The Banach–Tarski Paradox
An Exposition

Omar Elshinawy

Fall 2023

Abstract

Contrary to common belief, the result by Stefan Banach and Alfred Tarski is not at all paradoxical, and rather challenges our intuition about the notion of infinity. In this article, we introduce a paradoxical decomposition of F_2, the free group on two generators, then extend this notion by the embedding of F_2 in Euclidean space $\mathbb{R}^3$. We conclude with a few notes on non-measurability, and reflect on criticism of the Axiom of Choice.

OVERVIEW

Contents

1 On the Free Group

1.1 A Prelude

To familiarize ourselves with the concept of a free group, it may be necessary to first discuss what a group is.

Definition 1.1 (Group Axioms)**.** Let G be a set. Endowed with a binary operation $\cdot$, if

1. $(\forall r, s, t \in G): \ (r \cdot s) \cdot t = r \cdot (s \cdot t),$ Associativity
2. $(\exists\ e \in G)(\forall s \in G) : e \cdot s = s \cdot e = s,$ Identity Element
3. $(\forall s \in G)(\exists\ s^{-1} \in G) : s \cdot s^{-1} = e,$ Inverse Element

then we say that G is a group.

Remark. We adopt the convention of writing $s \cdot t$ as st.

Next, to define a free group, we start with an arbitrary set S, and define S^{-1} to be the set of inverses of $s \in S$ as in **the third axiom.** This allows us to define an alphabet $T := S \cup S^{-1}$.

Definition 1.2 (Free Group)**.** Let S be an arbitrary set. Then, the free group generated by S is

$$F_{\,|S|} = \langle S \rangle := \{t_1 t_2 \dots t_n : \ t_i t_{i+1} \neq e, \ t_i \in T, \ n \in \mathbb{N}_0\}$$

and is of rank $|S|$. We say that $w \in \langle S \rangle$ is a reduced word of finite length, and define $t_0 := e$ to be the word of zero length.

In order for $\langle S \rangle$ to become a free group, any two words must be different unless their equality follows from **the second and third group axioms**. When considering rotations in $\mathbb{R}^3$, we want to make a choice of rotations σ, τ that does not allow for further reduction of words.

In the following section, we discuss the free group of two elements, explain what a reduced word is (not), and emphasize why we need a unique representation of the identity rotation e.

1.2 The Free Group F_2

Let $S = \{\sigma, \tau\}$. Then, $S^{-1} = \{\sigma^{-1}, \tau^{-1}\}$, and $T = S \cup S^{-1} = \{\sigma, \sigma^{-1}, \tau, \tau^{-1}\}$ is our alphabet. We have that

- $\tau\tau\sigma^{-1}\tau^{-1} \in F_2 = \langle \sigma, \tau \rangle$ *No two letters are inverses of each other*
- $\tau^{-1}\sigma\sigma^{-1}\tau \notin F_2$ *This word should be reduced to e.*

We impose this condition to uniquely represent each word. A nice exercise is to convince yourself that there is an infinite class of non-reduced, non-trivial represntations of the identity element e.

This is analogous to choosing two rotations σ, τ (and σ^{-1}, τ^{-1}) that act freely on the group.

1.3 An Interesting Decomposition of F_2

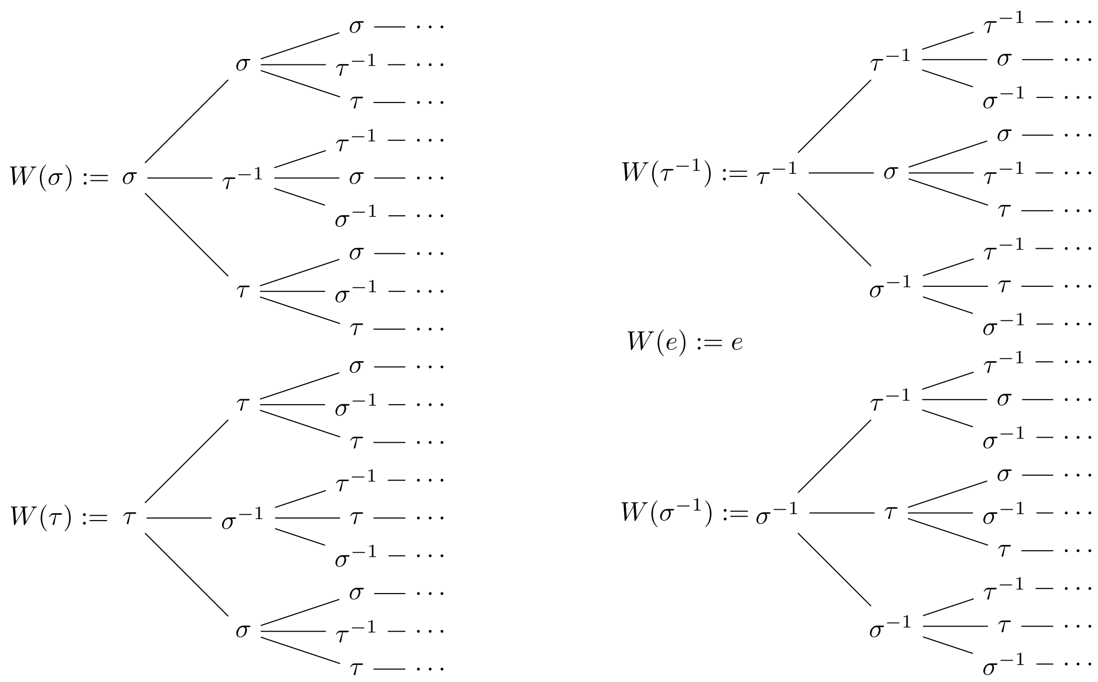
Let us study the anatomy of F_2.

