoint. Let $a \in x\left(S(x^{-1}) - G\right) \cap S(x)$. In particular $a \in xS(x^{-1})$ and $a \in S(x)$ which contradicts their disjointness. Now let $a \in x\left(S(x^{-1}) - G\right) \cap (f(\mathcal{A}) \cup G)$. Should $a \in f(\mathcal{A})$ then we'd have the contradiction $a \in xS(x^{-1}) \cap f(\mathcal{A})$, so $a \in G$. However, as $xG = f(\mathcal{A}) \cup G$ it follows $a \notin xG$ which gives the contradiction $a \notin G$. □

To handle the decomposition of D, we select a nice enough rotation ρ.

Lemma 5.6. *There is a rotation ρ on S^2 such that the rotation ρ^n, for all $n \in \mathbb{N}$, sends points in D to D^c.*

Proof. We proceed by cardinality arguments. Since D^c is nonempty we fix an axis of rotation by picking a point in D^c. The only two nontrivial fixed points for rotations about this axis, are elements of D^c. Let J be the set of all rotations about this axis which don't satisfy the desired property. I.e. J consists of rotations ψ about the axis for which ψ^n, for some $n \in \mathbb{N}$, maps at least one point in D to D. For any pair of elements in D, there is at most one rotation about the axis which maps the first point to the other. Take integer divisors of the corresponding angle for such rotations, modulo 2π. This set of angles corresponds exactly to rotations in J. Since D is countable (Lemma 5.1), it follows J is countable, and so we can pick an angle not in J. Name the corresponding rotation ρ. □

Lemma 5.7. *S^2 is equidecomposable with D^c.*

Proof. Using the ρ from above, define $E := D \cup \rho D \cup \rho^2 D \cup \ldots$. The idea is to first break S^2 into two pieces: E and E^c; then rotate E by ρ; and finally reassemble ρE with E^c. Indeed the resulting set is D^c. Our proof formalizes this intuition.

We partition S^2 into the sets E and E^c. The details for why they form a partition are, hopefully, straightforward. It's clear they're disjoint and nonempty.

We rotate E by ρ to form ρE which equals, by design in Lemma 5.6, the set $E - D$. Therefore the family $\{\rho E, E^c\}$ partitions D^c, as desired. $\square$

We can visualize the process of extending the paradox for the ball D^3 without its center. Imagine the set as a spherical cake. From earlier, we know we can disassemble the surface into finitely many parts, which can then be rearranged, using only translations and rotations, to form two identical copies. Now consider extending these surface dissections into the cake's volume, by means of a sharp knife, carving along radial lines to the center. This yields finitely many volumetric pieces of the cake. We then apply the same reassembly process, using the same rigid transformations as before, to form two identical copies of centerless balls.

Even provided a blueprint, which alone is impossible to design since it uses the axiom of choice, the knife-work involved is beyond human ability. And even with divine skill, the volumetric pieces bear no resemblance to cake. They'd be extremely prickly, porous, and crumbly, like fragile radial needles emitting from the core — highly unstable, teetering on the brink of collapse under its own weight.

Lemma 5.8. *D^3 without its center is equidecomposable with two identical copies of D^3 without their centers.*

Proof. From Lemmas 5.5 and 5.7, the sphere S^2 is equidecomposable with two identical copies of S^2. Denote the finite partition in the disassembly as $\{A_1, A_2, \ldots, A_n\}$. Then there are an equal number of rigid transformations $\{T_1, T_2, \ldots, T_n\}$ such that $\{T_1 A_1, T_2 A_2, \ldots, T_n A_n\}$ partition the two copies of S^2. We can extend A_i's, denoted A_i', radially inward. More formally, define A_i' as the set of all points in R^3 except the origin, that lives in the finite line segment which connects the origin to a point in A_i. Then $\{A_1', A_2', \ldots, A_n'\}$ partitions D^3 without the origin. Indeed $\{T_1 A_1', T_2 A_2', \ldots, T_n A_n'\}$ partitions two copies of centerless D^3s. $\square$

To accommodate the center we can use a similar trick, as in Lemma 5.7, to vacate just one of the points in the ball. For example, the ball without the center equidecomposes to the ball without another point, say $(0, 0, 1)$. Consider the orbit of this point by just rotations x. The strategy is to decompose the ball into this partial orbit and its complement. That way when we rotate the orbit by x, we vacate $(0, 0, 1)$.

Lemma 5.9. *D^3 equidecomposes to D^3 without its center.*

Proof. The ball without its center equidecomposes to the ball without $(0,0,1)$; after all, we can juggle between two points. Now it remains to show D^3 equidecomposes to D^3 without $(0,0,1)$. Put $F := \{(0,0,1), x(0,0,1), x^2(0,0,1), \dots\}$. Indeed $\{F, F^c\}$ partitions D^3 so consider $\{xF, F^c\}$. Let $a \in xF \cap F^c$. But as $a \in xF$, we have $a = x^n(0,0,1)$ for some $n \in \mathbb{N}$, which contradicts $a \notin F$. Hence, xF and F^c are disjoint.

Now let a be any point in D^3 expect $(0,0,1)$. Either a equals $x^n(0,0,1)$, for some n, in which case $a \in xF$, or a doesn't in which case $a \in F^c$. Conversely, let $a \in xF \cup F^c$. If $a \in F^c$ then a cannot equal $(0,0,1)$. On the other hand when $a \in xF$, observe that $(0,0,1)$ isn't a fixed point of x^n for any $n \in \mathbb{N}$, as θ is an irrational multiple of π. It follows $a \neq (0,0,1)$. Thus, $xF \cup F^c = D^3 - \{(0,0,1)\}$.

