PARADOXICAL DECOMPOSITIONS OF GROUPS AND THE SETS THEY ACT UPON

A Thesis

Presented to the

Department of Mathematics

and the

Faculty of the Graduate College

In Partial Fulfillment

of the Requirements for the Degree of

Master's of Arts in Mathematics

University of Nebraska at Omaha

by

Joseph Lee

May 2010

Supervisory Committee:

Dr. Andrzej Roslanowski

Dr. Griff Elder

Dr. Valentin Matache

Dr. Wai-Ning Mei

PARADOXICAL DECOMPOSITIONS OF GROUPS AND THE SETS THEY ACT UPON

Joseph Lee, MA

University of Nebraska, 2010

Advisor: Dr. Andrzej Roslanowski

The Banach-Tarski paradox states that a solid ball in $\mathbb{R}^3$ can be partitioned in such a way that by simply rotating and translating those pieces, you end up with two exact copies of the original ball you started with. We give a detailed decomposition of nine pieces and nine different isometries that can be used to produce two copies of the unit ball. We also show this paradox is specifically due to the fact that the group of isometries in $\mathbb{R}^3$ has a free subgroup with two generators. In the second half of this paper, we introduce the property that prevents this paradox from occuring and show the group of isometries in $\mathbb{R}$ and $\mathbb{R}^2$ possess this property.

Common Symbols

A, B, X, Y, Z	sets
$\mathcal{A}$	an algebra of sets
$\mathcal{A}(X)$	the algebra induced by a set X
$fr(X)$	the free group generated by a set X
G, H	groups
G_n	the group of isometries of $\mathbb{R}^n$
g, h	elements of a group G
$H \leq G$	H is a subgroup of G
$H \cong G$	H is isomorphic to G
μ	a measure
$\mathbb{N}$	the set of natural numbers, $\{1, 2, 3, \ldots\}$
$P(\mathbb{X})$	the power set of $\mathbb{X}$
R	a relation
R^*	the transitive closure of a relation R
$\mathbb{R}^n$	the n-dimensional Euclidean space
$\mathcal{R}$	a ring of sets
T_g	a translation of a G-set by an element $g \in G$
T_n	the group of translations of $\mathbb{R}^n$
SG_n	the group of orientation-preserving isometries of $\mathbb{R}^n$
SO_n	the group of rotations about the origin of $\mathbb{R}^n$
$[x]_R$	the equivalence class of x with respect to a relation R
$\mathbb{X}$	a space
$\mathcal{X}$	a class of subsets of $\mathbb{X}$
$\triangle$	the symmetric difference

Contents

1. Introduction

1.1. **On the Banach-Tarski Paradox.** In 1638, Galileo Galilei made an interesting observation: there is a one-to-one correspondence between the natural numbers, $\{1, 2, 3, 4, 5, \ldots\}$, and the squares of natural numbers, $\{1, 4, 9, 16, 25, \ldots\}$, i.e., for any natural number n, there is another natural number n^2, and for any natural number n^2, there is a number n that is a natural number [3]. While this perhaps seems obvious, note what this implies. The set of all natural numbers, $\{1, 2, 3, 4, 5, \ldots\}$ has the same number of elements as the set of all squares, $\{1, 4, 9, 16, 25, \ldots\}$. Observe any number that is a square is a natural number, but not every natural number is a square. Moreover, there is also a one-to-one correspondence between the set of natural numbers, $\{1, 2, 3, 4, 5, \ldots\}$, and the set of natural numbers that are not squares, $\{2, 3, 5, 6, 7, \ldots\}$, by the following assignments:

$$\begin{array}{ccc} 1 \to 2 & 4 \to 6 & 7 \to 10 \\ 2 \to 3 & 5 \to 7 & 8 \to 11 \\ 3 \to 5 & 6 \to 8 & \text{etc.} \end{array}$$

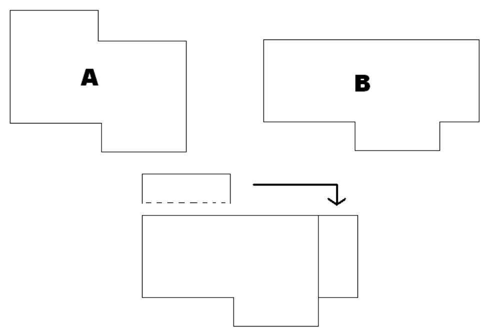
Thus, the set of natural numbers can be split apart into two completely separate sets that in effect have the same size as the original set.

We may observe that even though we split the natural numbers into two separate sets, because both of those sets also had an infinite number of elements, we were able to do the following. The natural numbers were split into two sets, squares and non-squares. Then each of those sets were "shrunk" back to the original set:

Squares	Non-squares
$\{1, 4, 9, 16, 25, \ldots\}$	$\{2, 3, 5, 6, 7, \ldots\}$
$1 \to 1$	$2 \to 1$
$4 \to 2$	$3 \to 2$
$9 \to 3$	$5 \to 3$
$16 \to 4$	$6 \to 4$
$25 \to 5$	$7 \to 5$

In fact, any way you divide up the natural numbers, assuming each division has an infinite amount of numbers, each division can be “shrunk” back to the original set of natural numbers. Note we say “shrunk” because some numbers get shifted more than others. Shifting is an entirely different process. When we “shrink,” it changes the distances between our numbers. If, however, we were to simply shift these sets of numbers, the distance between each number would be preserved.

Since ancient times, the ability to shift something has been important in mathematics. Consider the following figures.



Here, we can say figure A and figure B have the same area because we can cut a piece off figure A and move it over to the side to form figure B, as depicted. This is an intuitive way to determine the size of many different types of figures. As we move the pieces by translation and rotation, this should not effect the area.

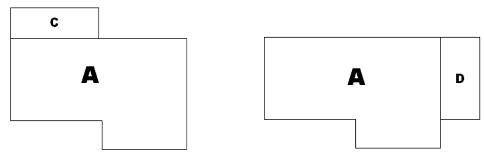
In 1924, Stefan Banach and Alfred Tarski did something similar to Galileo but without shrinking. They were able to cut a solid ball into pieces, move those pieces around, as mentioned above, but they ended up with two whole copies of the original sphere. Hence, after cutting the pieces and moving them around, the volume doubled. The theorem they published was appropriately named the Banach-Tarski paradox. Intuitively, it should not be possible to take something, cut it apart, move it around without shrinking or stretching, and end up with twice what you started with.

When examining this result, there is an obvious question to consider: is there some mistake in the proof, or in other words, is there an assumption made that is actually not correct? For years after its publication, many mathematicians believed that this was in fact the case. To understand this possible flaw, we must introduce the development of axiomatic set theory. In mathematics, it is not possible to prove everything—meaning, there are things that are so basic one is forced to simply believe that they are true. Through the work of two mathematicians, Ernst Zermelo and Abraham Fraenkel, these facts which must be accepted were reduced to nine axioms. All nine were easily accepted by mathematicians, that is until the publication of the Banach-Tarski paradox. Banach and Tarski relied heavily on the use of the ninth axiom, the axiom of choice. The axiom states: if we are given any number, possibly infinite, of collections of objects, it is possible to make a selection from each collection. Simply put, it is always possible to make a choice. This is often stated in the sock and shoe problem.

If given an infinite number of pairs of shoes, it is always possible to name a selection of one shoe from each pair, i.e., every right show. However, if you have an infinite number of pairs of socks, there is no way to name a selection from each pair. The axiom of choice, in this case, tells us even though it is impossible to name a selection of one sock from each pair, it is nonetheless possible to choose one sock from each pair. By accepting this axiom, while we cannot give an explicit definition of the set of our choices, we assume that such a set exists.

The Banach-Tarski paradox relies heavily on the axiom of choice, so for years, mathematicians used this proof as the reason this axiom should not be accepted. After some time passed however, the axiom of choice once again became widely accepted by mathematicians as a whole. The axiom is necessary for many areas of mathematics today, and ultimately, it would be even more counter-intuitive to reject the axiom. It would be hard to accept that, even if you cannot name a choice, it is impossible to make one.

Accepting the axiom, however, means that we are still left with this paradoxical theorem. Ultimately, we resolve this paradox by returning to our discussion of determining the area, or in this case volume, of figures. Recall these two figures had the same area.



We argued this by moving C to D. In fact, we can say the area of this figure is the area of A plus the area of C (or D). This, however, does not work with the Banach-Tarski paradox. In this paper, we divide the sphere into nine separate pieces, but unlike in the picture above, if we move those nine pieces around, we do not get the same volume. The reason for this paradox, and the ultimate unravelling of our mystery, is that the nine pieces we divided the ball into are actually not measurable.

This explanation, of course, is also somewhat counter-intuitive, as it suggests that there are things that we simply cannot measure. In this paper, we will formalize what is necessary for something to be measurable, but basically it requires two of the concepts mentioned already:

- when you combine two separate figures that are measurable, the union of the two is also measurable,
- when a figure that is measurable is moved, it has the same measure as it had before it was moved.

To be correct mathematically, the latter is for a stricter form of measurement called an invariant measurement, but this stricter form will be necessary for much of our work.

Our conclusion is that paradoxical sets, sets that can be divided up and moved around to form two copies of the original set, are directly linked to the fact that their paradoxical decompositions are not measurable.

1.2. **Timeline.** The following are some of the contributions relevant to the field of paradoxical groups.

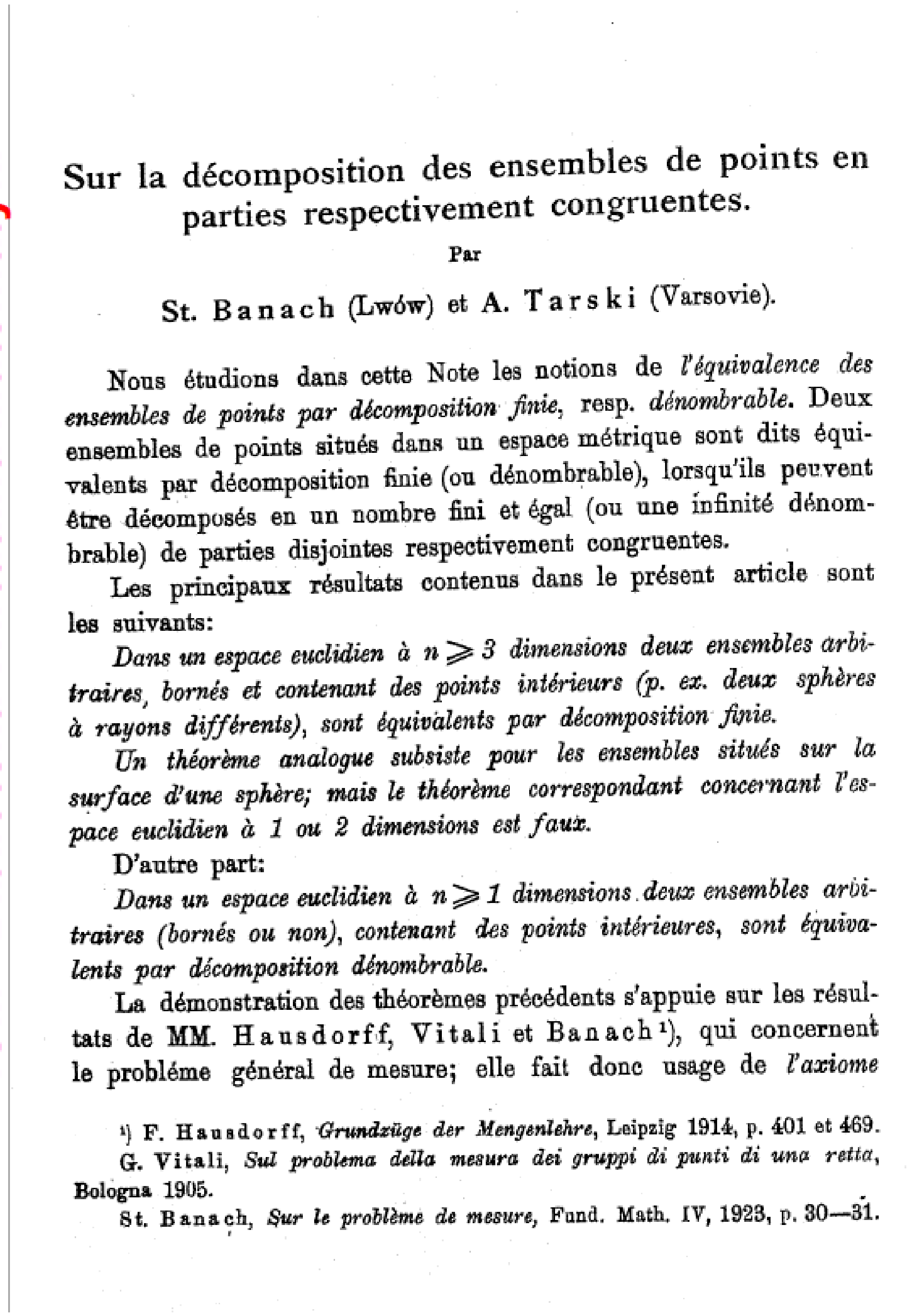
- In 1638, Galileo Galilei showed that the natural numbers could be partitioned into two pieces which were both equinumerous with all natural numbers [3].
- In 1905, Giuseppe Vitali showed that there existed a non-measurable subset of the real number line [16] (compare with Proposition 7.6).
- In 1914, Hausdorff showed that the sphere, minus a countable number of points, was paradoxical [5].
- Also in 1914, Wacław Sierpiński and Stefan Mazurkiewicz gave an example of a paradoxical subset of the plane [9].
- In 1924, Stefan Banach and Alfred Tarski showed a ball in $\mathbb{R}^3$ was paradoxical [1].
- In 1991, Janusz Pawlikowski showed that the Hahn-Banach theorem is enough to construct a paradoxical decomposition of the ball [14].
- In 1994, Randall Dougherty and Matthew Foreman showed the Banach-Tarski paradox could be done with pieces that possess the Property of Baire [2].

1.3. **Mathematicians.** In the study of paradoxical groups, there are two results that are quintessential to this topic: the Hausdorff paradox and the Banach-Tarski paradox. Here we introduce three mathematicians who were crucial to the development of paradoxical groups. The following bibliographies are paraphrased from the respected University of St. Andrews mathematics history website "The MacTutor History of Mathematics archive" [11].

Felix Hausdorff was born on November 8, 1868, in Breslau, Germany (present day Wrocław, Poland). He was born into a rich Jewish family, so Hausdorff did not have to work to support himself and was encouraged to pursue academics ranging from literature and music to mathematics. In 1891, Hausdorff graduated with his doctorate thesis on a mathematical topic in astronomy. He continued to publish papers in this field, but he began to focus more on literary works. Hausdorff went on to publish multiple literary works until, when in 1904, he began his preeminent work in set theory. In 1914, Hausdorff published *Grundzüge der Mengenlehre* which contained his paradoxical decomposition of almost the entire sphere. Hausdorff is famous for numerous other contributions in mathematics, including the generalized continuum hypothesis, Hausdorff spaces, Hausdorff measure, and Hausdorff dimension. In 1942, Hausdorff took his own life to avoid being sent to a concentration camp by the Nazis.

Stefan Banach was born on March 30, 1892, in Kraków, Poland (then Austria-Hungary). Banach was raised by his father, as his mother left him right after his birth. During his secondary education, Banach mathematical prowess became evident and math was seemingly his only interest. He became friends with a classmate, Witold Wilkosz, who also went on to become a professor of mathematics. Banach went on to an engineering university, where he supported himself by tutoring young students. In 1916, by a chance meeting with Hugo D. Steinhaus, Banach began to work in mathematical research. He published his first paper in 1918 and shortly thereafter continued to publish many papers. Banach became part of and helped found the Mathematic Society of Kraków, which would later become the Polish Mathematical Society. In 1920, Banach was offered an assistantship at the Jan Kazimierz University of Lwów. At this time, Banach submitted a doctorate thesis and was awarded his degree with no other university course work. This thesis has since been called the birth of functional analysis. In 1924, Banach was promoted to full professor, and this same year, he published the joint paper with Tarski, *Sur la décomposition des ensembles de points en partiens respectivement congruent* [1], on the paradoxical decomposition of the sphere. Banach would continue produce many works until the start of World War II. However, due to the Nazi occupation in 1941, Banach was forced to work in a German lab feeding lice. He died of lung cancer in 1945.

Alfred Tarski was born on January 14, 1902, in Warsaw, Poland (then Russian Empire). He was born into a Jewish family and his name at birth was Alfred



Sur la décomposition des ensembles de points en parties respectivement congruentes.

Par

St. Banach (Lwów) et A. Tarski (Varsovie).

Nous étudions dans cette Note les notions de *l'équivalence des ensembles de points par décomposition finie*, resp. *dénombrable*. Deux ensembles de points situés dans un espace métrique sont dits équivalents par décomposition finie (ou dénombrable), lorsqu'ils peuvent être décomposés en un nombre fini et égal (ou une infinité dénombrable) de parties disjointes respectivement congruentes.

Les principaux résultats contenus dans le présent article sont les suivants:

Dans un espace euclidien à $n \geqslant 3$ dimensions deux ensembles arbitraires, bornés et contenant des points intérieurs (p. ex. deux sphères à rayons différents), sont équivalents par décomposition finie.

Un théorème analogue subsiste pour les ensembles situés sur la surface d'une sphère; mais le théorème correspondant concernant l'espace euclidien à 1 ou 2 dimensions est faux.

D'autre part:

Dans un espace euclidien à $n \geqslant 1$ dimensions deux ensembles arbitraires (bornés ou non), contenant des points intérieures, sont équivalents par décomposition dénombrable.

La démonstration des théorèmes précédents s'appuie sur les résultats de MM. Hausdorff, Vitali et Banach[1]), qui concernent le probléme général de mesure; elle fait donc usage de *l'axiome*

[1]) F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig 1914, p. 401 et 469. G. Vitali, *Sul problema della mesura dei gruppi di punti di una retta*, Bologna 1905. St. Banach, *Sur le problème de mesure*, Fund. Math. IV, 1923, p. 30—31.

Figure 1. Banach, Tarski. Sur la décomposition des ensembles de points en partiens respectivement congruent. *Fundamenta Mathematicae* 6: 244-277, 1924. Available online at http://matwbn.icm.edu.pl/ksiazki/fm/fm6/fm6127.pdf.

Teitelbaum. In 1915, Russia withdrew from Warsaw and a vibrant math community was formed at the re–founded Warsaw University. In 1918, Teitelbaum began studying at the Warsaw University, where he would come to study under Stanislaw Leśniewski, Jan Łukasiewicz, Wacław Sierpiński, and Stefan Mazurkiewicz. At age 19, Teitelbaum published his first paper on set theory. In 1923, Teitelbaum converted to Roman Catholicism and changed his last name to Tarski. These changes reflected both a strong Polish identity and a desire to receive a university appointment, which would have been difficult due to anti-Semitic sentiments. In 1924, Tarski published, along with Stefan Banach, *Sur la décomposition des ensembles de points en partiens respectivement congruent.* Still, Tarski had difficulties getting appointments due to his Jewish heritage. In 1939, Tarski traveled to the United States. When Germany invaded Poland two weeks later, Tarski was allowed to stay in the United States and was eventually able to get his family to join him. In 1943, Tarski received an appointment at University of California, Berkeley and would continue to teach until 1973. He died on October 26, 1983. Tarski has been called the founder of the American school of mathematical logic and philosophy of mathematics.

1.4. **Contents of the Thesis.** Section 2 provides some of the necessary background needed to develop the theory of paradoxical groups. Subsection 2.1 covers the theory of relations which is used in Subsection 2.3 to define free groups. Subsection 2.2 introduces groups for Subsection 2.3 and Section 7. Subsection 2.4 introduces group actions which we use to move around the pieces to form our paradoxical decompositions. Section 3 introduces the main topic of the thesis, paradoxical groups. Section 4 shows a group of rotations in $\mathbb{R}^3$ is isomorphic to the paradoxical group introduced in Section 3. Section 5 ties everything together to form a paradoxical decomposition of almost the entire sphere (the Hausdorff paradox) and the ball (the Banach-Tarski paradox).

The second half of this thesis focuses on showing that this paradox is not possible in $\mathbb{R}$ and $\mathbb{R}^2$. Section 6 defines measure and integration. Section 7 introduces amenable groups—groups that are measurable in a strict manner that do not allow for such paradoxes that were shown in Section 5. Section 8 concludes the second half of the paper by showing that no bounded subset of $\mathbb{R}$ and $\mathbb{R}^2$ with non-empty interior is paradoxical.

2. Prelude: Relations, Free Groups, and Group Actions

2.1. **Relations.** We will review the basic concepts on the theory of relations which we will use to establish equivalence classes.

Definition 2.1. (1) *A relation R on a set X is a set of ordered pairs of elements of X. By convention, if an ordered pair $(a, b) \in R$, then we write aRb.*

(2) *A relation R is said to be*
- *reflexive if aRa for every $a \in X$.*
- *symmetric if aRb implies bRa for any $a, b \in X$.*
- *transitive if aRb and bRc implies aRc for any $a, b, c \in X$.*
- *an equivalence relation if R is reflexive, symmetric and transitive.*

Definition 2.2. *The transitive closure R^* of a relation R is the relation introduced as follows: aR^*b if and only if there exists a finite sequence a_1, a_2, …, a_n such that aRa_1, a_1Ra_2, …, a_nRb.*

Proposition 2.3. *The transitive closure R^* of a relation R on a set X is transitive.*

Proof. Suppose aR^*b and bR^*c for some $a, b, c \in X$. Then there exist finite sequences $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_m$ such that

$$aRa_1, a_1Ra_2, \ldots, a_nRb \qquad \text{and} \qquad bRb_1, b_1Rb_2, \ldots, b_mRc.$$

When we combine these two finite sequences, we get another finite sequence a_1, a_2, …, a_n, b, b_1, b_2, …, b_m such that

$$aRa_1, a_1Ra_2, \ldots, a_nRb, bRb_1, b_1Rb_2, \ldots, b_mRc,$$

showing that aR^*c. Thus, R^* is transitive. □

Proposition 2.4. *The transitive closure R^* of a reflexive and symmetric relation R on a set X is an equivalence relation.*

Proof. We have already shown that R^* is transitive, so we must now show that R^* is reflexive and symmetric. Let R be a relation that is reflexive and symmetric, and let R^* be its transitive closure. Since for every $a \in X$, aRa, the finite sequence of just a witnesses the fact that aR^*a. Thus, R^* is reflexive. Now suppose aR^*b for some $a, b \in X$. Then there exists a finite sequence such that

$$aRa_1, a_1Ra_2, \ldots, a_nRb.$$

But R is symmetric, so we also have

$$a_1Ra, a_2Ra_1, \ldots, bRa_n,$$

or if we rearrange our sequence,

$$bRa_n, a_nRa_{n-1}, \ldots, a_1Ra.$$

Thus, R^* is symmetric, so R^* is an equivalence relation. □

Definition 2.5. *An operation (or a binary operation) $\circ$ on a set X is a mapping from $X \times X$ into X.*

Definition 2.6. *Let $\circ$ be an operation on a set X, and let R be a relation on X.*

- *We say R is pre-normal with respect to the operation $\circ$ if the following condition is satisfied:*

  $$\text{if } aRa' \text{ and } b \in X, \text{ then } b \circ aRb \circ a' \text{ and } a \circ bRa' \circ b.$$

- *We say R is normal with respect to the operation $\circ$ if the following condition is satisfied:*

  $$\text{if } aRa' \text{ and } bRb', \text{ then } b \circ aRb' \circ a' \text{ and } a \circ bRa' \circ b'.$$

Theorem 2.7. *Let $\circ$ be an operation on a set X, and let R be a relation on X. Then*

(1) *if R is transitive and pre-normal, then R is normal,*

(2) *if R is reflexive and normal, then R is pre-normal,*

(3) *if a R is pre-normal, then its transitive closure R^* is normal.*

Proof. (1) Let R be transitive and pre-normal. Suppose aRa' and bRb' for some $a, a', b, b' \in X$. As R is pre-normal, we have

$$b \circ aRb \circ a' \qquad \text{and} \qquad b \circ a'Rb' \circ a',$$

and as R is transitive, we have $b \circ aRb' \circ a'$. Similarly, as R is pre-normal, we have

$$a \circ bRa' \circ b \qquad \text{and} \qquad a' \circ bRa' \circ b',$$

and as R is transitive, we have $a \circ bRa' \circ b'$. Thus R is normal.