Definition 1.3 (Classification of Reduced Words.)**.** Let T be an alphabet. For $t \in T$, we define

$$W(t) := \{w \in \langle \sigma, \tau \rangle : t_1 = \sigma\}$$

to be all reduced words that start with $t \in T$.

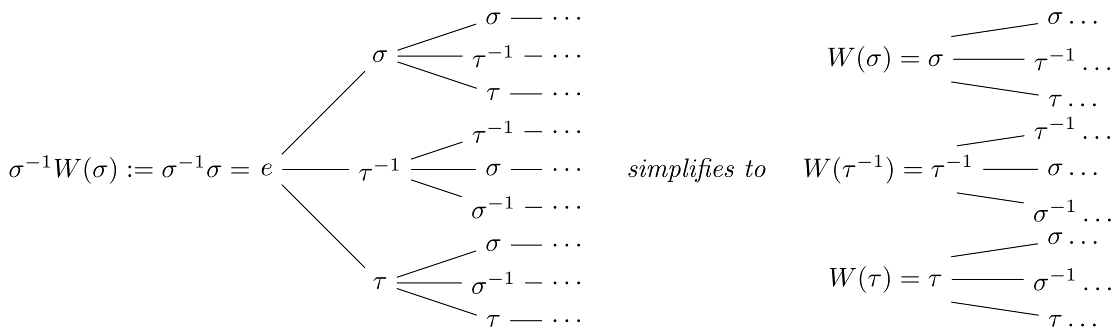
A free group is truly free, without any restriction beyond the group axioms. To illustrate this, we now endow the reader with the following nice visual images.



Observe that a word that starts with $t \in T$ cannot continue with $t^{-1} \in T$, leaving three options at each node. We can write that

$$F_2 := W(\sigma) \ \sqcup W(\tau) \sqcup W(\tau^{-1}) \sqcup W(\sigma_{-1}) \sqcup W(e) \tag{1}$$

with $\sqcup$ as the disjoint union. We remind the reader that all words are finite, and will terminate at some given node. Next, let us consider $W(\sigma)$. Apply σ^{-1} from the left, then we get that



Indeed, we have

$$\sigma^{-1}W(\sigma) = W(\sigma) \sqcup W(\tau^{-1}) \sqcup W(\tau) \sqcup W(e) = F_2 - W(\sigma). \tag{2}$$

We note that the choice of σ is arbitrary, and the statement is generally true for any $t \in T$.

It is rather interesting (and perhaps paradoxical) that we "almost" obtain the whole free group by applying one group operation on a subset. To exploit this even further, observe that

$$F_2 = \sigma^{-1}W(\sigma) \sqcup W(\sigma) = \tau^{-1}W(\tau) \sqcup W(\tau). \tag{3}$$

The goal of this demonstration is to refine the intuition behind the paradoxical decomposition. Note, however, that our copies generate the identity e twice, and we only have one copy of the identity. This hints towards a perhaps more complicated decomposition of F_2, which we rigorously treat with the following theorem.