We've shown $\{xF, F^c\}$ partitions D^3 without its center, as desired. □

This proof of equidecomposition of D^3 to D^3 without $(0,0,1)$ is akin to Hilbert's Hotel. Imagine the points $(0,0,1)$, $x(0,0,1)$, $x^2(0,0,1)$, …, etc. as booked rooms in an infinite hotel. Since each of these points are distinct, we can enumerate the rooms by numbers 0, 1, …, etc. using the unique power associated to x. The set F, in the above proof, represents the hotel; it's the collection of all occupied rooms. The action of rotating F by x, forming xF, represents relocating each guest to the next numbered room. This process vacates room 0 and, thus, $xF = F - \{(0,0,1)\}$. In just one piece, by a single rotation or *shift*, we can equidecompose F into F without $(0,0,1)$. In extension, we can equidecompose D^3 into D^3 without $(0,0,1)$.

This operation is not limited to singleton points. We used a similar trick in the proof of Lemma 5.1. The occupants of the hotel were themselves sets, namely D, ρD, $\rho^2 D$, …, etc. The hotel, symbolized by E, is the union of all these sets. Rotating E by ρ shifts the occupants, in the same manner, so that room 0 is empty. The difference between xF and ρE is the scale of these operations: the first frees a singletons' worth of points whereas the second, a whole D's worth.

Theorem 5.1 (Banach-Tarski's Paradox). *D^3 is equidecomposable with two identical copies of D^3.*

Proof. Straightforward by chaining equidecompositions in Lemmas 5.8 and 5.9. □

Theorem 1.3 is an immediate corollary of Theorem 5.1.

6. Conclusion

In this paper, we've explored three paradoxes of infinity, involving the axiom of choice, by analyzing Hardin and Taylor's paper [2], Nelson's textbook [5], and Daniel's formalized proof [1]. Specifically, we examined Theorems 3.2 which solve the infinite hat riddle, 1.2 which forms an immeasurable subset in $\mathbb{R}^d$, and ultimately 5.1 which equidecomposes a ball into two identical copies.

It's impossible to profit from the Banach-Tarski's paradox, by (foolishly) attempting to double a ball of gold; after all, the proofs of these theorems invoke the axiom of choice. Showing that the axiom is necessary is an advanced topic in model theory beyond the scope of this paper. We hope our analogies — likening these highly artificial and bizarre sets to combs, bouquets, cakes, and hotels — provide a colorful insight into the axiom of choice.

There is a clear connection between Theorems 1.2 and 5.1; rational translations of representatives in $[0, 1]$ correspond to orbits of initial points in S^2. The same proof of equidecomposability of a ball in $\mathbb{R}^3$ doesn't work for $[0, 1]$ in $\mathbb{R}^1$. Higher dimensions enjoy extra degrees of freedom: $\mathbb{R}^1$ has no genuine rotation, $\mathbb{R}^2$ has just one, but $\mathbb{R}^3$ has two rotations. Having a second rotation is key for Lemma 5.4.

However, the connection between Theorems 5.1 and 3.2 seems more elusive. The equivalence relation for the Banach-Tarski paradox identifies finite sequences, whereas for the infinite hat riddle, identifies infinite sequences. Perhaps there is some deeper connection, which allows one to draw an equivalence between the two. That way a solution of one can be transformed into a strategy for the other.

Acknowledgments

We would like to thank Will Hoffer for his exposition of Vitali's set and suggestion of taking cartesian products of Vitali sets, to obtain an immeasurable subset of $\mathbb{R}^d$. We are grateful for the opportunity offered by Professor Michel Lapidus to explore these remarkable results of the axiom of choice. Special thanks to Daniel's insightful paper on a formalized proof of of Banach-Tarski's paradox [1]. This significantly aided the author's understanding.

References

1. Daniel de Rauglaudre, (2017). *Formal Proof of Banach-Tarski Paradox*. Journal of Formalized Reasoning, 10 (1), pp. 37-49. Available at `https://doi.org/10.6092/issn.1972-5787/6927`.
2. Hardin, C.S., Taylor, A.D (2008). *An introduction to infinite hat problems*. The Mathematical Intelligencer 30, 20–25. Available at: `https://doi.org/10.1007/BF03038092`.
3. Jech, T. (2006). *Set Theory: The Third Millennium Edition*, Chapter 3: Cardinals. Springer.
4. Karagila, A. (2020), *Zornian Functional Analysis or: How I Learned to Stop Worrying and Love the Axiom of Choice*. ArXiv preprint, Chapter 1. Available at: `https://arxiv.org/abs/2010.15632`.
5. Nelson, G. S. (2015). *A User-Friendly Introduction to Lebesgue Measure and Integration*, Chapter 1: Lebesgue Measure. *Student Mathematical Library*, Vol. 59, pp. 48. American Mathematical Society.
6. Palais, R.S. (2009). *The Visualization of Mathematics: Towards a Mathematical Exploratorium*, 116:10, pp. 892. The American Mathematical Monthly. Available at: `https://www.maa.org/sites/default/files/pdf/upload_library/22/Ford/Palais.pdf`.
7. Wagon, S. (1985). *The Banach-Tarski Paradox*, Chapter 8: Paradoxical Decompositions in $\mathbb{R}^n$. Cambridge University Press.