(2) Let R be a reflexive and normal. As R is reflexive, if $b \in X$ then bRb. Now suppose aRa'. Then

$$a \circ bRa' \circ b \qquad \text{and} \qquad b \circ aRb \circ a',$$

so R is pre-normal.

(3) As R^* is transitive by Proposition 2.3, we simply must show R^* is pre-normal. Suppose aR^*a' and $b \in X$. Then there exists a finite sequence $a_1, a_2, \ldots, a_n$ such that

$$aRa_1, a_1Ra_2, \ldots, a_nRa'.$$

Since R is pre-normal, we have

$$b \circ aRb \circ a_1, b \circ a_1Rb \circ a_2, \ldots, b \circ a_nRb \circ a'$$

and

$$a \circ bRa_1 \circ b, a_1 \circ bRa_2 \circ b, \ldots, a_n \circ bRa' \circ b.$$

So

$$b \circ aR^*b \circ a' \qquad \text{and} \qquad a \circ bR^*a' \circ b.$$

Thus, R^* is pre-normal, and by (1), R^* is normal.

□

Definition 2.8. *Let R be a reflexive and symmetric relation on a set X, and let R^* be its transitive closure. We define $[x]_{R^*} = \{a : aR^*x\}$ and call it the equivalence class of x. If the equivalence relation R^* is clear, we may simply write $[x]$.*

2.2. **Groups.** In this subsection, we introduce groups for the following subsection, as well as Section 7.

Definition 2.9. *Let $\circ$ be an operation on a set G and let $g, h, k \in G$. Then*

- *$\circ$ is associative if and only if $g \circ (h \circ k) = (g \circ h) \circ k$,*
- *$\circ$ has an identity if and only if there exists $e \in G$ such that $g \circ e = e \circ g = g$ for all $g \in G$,*
- *$\circ$ has inverses if and only if for any $g \in G$, there exists $g^{-1} \in G$ such that $g \circ g^{-1} = g^{-1} \circ g = e$.*

If $\circ$ is associative, has an identity, and has inverses, then the set G with the operation $\circ$ is called a group. If the operation is clear, then we simply say that G is a group.

Definition 2.10. *Let G be a group. If $g \circ h = h \circ g$ for any $g, h \in G$, then G is called Abelian.*

Definition 2.11. *Let G be a group and $H \subseteq G$. If H is also a group with the same operation as G, then H is called a subgroup of G.*

Definition 2.12. *Let G be a group, $g \in G$, and H be a subgroup of G.*

(1) *We define the set $gH = \{g \circ h : h \in H\}$ and call it a left coset of G. Similarly, we define $Hg = \{h \circ g : h \in H\}$ and call it a right coset of G.*
(2) *If $gH = Hg$ for every $g \in G$, H is called a normal subgroup of G.*

Proposition 2.13. *Let H be a normal subgroup of a group G. Define the set $G/H = \{gH : g \in G\}$, and define the operation $\odot : G/H \times G/H \longrightarrow G/H$ by $g_1H \odot g_2H = (g_1 \circ g_2)H$. Then G/H with the operation $\odot$ is a group and is called a factor group of G.*

Proof. First we will note the operation $\odot$ is well-defined, i.e., for $g_1H = g_2H$ and $g_3H = g_4H$ that $g_1H \odot g_3H = g_2H \odot g_4H$. So let $g_1H = g_2H$ and $g_3H = g_4H$.

$$\begin{aligned} g_1H \odot g_3H &= (g_1 \circ g_3)H = g_1(g_3H) = g_1(g_4H) = (g_1 \circ g_4)H \\ &= H(g_1 \circ g_4) = (Hg_1)g_4 = (Hg_2)g_4 = H(g_2 \circ g_4) \\ &= (g_2 \circ g_4)H = g_2H \odot g_4H. \end{aligned}$$

To show $\odot$ is associative, observe

$$\begin{aligned} g_1H \odot (g_2H \odot g_3H) &= g_1H \odot (g_2 \circ g_3)H \\ &= (g_1 \circ (g_2 \circ g_3))H \\ &= ((g_1 \circ g_2) \circ g_3)H \\ &= (g_1 \circ g_2)H \odot g_3H \\ &= (g_1H \odot g_2H) \odot g_3H. \end{aligned}$$

Note H is the identity of G/H as $gH \odot H = H \odot gH = gH$ for any $gH \in G/H$. Also observe $(gH)^{-1} = g^{-1}H$ as $gH \odot g^{-1}H = g^{-1}H \odot gH = eH = H$. Thus G/H is a group. □

2.3. **Free Groups.** In this subsection, we define a free group generated by a set X. Free groups are key to our paradoxical decomposition of the sphere. Specifically, in subsequent sections, we will focus on the free group with two generators, i.e., $X = \{a, b\}$.

Definition 2.14. *Let X be a non-empty set where elements of X are characters in some alphabet. For every $x \in X$, we consider a formal object $x^{-1} \notin X$ and let $X^{-1} = \{x^{-1} : x \in X\}$.*

- *A word from X is a finite string $x_1x_2x_3 \ldots x_n$ where $x_i \in X \cup X^{-1}$ for all $i \leq n$. Let $W(X)$ be the set of all possible words of X including the empty word, e.*
- *We define $lh : W(X) \longrightarrow \mathbb{Z}$ by $lh(x_1x_2x_3 \ldots x_n) = n$ and we call n the length of $x_1x_2x_3 \ldots x_n$.*

- *We define the binary operation $\cdot$ on $W(X)$ to be concatenation, i.e., if $x_1x_2x_3\dots x_n \in W(X)$ and $y_1y_2y_3\dots y_m \in W(X)$ then*

$$(x_1x_2x_3\dots x_n)\cdot(y_1y_2y_3\dots y_n) = x_1x_2x_3\dots x_ny_1y_2y_3\dots y_m.$$

This operation is well-defined, associative, and has an identity (the empty word). However, we do not have inverses. Note that a^{-1} is not the inverse of a as aa^{-1} is not the empty word. To correct this problem, we will identify some words, in particular, $aa^{-1} = a^{-1}a = e$.

Let $u, v \in W(X)$. We define the relation R on $W(X)$ as follows. Let $u, v \in W(X)$. We say uRv if and only if $u = v$ or v can be obtained from u by adding or removing a word of the form xx^{-1} or $x^{-1}x$, where $x \in X$.

Let R^* be the transitive closure of R (see Definition 2.2).

Lemma 2.15. *The transitive closure R^* forms an equivalence relation on $W(X)$.*

Proof. From Proposition 2.4, we only need to show R is reflexive and symmetric. By definition, R is reflexive, i.e., uRu for all $u \in W(X)$. Now suppose uRv, i.e. v can be obtained from u by adding or removing a word of the form xx^{-1} or $x^{-1}x$, where $x \in X$. Then u can be formed from v by doing the opposite addition or removal of that same word xx^{-1} or $x^{-1}x$. Thus uRv, and R is symmetric. □

We will denote the set of equivalence classes of $W(X)$ with respect to R^* by $fr(X)$. We will denote the R^* equivalence class of $x \in X$ by $[x]$.

Definition 2.16. *Let $u \in W(X)$. We say u is a reduced word if and only if for all $v \in W(X)$ such that $[u] = [v]$, $lh(u) \leq lh(v)$.*

Let $[u], [v] \in fr(X)$. Define the binary operation $\circ$ on $fr(X)$ by $[u] \circ [v] = [u \cdot v]$.

Lemma 2.17. *The relation R^* is normal with respect to the operation $\cdot$. Consequently, the operation $\circ$ defined above is well-defined.*

Proof. First, we will show that R^* is pre-normal with respect to $\cdot$. Let $u, v, w \in W(X)$ and suppose uR^*v. Then there exists a finite sequence uRu_1, u_1Ru_2, …, u_nRv. Since uRu_1, u_1 can be formed from u by adding or removing a word of the form xx^{-1} or $x^{-1}x$, where $x \in X$. Then $u_1 \cdot w$ can be formed by doing the corresponding addition or removal from $u \cdot w$. Thus $u \cdot wRu_1 \cdot w$, and by an analogous argument, we have

$$u \cdot wRu_1 \cdot w, u_1 \cdot wRu_2 \cdot w, \dots, u_n \cdot wRv \cdot w.$$

Thus, $u \cdot wR^*v \cdot w$. By the same argument, we may show that $w \cdot uR^*w \cdot v$. Thus, R^* is pre-normal with respect to the operation $\cdot$. As R^* is an equivalence relation, R^* is transitive, and by Proposition 2.7, R^* is normal. □

Theorem 2.18. *The set $fr(X)$ with the operation $\circ$ is a group and is called the free group generated by X.*

Proof. First, note that $[e]$, the equivalence class of the empty word, is the identity element, as for any $[u] \in fr(X)$,

$$[u] \circ [e] = [u \cdot e] = [u] = [e \cdot u] = [e] \circ [u].$$

To show the operation $\circ$ is associative, recall the operation $\cdot$ was associative, so for any $[u], [v], [w] \in fr(X)$,

$$[u] \circ ([v] \circ [w]) = [u \cdot (v \cdot w)] = [(u \cdot v) \cdot w] = ([u] \circ [v]) \circ [w].$$

Finally, to show any element of $fr(X)$ has an inverse with respect to $\circ$, let $[u] \in fr(X)$. Then $u = (x_1)^{t_1}(x_2)^{t_2} \dots (x_n)^{t_n}$, where $x_i \in X$, and

$$(x_i)^{t_i} = \begin{cases} x_i^{-1} & \text{if} \quad t = -1 \\ x_i & \text{if} \quad t = 1 \end{cases}$$

Let $(u)^{-1} = (x_n)^{-t_n}(x_{n-1})^{-t_{n-1}} \dots (x_1)^{-t_1}$. Then

$$\begin{aligned}
[u] \circ [(u)^{-1}] &= [u \cdot (u)^{-1}] \\
&= [(x_1)^{t_1}(x_2)^{t_2} \dots (x_n)^{t_n} \cdot (x_n)^{-t_n}(x_{n-1})^{-t_{n-1}} \dots (x_1)^{-t_1}] \\
&= [(x_1)^{t_1}(x_2)^{t_2} \dots (x_{n-1})^{t_{n-1}} \cdot (x_{n-1})^{-t_{n-1}}(x_{n-2})^{-t_{n-2}} \dots (x_1)^{-t_1}] \\
&= \dots = [e] \\
&= [(x_1)^{-t_1} \cdot (x_1)^{t_1}] \\
&= [(x_1)^{-t_1}(x_2)^{-t_2} \cdot (x_2)^{t_2}(x_1)^{t_1}] \\
&= \dots = [(x_1)^{-t_1}(x_2)^{-t_2} \dots (x_n)^{-t_n} \cdot (x_n)^{t_n}(x_{n-1})^{t_{n-1}} \dots (x_1)^{t_1}] \\
&= [(u)^{-1} \cdot u] = [(u)^{-1}] \circ [u].
\end{aligned}$$

Thus, $[u]^{-1} = [(u)^{-1}] \in fr(X)$.

Hence, $fr(X)$ is a group. □

The following theorem is the main property of free groups, which can be used to define free groups in an abstract way.

Theorem 2.19. *If G is a group, X is a set, and $f : X \longrightarrow G$ is a function, then there exists a unique homomorphism $\hat{f} : fr(X) \longrightarrow G$ such that $\hat{f}([x]) = f(x)$ for all $x \in X$.*

Proof. Let G be a group, X be a set, and $f : X \longrightarrow G$ be a function. Now define $\hat{f} : fr(X) \longrightarrow G$ by

$$\hat{f}([x_1x_2 \dots x_n]) = f(x_1)f(x_2) \dots f(x_n) \qquad (*)$$

where $f(x^{-1}) = (f(x))^{-1}$ and $(f(x))^{-1}$ is the inverse of $f(x)$ in G. First, we must show $\hat{f}$ is well-defined. To show $\hat{f}$ is well-defined, recall $[x_1x_2 \dots x_n] = [y_1y_2 \dots y_m]$ if and only if $y_1y_2 \dots y_n$ can be formed from $x_1x_2 \dots x_n$ by the addition and removal of words of the form zz^{-1} or $z^{-1}z$. However, when we apply $\hat{f}$ to $[x_1x_2 \dots x_n]$ or $[y_1y_2 \dots y_m]$, any words of the form zz^{-1} or $z^{-1}z$ have no effect on the output, as

$$f(z)f(z^{-1}) = f(z)f(z)^{-1} = e_G = f(z)^{-1}f(z) = f(z^{-1})f(z).$$

Therefore, if $[x_1x_2\dots x_n] = [y_1y_2\dots y_m]$, then $\hat{f}([x_1x_2\dots x_n]) = \hat{f}([y_1y_2\dots y_m])$. Thus, $\hat{f}$ is well-defined.

To show $\hat{f}$ is a homomorphism, observe

$$\begin{aligned}\hat{f}([x_1x_2\dots x_n] \circ [y_1y_2\dots y_m]) &= \hat{f}([x_1x_2\dots x_n \cdot y_1y_2\dots y_m]) \\ &= f(x_1)f(x_2)\dots f(x_n)f(y_1)f(y_2)\dots f(y_m) \\ &= \hat{f}([x_1x_2\dots x_n])\hat{f}([y_1y_2\dots y_m]).\end{aligned}$$

Thus, $\hat{f}$ is a homomorphism.

By its definition, $\hat{f}([x]) = f(x)$. Moreover, since $[x_1x_2\dots x_n] = [x_1] \circ [x_2] \circ \dots \circ [x_n]$, it follows that any homomorphism with this property is defined by $(*)$. □

2.4. **Group Actions.** In subsequent sections, we will view rotations as group actions acting on a sphere. To that end, we develop the necessary theory of group actions.

Definition 2.20. *Let G be a group and X a set. We say G operates on X (on the left) when there is a mapping $a : G \times X \longrightarrow X$ such that for all $g, h \in G$ and $x \in X$, $a(gh, x) = a(g, a(hx))$ and $a(e_G, x) = x$, where e_G is the identity of G. If G operates on X, we call X a G-set. We abuse the notation by writing gx for $a(g, x)$. Typically, the group G is a set of permutations of X so this abuse should cause no confusion.*

Let X be a G-set, $g \in G$, and $x, y \in X$. Then

- *the mapping $T_g : X \longrightarrow X$ defined by $T_g(x) = gx$ is called a translation,*
- *the set $orb(x) = \{T_g(x) : g \in G\}$ is called the orbit of x,*
- *define the relation R^G on X by xR^Gy if and only if $x \in orb(y)$.*

For the remainder of this section, assume that G is a group and X is a G-set.

Example 2.21. *The group G is a G-set.*

Proof. Let $g, h, x \in G$. Then $g(hx) = (gh)x$ as the group is associative and $e_Gx = x$ by the definition of the identity of G. Thus, G is a G-set. □

As G is a G-set, we will use the group action notation in groups, in particular $T_g : G \longrightarrow G$ such that $T_g(h) = gh$, and we will call T_g a group translation. Also, we will freely switch between the notation $T_g(x)$ and gx for a translation by group element g.

Proposition 2.22. *The relation R^G defined in Definition 2.20 forms an equivalence relation on X.*

Proof. We must show R^G is reflexive, symmetric, and transitive. First, note $x \in orb(x)$ so xR^Gx. Thus R^G is reflexive. To show R^G is symmetric, assume xR^Gy. In other words, $x = T_e(x) \in orb(y)$, so there exists a $g \in G$ such that $T_g(y) = gy = x$. Then

$$y = e_G y = (g^{-1}g)y = g^{-1}(gy) = g^{-1}x = T_{g^{-1}}(x).$$

Hence, $y \in orb(x)$ so yR^Gx. Thus, R^G is symmetric. To show R^G is transitive, we suppose xR^Gy and yR^Gz, that is $x \in orb(y)$ and $y \in orb(z)$. Thus, there exist $g, h \in G$ such that

$$T_g(y) = gy = x \qquad \text{and} \qquad T_h(z) = hz = y.$$

Then

$$T_{gh}(z) = (gh)z = g(hz) = g(y) = x,$$

and thus, $x \in orb(z)$. Hence xR^Gz and R^G is transitive. As R^G is reflexive, symmetric, and transitive, R^G forms an equivalence relation on X. □

Definition 2.23. *Let X be a G-set and $x \in X$. Then the set $G_{stab}(x) = \{g \in G : gx = x\}$ is called the stabilizer of x.*

One easily checks that $G_{stab}(x)$ is a subgroup of G.

Proposition 2.24. *Let $g, h \in G$. If $gG_{stab}(x) = hG_{stab}(x)$ then $gx = hx$.*

Proof. Let $x \in X$ and $g, h \in G$. Suppose $gG_{stab}(x) = hG_{stab}(x)$. Since, $g \in gG_{stab}(x)$, $g \in hG_{stab}(x)$, so there exists a $g' \in G_{stab}(x)$ such that $hg' = g$. Then $gx = hg'x = hx$. □

Proposition 2.25. *If $G_{stab}(x) = \{e_G\}$ and $gx = hx$ for $g, h \in G$, then $g = h$.*

Proof. Let $gx = hx$. Then $x = g^{-1}hx$ so $g^{-1}h \in G_{stab}(x)$. Thus $g^{-1}h = e_G$ so $g = h$. □

Definition 2.26. *Let X be a G-set and $x \in X$. Define the mapping $\pi_x : G \longrightarrow orb(x)$ by $\pi_x(g) = gx$.*

Definition 2.27. *Let X be a set and let $X_1, X_2, \ldots, X_n$ be subsets of X such that*

- $X_1 \cup X_2 \cup \ldots \cup X_n = X$,
- *for all $i \neq j$, $X_i \cap X_j = \emptyset$. Then $X_1, X_2, \ldots, X_n$ is called a partition of X.*

Proposition 2.28. *Let $x \in X$ such that $G_{stab}(x) = \{e_G\}$. If A_1, A_2, $\ldots$, A_n is a partition of G, then $\pi_x[A_1]$, $\pi_x[A_2]$, $\ldots$, $\pi_x[A_n]$ is a partition of $orb(x)$.*

Proof. Suppose $G_{stab}(x) = \{e_G\}$ for some $x \in X$ and A_1, A_2, $\ldots$, A_n is a partition of G. First, we will show

$$\pi_x[A_1] \cup \pi_x[A_2] \cup \ldots \cup \pi_x[A_n] = orb(x).$$

Suppose $y \in orb(x)$. Then there exists $g \in G$ such that $gx = y$, and as $A_1 \cup A_2 \cup \ldots \cup A_n = G$, $g \in A_i$ for some $i \in \{1, 2, \ldots n\}$. Therefore,

$$y = \pi_x(g) \in \pi_x[A_i] \subseteq \pi_x[A_1] \cup \pi_x[A_2] \cup \ldots \cup \pi_x[A_n].$$

The other inclusion follows immediately. Now we want to show $\pi_x[A_i] \cap \pi_x[A_j] = \emptyset$ for $i \neq j$. Suppose toward contradiction that $y \in \pi_x[A_i] \cap \pi_x[A_j]$ for $i \neq j$. Then there exist $g \in A_i$ and $h \in A_j$ such that $T_g(x) = gx = y$ and $T_h(x) = hx = y$. But then $gx = hx$, so $(h^{-1}g)x = x$. Thus $h^{-1}g = e_G$ so $g = h$, by our assumption that

$G_{stab}(x) = \{e_G\}$. Hence, $g = h \in A_i \cap A_j \neq \emptyset$, contradicting $i \neq j$ and A_1, A_2, ..., A_n being a partition.