Theorem 1.4 (Paradoxical Decomposition of F_2)**.** *For the free group $F_2 = \langle \sigma, \tau \rangle$ on two generators, we can find a partition*

$$\langle \sigma, \tau \rangle = G_1 \sqcup G_2 \sqcup G_3 \sqcup G_4$$

such that one can write

$$\langle \sigma, \tau \rangle = G_1 \sqcup \sigma G_2 = G_3 \sqcup \tau G_4.$$

Proof. We make the following clever choices,

$$G_1 = W(\sigma); \quad G_2 = W(\sigma^{-1}); \quad G_3 = W(\tau) \sqcup \{e, \tau^{-1}, \tau^{-2}, \dots\}; \quad G_4 = W(\tau^{-1}) - \{\tau^{-1}, \tau^{-2}, \dots\}.$$

– $G_4 \subset W(\tau^{-1})$, meaning $G_1,\ G_2,\ G_4$ are disjoint. It suffices to compare these sets to G_3.

– G_3 has no words starting with σ or σ^{-1}, therefore $G_1,\ G_2$ are disjoint compared with G_3.

– $W(\tau),\ W(\tau^{-1})$ are disjoint, and G_3, G_4 do not intersect on $\{\tau^{-1}, \tau^{-2}, \dots\}$, nor on e.

The first statement immediately follows. For the following statement, we observe that

$$G_1 \sqcup \sigma G_2 = W(\sigma) \sqcup \sigma W(\sigma^{-1}) =^{\mathbf{(3)}} F_2.$$

Next, we check that $G_3 \sqcup \tau G_4$ indeed generates F_2.

$$\begin{aligned} G_3 \sqcup \tau G_4 = & \Big(W(\tau) \sqcup \{e, \tau^{-1}, \tau^{-2}, \dots\}\Big) \sqcup \tau\Big(W(\tau^{-1}) - \{\tau^{-1}, \tau^{-2}, \dots\}\Big) \\ =^{\mathbf{(3)}} & \Big(W(\tau) \sqcup \{e, \tau^{-1}, \tau^{-2}, \dots\}\Big) \sqcup \Big(\Big(F_2 - W(\tau)\Big) - \{e, \tau^{-1}, \tau^{-2}, \dots\}\Big) \\ = & \Big(W(\tau) \sqcup \{e, \tau^{-1}, \tau^{-2}, \dots\}\Big) \sqcup \Big(F_2 - \Big(W(\tau) \sqcup \{e, \tau^{-1}, \tau^{-2}, \dots\}\Big)\Big) = F_2. \end{aligned}$$

As promised in **(3)**, we obtain exactly one copy of e. □

We have used **one copy** of $\langle \sigma, \tau \rangle$ to generate two copies of $\langle \sigma, \tau \rangle$, by the means of rotations σ, τ. To extend this notion to S^2, we would like to argue the existence of these two rotations that behave similarly to the generators of F_2. This is precisely what we dedicate the next section to.

2 From Groups to Spheres

We show that there exists σ, τ such that $\langle \sigma, \tau \rangle$ is free. It turns out there is an infinite class of such pairs of rotations, but we do not delve into that at all. We will however give explicit rotation matrices for our purposes. We then proceed immediately with paradoxically decomposing S^2.

2.1 A Free Group of Rotations

Let us start by choosing $\theta = \arccos \frac{1}{3}$. Next, we set σ be a rotation by θ about the $x-$axis, and τ a rotation by θ about the $z-$axis. We give explicit matrices,

$$\sigma = \frac{1}{3}\begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & -2\sqrt{2} \\ 0 & 2\sqrt{2} & 1 \end{bmatrix} \quad \& \quad \sigma^{-1} = \frac{1}{3}\begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2\sqrt{2} \\ 0 & -2\sqrt{2} & 1 \end{bmatrix}$$

$$\tau = \frac{1}{3}\begin{bmatrix} 1 & -2\sqrt{2} & 0 \\ 2\sqrt{2} & 1 & 0\sqrt{2} \\ 0 & 0 & 3 \end{bmatrix} \quad \& \quad \tau^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 2\sqrt{2} & 0 \\ -2\sqrt{2} & 1 & 0\sqrt{2} \\ 0 & 0 & 3 \end{bmatrix}$$

Some motivation behind this choice is the fact that $\arccos \frac{1}{3}$ is an irrational rotation. A rational rotation is periodic, and that is of course not suitable for our purposes. Now we show why $\arccos \frac{1}{3}$ is indeed a more suitable choice.

Lemma 2.1. *For $\rho \in \langle \sigma, \tau \rangle$ we find that*

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

for $a, b, c \in \mathbb{Z}$

Proof. Induction on the length n of ρ. Base case is trivial, suppose the lemma holds for all $k < n$. Then $\rho_n = r\rho_{n-1}$ for $r \in \{\sigma, \tau, \sigma^{-1}, \tau^{-1}\}$. Applying the hypothesis for ρ_{n-1} yields

$$(\sigma \circ \rho')(0,1,0) = \frac{1}{3}^n \big(3a\sqrt{2}, b - 4c, (c + 2b)\sqrt{2}\big)$$
$$(\sigma^{-1} \circ \rho')(0,1,0) = \frac{1}{3}^n \big(3a\sqrt{2}, b + 4c, (c - 2b)\sqrt{2}\big)$$
$$(\tau \circ \rho')(0,1,0) = \frac{1}{3}^n \big((a - 2b)\sqrt{2}, b + 4a, 3c\sqrt{2}\big)$$
$$(\tau^{-1} \circ \rho')(0,1,0) = \frac{1}{3}^n \big((a + 2b)\sqrt{2}, b - 4a, 3c\sqrt{2}\big)$$

This concludes the argument. □

Theorem 2.2. *$\langle \sigma, \tau \rangle$ is a free group such that there is no non-trivial identity.*

Proof. Suppose there exists a non-trivial identity rotation ρ. Then $\rho(0,1,0) = (0,1,0)$. By **2.1** we have that

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

so we infer that $a = c = 0$ and $b = 3^n$ for $n \in \mathbb{N}$, meaning $a \equiv^3 b \equiv^3 c$. This yields a contradiction, which for the sake of brevity, is shown in **Proposition 3.1**, Tom Weston's expository. [4]. □

2.2 Decomposing S_2

There is a small subtlety with rotations, that is every axis of rotation fixes two points. If we want to have a true free action, then we ought to eliminate these two points for every rotation. For now, let us denote the set of all such points by P, and proceed our discussion on S^2/P.

Definition 2.3 (Problematic Poles)**.** We define

$$P := \{x \in S^2 \mid r(x) = x \text{ for } r \in \langle\sigma, \tau\rangle\}$$

to be points $x \in S^2$ that are fixed by rotations $r \in \langle\sigma, \tau\rangle$.

Remark. Clearly $|P| := 2|F_2|$ as for each rotation we have two fixed points.

For the sake of our construction, we partition the sphere into equivalence classes.

Lemma 2.4 (Partition of S^2)**.** *The the following relation,*

$$x \sim y \iff \exists \rho \in \langle\sigma, \tau\rangle \ : \rho(x) = y.$$

defines an equivalence relation on $x, y \in S^2$.

Proof. One only checks that $\sim$ is symmetric, transitive and reflexive, a simple exercise. □

Using the lemma above, we define

$$\big(S^2 - P\big)/\sim := \{[x] \mid x \in S^2 - P\}$$

to be the set of equivalence classes of points in $S^2 - P$ under $\sim$, which we use in the next proof.

Theorem 2.5 (Hausdorff Paradox)**.** *There exists a countable subset P of the sphere S^2, and a decompostion*

$$S^2 - P = \Omega_1 \sqcup \Omega_2 \sqcup \Omega_3 \sqcup \Omega_4$$

such that

$$S^2 - P = \Omega_1 \sqcup \sigma\big(\Omega_2\big) = \Omega_3 \sqcup \tau\big(\Omega_4\big)$$

Proof. We summon the axiom of choice to pick a representative from each equivalence class of points $[x] \in \big(S^2 - P\big)/\sim$. Define that set to be

$$X = \{x \mid x \text{ is a representative of } [x]\}.$$

Since X contains a point from each equivalence class, we can re-apply all rotations to $x \in X$ to generate all the elements back. If we do this for all x, we generate $S^2 - P$ once again **if we note that $\langle\sigma, \tau\rangle$ acts freely on $S^2 - P$.**

$$\langle\sigma, \tau\rangle(X) := \bigsqcup_{x \in X} \langle\sigma, \tau\rangle(x) = S^2 - P$$

The claim follows from the paradoxical decomposition **1.4** of $\langle\sigma, \tau\rangle$ with $\Omega_i = G_i$. □

Let us now deal with the problematic poles P.