Thus, $\pi_x[A_1]$, $\pi_x[A_2]$, ..., $\pi_x[A_n]$ is a partition of $orb(x)$. $\square$

3. Paradoxical Groups

Consider the free group with two generators, $F = fr(\{a, b\})$. We want to partition F into pieces which are in some sense similar. Define

$$F_a = \{[w] \in F : w \text{ is a reduced word and } w \text{ begins, on the left, with } a\},$$

(see Definition 2.16 for reduced word). Similarly, define $F_{a^{-1}}$, F_b, and $F_{b^{-1}}$. Note that this way, we have formed a partition of F,

$$F = F_a \cup F_{a^{-1}} \cup F_b \cup F_{b^{-1}} \cup \{[e]\}.$$

We now consider group translations of the parts of this partition. First, we will examine $T_{[a^{-1}]}[F_a]$. Let $[ax_1x_2 \ldots x_n] \in F_a$, where $ax_1x_2 \ldots x_n$ is a reduced word, so in particular $x_1 \neq a^{-1}$. Then

$$T_{[a^{-1}]}([ax_1x_2 \ldots x_n]) = [a] \circ [x_1x_2 \ldots x_n] = [x_1x_2 \ldots x_n] \in F \setminus F_{a^{-1}}.$$

Conversely, consider $[x_1x_2 \ldots x_n] \in F \setminus F_{a^{-1}}$, where $x_1x_2 \ldots x_n$ is a reduced word so in particular $x_1 \neq a^{-1}$. Then $[ax_1x_2 \ldots x_n] \in F_a$ so

$$[x_1x_2 \ldots x_n] = [a^{-1}ax_1x_2 \ldots x_n] \in T_{[a^{-1}]}[F_a].$$

Thus, we have shown

$$T_{[a^{-1}]}[F_a] = F \setminus F_{a^{-1}}.$$

Clearly, as $T_{[e]}$ is the identity mapping, we have

$$T_{[e]}[F_{a^{-1}}] = F_{a^{-1}}.$$

Similarly, we have $T_{[b^{-1}]}[F_b] = F \setminus F_{b^{-1}}$ and $T_{[e]}[F_{b^{-1}}] = F_{b^{-1}}$. Thus, sets F_a and $F_{a^{-1}}$, after being mapped by $T_{[a^{-1}]}$ and $T_{[e]}$ respectively, create a copy of F. Also, sets F_b and $F_{b^{-1}}$ get mapped by group translation to a copy of F. In other words, we have divided F into five pieces and after using group translations on four of those pieces, ended up with two copies of F. However, note any translation of the last part of our partition, $\{[e]\}$, is non-empty. It is our goal, however, to get exactly two copies of F, so we will add the fifth piece, $\{[e]\}$, to one of the other four and make the following modification:

$$A_1 = F_a \setminus \{[a], [aa], [aaa], \ldots\}$$
$$A_2 = F_{a^{-1}} \cup \{[e]\} \cup \{[a], [aa], [aaa], \ldots\}$$
$$B_1 = F_b$$
$$B_2 = F_{b^{-1}}.$$

Now we examine $T_{[a^{-1}]}[A_1]$. Let $[ax_1x_2 \ldots x_n] \in A_1$ where $ax_1x_2 \ldots x_n$ is a reduced word, so in particular $x_1 \neq a^{-1}$ and $x_i \neq a$ for some $i \in \{1, 2, \ldots, n\}$. Then

$$T_{[a^{-1}]}([ax_1x_2 \ldots x_n]) = [a^{-1}] \circ [ax_1x_2 \ldots x_n] = [x_1x_2 \ldots x_n],$$

a reduced word such that $x_1 \neq a^{-1}$ and $x_i \neq a$ for some $i \in \{1, 2, \ldots, n\}$. Also,

$$T_{[a^{-1}]}([ax_1x_2 \ldots x_n]) \neq [e] \text{ as } [a] \notin A_1.$$

Thus,

$$T_{[a^{-1}]}([ax_1x_2 \ldots x_n]) \in F \setminus A_2.$$

Conversely, let $[x_1x_2 \ldots x_n] \in F \setminus A_2$, where $x_1x_2 \ldots x_n$ is a non-empty reduced word such that $x_1 \neq a^{-1}$ and $x_i \neq a$ for some $i \in \{1, 2, \ldots, n\}$. Then $[ax_1x_2 \ldots x_n] \in A_1$ so

$$[x_1x_2 \ldots x_n] = [a^{-1}ax_1x_2 \ldots x_n] = [a^{-1}] \circ [ax_1x_2 \ldots x_n] \in T_{[a^{-1}]}[A_1].$$

This completes the argument that

$$T_{[a^{-1}]}[A_1] = F \setminus A_2.$$

Also, from our previous arguments, we have all of the following:

$$T_{[e]}[A_2] = A_2$$
$$T_{[b^{-1}]}[B_1] = F \setminus B_2$$
$$T_{[b^{-1}]}[B_2] = B_2.$$

We have now partitioned F into four pieces and moved A_1 and A_2 by group translations to get the entire group F, and we have also moved B_1 and B_2 by group translations to get the entire group F. Also, these two copies of F do not overlap. This effect is formalized in the following definition.

Definition 3.1. *We say that a group G is n–paradoxical if there exist a partition A_1, A_2, …, A_i, … A_n of G and translations T_1, T_2, …,T_i, …, T_n such that*

$$T_1[A_1], T_2[A_2], \ldots, T_i[A_i] \qquad \textit{and} \qquad T_{i+1}[A_{i+1}], T_{i+2}[A_{i+2}], \ldots, T_n[A_n]$$

are each partitions of G.

Corollary 3.2. *The free group on two generators is 4–paradoxical.*

Using Proposition 2.28, we may transfer the paradoxical decomposition of a group G onto a G–set X, getting a G–paradoxical decomposition of X. To describe this effect fully, we need the following definition.

Definition 3.3. *Suppose that a group G acts on a space $\mathbb{X}$. We say that the set $X \subseteq \mathbb{X}$ has a G–paradoxical decomposition if there exist*

- *a partition $A_1, A_2, \ldots, A_i, A_{i+1}, \ldots, A_n$ of X, and*
- *elements $g_1, g_2, \ldots, g_i, g_{i+1}, \ldots, g_n$ of G*

such that for some $i < n$,

- *$g_1[A_1], g_2[A_2], \ldots, g_i[A_i]$ is a partition of X, and*
- *$g_{i+1}[A_{i+1}], g_{i+2}[A_{i+2}], \ldots, g_n[A_n]$ is a partition of X.*

Recall $gx = T_g(x)$ so $g[A] = T_g[A]$. Also, if G is understood to be our group acting on $\mathbb{X}$, we may simply say X is paradoxical.

Theorem 3.4. *If G is paradoxical and operates on a set X in such a way that for all $x \in X$, $G_{stab}(x) = \{e_G\}$, then X has a G–paradoxical decomposition.*

Proof. Let G be paradoxical. Then we have a partition

$$B_1, B_2, \ldots, B_i, B_{i+1}, \ldots, B_n$$

of G and elements $g_1, g_2, \ldots, g_i, g_{i+1}, \ldots, g_n$ of G such that both

$$g_1[B_1], g_2[B_2], \ldots, g_i[B_i] \qquad \text{and} \qquad g_{i+1}[B_{i+1}], g_{i+2}[B_{i+2}], \ldots, g_n[B_n]$$

are partitions of G. Also, assume for any $x \in X$ that $G_{stab}(x) = \{e_G\}$. Then $gx = hx$ only if $g = h$ by Proposition 2.25.

Let C be a selector from the orbits of X, that is $C \subseteq X$ and $|C \cap orb(x)| = 1$ for all $x \in X$. For any $k \in \{1, 2, \ldots, n\}$ and $x \in C$, define

$$A_k^x = \{\pi_x(g) : g \in B_k\}$$

(see 2.26 for the definition of π_x). Also define

$$A_k = \bigcup \{A_k^x : x \in C\}.$$

Claim 3.4.1. *For each $x \in C$, $A_1^x, A_2^x, \ldots, A_n^x$ is a partition of $orb(x)$.*

Proof of the Claim. Immediate by Proposition 2.28. □

Claim 3.4.2. *$A_1, A_2, \ldots, A_n$ is a partition of X.*

Proof of the Claim. First, we will show $A_1 \cup A_2 \cup \ldots \cup A_n = X$. Let $y \in X$. Then $y \in orb(x)$ for some $x \in C$, as $\{orb(x) : x \in C\}$ is a partition of X (as C was a selector of all the orbits). But $A_1^x, A_2^x, \ldots, A_n^x$ is a partition of $orb(x)$ (by Claim 3.4.1), so $y \in A_1^x \cup A_2^x \cup \ldots \cup A_n^x$. Then for some $i \in \{1, 2, \ldots, n\}$, $y \in A_i^x \subseteq A_i \subseteq A_1 \cup A_2 \cup \ldots \cup A_n$. The other inclusion follows immediately.

Now we want to show $A_j \cap A_l = \emptyset$ for $j \neq l$. Suppose towards contradiction that $y \in A_j \cap A_l$ and $j \neq l$. Then for some $x_1, x_2 \in C$, $y \in A_j^{x_1} \cap A_l^{x_2}$. But then $y \in orb(x_1) \cap orb(x_2)$, so $orb(x_1) = orb(x_2)$. Thus $x_1 = x_2$ (by the definition of C). Then $y \in A_j^{x_1} \cap A_l^{x_1}$, so for some $g_1 \in B_j$ and $g_2 \in B_l$, $y = g_1 x_1 = g_2 x_2$. By Proposition 2.25, $g_1 = g_2$, so $B_j \cap B_l \neq \emptyset$, a contradiction.

Thus, $A_1, A_2, \ldots, A_n$ is a partition of X. □

Claim 3.4.3. *Let $x \in C$. Then $g_1[A_1^x], g_2[A_2^x], \ldots, g_i[A_i^x]$ is a partition of $orb(x)$.*

Proof of the Claim. First note, for any $k \in \{1, 2, \ldots, i\}$,

$$g_k[A_k^x] = \{g_k g x : g \in B_k\} = \{g' x : g' \in g_k[B_k]\}.$$

So

$$\begin{aligned} & g_1[A_1^x] \cup g_2[A_2^x] \cup \ldots \cup g_i[A_i^x] \\ & \quad = \{g'x : g' \in g_1[B_1]\} \cup \{g'x : g' \in g_2[B_2]\} \cup \ldots \cup \{g'x : g' \in g_i[B_i]\} \\ & \quad = \{g'x : g' \in g_1[B_1] \cup g_2[B_2] \cup \ldots \cup g_i[B_i]\} \\ & \quad = \{g'x : g' \in G\} = orb(x). \end{aligned}$$

Also, if $y \in g_k[A_k^x] \cap g_l[A_l^x]$, then $y = g_1 x = g_2 x$ for some $g_1 \in g_k[B_k]$ and $g_2 \in g_l[B_l]$. By Proposition 2.25, $g_1 = g_2$, so $g_k[A_k] \cap g_l[A_l] \neq \emptyset$. Hence, $k = l$. □

Claim 3.4.4. *Let $x \in C$. Then $g_{i+1}[A_{i+1}^x], g_{i+2}[A_{i+2}^x], \ldots, g_n[A_n^x]$ is a partition of $orb(x)$.*

Proof of the Claim. Similar to Claim 3.4.3. □

Claim 3.4.5. *$g_1[A_1], g_2[A_2], \ldots, g_i[A_i]$ is a partition of X.*

Proof of the Claim. First, we will show

$$g_1[A_1] \cup g_2[A_2] \cup \ldots \cup g_i[A_i] = X.$$

Let $y \in X$. Then $y \in orb(x)$ for some $x \in C$. As $g_1[A_1^x], g_2[A_2^x], \ldots, g_i[A_i^x]$ is a partition of $orb(x)$, $y \in g_k[A_k^x]$ for some $k \in \{1, 2, \ldots, i\}$. But

$$g_k[A_k^x] \subseteq g_k[A_k] \subseteq g_1[A_1] \cup g_2[A_2] \cup \ldots \cup g_i[A_i],$$

so $y \in g_1[A_1] \cup g_2[A_2] \cup \ldots \cup g_i[A_i]$. The other inclusion follows immediately.

Now we want to show $g_j[A_j] \cap g_l[A_l] = \emptyset$ for $j, l \in \{1, 2, \ldots, i\}$ and $j \neq l$. Suppose towards contradiction that $y \in g_j[A_j] \cap g_l[A_l]$, for $j \neq l$. Also, $y \in orb(x)$ for some $x \in C$. Thus,

$$y \in g_j[A_j] \cap orb(x) \cap g_l[A_l].$$

Note for any k,

$$g_k[A_k] \cap orb(x) = g_k[A_k] \cap g_k[orb(x)] = g_k[A_k \cap orb(x)] \subseteq g_k[A_k^x].$$

Thus

$$y \in g_j[A_j^x] \cap g_l[A_l^x],$$

so $g_j[A_j^x] \cap g_l[A_l^x] \neq \emptyset$, contradicting Claim 3.4.3.

Thus, $g_1[A_1], g_2[A_2], \ldots, g_i[A_i]$ is a partition of X. □

Claim 3.4.6. *$g_{i+1}[A_{i+1}], g_{i+2}[A_{i+2}], \ldots, g_n[A_n]$ is a partition of X.*

Proof of the Claim. Similar to Claim 3.4.5. □

By Claims 3.4.2, 3.4.5, and 3.4.6, X is G-paradoxical.

□

4. A Free Group of Rotations

We now will construct a free group of rotations in $\mathbb{R}^3$. Let θ and ρ be rotations by an angle of $\arccos \frac{1}{3}$ about the z and x axises respectively. Then the linear transformations $\theta, \theta^{-1}, \rho, \rho^{-1}$ have the following representation in the standard basis.

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

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

We should also note that while we have chosen these angles to work with, any angles chosen that are irrational with respect to the circle would form a free group of rotations. Every composition of transformations $\theta, \theta^{-1}, \rho, \rho^{-1}$ can be thought of as a word in $W(\{\theta, \rho\})$. We will identify the word in $W(\{\theta, \rho\})$ by the corresponding rotations.

Lemma 4.1. *Let w be a non-empty reduced word that ends with the letter θ or θ^{-1}. Then $w(1,0,0) \neq (1,0,0)$.*

Proof. We will argue by induction on the length, n, of w that

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

for some $a, b, c \in \mathbb{Z}$ and that 3 does not divide b. Once that is shown, this will imply $w(1,0,0) \neq (1,0,0)$ for every non-empty word, as 0 is divisible by 3.

For $n = 1$, observe the following

$$\theta(1,0,0) = \tfrac{1}{3}(1, 2\sqrt{2}, 0)$$
$$\theta^{-1}(1,0,0) = \tfrac{1}{3}(1, -2\sqrt{2}, 0).$$

For $n = 2$, observe the following:

$$\begin{aligned}
\theta\theta(1,0,0) &= \tfrac{1}{9}(-7, 4\sqrt{2}, 0)\\
\rho\theta(1,0,0) &= \tfrac{1}{9}(3, 2\sqrt{2}, 8)\\
\rho^{-1}\theta(1,0,0) &= \tfrac{1}{9}(3, 2\sqrt{2}, -8)\\
\theta^{-1}\theta^{-1}(1,0,0) &= \tfrac{1}{9}(-7, -4\sqrt{2}, 0)\\
\rho\theta^{-1}(1,0,0) &= \tfrac{1}{9}(3, -2\sqrt{2}, -8)\\
\rho^{-1}\theta^{-1}(1,0,0) &= \tfrac{1}{9}(3, -2\sqrt{2}, 8).
\end{aligned}$$

Now suppose

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

for every reduced word w of length $k \leq n-1$ that ends with the letter θ or θ^{-1}. We will keep the convention that for $k = n-1$,

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

for $k = n-2$,

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

etc.

We want to show that for word w of length n, $w(1,0,0)$ is of the form

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

where $a, b, c \in \mathbb{Z}$. To show this, we consider 4 cases:

(A1) $w = \theta w'$, where w' is a reduced word of length $n-1$ ending with the letter θ or θ^{-1}. By the inductive hypothesis,

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

where $a', b', c' \in \mathbb{Z}$. Then

$$w(1,0,0) = \frac{1}{3^n}(a' - 4b', (b' + 2a')\sqrt{2}, 3c').$$

Then as $a', b', c' \in \mathbb{Z}$, so are $(a' - 4b')$, $(b' + 2a')$, and $(3c')$.

(A2) $w = \theta^{-1}w'$. Then

$$w(1,0,0) = \frac{1}{3^n}(a' + 4b', (b' - 2a')\sqrt{2}, 3c').$$

Then as $a', b', c' \in \mathbb{Z}$, so are $(a' + 4b'), (b' - 2a'),$ and $(3c')$.

(A3) $w = \rho w'$. Then

$$w(1,0,0) = \frac{1}{3^n}(3a', (b' - 2a')\sqrt{2}, c' + 4b').$$

Then as $a', b', c' \in \mathbb{Z}$, so are $(3a'), (b' - 2a'),$ and $(c' + 4b')$.

(A4) $w = \rho^{-1}w'$. Then

$$w(1,0,0) = \frac{1}{3^n}(3a', (b' + 2a')\sqrt{2}, c' - 4b').$$

Then as $a', b', c' \in \mathbb{Z}$, so are $(3a'), (b' + 2a'),$ and $(c' - 4b')$.

In any case now, $w(1,0,0)$ is of the form $\frac{1}{3^n}(a, b\sqrt{2}, c)$ where $a, b, c \in \mathbb{Z}$.

Now to argue that 3 does not divide b, we must consider 12 cases:

(B1) $w = \theta\rho w''$ where w'' is a reduced word of length $n - 2$ ending with the letter θ or θ^{-1}. By the inductive hypothesis,

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

where $a'', b'', c'' \in \mathbb{Z}$. Then

$$\begin{aligned} w(1,0,0) = \theta\rho w''(1,0,0) &= \theta\tfrac{1}{3^{n-1}}(3a'', (b'' - 2c'')\sqrt{2}, c'' + 4b'') \\ &= \tfrac{1}{3^n}(x_1, [(b'' - 2c'') + 6a'']\sqrt{2}, z_1). \end{aligned}$$

Note that $\rho w''$ is a reduced word of length $n - 1$, so by the inductive hypothesis, $b'' - 2c'' = b'$, and 3 does not divide b'. So

$$b = (b'' - 2c'') + 6a'' = b' + 6a'',$$

and as 3 divides $6a''$ but does not divide b', 3 does not divide $b = b' + 6a''$.

(B2) $w = \rho\rho w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_2, (-7b'' - 4c'')\sqrt{2}, z_2) = \frac{1}{3^n}(x_2, (2b' - 9b'')\sqrt{2}, z_2).$$

So $b = 2b' - 9b''$, and as 3 divides $-9b''$ but does not divide $2b'$, 3 does not divide $b = 2b' - 9b''$.

(B3) $w = \theta^{-1}\rho w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_3, (b' - 6a'')\sqrt{2}, z_3).$$

So $b = b' - 6a''$, and as 3 divides $6a''$ but does not divide b', 3 does not divide $b = b' - 6a''$.

(B4) $w = \theta\rho^{-1}w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_4, (b' + 6a'')\sqrt{2}, z_4).$$

So $b = b' + 6a''$, and as 3 divides $6a''$ but does not divide b', 3 does not divide $b = b' + 6a''$.

(B5) $w = \theta^{-1}\rho^{-1}w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_5, (b' - 6a'')\sqrt{2}, z_5).$$

So $b = b' - 6a''$, and as 3 divides $6a''$ but does not divide b', 3 does not divide $b = b' - 6a''$.

(B6) $w = \rho^{-1}\rho^{-1}w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_6, (-7b'' - 4c'')\sqrt{2}, z_6) = \frac{1}{3^n}(x_6, (2b' - 9b'')\sqrt{2}, z_6).$$

So $b = 2b' - 9b''$, and as 3 divides $-9b''$ but does not divide $2b'$, 3 does not divide $b = 2b' - 9b''$.

(B7) $w = \theta\theta w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_7, (-7b'' + 4a'')\sqrt{2}, z_7) = \frac{1}{3^n}(x_7, (2b' - 9b'')\sqrt{2}, z_7).$$

So $b = 2b' - 9b''$, and as 3 divides $-9b''$ but does not divide $2b'$, 3 does not divide $b = 2b' - 9b''$.

(B8) $w = \rho\theta w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_8, (b' - 6c'')\sqrt{2}, z_8).$$

So $b = b' - 6c''$, and as 3 divides $6c''$ but does not divide b', 3 does not divide $b = b' - 6c''$.

(B9) $w = \rho^{-1}\theta^{-1}w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_9, (b' + 6c'')\sqrt{2}, z_9).$$

So $b = b' + 6c''$, and as 3 divides $6c''$ but does not divide b', 3 does not divide $b = b' + 6c''$.

(B10) $w = \rho\theta^{-1}w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_{10}, (b' - 6c'')\sqrt{2}, z_{10}).$$

So $b = b' - 6c''$, and as 3 divides $6c''$ but does not divide b', 3 does not divide $b = b' - 6c''$.

(B11) $w = \theta^{-1}\theta^{-1}w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_{11}, (-7b'' - 4a'')\sqrt{2}, z_{11}) = \frac{1}{3^n}(x_{11}, (2b' - 9b'')\sqrt{2}, z_{11}).$$

So $b = 2b' - 9b''$, and as 3 divides $-9b''$ but does not divide $2b'$, 3 does not divide $b = 2b' - 9b''$.

(B12) $w = \rho^{-1}\rho^{-1}w''$. Then

$$w(1,0,0) = \frac{1}{3^n}(x_{12}, (-7b'' + 4c'')\sqrt{2}, z_{12}) = \frac{1}{3^n}(x_{12}, (2b' - 9b'')\sqrt{2}, z_{12})$$

So $b = b' + 6c''$, and as 3 divides $-9b''$ but does not divide $2b'$, 3 does not divide $b = 2b' - 9b''$.

In any case now, 3 does not divide b. □

Lemma 4.2. *Let w be a non-empty reduced word that ends with the letter ρ or ρ^{-1}. Then $w(0,0,1) \neq (0,0,1)$.*

Proof. A very similar argument will work for all non-empty reduced words w that end with the letter ρ or ρ^{-1}. We will argue by induction on the length, n, of w that

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

for some $a, b, c \in \mathbb{Z}$ and that 3 does not divide b. Once that is shown, this will imply $w(0,0,1) \neq (0,0,1)$ for every non-empty word, as 0 is divisible by 3.
For $n = 1$, observe the following

$$\rho(0,0,1) = \tfrac{1}{3}(0, -2\sqrt{2}, 1)$$
$$\rho^{-1}(0,0,1) = \tfrac{1}{3}(0, 2\sqrt{2}, 1).$$

For $n = 2$, observe the following:

$$\theta\rho(0,0,1) = \tfrac{1}{9}(8, -2\sqrt{2}, 3)$$
$$\rho\rho(0,0,1) = \tfrac{1}{9}(0, -4\sqrt{2}, -7)$$
$$\theta^{-1}\rho(0,0,1) = \tfrac{1}{9}(-8, -2\sqrt{2}, 3)$$
$$\theta\rho^{-1}(0,0,1) = \tfrac{1}{9}(-8, 2\sqrt{2}, 3)$$
$$\rho^{-1}\rho^{-1}(0,0,1) = \tfrac{1}{9}(0, 4\sqrt{2}, -7)$$
$$\theta^{-1}\rho^{-1}(0,0,1) = \tfrac{1}{9}(8, 2\sqrt{2}, 3).$$

The inductive step of this proof is the same as Lemma 4.1, replacing $w''(1,0,0)$ with $w''(0,0,1)$. Thus, suppose

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

for some $a'', b'', c'' \in \mathbb{Z}$ and that 3 does not divide b'' for all non-empty reduced words w'' of length $n-2$ that end with the letter ρ or ρ^{-1}, and

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

for some $a', b', c' \in \mathbb{Z}$ and that 3 does not divide b' for all non-empty reduced words w' of length $n-1$ that end with the letter ρ or ρ^{-1}.