Lemma 2.6. *There exists a rotation φ such that for all $n \geq 1$ we have that $\varphi^n(P) \cap P = \phi$. Hence, $\varphi^n(D) \cap \varphi^m(D) = \phi$ for distinct m, n.*

Proof. We use the remark from 2.3. Choose an axis of rotation l that does not intersect P. Thus, no rotation of l will fix a point in P. Next, choose a random pair $(p, q) \in P \times P$ of which there are countably many, and consider rotations around l which take p to q. It turns out we can find countably many, since p, q must be on the same plane orthogonal to l. If there are only countably many rotations φ around l such that $\varphi^n(D) \cap D \neq \phi$, then we can choose φ around l such that $\varphi^n(p) \neq q$ for all $(p, q) \in D \times D$. The second statement immediately follows. □

Proposition 2.7. *Let D be a countable subset of S^2. Then*

$$S^2 = \Sigma_1 \sqcup \Sigma_2$$

such that

$$S^2 - P = \Sigma_1 \sqcup \varphi\big(\Sigma_2\big)$$

for some $\varphi \in \langle \sigma, \tau \rangle$.

Proof. By 2.6 we can choose φ such that

$$\Sigma_2 := P \sqcup \varphi(P) \sqcup \varphi^2(P) \sqcup \dots \quad ; \quad \Sigma_1 := S^2 - \Sigma_2$$

thus concluding the proof. □

Remark. Applying φ to Σ_2 precisely yields $\Sigma_2 - P$. Infinity is cool. □

Using the machinery developed in the **Hausdorff paradox** 2.5 and 3.3, we conclude that

Corollary 2.8. *There exists a partition*

$$S^2 = \Gamma_1 \sqcup \dots \sqcup \Gamma_8$$

and rotations $\rho_1, \dots, \rho_8$ such that

$$S^2 = \bigsqcup_{i=1}^{4} \rho_i(\Gamma_i) = \bigsqcup_{i=5}^{8} \rho_i(\Gamma_i)$$

Proof. Simple algebra. Note that we can write

1. $S^2 = (S^2 - \Sigma_2) \sqcup \Sigma_2$
2. $P \subset \Sigma_2 \implies S^2 - \Sigma_2 = (S^2 - P) \cap (S^2 - \Sigma_2)$
3. $\varphi(\Sigma_2) = \Sigma_2 - P \implies \Sigma_2 = \varphi^{-1}(\Sigma_2 - P) = \varphi^{-1}\big((S^2 - P) \cap \Sigma_2\big)$ *De Morgan's Law*

Next, for X the set of representatives $[x]$, define

$$X_\sigma = W(\sigma)(X) \qquad X_{\sigma^{-1}} = W(\sigma^{-1})(X) \qquad X_\tau = W(\tau)(X) \qquad X_{\tau^{-1}} = W(\tau^{-1})(\tau)$$

Then,

4.1 $\tau^{-1}(X_\tau) \sqcup X_{\tau^{-1}} = S^2 - P$

4.2 $\sigma^{-1}(X_\sigma) \sqcup X_{\sigma^{-1}} = S^2 - P$

by the paradoxical decomposition. We additionally use the following identities,

5. $\big(\tau^{-1}(X_\tau) \sqcup X_{\tau^{-1}}\big) \cap (S^2 - \Sigma_2) = \big(\tau^{-1}(X_\tau) \cap (S^2 - \Sigma_2)\big) \sqcup \big(X_{\tau^{-1}} \cap (S^2 - \Sigma_2)\big)$
6. $\tau^{-1}(X_\tau) \cap (S^2 - \Sigma_2) = \tau^{-1}\big(X_\tau \cap \tau(S^2 - \Sigma_2)\big)$
7. $\varphi^{-1}\Big(\tau^{-1}(X_\tau) \sqcup X_{\tau^{-1}}) \cap \Sigma_2\Big) = \varphi^{-1}\Big(\big(\tau^{-1}(X_\tau) \cap \Sigma_2\big) \sqcup \big(X_{\tau^{-1}} \cap \Sigma_2\big)\Big) =$
8. $= \varphi^{-1} \circ \tau^{-1}\big(X_\tau \cap \tau(\Sigma_2)\big) \sqcup \varphi^{-1}\big(X_{\tau^{-1}} \cap \Sigma_2\big)$

and make the following choices.

$\Gamma_1 = (X_\tau) \cap (S^2 - \Sigma_2) \quad \Gamma_2 = X_{\tau^{-1}} \cap (S^2 - \Sigma_2) \quad \Gamma_3 = X_\tau \cap \tau(\Sigma_2) \quad \Gamma_4 = X_{\tau^{-1}} \cap \Sigma_2$

This yields

$$\begin{aligned}
S^2 &=^{\mathbf{1}} (S^2 - \Sigma_2) \sqcup \Sigma_2 \\
&=^{\mathbf{2}} \Big((S^2 - P) \cap (S^2 - \Sigma_2)\Big) \sqcup \Sigma_2 \\
&=^{\mathbf{3}} \Big((S^2 - P) \cap (S^2 - \Sigma_2)\Big) \sqcup \varphi^{-1}\Big((S^2 - P) \cap \Sigma_2\Big). \\
&=^{\mathbf{4.1}} \Big(\big(\tau^{-1}(X_\tau) \sqcup X_{\tau^{-1}}\big) \cap (S^2 - \Sigma_2)\Big) \sqcup \varphi^{-1}\Big((\tau^{-1}(X_\tau) \sqcup X_{\tau^{-1}}) \cap \Sigma_2\Big) \\
&=^{\mathbf{5}} \bigg(\Big(\tau^{-1}(X_\tau) \cap (S^2 - \Sigma_2)\Big) \sqcup \Big(X_{\tau^{-1}} \cap (S^2 - \Sigma_2)\big)\bigg) \sqcup \varphi^{-1}\Big((\tau^{-1}(X_\tau) \sqcup X_{\tau^{-1}}) \cap \Sigma_2\Big) \\
&=^{\mathbf{6,7}} \bigg(\tau^{-1}\Big(X_\tau \cap \tau(S^2 - \Sigma_2)\Big) \sqcup \Gamma_2\bigg) \sqcup \varphi^{-1}\Big(\big(\tau^{-1}(X_\tau) \cap \Sigma_2\big) \sqcup \big(X_{\tau^{-1}} \cap \Sigma_2\big)\Big) \\
&=^{\mathbf{8}} \Big(\tau^{-1}(\Gamma_1) \sqcup \Gamma_2\Big) \sqcup \Big(\varphi^{-1} \circ \tau^{-1}\big(X_\tau \cap \tau(\Sigma_2)\big) \sqcup \varphi^{-1}\big(X_{\tau^{-1}} \cap \Sigma_2\big)\Big) \\
&= \tau^{-1}(\Gamma_1) \sqcup \Gamma_2 \sqcup \varphi^{-1} \circ \tau^{-1}(\Gamma_3) \sqcup \varphi^{-1}(\Gamma_4). \qquad (1)
\end{aligned}$$

Note that we could have proceeded similarly with 4.2 (σ) instead of 4.1 (τ),

$\Gamma_5 = (X_\sigma) \cap (S^2 - \Sigma_2) \quad \Gamma_6 = X_{\sigma^{-1}} \cap (S^2 - \Sigma_2) \quad \Gamma_7 = X_\tau \cap \sigma(\Sigma_2) \quad \Gamma_8 = X_{\sigma^{-1}} \cap \Sigma_2$

with

$$\sigma^{-1}(\Gamma_5) \sqcup \Gamma_6 \sqcup \varphi^{-1} \circ \sigma^{-1}(\Gamma_7) \sqcup \varphi^{-1}(\Gamma_8). \qquad (2)$$

The proof concludes with (1), (2), and the simple observation that all Γ_i are disjoint. □

By this construction, we (almost) have the Banach-Tarski paradox. We address this statement in more detail in the following section.

3 The Unit Ball B^3

3.1 S^2 to B^3

Indeed, we have constructed a paradoxical decomposition of S^2. Note, however, that our goal is to transfer it to the whole unit ball B^3. This itself proves to be a simple task.

Definition 3.1 (The Unit Ball)**.** Define

$$B^3 := \{x \in \mathbb{R}^3 \mid ||x|| \leq 1\}$$

to be the unit ball in $\mathbb{R}^3$, with $||\cdot||$ the Euclidean norm.

The punctured ball $B^3 - \{0\}$ is thought of as the product of the sphere S^2 and the interval $(0, 1]$.

Lemma 3.2. *There exists a paradoxical decomposition of* $B^3 - \{0\}$.

Proof. This builds upon the work demonstrated in 2.8. Similar to Γ_i, define

$$\Gamma'_i = \{x \in B^3 - \{0\} \mid \exists p \in \Gamma_i \ , r \in (0, 1] \ x = rp\}$$

and therefore

$$B^3 - \{0\} = \bigsqcup_{i=1}^{4} \rho_i(\Gamma'_i) = \bigsqcup_{i=5}^{8} \rho_i(\Gamma'_i).$$

This concludes the argument. □

Remark. A good visual image is to think of the extension of points on the boundary to the origin. This is the "fiber" of the ball B^3, and if the surface is decomposed paradoxically then we can clearly extend that to B^3 (without the origin).

3.2 Why We Need 8 Partitions

A is simply amendable detail is that we are still missing the origin point.