We want to show that $w(0,0,1)$ is of the form

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

where $a, b, c \in \mathbb{Z}$. Note this is shown in cases (A1 - A4) in 4.1.

Also, 3 does not divide b by cases (B1 - B12) in 4.1. □

Definition 4.3. *Let G^* be the subgroup of the group of rotations of the sphere $S_1 = \{(x,y,z) \in R^3 : x^2 + y^2 + z^2 = 1\}$ generated by $\{\theta, \rho\}$.*

Theorem 4.4. *G^* is isomorphic to $fr(\{a,b\})$.*

Proof. Define $f : \{a, a^{-1}, b, b^{-1}\} \longrightarrow G^*$ by $f(a) = \theta$, $f(a^{-1}) = \theta^{-1}$, $f(b) = \rho$, and $f(b^{-1}) = \rho^{-1}$. By Theorem 2.19, there exists a homomorphism $\hat{f} : fr(\{a,b\}) \longrightarrow G^*$ such that

$$\hat{f}([x_1 x_2 \ldots x_n]) = f(x_1) f(x_2) \ldots f(x_n).$$

We claim $\hat{f}$ is an isomorphism. We will first show $\hat{f}$ is onto G^*. Note that $\theta, \rho \in rng(\hat{f}) = \{\hat{f}([w]) : [w] \in fr(\{a,b\})\}$. As $rng(\hat{f})$ is a subgroup of G^* (as $\hat{f}$ is a homomorphism), $rng(\hat{f}) = G^*$ (as G^* is the group generated by $\{\theta, \rho\}$). Next we show $\hat{f}$ is injective. We will examine the kernel of $\hat{f}$,

$$ker(\hat{f}) = \{[w] : \hat{f}([w]) = e_{G^*}\}.$$

By Lemma 4.1, if w is a reduced word that ends with a or a^{-1}, $w(1,0,0) \neq (1,0,0)$ so $[w]$ is not the identity and hence, $[w] \notin ker(\hat{f})$. By Lemma 4.2, if w is a reduced word that ends with b or b^{-1}, $w(0,0,1) \neq (0,0,1)$ so again $[w]$ is not the identity and therefore, $[w] \notin ker(\hat{f})$. Hence, $ker(\hat{f}) = \{e_{G^*}\}$. As the kernel of the homomorphism $\hat{f}$ is trivial, $\hat{f}$ is injective. □

5. Hausdorff and Banach-Tarski Paradoxes

5.1. **The Hausdorff Paradox.** Let $S_1 = \{(x, y, z) \in R^3 : x^2 + y^2 + z^2 = 1\}$ and let G^* be our free group of rotations defined in Definition 4.4. Our goal is to choose $X_1 \subseteq S_1$ to be a G^*-set under the natural action in such a way that

$$G^*_{stab}(x) = \{e_{G^*}\} \text{ for all } x \in X_1. \qquad (\oplus)$$

Note that $\theta((0, 0, 1)) = (0, 0, 1)$, so $\theta \in G^*_{stab}((0, 0, 1)) \neq \{e_{G^*}\}$ and thus we cannot set $X_1 = S_1$. In fact, for any rotation $g \in G^*$, both of the poles, $p_{g,1}$ and $p_{g,2}$, will have non-trivial stabilizers. Then let $P_1 = \{p_{g,1}, p_{g,2} : g \in G^*\}$. Thus, our demand $\oplus$ requires $X_1 \subseteq S_1 \setminus P_1$.

As we have taken points out of S_1, we must be concerned that G^* operating on X_1 is well-defined. To ensure $gx \in X_1$ we may simply set $\bar{P}_1 = \{q \in orb(p) : p \in P_1\}$ and then define $X_1 = S_1 \setminus \bar{P}_1$. Let us argue that for any $x \in X_1$ and $g \in G^*$, $gx \in X_1$. Plainly for any $s \in S_1$, $gs \in S_1$. Therefore as $x \in S_1$, $gx \in S_1$. As $x \notin orb(p)$ for any $p \in P_1$, $orb(x) \cap orb(p) = \emptyset$, and thus, $gx \notin orb(p)$ for any $p \in P_1$. Thus $gx \in X_1$. This proves that the action of G^* on X_1 is well-defined.

Let us show that X_1 is not something trivial. First note that S_1 is uncountable. Now, we will argue that $\bar{P}_1$ is countable. Define p_i to be the i-th prime number, so $p_1 = 2$, $p_2 = 3$, etc. Define $\alpha : \{\theta, \theta^{-1}, \rho, \rho^{-1}\} \longrightarrow \mathbb{N}$ by

$$\alpha(a) = \begin{cases} 1 & \text{if } a = \theta \\ 2 & \text{if } a = \theta^{-1} \\ 3 & \text{if } a = \rho \\ 4 & \text{if } a = \rho^{-1}. \end{cases}$$

Now define $f : G^* \longrightarrow \mathbb{N}$ by

$$f(g) = \begin{cases} 1 & \text{if } g = e_{G^*} \\ p_1^{\alpha(a_1)} p_2^{\alpha(a_2)} p_3^{\alpha(a_3)} \dots p_n^{\alpha(a_n)} & \text{if } g = a_1 a_2 a_3 \dots a_n \text{ is a reduced word and} \\ & \quad a_i \in \{\theta, \theta^{-1}, \rho, \rho^{-1}\}. \end{cases}$$

This defines an injective function f, so G^* is equinumerous with a subset of the natural numbers, so G^* is countable.

As $P_1 = \{p_{g,1}, p_{g,2} : g \in G^*\} = \{p_{g,1} : g \in G^*\} \cup \{p_{g,2} : g \in G^*\}$, P_1 is a union of two countable sets, P_1 is countable.

Finally,

$$\bar{P}_1 = \{q \in orb(p) : p \in P_1\} = \bigcup_{p \in \bar{P}_1} orb(p) = \bigcup_{p \in \bar{P}_1} \{gp : g \in G^*\}$$

is a countable union of countable sets, so $\bar{P}_1$ is countable. Therefore, $X_1 = S_1 \setminus \bar{P}_1$ is uncountable, as we have simply removed a countable number of points from an uncountable set.

To summarize, we have formed X_1 which is almost the entire sphere such that G^* acts on X_1 and each $x \in X_1$ has a trivial stabilizer. Therefore, any non-trivial rotation moves every point of X_1 and any partition of our group G^* creates a partition of X_1.

Corollary 5.1 (Felix Hausdorff, [5]). *X_1 has a G^* paradoxical decomposition into 4 pieces.*

Proof. By Corollary 3.2, Theorem 4.4, and Theorem 3.4. □

5.2. **A Basis for the Paradoxical Decomposition of the Sphere.** While it is remarkable in itself that the surface of the sphere, minus a countable number of points, $X_1 = S_1 \setminus \bar{P}_1$, is paradoxical, we really want to show that the entire surface of the sphere S_1 is paradoxical. Recall how X_1 was paradoxically decomposed. We identify the group of rotations generated by $\{\theta, \rho\}$ with $F = fr\{a, b\}$. Then $A_1, A_2, B_1, B_2 \subseteq$

G^* were defined on page 23. This creates a partition X_1 into $T_1^{A_1}, T_1^{A_2}, T_1^{B_1}, T_1^{B_2}$ where

$$T_1^{A_1} = \{gx : x \in C_1, g \in A_1\},$$
$$T_1^{A_2} = \{gx : x \in C_1, g \in A_2\},$$
$$T_1^{B_1} = \{gx : x \in C_1, g \in B_1\},$$
$$T_1^{B_2} = \{gx : x \in C_1, g \in B_2\},$$

and C_1 is a selector from all of the orbits in X_1.

Thus, we have a partition of S_1 into $T_1^{A_1}, T_1^{A_2}, T_1^{B_1}, T_1^{B_2}, \bar{P}_1$. Then

$$\theta^{-1}[T_1^{A_1}], T_1^{A_2}, \bar{P}_1$$

is a partition of S_1 and

$$\rho^{-1}[T_1^{B_1}], T_1^{B_2}$$

is a partition of $S_1 \setminus \bar{P}_1$. Now, we only must fill in $\bar{P}_1$ of the second copy of S_1.

As $\bar{P}_1$ is countable and $S_1 \setminus \bar{P}_1$ is uncountable, we may find an axis of rotation through the origin such that its poles on the sphere, q_1, q_2, are elements of $S_1 \setminus \bar{P}_1$. We want to choose a rotation σ about this axis so that no $p \in \bar{P}_1$ gets moved back to itself or to another point $p' \in \bar{P}_1$ no matter how many times σ is applied to it. As $\bar{P}_1$ is countable, there is only a countable number of points σ must avoid, and there are only a countable number of ways to reach each point, $\sigma^n(p) = p'$ where $n \in \mathbb{N}$. As the number of rotations is uncountable, we may choose a desired rotation σ. Now let

$$P_1^* = \{x \in X_1 : (\exists n \in \mathbb{N}, p \in \bar{P}_1)(\sigma^n(p) = x)\} \subseteq S_1 \setminus \bar{P}_1.$$

Then

$$\sigma^{-1}[P_1^*] = P_1^* \cup \bar{P}_1 \stackrel{\text{def}}{=} \hat{P}_1.$$

Finally, $\rho^{-1}[T_1^{B_1}], T_1^{B_2}$ is a partition of $S_1 \setminus \bar{P}_1 = (S_1 \setminus \hat{P}_1) \cup P_1^*$. Then

$$(S_1 \setminus \hat{P}_1) \cup \sigma^{-1}[P_1^*] = (S_1 \setminus \hat{P}_1) \cup \hat{P}_1 = S_1.$$

This gives us our second copy of S_1 and will give us the basis for our paradoxical decomposition the entire surface of the sphere.

5.3. **A Decomposition of the Sphere.** To specify the six pieces used for this decomposition of the sphere, we will use the following notation. For a positive number r, let

$$S_r = \{(x, y, z) : \sqrt{x^2 + y^2 + z^2} = r\}$$
$$P_r = \{p : p \text{ is a pole of } S_r \text{ for some rotation } g \in G^*\}$$
$$\bar{P}_r = \{q \in orb(p) : p \in P_r\}$$
$$X_r = S_r \setminus \bar{P}_r$$
$$P^*_r = \{x \in X_r : (\exists n \in \mathbb{N}, p \in \bar{P}_r)(\sigma^n(p) = x)\}$$
$$\hat{P}_r = P^*_r \cup \bar{P}_r$$
C_r is a selector from the orbits of X_r
$$T_r^{A_1} = \{gx : x \in C_r, g \in A_1\}$$
$$T_r^{A_2} = \{gx : x \in C_r, g \in A_2\}$$
$$T_r^{B_1} = \{gx : x \in C_r, g \in B_1\}$$
$$T_r^{B_2} = \{gx : x \in C_r, g \in B_2\}$$
σ is a rotation with poles in X_r such that $(\forall n \in \mathbb{N}, p \in \bar{P}_r)(\sigma^n(p) \notin \bar{P}_r)$.

For now we consider r to be 1.

Our decomposition of S_1, the surface of the sphere of radius 1, is

$$\mathbf{A}_1 = T_1^{A_1}$$
$$\mathbf{A}_2 = T_1^{A_2} \cup \bar{P}_1$$
$$\mathbf{A}_3 = \rho[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P^*_1]$$
$$\mathbf{A}_4 = T_1^{B_1} \setminus \rho[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P^*_1]$$
$$\mathbf{A}_5 = T_1^{B_2} \cap P^*_1$$
$$\mathbf{A}_6 = T_1^{B_2} \setminus P^*_1.$$

Note that $\rho[T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}] \subseteq T_1^{B_1}$ so $\rho[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P_1^*] \subseteq T_1^{B_1}$. The following rotations will witness that $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4, \mathbf{A}_5, \mathbf{A}_6$ is a paradoxical decomposition of S_1:

$$\begin{aligned} g_1 &= \theta^{-1} \\ g_2 &= e_{G^*} \\ g_3 &= \sigma^{-1}\rho^{-1} \\ g_4 &= \rho^{-1} \\ g_5 &= \sigma^{-1} \\ g_6 &= e_{G^*}. \end{aligned}$$

As we have mentioned before,

$$g_1[\mathbf{A}_1] = \theta^{-1}[T_1^{A_1}] = X_1 \setminus T_1^{A_2} \qquad \text{and} \qquad g_2[\mathbf{A}_2] = e_{G^*}[T_1^{A_2} \cup \bar{P}_1] = T_1^{A_2} \cup \bar{P}_1$$

is a partition of S_1.

Next observe for $\mathbf{A}_3$,

$$\begin{aligned} g_3[\mathbf{A}_3] &= \sigma^{-1}\rho^{-1}[\rho[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P_1^*] \\ &= \sigma^{-1}[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P_1^*]. \end{aligned}$$

For $\mathbf{A}_4$,

$$\begin{aligned} g_4[\mathbf{A}_4] &= \rho^{-1}[T_1^{B_1} \setminus \rho[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P_1^*]] \\ &= \rho^{-1}[T_1^{B_1}] \setminus \rho^{-1}[\rho[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P_1^*]] \\ &= (T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \setminus ((T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P_1^*) \\ &= (T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \setminus P_1^*. \end{aligned}$$

For $\mathbf{A}_5$,

$$g_5[\mathbf{A}_5] = \sigma^{-1}[T_1^{B_2} \cap P_1^*].$$

And simply

$$g_6[\mathbf{A}_6] = T_1^{B_2} \setminus P_1^*.$$

Observe

$$\begin{aligned}
g_3[\mathbf{A}_3] \cup g_5[\mathbf{A}_5] &= \sigma^{-1}[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \cap P_1^*] \cup \sigma^{-1}[T_1^{B_2} \cap P_1^*] \\
&= \sigma^{-1}[(T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1} \cup T_1^{B_2}) \cap P_1^*] \\
&= \sigma^{-1}[P_1^*] = \hat{P}_1.
\end{aligned}$$

Clearly, $g_3[\mathbf{A}_3] \cap g_5[\mathbf{A}_5] = \emptyset$.

Also observe

$$\begin{aligned}
g_4[\mathbf{A}_4] \cup g_6[\mathbf{A}_6] &= ((T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1}) \setminus P_1^*) \cup (T_1^{B_2} \setminus P_1^*) \\
&= (T_1^{A_1} \cup T_1^{A_2} \cup T_1^{B_1} \cup T_1^{B_2}) \setminus P_1^* \\
&= X_1 \setminus P_1^*.
\end{aligned}$$

Clearly, $g_4[\mathbf{A}_4] \cap g_6[\mathbf{A}_6] = \emptyset$. Moreover,

$$(g_3[\mathbf{A}_3] \cup g_5[\mathbf{A}_5]) \cap (g_4[\mathbf{A}_4] \cup g_6[\mathbf{A}_6]) = \emptyset,$$

and consequently, $g_3[\mathbf{A}_3]$, $g_4[\mathbf{A}_4]$, $g_5[\mathbf{A}_5]$, $g_6[\mathbf{A}_6]$ are all pairwise disjoint.

Finally, we have

$$g_3[\mathbf{A}_3] \cup g_4[\mathbf{A}_4] \cup g_5[\mathbf{A}_5] \cup g_6[\mathbf{A}_6] = \hat{P}_1 \cup (X_1 \setminus P_1^*) = S_1.$$

Thus $g_3[\mathbf{A}_3]$, $g_4[\mathbf{A}_4]$, $g_5[\mathbf{A}_5]$, $g_6[\mathbf{A}_6]$ is a partition of S_1, and since $g_1[\mathbf{A}_1]$, $g_2[\mathbf{A}_2]$ is a partition of S_1, we have paradoxically decomposed S_1.

5.4. **The Banach-Tarski Paradox.** We can now paradoxically decompose the unit ball

$$B = \{(x, y, z) : \sqrt{x^2 + y^2 + z^2} \leq 1\}.$$

Define

$$\begin{aligned}
X &= \textstyle\bigcup_{r \in (0,1]} X_r \\
P &= \textstyle\bigcup_{r \in (0,1]} P_r \\
\bar{P} &= \textstyle\bigcup_{r \in (0,1]} \bar{P}_r \\
P^* &= \textstyle\bigcup_{r \in (0,1]} P_r^*
\end{aligned}$$

$\hat{P} = \bigcup_{r\in(0,1]} \hat{P}_r$

$C = \bigcup_{r\in(0,1]} C_r$

$T^\alpha = \bigcup_{r\in(0,1]} T^\alpha_r$ (for $\alpha \in \{A_1, A_2, B_1, B_2\}$).

Choose $x_0 \in T_1^{A_1} \subseteq X_1 = S_1 \setminus \bar{P}_1$. In a similar manner that we found a rotation σ, we will argue we may find a rotation γ such that for all $n \in \mathbb{N}$

(1) $\gamma^n(x_0) \in X$,

(2) $\gamma^n(x_0) \neq x_0$.

Fix two poles in $S_1 \setminus \{x_0\}$. As there are only countably many ways to reach a point in $\bar{P}_1$ and only countably many points in $\bar{P}_1$ that we must avoid, there are only a countable number of rotations that do not work for demand (1). For any $n \in \mathbb{N}$, there is at most a finite number of rotations such that $\gamma^n(x_0) = x_0$. Therefore, there are only countably many rotations such that $\gamma^n(x_0) = x_0$ for some $n \in \mathbb{N}$, so there is only a countable number of rotations that do not work for demand (2). But as the number of rotations about the axis through the fixed poles is uncountable, we may choose a desired rotation γ.

Define the set $Y = \{\gamma^n(x_0) : n \in \mathbb{N}\}$, and let the translation τ be such that $\tau(x_0) = (0, 0, 0)$.

Our decomposition of B is

$$\mathbf{B}_1 = \theta[(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \cap Y]$$
$$\mathbf{B}_2 = T^{A_1} \setminus \theta[[(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \cap Y] \cup \{x_0\}]$$
$$\mathbf{B}_3 = T^{A_2} \cap Y$$
$$\mathbf{B}_4 = (T^{A_2} \setminus Y) \cup \bar{P} \cup \{(0, 0, 0)\}$$
$$\mathbf{B}_5 = \theta[\{x_0\}]$$
$$\mathbf{B}_6 = \rho[(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \cap P^*]$$
$$\mathbf{B}_7 = T^{B_1} \setminus \rho[(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \cap P^*]$$
$$\mathbf{B}_8 = T^{B_2} \cap P^*$$
$$\mathbf{B}_9 = T^{B_2} \setminus P^*.$$

Here we note that $\theta[T^{A_1} \cup T^{B_1} \cup T^{B_2}] \subseteq T^{A_1}$. Easily, $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_5$ is a partition of T^{A_1}; $\mathbf{B}_3, \mathbf{B}_4$ is a partition of $T^{A_2} \cup \bar{P} \cup \{(0,0,0)\}$; $\mathbf{B}_6, \mathbf{B}_7$ is a partition of T^{B_1}; and $\mathbf{B}_8, \mathbf{B}_9$ is a partition of T^{B_2}. On the other hand, T^{A_1}, $T^{A_2} \cup \bar{P} \cup \{(0,0,0\}$, T^{B_1}, T^{B_2} is a partition of B. Therefore, $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_9$ is a partition of B as well.

Using the following rotations and translation, we will show that our partition into 9 pieces is a paradoxical decomposition of B:

$$\begin{aligned}
h_1 &= \gamma^{-1}\theta^{-1} \\
h_2 &= \theta^{-1} \\
h_3 &= \gamma^{-1} \\
h_4 &= e_{G^*} \\
h_5 &= \tau\theta^{-1} \\
h_6 &= \sigma^{-1}\rho^{-1} \\
h_7 &= \rho^{-1} \\
h_8 &= \sigma^{-1} \\
h_9 &= e_{G^*}.
\end{aligned}$$

Observe for $\mathbf{B}_1$,

$$\begin{aligned}
h_1[\mathbf{B}_1] &= \gamma^{-1}\theta^{-1}[\theta[(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \cap Y]] \\
&= \gamma^{-1}[(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \cap Y].
\end{aligned}$$

For $\mathbf{B}_2$,

$$\begin{aligned}
h_2[\mathbf{B}_2] &= \theta^{-1}[T^{A_1} \setminus \theta[[(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \cap Y] \cup \{x_0\}]] \\
&= \theta^{-1}[T^{A_1}] \setminus \theta^{-1}[\theta[[(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \cap Y] \cup \{x_0\}]]] \\
&= (T^{A_1} \cup T^{B_1} \cup T^{B_2}) \setminus [(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \cap Y] \cup \{x_0\}] \\
&= (T^{A_1} \cup T^{B_1} \cup T^{B_2}) \setminus (Y \cup \{x_0\}).
\end{aligned}$$

For $\mathbf{B}_3$,

$$h_3[\mathbf{B}_3] = \gamma^{-1}[T^{A_2} \cap Y].$$

For $\mathbf{B}_4$, simply

$$h_4[\mathbf{B}_4] = (T^{A_2} \setminus Y) \cup \bar{P} \cup \{(0,0,0)\}.$$

Observe

$$\begin{aligned} h_1[\mathbf{B}_1] \cup h_3[\mathbf{B}_3] &= \gamma^{-1}[(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \cap Y] \cup \gamma^{-1}[T^{A_2} \cap Y] \\ &= \gamma^{-1}[(T^{A_1} \cup T^{A_2} \cup T^{B_1} \cup T^{B_2}) \cap Y] \\ &= \gamma^{-1}[X \cap Y] = \gamma^{-1}[Y] = Y \cup \{x_0\}. \end{aligned}$$

Clearly, $h_1[\mathbf{B}_1] \cap h_3[\mathbf{B}_3] = \emptyset$.

Also observe

$$\begin{aligned} h_2[\mathbf{B}_2] \cup h_4[\mathbf{B}_4] &= \big[(T^{A_1} \cup T^{B_1} \cup T^{B_2}) \setminus (Y \cup \{x_0\})\big] \cup \big[(T^{A_2} \setminus Y) \cup \bar{P} \cup \{(0,0,0)\}\big] \\ &= \big[(T^{A_1} \cup T^{A_2} \cup T^{B_1} \cup T^{B_2})\big] \setminus \big[(Y \cup \{x_0\}) \cup \bar{P} \cup \{(0,0,0)\}\big] \\ &= [X \setminus (Y \cup \{x_0\})] \cup \bar{P} \cup \{(0,0,0)\}. \end{aligned}$$

Clearly, $h_2[\mathbf{B}_2] \cap h_4[\mathbf{B}_4] = \emptyset$. Moreover,

$$(h_1[\mathbf{B}_1] \cup h_3[\mathbf{B}_3]) \cap (h_2[\mathbf{B}_2] \cup h_4[\mathbf{B}_4]) = \emptyset,$$

and consequently, $h_1[\mathbf{B}_1]$, $h_2[\mathbf{B}_2]$, $h_3[\mathbf{B}_3]$, $h_4[\mathbf{B}_4]$ are all pairwise disjoint.

Finally, we have

$$\begin{aligned} & h_1[\mathbf{B}_1] \cup h_2[\mathbf{B}_2] \cup h_3[\mathbf{B}_3] \cup h_4[\mathbf{B}_4] \\ &\quad = (Y \cup \{x_0\}) \cup [X \setminus (Y \cup \{x_0\}) \cup \bar{P} \cup \{(0,0,0)\}] \\ &\quad = X \cup \bar{P} \cup \{(0,0,0)\} \\ &\quad = B. \end{aligned}$$

Thus $h_1[\mathbf{B}_1]$, $h_2[\mathbf{B}_2]$, $h_3[\mathbf{B}_3]$, $h_4[\mathbf{B}_4]$ is a partition of B.

Now observe for $\mathbf{B}_5$,

$$h_5[\mathbf{B}_5] = \tau\theta^{-1}[\theta[\{x_0\}]] = \tau[\{x_0\}] = \{(0,0,0)\}.$$

For $\mathbf{B}_6$,

$$\begin{aligned} h_6[\mathbf{B}_6] &= \sigma^{-1}[\rho^{-1}[\rho[(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \cap P^*]]] \\ &= \sigma^{-1}[(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \cap P^*]. \end{aligned}$$

For $\mathbf{B}_7$,

$$\begin{aligned} h_7[\mathbf{B}_7] &= \rho^{-1}[T^{B_1} \setminus \rho[(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \cap P^*]] \\ &= \rho^{-1}[T^{B_1}] \setminus \rho^{-1}[\rho[(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \cap P^*]] \\ &= (T^{A_1} \cup T^{A_2} \cup T^{B_1}) \setminus [(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \cap P^*] \\ &= (T^{A_1} \cup T^{A_2} \cup T^{B_1}) \setminus P^*. \end{aligned}$$

For $\mathbf{B}_8$,

$$h_8[\mathbf{B}_8] = \sigma^{-1}[T^{B_2} \cap P^*].$$

And simply

$$h_9[\mathbf{B}_9] = T^{B_2} \setminus P^*.$$

Observe

$$\begin{aligned} h_6[\mathbf{B}_6] \cup h_8[\mathbf{B}_8] &= \sigma^{-1}[(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \cap P^*] \cup \sigma^{-1}[T^{B_2} \cap P^*] \\ &= \sigma^{-1}[(T^{A_1} \cup T^{A_2} \cup T^{B_1} \cup T^{B_2}) \cap P^*] \\ &= \sigma^{-1}[X \cap P^*] = \sigma^{-1}[P^*] = \hat{P}. \end{aligned}$$

Clearly, $h_6[\mathbf{B}_6] \cap h_8[\mathbf{B}_8] = \emptyset$.

Also observe

$$\begin{aligned} h_7[\mathbf{B}_7] \cup h_9[\mathbf{B}_9] &= [(T^{A_1} \cup T^{A_2} \cup T^{B_1}) \setminus P^*] \cup (T^{B_2} \setminus P^*) \\ &= X \setminus P^*. \end{aligned}$$

Clearly, $h_7[\mathbf{B}_7] \cap h_9[\mathbf{B}_9] = \emptyset$. Moreover,

$$(h_6[\mathbf{B}_6] \cup h_8[\mathbf{B}_8]) \cap (h_7[\mathbf{B}_7] \cup h_9[\mathbf{B}_9]) = (P^* \cup \bar{P}) \cap [B \setminus (P^* \cup \bar{P})] = \emptyset,$$

and therefore, as $h_5[\mathbf{B}_5] = \{(0,0,0)\}$, $h_5[\mathbf{B}_5]$, $h_6[\mathbf{B}_6]$, $h_7[\mathbf{B}_7]$, $h_8[\mathbf{B}_8]$, and $h_9[\mathbf{B}_9]$ are pairwise disjoint.

Finally, we have

$$h_5[\mathbf{B}_5] \cup h_6[\mathbf{B}_6] \cup h_7[\mathbf{B}_7] \cup h_8[\mathbf{B}_8] \cup h_9[\mathbf{B}_9] = \{(0,0,0)\} \cup \hat{P} \cup (X \setminus P^*) = B.$$

Thus $h_5[\mathbf{B}_5]$, $h_6[\mathbf{B}_6]$, $h_7[\mathbf{B}_7]$, $h_8[\mathbf{B}_8]$, $h_9[\mathbf{B}_9]$ is a partition of B, and since $h_1[\mathbf{B}_1]$, $h_2[\mathbf{B}_2]$, $h_3[\mathbf{B}_3]$, $h_4[\mathbf{B}_4]$ is a partition of B, we have paradoxically decomposed B.

6. Interlude: Measure and Integration

In order to explain why there is no parallel of the Banach-Tarski paradox on the plane, we need to study measures invariant under isometries. To introduce and develop the concept, we need some background knowledge of measures and integration. We will develop the theory of integration for finitely additive measures. For a more standard approach, see [4] or [15].

Definition 6.1. (1) *A ring of sets $\mathcal{R}$ is a class of subsets of a space $\mathbb{X}$ such that $\emptyset \in \mathcal{R}$ and for any $X, Y \in \mathcal{R}$,*

$$X \cup Y \in \mathcal{R} \text{ and } X \setminus Y \in \mathcal{R}.$$

(2) *An algebra of sets $\mathcal{A}$ is a ring of subsets of a space $\mathbb{X}$ such that $\mathbb{X} \in \mathcal{A}$.*

(3) *A σ-ring (or σ-algebra) $\mathcal{S}$ is a ring (respectively, algebra) of sets such that if $X_i \in \mathcal{S}$ for all $i \in \mathbb{N}$, then*

$$\bigcup_{i=1}^{\infty} X_i \in \mathcal{S}.$$

Example 6.2. *Let $\mathcal{R}$ be the family of all bounded subsets of $\mathbb{R}^2$. Then $\mathcal{R}$ is a ring but not an algebra. Also, $\mathcal{R}$ is not a σ-ring.*

Definition 6.3. *Let $\mathcal{A}$ be an algebra of subsets of $\mathbb{X}$ and let $C \in \mathcal{A}$. Then $\mathcal{A}(C)$ defined by*

$$\mathcal{A}(C) = \{X \in \mathcal{A} : X \subseteq C\}.$$

One notices that $\mathcal{A}(C)$ is an algebra of subsets of C and is called the algebra induced by C.

Definition 6.4. *A finitely additive measure is a function $\mu : \mathcal{R} \longrightarrow [0, \infty]$ such that*

- *$\mathcal{R}$ is a ring*
- $\mu(\emptyset) = 0$

- *if* $X_i \in \mathcal{R}$ *for all* $i \in \{1, 2, \ldots, n\}$ *and* $X_i \cap X_j = \emptyset$ *for* $j \neq k$*, then*

$$\mu(\bigcup_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mu(X_i)$$

A σ*-additive (or countably additive) measure is a finitely additive measure* $\mu : \mathcal{R} \longrightarrow [0, \infty]$ *such that* $\mathcal{R}$ *is a* σ*-ring and if* $X_i \in \mathcal{A}$ *for all* $i \in \mathbb{N}$ *and* $X_i \cap X_j = \emptyset$ *for* $j \neq k$*, then*

$$\mu(\bigcup_{i=1}^{\infty} X_i) = \sum_{i=1}^{\infty} \mu(X_i).$$

Definition 6.5. *A finitely subadditive outer measure is a function* $\mu^* : P(\mathbb{X}) \longrightarrow [0, \infty]$ *defined on all subsets of a set* $\mathbb{X}$ *such that*

- $\mu^*(\emptyset) = 0$
- *for any* $X, Y \in P(\mathbb{X})$ *where* $X \subseteq Y$*,* $\mu^*(X) \leq \mu^*(Y)$
- *for any* $X_i \in P(\mathbb{X})$*,* $\bigcup_{i=1}^{n} \mu^*(X_i) \leq \sum_{i=1}^{n} \mu^*(X_i)$

A σ*-subadditive outer measure is a finitely subadditive* $\mu^* : P(\mathbb{X}) \longrightarrow [0, \infty]$ *such that for any* $X_i \in P(\mathbb{X})$*,*

$$\bigcup_{i=1}^{\infty} \mu(X_i) \leq \sum_{i=1}^{\infty} \mu^*(X_i).$$

For our purposes, we are more concerned with finitely additive measures (see Proposition 7.6), so when we say μ is a measure, we mean μ is a finitely additive measure. However, unless otherwise specified, the same definitions could be stated with a σ-additive measure as well.

Definition 6.6. (1) *A measure space* $(\mathbb{X}, \mathcal{R}, \mu)$ *is a tuple such that* $\mathbb{X}$ *is a set,* $\mathcal{R}$ *is a ring of subsets of* $\mathbb{X}$*, and* μ *is a measure on* $\mathcal{R}$*.*

(2) *A subset* $X \subseteq \mathbb{X}$ *is called measurable if* $X \in \mathcal{R}$*.*

(3) *Define* $B(\mathbb{X})$ *to be the class of all real-valued bounded functions* $f : \mathbb{X} \longrightarrow \mathbb{R}$*.*

(4) *A function* $f \in B(\mathbb{X})$ *is called measurable if* $f^{-1}[I]$ *is measurable for any open interval* $I \subseteq \mathbb{R}$*.*

Typically, integration theory is developed for σ-additive measures. When dealing with finitely additive measures, one of the difficulties one encounters is dealing with the measurability of some sets. We will ameliorate this difficulty by working with measures that measure any set. Having found no good source covering integration with finitely additive measures, we will present the details fully.

> *From now on in this section, we will assume $(\mathbb{X}, P(\mathbb{X}), \mu)$ is our measure space, where μ is a finitely additive measure of total measure 1. In this setting, every function $f : \mathbb{X} \longrightarrow \mathbb{R}$ will be measurable.*

Definition 6.7. (1) *For any subset $X \subseteq \mathbb{X}$, the function $\chi_X : \mathbb{X} \longrightarrow \{0,1\}$ defined by*

$$\chi_X(x) = \begin{cases} 0 & \text{if } x \notin X \\ 1 & \text{if } x \in X \end{cases}$$

is called the characteristic function of X.

(2) *A simple function is a function $f : \mathbb{X} \longrightarrow \mathbb{R}$ defined by*

$$f(x) = \sum_{i=1}^{n} \alpha_i \chi_{X_i}(x),$$

where $\alpha_i \in \mathbb{R}$ and $X_1, X_2, \ldots, X_n$ is a partition of $\mathbb{X}$.

(3) *The integral of a simple function $f = \sum_{i=1}^{n} \alpha_i \chi_{X_i}$ is defined by*

$$\int f(x)\, \mathrm{d}\mu(x) = \sum_{i=1}^{n} \alpha_i \mu(X_i).$$

Note if f is a simple function, then $f \in B(\mathbb{X})$.

We now prove some basic facts about integrals of simple functions, which we will later show are also true for all bounded functions in general.

Lemma 6.8. *Let $f, g \in B(\mathbb{X})$ be simple functions and $\alpha, \beta \in \mathbb{R}$. Then $\alpha f + \beta g$ is a simple function and*

$$\int (\alpha f + \beta g)(x)\, \mathrm{d}\mu(x) = \alpha \int f(x)\, \mathrm{d}\mu(x) + \beta \int g(x)\, \mathrm{d}\mu(x).$$

Also, $|f|$ is a simple function and $|\int f(x)\,\mathrm{d}\mu(x)| \le \int |f(x)|\,\mathrm{d}\mu(x)$. Hence, if $f(x) \ge 0$, then also $\int f(x)\,\mathrm{d}\mu(x) \ge 0$. Finally, if $f(x) \le g(x)$ for all $x \in \mathbb{X}$, then $\int f(x)\,\mathrm{d}\mu(x) \le \int g(x)\,\mathrm{d}\mu(x)$.

Proof. Let $f, g \in B(\mathbb{X})$ be simple functions, i.e.

$$f(x) = \sum_{i=1}^{n} \alpha_i \chi_{X_i}(x) \qquad \text{and} \qquad g(x) = \sum_{i=1}^{m} \beta_i \chi_{Y_i}(x),$$

where X_1, X_2, …, X_n and Y_1, Y_2, …, Y_m are partitions of $\mathbb{X}$. But then there exists a partition Z_1, Z_2, …, Z_N of $\mathbb{X}$, such that

$$f(x) = \sum_{i=1}^{N} \alpha_i' \chi_{Z_i}(x) \qquad \text{and} \qquad g(x) = \sum_{i=1}^{N} \beta_i' \chi_{Z_i}(x).$$

Now let $\alpha, \beta \in \mathbb{R}$ and observe

$$\begin{aligned}(\alpha f + \beta g)(x) &= \alpha \textstyle\sum_{i=1}^{N} \alpha_i' \chi_{Z_i}(x) + \beta \sum_{i=1}^{N} \beta_i' \chi_{Z_i}(x) \\ &= \textstyle\sum_{i=1}^{N} [(\alpha\alpha_i' + \beta\beta_i')\chi_{Z_i}(x)].\end{aligned}$$

Consequently, $\alpha f + \beta g$ is a simple function and

$$\begin{aligned}\textstyle\int (\alpha f + \beta g)(x)\,\mathrm{d}\mu(x) &= \textstyle\sum_{i=1}^{N} (\alpha\alpha_i' + \beta\beta_i')\mu(Z_i) \\ &= \alpha \textstyle\sum_{i=1}^{N} \alpha_i' \mu(Z_i) + \beta \sum_{i=1}^{N} \beta_i' \mu(Z_i) \\ &= \alpha \textstyle\int f(x)\,\mathrm{d}\mu(x) + \beta \int g(x)\,\mathrm{d}\mu(x).\end{aligned}$$

Concerning $|f|$, note that, as $X_1, X_2, \ldots, X_n$ are pairwise disjoint,

$$|f(x)| = |\sum_{i=1}^{n} \alpha_i \chi_{X_i}(x)| = \sum_{i=1}^{n} |\alpha_i| \chi_{X_i}(x),$$

so $|f|$ is a simple function. Furthermore,

$$|\int f(x)\,\mathrm{d}\mu(x)| = |\sum_{i=1}^{N} \alpha_i \mu(X_i)| \le \sum_{i=1}^{N} |\alpha_i| \mu(X_i) = \int |f(x)|\,\mathrm{d}\mu(x).$$

Finally, for $f(x) \leq g(x)$ for all $x \in \mathbb{X}$,

$$\int g(x)\,\mathrm{d}\mu(x) - \int f(x)\,\mathrm{d}\mu(x) = \int g(x) - f(x)\,\mathrm{d}\mu(x) \geq 0.$$

□

Definition 6.9. (1) *Let f and g be simple functions. We define the function ρ by*

$$\rho(f,g) = \int |f-g| d\mu.$$

(2) *A sequence $\{f_n\}$ of simple functions is said to be fundamental in the mean if*

$$\rho(f_n, f_m) \longrightarrow 0 \text{ as } n, m \longrightarrow \infty.$$

(3) *A sequence $\{f_n\}$ of functions $f_n \in B(\mathbb{X})$ is said to converge in measure to a function $f \in B(\mathbb{X})$ if for every $\epsilon > 0$,*

$$\lim_{n\to\infty} \mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) = 0.$$

(4) *A sequence $\{f_n\}$ of functions is said to be uniformly bounded if for some $M > 0$, we have for every $n \in \mathbb{N}$ and every $x \in \mathbb{X}$, $-M < f_n(x) < M$.*

Lemma 6.10. *Let $f \in B(\mathbb{X})$. Then there exists a uniformly bounded sequence $\{f_n\}$ of simple functions such that*

- *$\{f_n\}$ is fundamental in the mean.*
- *$\{f_n\}$ converges in measure to f.*

Proof. As f is bounded, $-M < f(x) < M$ for some $M \in \mathbb{R}$. For $n \in \mathbb{N}$ and for all $k \in \{0, 1, 2, \dots, 2^n - 1\}$, define

$$X_{n,k} = f^{-1}\left[\left[\frac{2M}{2^n}k - M, \frac{2M}{2^n}(k+1) - M\right)\right],$$

and put $\alpha_{n,k} = \frac{2M}{2^n}k - M$. Then, for all $n \in \mathbb{N}$, define

$$f_n(x) = \sum_{k=1}^{n} \alpha_{n,k}\chi_{X_{n,k}}(x).$$

Observe that $|f_n(x)| < M$ for all $n \in \mathbb{N}$ and $x \in \mathbb{X}$. Note that if $n < m$, then each interval $\left[\frac{2M}{2^m}k - M, \frac{2M}{2^m}(k+1) - M\right)$ is included in exactly one $\left[\frac{2M}{2^n}l - M, \frac{2M}{2^n}(l+1) - M\right)$. Consequently, each $X_{m,k}$ is included in one of the $X_{n,l}$'s and each $X_{n,l}$ is divided into 2^{m-n} many $X_{m,k}$'s. Moreover, if k, l are such that the inclusion holds, then

$$\left|\left(\frac{2M}{2^m}k - M\right) - \left(\frac{2M}{2^n}l - M\right)\right| \leq \frac{2}{2^n},$$

so $|\alpha_{m,k} - \alpha_{n,l}| \leq \frac{2}{2^n}$. Therefore, for each $x \in \mathbb{X}$, $|f_n(x) - f_m(x)| \leq \frac{2}{2^n}$. Hence,

$$\rho(f_n, f_m) = \int |f_n(x) - f_m(x)| \, \mathrm{d}\mu(x) \leq \int \frac{2}{2^n} \, \mathrm{d}\mu(x) = \frac{2}{2^n}.$$

Thus, as $n, m \longrightarrow \infty$, $\rho(f_n, f_m) \longrightarrow 0$, so $\{f_n\}$ is fundamental in the mean.

To show $\{f_n\}$ converges in measure, we note that $|f_n(x) - f(x)| \leq \frac{2}{2^n}$. So let $\epsilon_0 > 0$. For $n > \log_2\left(\frac{2}{\epsilon_0}\right)$, we have $\epsilon_0 > \frac{2}{2^n}$, so

$$\{x : |f_n(x) - f(x)| \geq \epsilon_0\} = \emptyset.$$

Therefore,

$$\mu(\{x : |f_n(x) - f(x)| \geq \epsilon_0\}) = 0,$$

so $\{f_n\}$ converges in measure to f. □

Lemma 6.11. *Let $\{f_n\}, \{g_n\}$ be uniformly bounded sequences of simple functions such that*

(1) *$\{f_n\}$ and $\{g_n\}$ are fundamental in the mean,*

(2) *$\{f_n\}$ and $\{g_n\}$ converge in measure to a function $f \in B(\mathbb{X})$.*

Then

$$\lim_{n\to\infty} \int f_n(x) \, \mathrm{d}\mu(x) = \lim_{n\to\infty} \int g_n(x) \, \mathrm{d}\mu(x).$$

Proof. Let $\{f_n\}, \{g_n\}$ be uniformly bounded sequences of simple functions fundamental in the mean convergent in measure to f. Choose $M > 0$ such that for all $n \in \mathbb{N}$,

$$-M < f_n(x) < M, \quad -M < g_n(x) < M, \quad \text{and} \quad -M < f(x) < M.$$

First, we will argue that $\lim_{n\to\infty} \int f_n(x)\,\mathrm{d}\mu(x)$ exists. Note, by Lemma 6.8, we have

$$\begin{aligned}|\textstyle\int f_n(x)\,\mathrm{d}\mu(x) - \int f_m(x)\,\mathrm{d}\mu(x)| &= |\textstyle\int f_n(x) - f_m(x)\,\mathrm{d}\mu(x)| \\ &\le \textstyle\int |f_n(x) - f_m(x)|\,\mathrm{d}\mu(x) = \rho(f_n, f_m).\end{aligned}$$

But $\rho(f_n, f_m) \to 0$ when $n \to \infty$ (as $\{f_n\}$ is fundamental in the mean), thus $\{\int f_n(x)\,\mathrm{d}\mu(x)\}$ is a Cauchy sequence, so it converges.

Now, let

$$\begin{aligned}L &= \lim_{n\to\infty} \textstyle\int f_n(x)\,\mathrm{d}\mu(x) - \lim_{n\to\infty} \int g_n(x)\,\mathrm{d}\mu(x) \\ &= \lim_{n\to\infty} (\textstyle\int f_n(x)\,\mathrm{d}\mu(x) - \int g_n(x)\,\mathrm{d}\mu(x)) \\ &= \lim_{n\to\infty} \textstyle\int (f_n(x) - g_n(x))\,\mathrm{d}\mu(x). \qquad (*)\end{aligned}$$

Suppose towards contradiction that $L \neq 0$. Set

$$\epsilon_0 = \frac{|L|}{8(M+1)} > 0.$$

Let

$$X_n = \{x : f_n(x) - f(x) \ge \epsilon_0\}, \qquad Y_n = \{x : g_n(x) - f(x) \ge \epsilon_0\}.$$

By assumption (1), we know

$$\lim_{n\to\infty} \mu(X_n) = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \mu(Y_n) = 0. \qquad (**)$$

By (*) and (**), we may choose $n \in \mathbb{N}$ such that

$$\mu(X_n) < \epsilon_0, \quad \mu(Y_n) < \epsilon_0, \quad \text{and} \quad |L - \int (f_n(x) - g_n(x))\,\mathrm{d}\mu(x)| < \epsilon_0.$$

For $x \in \mathbb{X} \setminus (X_n \cup Y_n)$, we have

$$|f_n(x) - f(x)| < \epsilon_0 \qquad \text{and} \qquad |g_n(x) - f(x)| < \epsilon_0,$$

so also

$$|f_n(x) - g_n(x)| < 2\epsilon_0.$$

For $x \in X_n \cup Y_n$, we have only

$$|f_n(x) - g_n(x)| < 2M,$$

but $\mu(X_n \cup Y_n) < 2\epsilon_0$. Hence, we may easily conclude

$$\begin{aligned} &|\textstyle\int (f_n(x) - g_n(x))\,\mathrm{d}\mu(x)| \le \int |f_n(x) - g_n(x)|\,\mathrm{d}\mu(x) \\ &\quad \le 2\epsilon_0 + 2M \cdot 2\epsilon_0 \\ &\quad = \tfrac{|L|}{4(M+1)} + \tfrac{M}{M+1} \cdot \tfrac{|L|}{2} \\ &\quad \le \tfrac{1}{4}|L| + \tfrac{1}{2}|L| = \tfrac{3}{4}|L|. \end{aligned}$$

Then

$$\begin{aligned} &|L - \textstyle\int (f_n(x) - g_n(x))\,\mathrm{d}\mu(x)| \\ &\quad \ge \big||L| - |\textstyle\int (f_n(x) - g_n(x))\,\mathrm{d}\mu(x)|\big| \\ &\quad \ge |L| - \tfrac{3}{4}|L| = \tfrac{1}{4}|L|. \end{aligned}$$

However,

$$|L - \textstyle\int (f_n(x) - g_n(x))\,\mathrm{d}\mu(x)| < \epsilon_0 = \tfrac{|L|}{8(M+1)} < \tfrac{1}{4}|L|,$$

a contradiction. □

Definition 6.12. *Let $f \in B(\mathbb{X})$. The integral of f is defined as*

$$\int f(x)\,\mathrm{d}\mu(x) = \lim_{n\to\infty} \int f_n(x)\,\mathrm{d}\mu(x),$$

for some (equivalently, all) uniformly bounded sequence $\{f_n\}$ of simple functions such that

- *$\{f_n\}$ is fundamental in the mean.*
- *$\{f_n\}$ converges in measure to f.*

Lemmas 6.10 and 6.11 verify that the above definition is correct and indeed does not depend on the choice of $\{f_n\}$.