Proposition 3.3. *There exists a decomposition of* B^3 *such that*

$$B^3 = B_1 \sqcup B_2$$

such that

$$B^3 - \{0\} = B_1 \sqcup \theta(B_2)$$

for some $\theta \in \langle \sigma, \tau \rangle$.

Proof. We shall use a simple trick as in 3.3 to shift the origin and its orbits. Convince yourself that we can find a rotation θ such that $\theta^n(0) \in B^3$ for all $n \geq 0$. Then, choose

$$B_2 = 0 \sqcup \theta(0) \sqcup \theta^2(0) \sqcup \ldots; \quad B_1 = B^3 - B_2$$

then the claim holds. □

Up next, the main result of this paper.

Theorem 3.4 (The Banach-Tarski Paradox)**.** *There exists a decomposition of B^3 such that*

$$B^3 = \bigsqcup_{i=1}^{16} \rho_i(\Lambda_i)$$

such that

$$B^3 = \bigsqcup_{i=1}^{8} \rho_i(\Lambda_i) = \bigsqcup_{i=9}^{16} \rho_i(\Lambda_i)$$

for $\rho_i \in \langle \sigma, \tau \rangle$.

Proof. First, $B_2 = \theta^{-1}\big((B^3 - \{0\}) \cap B_2\big)$. We then write

$$B^3 = (B^3 - B_2) \sqcup B_2 = \Big((B^3 - \{0\}) \cap B_1\Big) \sqcup \theta^{-1}\Big((B^3 - \{0\}) \cap B_2\Big).$$

We then use $B^3 - \{0\} = \tau^{-1}(\Gamma_1') \sqcup \Gamma_2' \sqcup \varphi^{-1} \circ \tau^{-1}(\Gamma_3') \sqcup \varphi^{-1}(\Gamma_4')$ to get that

$$\begin{aligned} B^3 =& \Big(\big[\tau^{-1}(\Gamma_1') \sqcup \Gamma_2' \sqcup \varphi^{-1} \circ \tau^{-1}(\Gamma_3') \sqcup \varphi^{-1}(\Gamma_4')\big] \cap B_1\Big) \\ \sqcup\, \theta^{-1}&\Big(\big[\tau^{-1}(\Gamma_1') \sqcup \Gamma_2' \sqcup \varphi^{-1} \circ \tau^{-1}(\Gamma_3') \sqcup \varphi^{-1}(\Gamma_4')\big] \cap B_2\Big) \\ =& \Big(\tau^{-1}\big[\Gamma_1' \cap \tau(B_1)\big] \sqcup \big[\Gamma_2' \cap B_1\big] \sqcup \varphi^{-1} \circ \tau^{-1}\big[\Gamma_3' \cap \varphi \circ \tau(B_1)\big] \sqcup \varphi^{-1}\big[\Gamma_4' \cap \varphi(B_1)\big]\Big) \\ \sqcup\, \theta^{-1}&\Big(\tau^{-1}\big[\Gamma_1' \cap \tau(B_2)\big] \sqcup \big[\Gamma_2' \cap B_2\big] \sqcup \varphi^{-1} \circ \tau^{-1}\big[\Gamma_3' \cap \varphi \circ \tau(B_2)\big] \sqcup \varphi^{-1}\big[\Gamma_4' \cap \varphi(B_2)\big]\Big) \end{aligned}$$

And now finally we can see the decomposition. We can define

$$\begin{array}{llll} \Lambda_1 = \Gamma_1' \cap \tau(B_1) & \Lambda_2 = \Gamma_2' \cap B_1 & \Lambda_3 = \Gamma_3' \cap \varphi \circ \tau(B_1) & \Lambda_4 = \Gamma_4' \cap \varphi(B_1) \\ \Lambda_5 = \Gamma_1' \cap \tau(B_2) & \Lambda_6 = \Gamma_2' \cap B_2 & \Lambda_7 = \Gamma_3' \cap \varphi \circ \tau(B_2) & \Lambda_8 = \Gamma_4' \cap \varphi(B_2) \end{array}$$

to be the first copy of B^3. By the exact same argument for τ we can proceed for σ with

$$\begin{array}{llll} \Lambda_9 = \Gamma_5' \cap \sigma(B_1) & \Lambda_{10} = \Gamma_6' \cap B_1 & \Lambda_{11} = \Gamma_7' \cap \varphi \circ \sigma(B_1) & \Lambda_{12} = \Gamma_8' \cap \varphi(B_1) \\ \Lambda_{13} = \Gamma_5' \cap \sigma(B_2) & \Lambda_{14} = \Gamma_6' \cap B_2 & \Lambda_{15} = \Gamma_7' \cap \varphi \circ \sigma(B_2) & \Lambda_{16} = \Gamma_8' \cap \varphi(B_2) \end{array}$$

and one can verify that all claims of the theorem hold! □

4 Closing Words

4.1 Note on Non-measurability

The Banach-Tarski theorem is a classical example of a non-measurable set. In **2.4**, we partitioned the sphere into equivalence classes of points in the same orbit. Using the Axiom of Choice, we constructed **(The Hausdorff Paradox, 2.5)**

$$X = \{x \mid x \text{ is a representative of } [x]\}.$$

that is a non-measurable set. Recall our earlier discussion:

Since X contains a point from each equivalence class, we can re-apply all rotations to $x \in X$ to generate all the elements back. If we do this for all x, we generate $S^2 - P$ once again.

$$\langle\sigma,\tau\rangle(X) := \bigsqcup_{x\in X} \langle\sigma,\tau\rangle(x) = S^2 - P$$

To see this, we attempt to assign a measure to e.g. $S^2 - P$. We write

$$S^2 - P = \bigsqcup_{x\in X} \langle\sigma,\tau\rangle(x) := \bigsqcup_{r\in\langle\sigma,\tau\rangle} r(X) \implies \mu(S^2 - P) = \mu\Big(\bigsqcup_{r\in\langle\sigma,\tau\rangle} r(X)\Big)$$

A measure μ should be rotation invariant, i.e. rotating a piece should not change its size. Further, the measure of two pieces (disjoint sets) should be their sum. Therefore, we get that

$$\mu(S^2 - P) = \mu\Big(\bigsqcup_{r\in\langle\sigma,\tau\rangle} r(X)\Big) = \mu\Big(\bigsqcup_{r\in\langle\sigma,\tau\rangle} X\Big) = \sum_{r\in\langle\sigma,\tau\rangle} \mu(X).$$

We cannot assign a measure to X. Recall that $\text{card}(\langle\sigma,\tau\rangle) = \aleph_0$. We have two options.

$$\mu(X)\begin{cases} = 0 \implies S^2 = 0 \\ > 0 \implies S^2 = \sum_{r\in\langle\sigma,\tau\rangle} \mu(X) = \infty \end{cases}$$

The second sequence diverges, as we must add $\mu(X)$ for every rotation $r \in \langle\sigma,\tau\rangle$; of which there are infinitely many. We see that both options are not at all viable.

Therefore the Banach-Tarski theorem must divide B^3 into non-measurable sets. Otherwise, any attempt to assign a measure to $\mu(X)$ immediately implies that the volume of two balls is the same as the volume of one.

4.2 Axiom of Choice and Criticism

This result has received a lot of criticism in the mathematical world. For instance, Émile Borel considered the Banach−Tarski a *reductio ad abdsurdum* of the Axiom of Choice by the construction we just saw. The reality is, however, that the statement is true and not paradoxical at all.

Briefly, what Banach and Tarski are trying to tell us is that **if,** and we emphasize the if, you can divide a ball in such an infinite pathological construction, then indeed the theorem holds. The theorem is logically consistent, as the necessary condition shall never be satisfied for our physical space, which is far from having a structure similar to $\mathbb{R}^3$.

This features a significant property of the continuum, that is: dividing a spatial region into disjoint pieces need not preserve volume. We need to rather accustom ourselves to the notion of infinity, and refine our intuition to the natural consequences of it. This concludes the paper. $\boxed{\xi}$

Acknowledgements

This paper was prepard as part of the last undergraduate seminar on Mathematics at Jacobs University; a great initiative that wouldn't have been possible if not for Professor Keivan Mallahi-Karai. I would like to personally thank him for giving me the chance to learn something new, and share with my colleagues my enthusiasm for the subject.

References

[1] Avery Robinson. The banach-tarski paradox. https://math.uchicago.edu/~may/REU2014/REUPapers/Robinson.pdf. UChicago, Expository.

[2] Terence Tao. The banach-tarski paradox. https://www.math.ucla.edu/~tao/preprints/Expository/banach-tarski.pdf.

[3] Teun van Wesel. Non-measurable sets and the banach-tarski paradox. https://teunvanwezel.nl/Wezel-van-Teun-Bacehlor-Thesis.pdf.

[4] Tom Weston. The banach-tarski paradox. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5ab85e37c6e1da73049610cf75c600068af7a839.