Theorem 6.13. *Let* $f, g \in B(\mathbb{X})$*. Then* $\alpha f + \beta g \in B(\mathbb{X})$ *and*

$$\int (\alpha f + \beta g)(x)\,\mathrm{d}\mu(x) = \alpha \int f(x)\,\mathrm{d}\mu(x) + \beta \int g(x)\,\mathrm{d}\mu(x).$$

Proof. Let $f, g \in B(\mathbb{X})$ and let $\alpha, \beta \in \mathbb{R}$. Then there exist uniformly bounded sequences $\{f_n\}, \{g_n\}$ fundamental in the mean which converge in measure to f, g respectively such that each f_n, g_n is simple a function. Choose $M > 0$ such that for all $n \in \mathbb{N}$ and $x \in \mathbb{X}$,

$$-M < f_n(x) < M \qquad \text{and} \qquad -M < g_n(x) < M.$$

Note that for all $n \in \mathbb{N}$ and $x \in \mathbb{X}$,

$$|\alpha f_n(x) + \beta g_n(x)| \leq (|\alpha| + |\beta|)M,$$

so $\{\alpha f_n + \beta g_n\}$ is uniformly bounded.

Now, we will show $\{\alpha f_n + \beta g_n\}$ is fundamental in the mean. Given $\epsilon_0 > 0$. As $\{f_n\}$, $\{g_n\}$ are both fundamental in the mean, there exists $N \in \mathbb{N}$ such that for any $n, m > N$,

$$\rho(f_n, f_m) < \frac{\epsilon_0}{2(|\alpha| + 1)} \qquad \text{and} \qquad \rho(g_n, g_m) < \frac{\epsilon_0}{2(|\beta| + 1)}.$$

Then for any $n, m > N$,

$$\begin{aligned}
\rho(\alpha f_n + \beta g_n, \alpha f_m + \beta g_m) &= \textstyle\int |(\alpha f_n + \beta g_n)(x) - (\alpha f_m + \beta g_m)(x)|\,\mathrm{d}\mu(x) \\
&= \textstyle\int |\alpha f_n(x) - \alpha f_m(x) + \beta g_n(x) - \beta g_m(x)|\,\mathrm{d}\mu(x) \\
&\leq \textstyle\int |\alpha f_n(x) - \alpha f_m(x)|\,\mathrm{d}\mu(x) + \int |\beta g_n(x) - \beta g_m(x)|\,\mathrm{d}\mu(x) \\
&= |\alpha| \textstyle\int |f_n(x) - f_m(x)|\,\mathrm{d}\mu(x) + |\beta| \int |g_n(x) - g_m(x)|\,\mathrm{d}\mu(x) \\
&= |\alpha|\rho(f_n, f_m) + |\beta|\rho(g_n, g_m) \\
&< |\alpha|\tfrac{\epsilon_0}{2(|\alpha|+1)} + |\beta|\tfrac{\epsilon_0}{2(|\beta|+1)} \\
&\leq \tfrac{\epsilon_0}{2} + \tfrac{\epsilon_0}{2} = \epsilon_0.
\end{aligned}$$

Thus, $\{\alpha f_n + \beta g_n\}$ is fundamental in the mean.

Next we will show $\{\alpha f_n + \beta g_n\}$ converges in measure to $\alpha f + \beta g$. Since both $\{f_n\}$, $\{g_n\}$ converge in measure to f and g respectively, we know

$$\lim_{n\to\infty} \mu(\{x : |f_n(x) - f(x)| \geq \tfrac{\epsilon_0}{2(|\alpha|+1)}\}) = 0,$$
$$\lim_{n\to\infty} \mu(\{x : |g_n(x) - g(x)| \geq \tfrac{\epsilon_0}{2(|\beta|+1)}\}) = 0. \qquad (\circ)$$

Note that if

$$|f_n(x) - f(x)| < \frac{\epsilon_0}{2(|\alpha|+1)} \qquad \text{and} \qquad |g_n(x) - g(x)| < \frac{\epsilon_0}{2(|\beta|+1)}$$

then

$$\begin{aligned}
&|(\alpha f_n(x) + \beta g_n(x)) - (\alpha f(x) + \beta g(x))| \\
&\quad \leq |\alpha||f_n(x) - f(x)| + |\beta||g_n(x) - g(x)| \\
&\quad < |\alpha|\tfrac{\epsilon_0}{2(|\alpha|+1)} + |\beta|\tfrac{\epsilon_0}{2(|\beta|+1)} \leq \tfrac{\epsilon_0}{2} + \tfrac{\epsilon_0}{2} = \epsilon_0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\{x : |(\alpha f_n(x) + \beta g_n(x)) - (\alpha f(x) + \beta g(x))| \geq \epsilon_0\} \\
&\quad \subseteq \{x : |f_n(x) - f(x)| \geq \tfrac{\epsilon_0}{2(|\alpha|+1)}\} \cup \{x : |g_n(x) - g(x)| \geq \tfrac{\epsilon_0}{2(|\beta|+1)}\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
0 &\leq \mu(\{x : |(\alpha f_n(x) + \beta g_n(x)) - (\alpha f(x) + \beta g(x))| \geq \epsilon_0\}) \\
&\leq \mu(\{x : |f_n(x) - f(x)| \geq \tfrac{\epsilon_0}{2(|\alpha|+1)}\}) + \mu(\{x : |g_n(x) - g(x)| \geq \tfrac{\epsilon_0}{2(|\beta|+1)}\}).
\end{aligned}$$

Hence, remembering $(\circ)$,

$$\lim_{n\to\infty} \mu(\{x : |(\alpha f_n(x) + \beta g_n(x)) - (\alpha f(x) + \beta g(x))| \geq \epsilon_0\}) = 0,$$

so $\{\alpha f_n + \beta g_n\}$ converges in measure to $\alpha f + \beta g$. Hence,

$$\begin{aligned}
\textstyle\int (\alpha f + \beta g)(x)\, \mathrm{d}\mu(x) &= \lim_{n\to\infty} \textstyle\int (\alpha f_n + \beta g_n)(x)\, \mathrm{d}\mu(x) \\
&\overset{6.8}{=} \alpha \lim_{n\to\infty} \textstyle\int f_n(x)\, \mathrm{d}\mu(x) + \beta \lim_{n\to\infty} \textstyle\int g_n(x)\, \mathrm{d}\mu(x) \\
&= \alpha \textstyle\int f(x)\, \mathrm{d}\mu(x) + \beta \textstyle\int g(x)\, \mathrm{d}\mu(x).
\end{aligned}$$

□

Lemma 6.14. *Let $f : \mathbb{X} \longrightarrow [0, \infty)$ be a bounded function. Then*

$$\int f(x)\,\mathrm{d}\mu(x) \geq 0.$$

Proof. Refer to the proof of Lemma 6.10. If $f(x) \geq 0$, then the sets $X_{n,k}$ determined there will be empty whenever $\alpha_{n,k} < 0$. Therefore, $f_n(x) \geq 0$ for all $x \in \mathbb{X}$. Thus, $\int f(x)\,\mathrm{d}\mu(x) \geq 0$.

□

7. Amenable Groups

We like to believe that we have invariant measures that can measure any set. However, we have shown this belief is not true in $\mathbb{R}^3$. In this section, we develop the theory of groups that do possess an invariant measure that can measure any set, amenable groups.

Definition 7.1. *Let G be a group. If there exists a finitely additive measure $\mu : P(G) \longrightarrow [0,1]$ such that $\mu(G) = 1$ and for any $A \subseteq G$ and $g \in G$, $\mu(g[A]) = \mu(A)$, then G is called amenable.*

Proposition 7.2. *If G is amenable, G cannot be paradoxical.*

Proof. Let G be amenable, i.e. there exists a finitely additive measure $\mu : P(G) \longrightarrow [0,1]$ such that $\mu(G) = 1$ and for any $A \subseteq G$ and $g \in G$, $\mu(gA) = \mu(A)$. Suppose now, arguing by contradiction, that G is paradoxical, i.e. there exists a partition A_1, A_2, …, A_n of G and group translations T_1, T_2, …, T_i, …, T_n such that

$$T_1[A_1] \cup T_2[A_2] \cup \ldots \cup T_i[A_i] = G$$

and

$$T_{i+1}[A_{i+1}] \cup T_{i+2}[A_{i+2}] \cup \ldots \cup T_n[A_n] = G.$$

Then

$$\begin{aligned}\mu(G) &= \mu(A_1) + \mu(A_2) + \ldots + \mu(A_n) \\ &= \mu(T_1[A_1]) + \mu(T_2[A_2]) + \ldots + T_i[A_i] + T_{i+1}[A_{i+1}] + \ldots + \mu(T_n[A_n]) \\ &= \mu(G) + \mu(G),\end{aligned}$$

a contradiction. Thus, G cannot be paradoxical. □

Definition 7.3. *Let G act on $\mathbb{X}$, and let $(\mathbb{X}, \bar{\mu}, \mathcal{R})$ be a measure space.*

(1) *We say that $\bar{\mu}$ is left-invariant with respect to G if*
- *$g[A] \in \mathcal{R}$ for each $A \in \mathcal{R}$ and $g \in G$, and*

- $\bar{\mu}(g[A]) = \bar{\mu}(A)$.

(2) *We say that $\mathcal{R}$ is left-invariant with respect to G if*

$$(g \in G \quad \& \quad A \in \mathcal{R}) \implies g[A] \in \mathcal{R}.$$

Lemma 7.4. *Let θ be a left-invariant measure on G. If $f, h \in B(G)$, $g \in G$, and $h(x) = f(gx)$ then*

$$\int f(x)\, \mathrm{d}\theta(x) = \int h(x)\, \mathrm{d}\theta(x).$$

Proof. First note if f is a simple function, then h is as well, and easily, $\int f(x)\, \mathrm{d}\theta(x) = \int h(x)\, \mathrm{d}\theta(x)$. In the general case, pick a uniformly bounded sequence $\{f_n\}$ fundamental in the mean and convergent in measure to f. Let $h_n(x) = f_n(gx)$. Then $\{h_n\}$ is uniformly bounded, fundamental in the mean, converges in measure to h, and $\int f_n(x)\, \mathrm{d}\theta(x) = \int h_n(x)\, \mathrm{d}\theta(x)$. □

Theorem 7.5. *Let G be an amenable group that acts on a space $\mathbb{X}$. If there exists a finitely-additive measure $\bar{\mu}$ on $P(\mathbb{X})$ that is left-invariant with respect to G, then for any $X \subseteq \mathbb{X}$ such that $\bar{\mu}(X) \in (0, \infty)$, X is not G-paradoxical.*

Proof. Let G be an amenable group and $\mathbb{X}$ be a G-set. Then there exists a left-invariant finitely additive measure $\mu : P(G) \longrightarrow [0, 1]$ of total measure 1. Also, let $\bar{\mu}$ be a finitely additive measure $\bar{\mu} : P(\mathbb{X}) \longrightarrow [0, \infty]$ that is left-invariant with respect to G.

Suppose now towards contradiction that $X \subseteq \mathbb{X}$ such that $\bar{\mu}(X) \in (0, \infty)$ is G-paradoxical, i.e. there exist

- a partition $A_1, A_2, \ldots, A_i, A_{i+1}, \ldots, A_n$ of X, and
- elements $g_1, g_2, \ldots, g_i, g_{i+1}, \ldots, g_n$ of G

such that for some $i < n$,

- $g_1[A_1], g_2[A_2], \ldots, g_i[A_i]$ is a partition of X, and
- $g_{i+1}[A_{i+2}], g_{i+1}[A_{i+2}], \ldots, g_n[A_n]$ is a partition of X.

Then

$$\begin{aligned}
\bar{\mu}(X) &= \bar{\mu}(A_1 \cup A_2 \cup \ldots \cup A_n) \\
&= \bar{\mu}(A_1) + \bar{\mu}(A_2) + \ldots + \bar{\mu}(A_n) \\
&= \bar{\mu}(g_1[A_1]) + \bar{\mu}(g_2[A_2]) + \ldots \bar{\mu}(g_i[A_i]) + \bar{\mu}(g_{i+1}[A_{i+1}]) + \ldots + \bar{\mu}(g_n[A_n]) \\
&= \bar{\mu}(X) + \bar{\mu}(X) \\
&= 2\bar{\mu}(X),
\end{aligned}$$

a contradiction, as $\bar{\mu}(X) \in (0, \infty)$. Thus, $X \subseteq \mathbb{X}$ such that $\bar{\mu}(X) \in (0, \infty)$ is not G-paradoxical. □

The following proposition explains why we focus on finitely additive measures and not σ-additive measures.

Proposition 7.6 (Giuseppe Vitali, [16]). *Let G be an infinite group. Then there exists no left-invariant σ-additive measure $\mu : P(G) \longrightarrow [0, 1]$ of total measure 1.*

Proof. Let G be an infinite group, and let $H \leq G$ be a countably infinite subgroup, so $H = \{h_1, h_2, \ldots\}$. For any $g', h' \in G$, define the relation $g'Rh'$ if and only if $g'H = h'H$. Now choose a selector $S \subseteq G$ such that $|S \cap gH| = 1$ for all $gH \in G/H$. Note for any $g' \in G$, $g' \in gH$ for some $g \in S$. Then

$$HS = \{hS : h \in H\} = G.$$

Now suppose, by way of contradiction, that a left-invariant σ-additive measure μ exists on $P(G)$ of total measure 1. Then

$$\mu(G) = \sum_{i=1}^{\infty} \mu(h_i S).$$

If $\mu(S) = 0$, then $\mu(G) = 0$. If $\mu(S) > 0$, then $\mu(G) = \infty$. In either case, we arrive at a contradiction, so there exists no left-invariant σ-additive measure on G with total measure 1. □

Theorem 7.7. *Let $\mathcal{A}_0$ be a subring of an algebra $\mathcal{A}$ and G be an amenable group acting on $\mathbb{X}$. Suppose also that $\mathcal{A}$ and $\mathcal{A}_0$ are left-invariant with respect to G. If there exists a left-invariant finitely additive measure μ on $\mathcal{A}_0$, then there exists a left-invariant finitely additive measure $\bar{\mu}$ on $\mathcal{A}$ that extends μ.*

Proof. Let μ be a finitely additive measure on $\mathcal{A}_0$ that is left-invariant with respect to G. First, we will show there exists a finitely additive measure on $\mathcal{A}$ that extends μ.

To start, we will consider the case if $\mathcal{A}$ is finite and proceed by induction on the number of atoms of $\mathcal{A}$. If $\mathcal{A}$ has one atom, then $\mathcal{A} = \{\emptyset, \mathbb{X}\}$, so $\mathcal{A}_0 = \mathcal{A}$ and $\mu = \bar{\mu}$. Now let $\mathcal{A}$ have n atoms and assume for any algebra with less than n atoms that we can extend its measure. Let $B \in \mathcal{A}_0$ be an atom of $\mathcal{A}_0$. Define $C = B' = \mathbb{X} \setminus B \in \mathcal{A}$. Choose $A_0 \subseteq B$ such that A_0 is an atom of $\mathcal{A}$. Then $A_0 \notin \mathcal{A}(C)$ (see Definition 6.3 and recall $\mathcal{A}(C)$ is an algebra of subsets of C), so $\mathcal{A}(C)$ has less than n atoms, and we can find a finitely additive measure $\nu : \mathcal{A}(C) \longrightarrow [0,1]$ that extends $\mu|_{\mathcal{A}(C)\cap\mathcal{A}_0}$ by our inductive hypothesis (note $\mathcal{A}(C) \cap \mathcal{A}_0$ is a subring of $\mathcal{A}(C)$). Note that every atom $A \in \mathcal{A}$, $A \subseteq C$ or $A \subseteq B$. So for any atom $A \in \mathcal{A}$, define

$$\bar{\mu}_{atom}(A) = \begin{cases} \nu(A) & \text{if } A \in \mathcal{A}(C) \\ \mu(B) & \text{if } A = A_0 \\ 0 & \text{if } A \subseteq B \text{ but } A \neq A_0 \end{cases}$$

Finally, for any $X \in \mathcal{A}$, define

$$\bar{\mu}(X) = \sum \{\bar{\mu}_{atom}(A) : A \subseteq X, \text{ and } A \text{ is an atom of } \mathcal{A}\}$$

To show $\bar{\mu}$ is finitely additive, let $A_1, A_2, \ldots, A_k \in \mathcal{A}$ be pairwise disjoint. Then

$$\begin{aligned} \bar{\mu}(A_1) + \bar{\mu}(A_2) &= \sum \{\bar{\mu}_{atom}(X) : X \subseteq A_1, \text{ and } X \text{ is an atom of } \mathcal{A}\} \\ &\quad + \sum \{\bar{\mu}_{atom}(X) : X \subseteq A_2, \text{ and } X \text{ is an atom of } \mathcal{A}\} \\ &= \sum \{\bar{\mu}_{atom}(X) : X \subseteq A_1 \cup A_2, \text{ and } X \text{ is an atom of } \mathcal{A}\} = \bar{\mu}(A_1 \cup A_2). \end{aligned}$$

To show $\bar{\mu}$ extends μ, let $D \in \mathcal{A}_0$. Then either $D \cap B = \emptyset$ or $B \subseteq D$. If $D \cap B = \emptyset$, then $D \in \mathcal{A}(C)$, so $\bar{\mu}(D) = \nu(D) = \mu(D)$. If instead $B \subseteq D$, then

$$\begin{aligned}\bar{\mu}(D) &= \bar{\mu}(B) + \bar{\mu}(D \setminus B) = \bar{\mu}(A_0) + \bar{\mu}(B \setminus A_0) + \nu(D \setminus B)\\ &= \mu(B) + \mu(D \setminus B) = \mu(D).\end{aligned}$$

Thus $\bar{\mu}$ is a finitely additive measure on $\mathcal{A}$ that extends μ.

Consider the case when $\mathcal{A}$ is infinite. Let $[0, \infty]^{\mathcal{A}}$ be the set of all functions $f : \mathcal{A} \longrightarrow [0, \infty]$ and let it be equipped with the product space topology so that the basic open sets are of the form

$$\{f \in [0, \infty]^{\mathcal{A}} : f(A_1) \in U_1, f(A_2) \in U_2, \ldots, f(A_n) \in U_n\}$$

where $A_1, A_2, \ldots, A_n \in \mathcal{A}$ and $U_1, U_2, \ldots, U_n \subseteq [0, \infty]$ are open.

By Tychonoff's Theorem [10, p.234], we know that $[0, \infty]^{\mathcal{A}}$ is compact, and thus $[0, \infty]^{\mathcal{A}}$ has the finite intersection property, i.e. if $\mathcal{F}$ is a family of closed subsets of $[0, \infty]^{\mathcal{A}}$ and for any $F_1, F_2, \ldots, F_n \in \mathcal{F}$, $F_1 \cap F_2 \cap \ldots \cap F_n \neq \emptyset$, then $\bigcap \mathcal{F} \neq \emptyset$.

For all finite subalgebras $\mathcal{C}$ of $\mathcal{A}$, define

$$M(\mathcal{C}) = \{\nu \in [0, \infty]^{\mathcal{A}} : \nu|_{\mathcal{C}} \text{ is a measure and } \nu \text{ extends } \mu|_{\mathcal{C} \cap \mathcal{A}_0}\}$$

We will show that each $M(\mathcal{C})$ is a closed subset of $[0, \infty]^{\mathcal{A}}$ and that the family $\mathcal{F}_M$ of all of these sets satisfies the finite intersection property.

To show $M(\mathcal{C})$ is closed, let $\nu \in [0, \infty]^{\mathcal{A}} \setminus M(\mathcal{C})$. Then either ν does not extend $\mu|_{\mathcal{C} \cap \mathcal{A}_0}$ or $\nu|_{\mathcal{C}}$ is not a measure.

If ν does not extend $\mu|_{\mathcal{C} \cap \mathcal{A}_0}$, then for some $A \in \mathcal{C} \cap \mathcal{A}_0$, $\nu(A) \neq \mu(A)$. Then choose an open set $U \subseteq [0, \infty]$ such that $\nu(A) \in U$ but $\mu(A) \notin U$. Then $\{f \in [0, \infty]^{\mathcal{A}} : f(A) \in U\}$ is an open set disjoint from $M(\mathcal{C})$, and $\nu \in \{f \in [0, \infty]^{\mathcal{A}} : f(A) \in U\}$.

If $\nu|_{\mathcal{C}}$ is not a measure, then there exist $A_1, A_2 \in \mathcal{C}$ such that $A_1 \cap A_2 = \emptyset$ and $\nu(A_1) + \nu(A_2) \neq \nu(A_1 \cup A_2)$. We will define open sets $U_1, U_2, U_2 \subseteq [0, \infty]$ depending on one of the following three sub-cases.

If $\nu(A_1), \nu(A_2), \nu(A_1 \cup A_2) \in [0, \infty)$, let

$$\epsilon = |\nu(A_1 \cup A_2) - [\nu(A_1) + \nu(A_2)]|$$

and let

$$\begin{aligned} U_1 &= (\nu(A_1) - \epsilon/4, \nu(A_1) + \epsilon/4) \\ U_2 &= (\nu(A_2) - \epsilon/4, \nu(A_2) + \epsilon/4) \\ U_3 &= (\nu(A_1 \cup A_2) - \epsilon/4, \nu(A_1 \cup A_2) + \epsilon/4). \end{aligned}$$

If $\nu(A_1), \nu(A_2) \in [0, \infty)$, but $\nu(A_1 \cup A_2) = \infty$, let

$$\begin{aligned} U_1 &= (\nu(A_1) - 1/4, \nu(A_1) + 1/4) \\ U_2 &= (\nu(A_2) - 1/4, \nu(A_2) + 1/4) \\ U_3 &= (\nu(A_1) + \nu(A_2) + 1, \infty]. \end{aligned}$$

If $\nu(A_1 \cup A_2) \in [0, \infty)$, but $\nu(A_i) = \infty$ where $i \in \{1, 2\}$, let

$$\begin{aligned} U_i &= (\nu(A_1 \cup A_2) + 1, \infty] \\ U_{3-i} &= [0, \infty] \\ U_3 &= [0, \nu(A_1 \cup A_2) + 1/2). \end{aligned}$$

In any case, $\{f \in [0, \infty]^{\mathcal{A}} : f(A_1) \in U_1, f(A_2) \in U_2, f(A_1 \cup A_2) \in U_3\}$ is an open set disjoint from $M(\mathcal{C})$, and $\nu \in \{f \in [0, \infty]^{\mathcal{A}} : f(A_1) \in U_1, f(A_2) \in U_2, f(A_1 \cup A_2) \in U_3\}$.

Thus, $[0, \infty]^{\mathcal{A}} \setminus M(\mathcal{C})$ is open, and $M(\mathcal{C})$ is closed.

Now observe that $M(\mathcal{C})$ is non-empty as we have already shown that since $\mathcal{C}$ is finite, $\mu|_{\mathcal{C} \cap \mathcal{A}_0}$ can be extended to $\mathcal{C}$.

Also note that if $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n$ are finite subalgebras of $\mathcal{A}$, then $\mathcal{C}_1 \cup \mathcal{C}_2 \cup \ldots \cup \mathcal{C}_n$ generate a finite subalgebra $\mathcal{C}^*$ of $\mathcal{A}$. Then

$$\emptyset \neq M(\mathcal{C}^*) \subseteq M(\mathcal{C}_1) \cap M(\mathcal{C}_2) \cap \ldots \cap M(\mathcal{C}_n).$$

Thus, the family $\mathcal{F}_M$ satisfies the finite intersection property, so $\bigcap \mathcal{F}_M \neq \emptyset$. Hence, there exists a finitely additive measure $\bar{\mu}$ on $\mathcal{A}$ that extends μ.

Now that we have shown there exists a finitely additive measure on $\mathcal{A}$ that extends μ, we need to show there is a measure that is also left-invariant with respect to the amenable group G. Let ν be a finitely additive measure that extends μ and let θ be a finitely additive measure on G of total measure one. By our assumption, we know μ is left-invariant with respect to G, i.e. for any $A \in \mathcal{A}_0$ and $g \in G$, $\mu(g[A]) = \mu(A)$.

Now for any set $B \in \mathcal{A}$, define $f_B : G \longrightarrow [0, \infty]$ by

$$f_B(g) = \nu(g^{-1}B).$$

Then define

$$\bar{\mu}(B) = \int f_B d\theta.$$

Note $f_B \geq 0$ so $\bar{\mu}(B) \geq 0$. Also, for any $A_1, A_2 \in \mathcal{A}$ such that $A_1 \cap A_2 = \emptyset$,

$$\begin{aligned} &f_{A_1}(g) + f_{A_2}(g) = \nu(g^{-1}[A_1]) + \nu(g^{-1}[A_2]) = \nu(g^{-1}[A_1] \cup g^{-1}[A_2]) \\ &= \nu(g^{-1}[A_1 \cup A_2]) = f_{A_1 \cup A_2}(g) \end{aligned}$$

so

$$\begin{aligned} &\bar{\mu}(A_1) + \bar{\mu}(A_2) = \textstyle\int f_{A_1}\, \mathrm{d}\theta + \int f_{A_2}\, \mathrm{d}\theta = \int f_{A_1} + f_{A_2}\, \mathrm{d}\theta \\ &= \textstyle\int f_{A_1 \cup A_2}\, \mathrm{d}\theta = \bar{\mu}(A_1 \cup A_2). \end{aligned}$$

Thus, $\bar{\mu}$ is a measure.

To show $\bar{\mu}$ is left-invariant, let $A \in \mathcal{A}$ and let $h \in G$. Then

$$f_{h[A]}(g) = \nu(g^{-1}h[A]) = \nu((h^{-1}g)^{-1}[A]) = f_A(h^{-1}g)$$

so, remember Lemma 7.4,

$$\bar{\mu}(h[A]) = \int f_{h[A]}(g)\, d\theta(g) = \int f_A(h^{-1}g)\, d\theta(g)$$

$$= \int f_A(g)\, d\theta(g) = \bar{\mu}(A).$$

Thus, $\bar{\mu}$ is left-invariant.

Finally, to show $\bar{\mu}$ extends μ, let $A \in \mathcal{A}_0$. Then

$$f_A(g) = \nu(g^{-1}A) = \mu(g^{-1}A) = \mu(A)$$

so

$$\bar{\mu}(A) = \int f_A d\theta = \int \mu(A) d\theta = \mu(A).$$

Thus, μ is a left-invariant, finitely additive measure on $\mathcal{A}$ that extends μ. □

Theorem 7.8. *Let G be a group. If G is finite, then G is amenable.*

Proof. Let G be a group and $|G| = n$. For $A \subseteq G$, define $\mu : G \longrightarrow [0, 1]$ by

$$\mu(A) = \frac{|A|}{n}.$$

First note $\mu(G) = \frac{n}{n} = 1$. Now to show that μ is finitely additive, let $A_1, A_2 \subseteq G$ such that $A_1 \cap A_2 = \emptyset$. Then

$$\mu(A_1) + \mu(A_2) = \frac{|A_1|}{n} + \frac{|A_2|}{n} = \frac{|A_1 \cup A_2|}{n} = \mu(A_1 \cup A_2)$$

Finally, to show μ is left-invariant, let $A \subseteq G$ and $g \in G$. Then

$$\mu(gA) = \frac{|gA|}{n} = \frac{|A|}{n} = \mu(A)$$

So μ is a left-invariant finitely additive measure on G of total measure 1, and thus, G is amenable. □

Theorem 7.9. *Let H be a subgroup of a group G. If G is amenable, then H is amenable.*

Proof. Let H be a subgroup of a group G and let G be amenable. Then there exists a left-invariant finitely additive measure μ on G of total measure 1. Let M be a selector from the right cosets of H so that for all $g \in G$, $|M \cap Hg| = 1$. Then define

$\nu : P(H) \longrightarrow [0,1]$ by

$$\nu(A) = \mu\left(\bigcup_{g\in M} Ag\right).$$

Observe

$$\nu(H) = \mu\left(\bigcup_{g\in M} Hg\right) = \mu(G) = 1.$$

To show that ν is finitely additive, let $A_1, A_2 \subseteq H$ such that $A_1 \cap A_2 = \emptyset$. Then the sets $\bigcup_{g\in M} A_1 g$ and $\bigcup_{g\in M} A_2 g$ are disjoint, and

$$\nu(A_1) + \nu(A_2) = \mu\left(\bigcup_{g\in M} A_1 g\right) + \mu\left(\bigcup_{g\in M} A_2 g\right) = \mu\left(\bigcup_{g\in M} A_1 g + \bigcup_{g\in M} A_2 g\right)$$

$$= \mu\left(\bigcup_{g\in M} (A_1 g \cup A_2 g)\right) = \mu\left(\bigcup_{g\in M} (A_1 \cup A_2) g\right) = \nu(A_1 \cup A_2).$$

Finally, to show ν is left-invariant, let $A \subseteq H$ and $h \in G$. Then

$$\nu(hA) = \mu\left(\bigcup_{g\in M} (hA)g\right) = \mu\left(\bigcup_{g\in M} h(Ag)\right) = \mu\left(\bigcup_{g\in M} Ag\right) = \nu(A).$$

So ν is a left-invariant finitely additive measure on H of total measure 1, and thus, H is amenable. □

Theorem 7.10. *Let H be a normal subgroup of a group G. If G is amenable, then G/H is amenable.*

Proof. Let H be a normal subgroup of an amenable group G, and let μ be a left-invariant finitely additive measure on G of total measure 1. Define $\nu : P(G/H) \longrightarrow [0,1]$ by

$$\nu(\mathcal{B}) = \mu\left(\bigcup \mathcal{B}\right).$$

Observe

$$\nu(G/H) = \mu\left(\bigcup G/H\right) = \mu(G) = 1.$$

To show that ν is finitely additive, let $\mathcal{B}_1, \mathcal{B}_2 \subseteq G/N$ such that $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$. Then $\bigcup \mathcal{B}_1 \cap \bigcup \mathcal{B}_2 = \emptyset$ (as cosets form a partition of G), and

$$\begin{aligned}\nu(\mathcal{B}_1) + \nu(\mathcal{B}_2) &= \mu\left(\bigcup \mathcal{B}_1\right) + \mu\left(\bigcup \mathcal{B}_2\right) \\ &= \mu\left(\bigcup \mathcal{B}_1 + \bigcup \mathcal{B}_2\right) \\ &= \mu\left(\bigcup(\mathcal{B}_1 \cup \mathcal{B}_2)\right) = \nu(\mathcal{B}_1 \cup \mathcal{B}_2).\end{aligned}$$

Finally, to show ν is left-invariant, let $\mathcal{B} \subseteq G/H$ and $h \in G$. Let C be a selector from $\mathcal{B}$. Then, in particular, $\mathcal{B} = \{gH : g \in C\}$, and

$$\nu((hH)\mathcal{B}) = \mu\left(\bigcup (hH)\mathcal{B}\right) = \mu\left(\bigcup_{g \in C} hHg\right) = \mu\left(h \bigcup_{g \in C} Hg\right)$$

$$= \mu\left(h \bigcup_{g \in C} gH\right) = \mu\left(\bigcup_{g \in C} gH\right) = \mu\left(\bigcup \mathcal{B}\right) = \nu(\mathcal{B}).$$

So ν is a left-invariant finitely additive measure on G/H of total measure 1, and thus, G/N is amenable. ☐

Theorem 7.11. *Let H be a normal subgroup of a group G. If H and G/H are amenable, then G is amenable.*

Proof. Let H be a normal subgroup of a group G and suppose also that H and G/H are amenable. Then let ν_1 and ν_2 be left-invariant finitely additive measures of total measure 1 on H and G/H respectively. For $A \subseteq G$, define $f_A : G \longrightarrow [0, 1]$ by

$$f_A(g) = \nu_1(H \cap g^{-1}A).$$

Suppose $g_1 H = g_2 H$. Then $g_2^{-1} g_1 H = H$ so $g_2^{-1} g_1 = h \in H$. Then

$$\begin{aligned}f_A(g_2) &= \nu_1(H \cap g_2^{-1}A) = \nu_1(hH \cap hg_1^{-1}A) \\ &= \nu_1(h(H \cap g_1^{-1}A)) = \nu_1(H \cap g_1^{-1}A) = f_A(g_1).\end{aligned}$$

Thus, f_A is constant on each coset of H and we may define $f_A^* : G/H \longrightarrow [0, 1]$ by

$$f_A^*(gH) = f_A(g).$$

Now define $\mu : P(G/H) \longrightarrow [0,1]$ by

$$\mu(A) = \int f_A^*(gH)\, d\nu_2(gH).$$

Observe $f_G(g) = 1$ for all $g \in G$, and hence, $f_G^*(gH) = 1$ for all $gH \in G/H$. Then

$$\mu(G) = \int f_G^*(gH)\, d\nu_2(gH) = \int 1\, d\nu_2(gH) = \nu_2(G/H) = 1.$$

To show that ν is finitely additive, let $A_1, A_2 \subseteq G$ such that $A_1 \cap A_2 = \emptyset$. Note if $A_1 \cap A_2 = \emptyset$ then $g^{-1}A_1 \cup g^{-1}A_2 = \emptyset$, so

$$\begin{aligned} f_{A_1}(g) + f_{A_2}(g) &= \nu_1(N \cap g^{-1}A_1) + \nu_1(N \cap g^{-1}A_2) \\ &= \nu_1((N \cap g^{-1}A_1) \cup (N \cap g^{-1}A_2)) = \nu_1(N \cap (g^{-1}A_1 \cup g^{-1}A_2)) \\ &= \nu_1(N \cap g^{-1}(A_1 \cup A_2)) = f_{A_1 \cup A_2}(g). \end{aligned}$$

Then

$$\begin{aligned} \mu(A_1) + \mu(A_2) &= \textstyle\int f_{A_1}^*(gH)\, d\nu_2(gH) + \int f_{A_2}^*(gH)\, d\nu_2(gH) \\ &= \textstyle\int (f_{A_1}^*(gH) + f_{A_2}^*(gH))\, d\nu_2(gH) = \int f_{A_1 \cup A_2}^*(gH)\, d\nu_2(gH) = \mu(A_1 \cup A_2). \end{aligned}$$

Finally, to show ν is left-invariant, let $A \subseteq G$ and $h \in G$. Note

$$f_{hA}(g) = \nu_1(H \cap g^{-1}hA) = \nu_1(H \cap (g^{-1}g)^{-1}A) = f_A(h^{-1}g).$$

Then $f_{hA}^*(gH) = f_A^(((h^{-1}g)H) = f_A^*(h^{-1}H)(gH))$, and

$$\mu(hA) = \int f_{hA}^*(gH)\, d\nu_2(gH) = \int f_{hA}^*((h^{-1}H)(gH))\, d\nu_2(gH)$$

$$= \int f_A^*(gH)\, d\nu_2(gH) = \mu(A).$$

So μ is a left-invariant finitely additive measure on G of total measure 1, and thus, G is amenable. $\square$

Theorem 7.12. *If G is an Abelian group, then G is amenable.*

Proof. Let G be an Abelian group and X a finite subset of G. Let $[0,1]^{P(G)}$ be the space of all functions $f : P(G) \longrightarrow [0,1]$ equipped with the product topology. Thus, the basic open sets are of the form

$$\{f \in [0,1]^{P(G)} : f(A_1) \in U_1, f(A_2) \in U_2, \ldots, f(A_n) \in U_n\}$$

where $A_1, A_2, \ldots, A_n \in P(G)$ and $U_1, U_2, \ldots, U_n$ are open subsets of $[0,1]$. By Tychonoff's Theorem [10, p.234], the space $[0,1]^{P(G)}$ is compact, and therefore has the finite intersection property. For $\epsilon > 0$ and a finite set $X \subseteq G$, we define

$$\begin{aligned} M_{\epsilon,X} = \{f \in [0,1]^{P(G)} : \quad & f(G) = 1, \text{ for any } A_1, A_2 \subseteq G \text{ such that } A_1 \cap A_2 = \emptyset, \\ & f(A_1) + f(A_2) = f(A_1 \cup A_2), \\ & \text{and for any } A \subseteq G \text{ and } g \in X, |f(A) - f(gA)| \leq \epsilon\}. \end{aligned}$$

Let

$$\mathcal{F} = \{M_{\epsilon,X} : \epsilon > 0, X \subseteq G, X \text{ is finite}\}.$$

Our goal is to show that $\mathcal{F}$ satisfies the conditions for the finite intersection property.

To show $M_{\epsilon,X}$ is non-empty, choose $N \geq \frac{2}{\epsilon}$. Fix an enumeration of elements of $X = \{g_1, g_2, \ldots, g_n\}$ and for any subset $A \subseteq G$, define

$$Y_A = \{(i_1, i_2, \ldots, i_n) : i_j \in \{1, 2, \ldots, N\}, j \in \{1, 2, \ldots, n\}, g_1^{i_1} g_2^{i_2} \ldots g_n^{i_n} \in A\}.$$

Then define

$$\mu_\epsilon(A) = |Y_A|/N^n.$$

First, observe $Y_G = \{(i_1, i_2, \ldots, i_n) : i_j \in \{1, 2, \ldots, N\}, j \in \{1, 2, \ldots, n\}$

$$\mu_\epsilon(G) = \tfrac{|Y_G|}{N^n} = |\{(i_1, i_2, \ldots, i_n) : i_j \in \{1, 2, \ldots, N\}, j \in \{1, 2, \ldots, n\}\}|/N^n = \tfrac{N^n}{N^n} = 1.$$

To show that μ_ϵ is finitely additive, let $A_1, A_2 \subseteq G$ such that $A_1 \cap A_2 = \emptyset$. Then $Y_{A_1} \cup Y_{A_2} = Y_{A_1 \cup A_2}$, and hence

$$\begin{aligned} \mu_\epsilon(A_1) + \mu_\epsilon(A_2) &= \tfrac{|Y_{A_1}|}{N^n} + \tfrac{|Y_{A_2}|}{N^n} \\ &= \tfrac{|Y_{A_1 \cup A_2}|}{N^n} = \mu_\epsilon(A_1 \cup A_2). \end{aligned}$$

Finally, to show that μ_ϵ is almost left-invariant with respect to elements of X, let $A \subseteq G$ and $g_k \in X$. Observe now that, as G is Abelian,

$$\begin{aligned} &Y_{g_k A} = \big\{(i_1, i_2, \ldots, i_n) : i_j \in \{1, 2, \ldots, N\}, j \in \{1, 2, \ldots, n\}, g_1^{i_1} g_2^{i_2} \ldots g_n^{i_n} \in g_k A\big\} \\ &= \big\{(i_1, i_2, \ldots, i_n) : i_j \in \{1, 2, \ldots, N\}, j \in \{1, 2, \ldots, n\}, g_1^{i_1} g_2^{i_2} \ldots g^{i_k - 1} \ldots g_n^{i_n} \in A\big\}. \end{aligned}$$

Therefore,

$$\begin{aligned} &Y_A \triangle Y_{g_k A} = (Y_A \setminus Y_{g_k A}) \cup (Y_{g_k A} \setminus Y_A) = \\ &\{(i_1, i_2, \ldots, i_n) : i_j \in \{1, 2, \ldots, N\}, j \in \{1, \ldots, n\}, g_1^{i_1} g_2^{i_2} \ldots g_{k-1}^{i_{k-1}} g_k^N g_{k+1}^{i_{k+1}} \ldots g_n^{i_n} \in A\} \\ &\cup \{(i_1, i_2, \ldots, i_n) : i_j \in \{1, 2, \ldots, N\}, j \in \{1, \ldots, n\}, g_1^{i_1} g_2^{i_2} \ldots g_{k-1}^{i_{k-1}} g_{k+1}^{i_{k+1}} \ldots g_n^{i_n} \in A\}. \end{aligned}$$

Thus,

$$\begin{aligned} |\mu_\epsilon(A) - \mu_\epsilon(g_k A)| &= \left| \frac{|Y_A|}{N^n} - \frac{|Y_{g_k A}|}{N^n} \right| = \frac{\big||Y_A| - |Y_{g_k A}|\big|}{N^n} \\ &\leq \frac{|Y_A \triangle Y_{g_k A}|}{N^n} = \frac{2N^{n-1}}{N^n} = \frac{2}{N} < \epsilon. \end{aligned}$$

Thus, $\mu_\epsilon \in M_{\epsilon,X}$, so $M_{\epsilon,X}$ is not empty.

To show $M_{\epsilon,X}$ is closed, let $\nu \in [0,1]^{P(G)} \setminus M_{\epsilon,X}$. Then ν does not have total measure 1, ν is not finitely additive, or ν is not almost invariant with respect to elements of X.

If $\nu(G) < 1$, then

$$\left\{ f \in [0,1]^{P(G)} : f(G) \in \left[0, 1 - \frac{1 - \nu(G)}{2}\right) \right\}$$

is an open set disjoint from $M_{\epsilon,X}$ and ν belongs to it.

If ν is not finitely additive, then there exist $A_1, A_2 \subseteq G$ such that $A_1 \cap A_2 = \emptyset$ and $\nu(A_1) + \nu(A_2) \neq \nu(A_1 \cup A_2)$. Then let

$$\epsilon_0 = \Big|[\nu(A_1) + \nu(A_2)] - \nu(A_1 \cup A_2)\Big|.$$

Then

$$\Big\{ f \in [0,1]^{P(G)} : \begin{aligned} &\nu(A_1) \in \big(\nu(A_1) - \tfrac{\epsilon_0}{4}, \nu(A_1) + \tfrac{\epsilon_0}{4}\big)\,, \\ &\nu(A_2) \in \big(\nu(A_2) - \tfrac{\epsilon_0}{4}, \nu(A_2) + \tfrac{\epsilon_0}{4}\big)\,, \\ &\nu(A_1 \cup A_2) \in \big(\nu(A_1 \cup A_2) - \tfrac{\epsilon_0}{4}, \nu(A_1 \cup A_2) + \tfrac{\epsilon_0}{4}\big) \Big\} \end{aligned}$$

is an open set disjoint from $M_{\epsilon,X}$ and ν belongs to it.

If ν is not almost invariant with respect to elements of X, then there exist $A \subseteq G$ and $g_k \in X$ such that $|\nu(A) - \nu(g_k A)| > \epsilon$. Let $\epsilon_0 = |\nu(A) - \nu(g_k A)| - \epsilon > 0$. Then

$$\Big\{ f \in [0,1]^{P(G)} : \nu(A) \in \big(\nu(A) - \tfrac{\epsilon}{4}, \nu(A) + \tfrac{\epsilon}{4}\big)\,, \nu(g_k A) \in \big(\nu(g_k A) - \tfrac{\epsilon}{4}, \nu(g_k A) + \tfrac{\epsilon}{4}\big) \Big\}$$

is an open set disjoint from $M_{\epsilon,X}$ and ν belongs to it.

In any case, ν is contained in an open set disjoint from $M_{\epsilon,X}$, so $M_{\epsilon,X}$ is closed.

Finally, to show that any finite intersection is non-empty, let

$$M_{\epsilon_1,X_1}, M_{\epsilon_2,X_2}, \ldots, M_{\epsilon_n,X_n} \in \mathcal{F},$$

where $\epsilon_i > 0$ and $X_i \in X$ for all $i \in \{1, 2, \ldots, n\}$. Then for $\epsilon = \min\{\epsilon_1, \epsilon_2, \ldots, \epsilon_\mathrm{n}\}$ and $X = X_1 \cup X_2 \cup \ldots \cup X_n$,

$$\emptyset \neq M_{\epsilon,X} \subseteq M_{\epsilon_1,X_1} \cap M_{\epsilon_2,X_2} \cap \ldots \cap M_{\epsilon_n,X_n}.$$

Hence we have shown that the family $\mathcal{F}$ of all sets $M_{\epsilon,X}$ satisfies the conditions for the finite intersection property, so

$$\bigcap \mathcal{F} \neq \emptyset.$$

Since every element of $\bigcap \mathcal{F}$ is a left-invariant finitely additive measure μ on G of total measure 1, the group G is amenable. $\square$

During the preceding proof, it is noteworthy that we only had to consider finitely generated subgroups of G. The following theorem states that explicitly.

Theorem 7.13. *If $G = \bigcup_{\alpha \in I} G_\alpha$ where G_α is an amenable subgroup of G and for any $\alpha, \beta \in I$, there exists $\gamma \in I$ such that $G_\alpha \leq G_\gamma$ and $G_\beta \leq G_\gamma$, then G is amenable.*

Proof. Let $G = \bigcup_{\alpha \in I} G_\alpha$ where G_α is an amenable subgroup of G and for any $\alpha, \beta \in I$, there exists $\gamma \in I$ such that $G_\alpha \leq G_\gamma$ and $G_\beta \leq G_\gamma$. As each G_α is amenable, let μ_α be a left-invariant finitely additive measure on G_α of total measure 1. Let $[0,1]^{P(G)}$ be the space of all functions $f : P(G) \longrightarrow [0,1]$ equipped with the product topology so that the basic open sets are of the form

$$\{f \in [0,1]^{P(G)} : f(A_1) \in U_1, f(A_2) \in U_2, \ldots, f(A_n) \in U_n\}$$

where $A_1, A_2, \ldots, A_n \in P(G)$ and $U_1, U_2, \ldots, U_n$ are open subsets of $[0,1]$. By Tychonoff's Theorem, $[0,1]^{P(G)}$ is compact, and therefore has the finite intersection property. For $\alpha \in I$, let

$$\begin{aligned} M_\alpha = \{f \in [0,1]^{P(G)} : \quad & f(G) = 1, \text{ for any } A_1, A_2 \subseteq G \text{ such that } A_1 \cap A_2 = \emptyset, \\ & f(A_1) + f(A_2) = f(A_1 \cup A_2), \\ & \text{and for any } g \in G_\alpha \text{ and } A \subseteq G,\ f(gA) = f(A)\}. \end{aligned}$$

We must now show that the family $\mathcal{F}_M$ of all sets M_α satisfies the conditions for the finite intersection property. Let $A \subseteq G$. Then define $f_\alpha(A) = \mu_\alpha(A \cap G_\alpha)$. Clearly $f_\alpha(G) = \mu_\alpha(G_\alpha) = 1$. To show f_α is finitely additive, let $A_1, A_2 \subseteq G$ such that $A_1 \cap A_2 = \emptyset$. Then $(A_1 \cup A_2) \cap G_\alpha = (A_1 \cap G_\alpha) \cup (A_2 \cap G_\alpha)$ and $(A_1 \cap G_\alpha) \cap (A_2 \cap G_\alpha) = \emptyset$. So

$$f_\alpha(A_1 \cup A_2) = \mu_\alpha((A_1 \cup A_2) \cap G_\alpha = \mu_\alpha(A_1 \cap G_\alpha) + \mu_\alpha(A_2 \cap G_\alpha = f_\alpha(A_1) + f_\alpha(A_2).$$

To show f_α is left-variant with respect to G, for $g \in G_\alpha$,

$$f_\alpha(gA) = \mu_\alpha(gA \cap G_\alpha) = \mu_\alpha(gA \cap gG_\alpha)$$
$$= \mu_\alpha(g(A \cap G_\alpha)) = \mu_\alpha(A \cap G_\alpha) = f_\alpha(A).$$

Thus, $f_\alpha \in M_\alpha$, so M_α is non-empty.

To show M_α is closed, let $\nu \in [0,1]^{P(G)} \setminus M_\alpha$. Then either $\nu(G) < 1$, $\nu(A \cup B) \neq \nu(A) + \nu(B)$ for some disjoint $A, B \subseteq G$, or $\nu(gA) \neq \nu(A)$. If $\nu(G) < 1$, then

$$\left\{ f \in [0,1]^{P(G)} : \nu(G) \in \left[0, \nu(G) + \frac{1 - \nu(G)}{2}\right] \right\}$$

is an open set disjoint from M_α and ν belongs to it. If ν is not finitely additive, then there exist $A_1, A_2 \subseteq G$ such that $A_1 \cap A_2 = \emptyset$ and $\nu(A_1) + \nu(A_2) \neq \nu(A_1 \cup A_2)$. Then let

$$\epsilon_0 = \Big|[\nu(A_1) + \nu(A_2)] - \nu(A_1 \cup A_2)\Big|.$$

Then

$$\Big\{ f \in [0,1]^{P(G)} : \begin{array}{l} \nu(A_1) \in \left(\nu(A_1) - \frac{\epsilon_0}{4}, \nu(A_1) + \frac{\epsilon_0}{4}\right), \\ \nu(A_2) \in \left(\nu(A_2) - \frac{\epsilon_0}{4}, \nu(A_2) + \frac{\epsilon_0}{4}\right), \\ \nu(A_1 \cup A_2) \in \left(\nu(A_1 \cup A_2) - \frac{\epsilon_0}{4}, \nu(A_1 \cup A_2) + \frac{\epsilon_0}{4}\right) \Big\} \end{array}$$

is an open set disjoint from M_α and ν belongs to it. If for some $g \in G$ and $A \subseteq G$, $\nu(gA) \neq \nu(A)$, then let $\epsilon = |\nu(gA) - \nu(A)|$. Then

$$\left\{ f \in [0,1]^{P(G)} : \nu(gA) \in \left(\nu(gA) - \tfrac{\epsilon}{4}, \nu(gA) + \tfrac{\epsilon}{4}\right), \nu(A) \in \left(\nu(A) - \tfrac{\epsilon}{4}, \nu(A) + \tfrac{\epsilon}{4}\right) \right\}$$

is an open set disjoint from M_α and ν belongs to it. Thus M_α is closed.

Finally, for any $\alpha_1, \alpha_2, \ldots, \alpha_n \in I$ there exists a $\gamma \in I$ such that $G_{\alpha_1} \leq G_\gamma$, $G_{\alpha_2} \leq G_\gamma$, …, $G_{\alpha_n} \leq G_\gamma$, so

$$\emptyset \neq M_\gamma \subseteq M_{\alpha_1} \cap M_{\alpha_2} \cap \ldots \cap M_{\alpha_n},$$

so

$$\bigcap_{k\in\{1,2,\ldots,n\}} M_{\alpha_k} \neq \emptyset.$$

Thus, the family $\mathcal{F}_M$ satisfies the conditions for the finite intersection property, so

$$\bigcap \mathcal{F}_M \neq \emptyset.$$

Thus there exists a $\mu : G \longrightarrow [0,1]$ that is left-invariant, finitely additive, and has total measure 1, so G is amenable. $\square$

8. Examples of Amenable Groups

In this final section, we will show that the group of isometries in $\mathbb{R}$ and $\mathbb{R}^2$ are amenable, and by extending measures in $\mathbb{R}$ and $\mathbb{R}^2$, we will show that no bounded subset of in $\mathbb{R}$ and $\mathbb{R}^2$ with non-empty interior is paradoxical.

8.1. Isometries in $\mathbb{R}$ and $\mathbb{R}^2$.

Definition 8.1. *An isometry $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a distance-preserving function.*

Therefore, an isometry $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is of the form

$$f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

where the determinant of

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

is -1 or 1 [6, p.312].

Let G_n be the group of isometries in $\mathbb{R}^n$. Also, let T_n be the group of translations of G_n, i.e. those isometries of the form

$$t\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Furthermore, we will also be concerned with those isometries that preserve orientation. We will denote the group of operations that preserve orientation by SG_n, and we note

$$SG_n = \{f \in G_n : f(x) = Ax + t, det(A) = 1\}.$$

Example 8.2. *The group of isometries G_1 is amenable.*

Proof. First we note

$$G_1 = \{f : f(x) = ax + b, |a| = 1, b \in \mathbb{R}\}.$$

Also,

$$T_1 = \{f : f(x) = x + b\}.$$

We will argue that T_1 is a normal subgroup of G_1 by showing that $fT_1 = T_1 f$ for any $f \in G_1$. First, let $f' \in G_1$ so $f'(x) = a'x + b'$, and let $f_1 \in T_1$ so $f_1(x) = x + b$. Then

$$\begin{aligned} f'(f_1(x)) &= a'(x + b) + b' \\ &= (a'x + b') + a'b \\ &= f_2(f'(x)), \end{aligned}$$

where $f_2(x) = x + a'b$, so $f_2 \in T_1$. Thus $f'T_1 \subseteq T_1 f'$.

Similarly, let $f' \in G_1$ so $f'(x) = a'x + b'$ and let $f_1 \in T_1$ so $f_1(x) = x + b$. Then

$$\begin{aligned} f_1(f'(x)) &= (a'x + b') + b \\ &= a'(x + a'b) + b' \\ &= f'(f_2(x)), \end{aligned}$$

where $f_2(x) = x + a'b$, so $f_2 \in T_1$. Thus $T_1 f' \subseteq f'T_1$, and hence $f'T_1 = T_1 f'$, so T_1 is a normal subgroup of G_1.

Observe now that $T_1 \cong \mathbb{R}$ by the isomorphism $\phi : \mathbb{R} \longrightarrow T_1$ defined by $\phi(n) = x+n$. As $\mathbb{R}$ is an Abelian group, T_1 is also Abelian, so by Theorem 7.12, T_1 is amenable.

Next observe $G_1/T_1 = \{x, -x\}$. Therefore $G_1/T_1 \cong \mathbb{Z}_2$, so G_1/T_1 is also Abelian, and therefore, G_1/T_1 is amenable.

Finally, we may conclude G_1 is amenable by Theorem 7.11 as T_1 and G_1/T_1 are amenable. □

Example 8.3. *The group of isometries G_2 is amenable.*

Proof. To show G_2 is amenable, we will argue that SG_2 and G_2/SG_2 are amenable. First, observe

$$G_2 = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} : 0 \le \theta < 2\pi, a, b \in \mathbb{R} \right\}$$
$$\cup \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} : 0 \le \theta < 2\pi, a, b \in \mathbb{R} \right\}$$

[6, p.314].

Also note

$$SG_2 = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} : 0 \le \theta < 2\pi, a, b \in \mathbb{R} \right\}.$$

To show SG_2 is amenable, we will argue that T_2 and SG_2/T_2 are amenable. First we show that T_2 is a normal subgroup of SG_2 by showing that $fT_2 = T_2 f$ for any $f \in SG_2$. First, let $f' \in SG_2$ and $f_1 \in T_2$. Then

$$f' \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix},$$

and

$$f_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}.$$

Then

$$\begin{aligned} f' \circ f_1 \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} \right) + \begin{pmatrix} a' \\ b' \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix} \\ &= \left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix} \right) + \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = f_2 \circ f' \begin{pmatrix} x \\ y \end{pmatrix}, \end{aligned}$$

where

$$f_2\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}.$$

Thus $f'T_2 \subseteq T_2 f'$.

Similarly, let $f' \in SG_2$ and $f_1 \in T_2$, so

$$f'\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}a'\\b'\end{pmatrix}$$

and

$$f_1\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}a\\b\end{pmatrix}.$$

Then

$$\begin{aligned}
f_1\circ f'\begin{pmatrix}x\\y\end{pmatrix} &= \left(\begin{pmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}a'\\b'\end{pmatrix}\right)+\begin{pmatrix}a\\b\end{pmatrix}\\
&= \begin{pmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}+\begin{pmatrix}a'\\b'\end{pmatrix}\\
&= \begin{pmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{pmatrix}\left(\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}\right)+\begin{pmatrix}a'\\b'\end{pmatrix}\\
&= f'\circ f_2\begin{pmatrix}x\\y\end{pmatrix},
\end{aligned}$$

where

$$f_2\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}.$$

Thus $T_2 f' \subseteq f'T_2$, and hence $f'T_2 = T_2 f'$, so T_2 is a normal subgroup of SG_2.

We want to argue now that T_2 is amenable. If $f_1, f_2 \in T_2$, then

$$f_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$$
$$f_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}.$$

Then

$$\begin{aligned} f_1 \circ f_2 \begin{pmatrix} x \\ y \end{pmatrix} &= \left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \right) + \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \\ &= \left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \right) + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = f_2 \circ f_1 \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned}$$

Thus, T_2 is Abelian, and by Theorem 7.12, T_2 is amenable.

Next, we want to argue that SG_2/T_2 is amenable. Observe now that $SG_2/T_2 \cong SO_2$ where

$$SO_2 = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} : 0 \leq \theta < 2\pi \right\}.$$

by the isomorphism $\phi : SG_2/T_2 \longrightarrow SO_2$ defined by

$$\phi \left(\left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} \right) T_2 \right) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Observe now that SO_2 is Abelian as

$$\begin{aligned} &\begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 & -\cos\theta_1\sin\theta_2 - \sin\theta_1\cos\theta_2 \\ \sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2 & -\sin\theta_1\sin\theta_2 + \cos\theta_1\cos\theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix}. \end{aligned}$$

Thus, SG_2/T_2 is Abelian as well, and by Theorem 7.12, SG_2/T_2 is amenable.

Then, by Theorem 7.11, SG_2 is amenable, as T_2 and SG_2/T_2 are amenable.

Now, we will argue that SG_2 is a normal subgroup of G_2 by showing that $fSG_2 = SG_2f$ for any $f \in G_2$. So let $f \in G_2$. Then either

$$f \in SG_2 = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} : 0 \leq \theta < 2\pi, a, b \in \mathbb{R} \right\}$$

or

$$f \in \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} : 0 \leq \theta < 2\pi, a, b \in \mathbb{R} \right\}$$

As SG_2 is a group, if $f \in SG_2$, then $fSG_2 = SG_2f$. So suppose $f \notin SG_2$ and let $f' \in SG_2$. Then

$$\begin{aligned} f\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ \sin\theta_1 & -\cos\theta_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} \\ f'\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix}. \end{aligned}$$

Then

$$\begin{aligned} f \circ f' \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ \sin\theta_1 & -\cos\theta_1 \end{pmatrix} \left(\begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix} \right) + \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ \sin\theta_1 & -\cos\theta_1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a'' \\ b'' \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 & \sin\theta_1\cos\theta_2 - \cos\theta_1\sin\theta_2 \\ \sin\theta_1\cos\theta_2 - \cos\theta_1\sin\theta_2 & -(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a'' \\ b'' \end{pmatrix} \\ &= A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a'' \\ b'' \end{pmatrix}. \end{aligned}$$

Observe

$$\begin{aligned}
det(A) &= -(\cos^2\theta_1\cos^2\theta_2 + 2\cos\theta_1\cos\theta_2\sin\theta_1\sin\theta_2 + \sin^2\theta_1\sin^2\theta_2) \\
&\quad -(\sin^2\theta_1\cos^2\theta_2 - 2\cos\theta_1\cos\theta_2\sin\theta_1\sin\theta_2 + \cos^2\theta_1\sin^2\theta_2) \\
&= -(\cos^2\theta_1\cos^2\theta_2 + \cos^2\theta_1\sin^2\theta_2) + \sin^2\theta_1\cos^2\theta_2 + \sin^2\theta_1\sin^2\theta_2) \\
&= -[\cos^2\theta_1(\cos^2\theta_2 + \sin^2\theta_2) + \sin^2\theta_1(\cos^2\theta_2 + \sin^2\theta_2)] \\
&= -(\cos^2\theta_1 + \sin^2\theta_1) \\
&= -1.
\end{aligned}$$

Thus, $f \circ f' \in G_2 \setminus SG_2$, and as SO_2 is Abelian, $f' \circ f \in G_2 \setminus SG_2$. In any case, $fSG_2 = SG_2 f$, so SG_2 is a normal subgroup of G_2.

Finally, observe

$$G_2/SG_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \cong \mathbb{Z}_2.$$

Thus, G_2/SG_2 is Abelian as $\mathbb{Z}_2$ is Abelian, and therefore, by Theorem 7.12, G_2/SG_2 is amenable.

Thus, by Theorem 7.11, G_2 is amenable, as SG_2 and G_2/SG_2 are amenable. □

8.2. **No Paradoxes in $\mathbb{R}$ or $\mathbb{R}^2$.** Having shown that G_1 and G_2 are amenable, our goal now is to find left-invariant measures on $P(\mathbb{R})$ and $P(\mathbb{R}^2)$ to show no bounded subset with non-empty interior of $\mathbb{R}_1$ or $\mathbb{R}_2$ can be paradoxical with respect to G_1 and G_2 respectively.

Let $\mathcal{R}_1 \subseteq P(\mathbb{R})$ be the ring formed by $\emptyset$ and the union of any finite number of disjoint intervals (open, closed, or half open and closed).

Proposition 8.4. *$\mathcal{R}_1$ is a ring.*

Proof. Immediate, as the union of two intervals is an interval or a union of intervals, and the difference of two intervals is either an interval, the union of intervals, or empty. □

Definition 8.5. *Let $\mu_1 : \mathcal{R}_1 \longrightarrow [0, \infty]$ be defined by*

$$\mu_1(I_1 \cup I_2 \cup \ldots \cup I_n) = \sum_{i=1}^{n}(b_i - a_i),$$

where I_i has endpoints $a_i \leq b_i$ and $I_i \cap I_j = \emptyset$ for $i \neq j$.

Observe

- for any $X, Y \in \mathcal{R}_1$ such that $X \cap Y = \emptyset$,

$$\mu(X) + \mu(Y) = \mu(X \cup Y)$$

- for any $g \in G_1$ and $X \in \mathcal{R}_1$,

$$\mu(gX) = \mu(X)$$

Corollary 8.6. *No bounded subset $X \subseteq \mathbb{R}$ with non-empty interior is paradoxical.*

Proof. By Theorem 7.7, there exists a left-invariant finitely additive measure $\bar{\mu} : P(\mathbb{R}) \longrightarrow [0, \infty]$ that extends μ_1 defined in 8.5. Let $X \subseteq \mathbb{R}$ be a bounded set with non-empty interior. Then for some $a, b, c, d \in \mathbb{R}$,

$$[a, b] \subseteq X \subseteq [c, d].$$

Thus,

$$0 < \bar{\mu}_1([a, b]) \leq \bar{\mu}_1(X) \leq \bar{\mu}_1([c, d]) < \infty,$$

so $\bar{\mu}_1(X) \in (0, \infty)$. Hence, by Theorem 7.5, X is not paradoxical. □

Corollary 8.7. *No bounded subset $X \subseteq \mathbb{R}^2$ with non-empty interior is paradoxical.*

Proof. Let c be the Jordan measure, and let $M(c) \subseteq P(\mathbb{R}^2)$ be the class of all Jordan measurable sets. Then $M(c)$ forms a sub-ring of $P(\mathbb{X})$. Also, c is left-invariant with respect to G_2 (see [7]). Thus, by Theorem 7.7, there exists a left-invariant finitely additive measure $\bar{c} : P(\mathbb{R}^2) \longrightarrow [0, \infty]$ that extends c. So let $X \subseteq \mathbb{R}^2$ be a bounded set with non-empty interior. Then

$$[a_1, b_1] \times [a_2, b_2] \subseteq X \subset [c_1, d_1] \times [c_2, d_2].$$

Thus,

$$0 < \bar{c}([a_1, b_1] \times [a_2, b_2]) \leq \bar{c}(X) \leq \bar{\mu}_1([c_1, d_1] \times [c_2, d_2]) < \infty,$$

so $\bar{c}(X) \in (0, \infty)$. Hence, by Theorem 7.5, X is not paradoxical. $\square$

9. Conclusion

In this thesis, we have given a detailed decomposition of the sphere into two identical spheres. This decomposition of a sphere runs contrary to every intuitive notion we have concerning volume, and therefore is aptly named the Banach-Tarski paradox. We are able to reconcile this seemingly false result by acknowledging the existence of non-measurable sets. In 3-dimensional space, sets can be formed, as we have shown, that are not measurable. The second half of this thesis showed that a similar paradox is not possible on the number line or the plane because every set with non-empty interior can be measured.

References

[1] Banach, Stefan; Tarski, Alfred. Sur la décomposition des ensembles de points en partiens respectivement congruent. *Fundamenta Mathematicae* 6: 244-277, 1924. `http://matwbn.icm.edu.pl/ksiazki/fm/fm6/fm6127.pdf`.

[2] Dougherty, Randall. Foreman, Matthew. Banach-Tarski decompositions using sets with the property of Baire. *Journal of the American Mathematical Society.* Volume 7, 1994.

[3] Galilei, Galileo. *Dialogues concerning two new sciences.* Translated by Henry Crew and Alfonso de Salvio. Northwestern University Press, Evanston, 1950. Originally published 1638.

[4] Halmos, Paul. *Measure Theory.* Springer-Verlag, New York, 1978.

[5] Hausdorff, Felix. *Grundzüge der Mengenlehre.* Veit, Leipzig, 1914.

[6] Hausner, Melvin. *A Vector Space Approach to Geometry.* Dover Publications, New York, 1998.

[7] Johnson, Matt. *The Jordan Measure.* Kerrigan Research Project, University of Nebraska – Omaha, 2010.

[8] Lang, Serge. *Algebra.* Addison-Wesley, Massachusetts, 1971.

[9] Mazurkiewicz, Stefan. Sierpiński, Wacław. Sur un ensemble superposables deux a deux. *C. R. Acad. Sci. Paris.* 1914.

[10] Munkres, James. *Topology.* Prentice Hall, New Jersey, 2000.

[11] O'Conner, John; Robertson, Edmund. Stefan Banach. *The MacTutor History of Mathematics archive.* `http://www-history.mcs.st-andrews.ac.uk/history/Biographies/Banach.html`. 2009.

[12] —, —. Felix Hausdorff. *The MacTutor History of Mathematics archive.* `http://www-history.mcs.st-andrews.ac.uk/history/Biographies/Hausdorff.html`. 2009.

[13] —, —. Alfred Tarski. *The MacTutor History of Mathematics archive.* `http://www-history.mcs.st-andrews.ac.uk/history/Biographies/Tarski.html`. 2009.

[14] Pawlikowski, Janusz. The Hahn-Banach theorem implies the Banach-Tarski paradox. *Fundamenta Mathematicae* 138: 21-22, 1991.

[15] Royden. *Real Analysis*

[16] Vitali, Giuseppe. *Sul problema della mesura dei gruppi di punti di una retta.* Bologna, 1905.

[17] Wagon, Stan. *The Banach-Tarski Paradox.* Cambridge University Press, New York, 1994.

[18] Wapner, Leonard. *The Pea and the Sun: A Mathematical Paradox.* A K Peters, Massachusetts, 2007.