2001

The Banach-Tarski paradox in euclidean spaces

Kenneth R. Hoover
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

---

Recommended Citation

Hoover, Kenneth R., "The Banach-Tarski paradox in euclidean spaces" (2001). *Master's Theses*. 2182.
DOI: <https://doi.org/10.31979/etd.ttvv-t3b8>
https://scholarworks.sjsu.edu/etd_theses/2182

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

NOTE TO USERS

This reproduction is the best copy available.

UMI

**THE BANACH-TARSKI PARADOX
IN EUCLIDEAN SPACES**

A Thesis

Presented to

**The Faculty of the Department of Mathematics
and Computer Science
San Jose State University**

In Partial Fulfillment

**of the Requirements for the Degree
Master of Science**

by

Kenneth R. Hoover

May 2001

UMI Number: 1405504

UMI®

UMI Microform 1405504

Copyright 2001 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2001

Kenneth R. Hoover

ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF
MATHEMATICS AND COMPUTER SCIENCE

Dr. Donald Weddington

Dr. Brian Peterson

Dr. Richard Pfeifer

APPROVED FOR THE UNIVERSITY

ABSTRACT

THE BANACH-TARSKI PARADOX IN EUCLIDEAN SPACES

by Kenneth R. Hoover

Proven in 1924 by Stefan Banach and Alfred Tarski, the Banach-Tarski Paradox states that any solid ball in $\mathbb{R}^3$ is paradoxical. Thus, any solid ball can be decomposed in such a way that two separate copies of the ball can be obtained by rearranging its subsets. Furthermore, the theorem can be strengthened to state that any two bounded subsets of $\mathbb{R}^3$ with nonempty interior are equidecomposable, meaning that one set can be decomposed and its subsets rearranged to obtain the other set. Although the Banach-Tarski Paradox fails to hold in the spaces $\mathbb{R}$ and $\mathbb{R}^2$, an analogous statement can be proven for any Euclidean space of dimension three or higher, effectively proving that there is no finitely additive extension of Lebesgue measure to all subsets of $\mathbb{R}^n$, with n equal to or greater than three.

*For my loving wife,
with gratitude for her patience
over the years.*

CONTENTS

Preface: The Axiom of Choice	vii
<i>Part One: The Banach-Tarski Paradox In Higher Dimensions</i>	1
Chapter 1: Free Groups and Paradoxical Sets	2
Chapter 2: The Hausdorff Paradox	13
Chapter 3: The Banach-Tarski Paradox	22
Chapter 4: The Banach-Tarski Paradox In Euclidean Spaces Higher Than $\mathbb{R}^3$	36
<i>Part Two: The Banach-Tarski Paradox In Lower Dimensions</i>	44
Chapter 5: Lebesgue Measure	45
Chapter 6: Lebesgue Measure In $\mathbb{R}^2$	57
Chapter 7: The Solvability of G_1 and G_2	69
Chapter 8: The Banach-Tarski Paradox In $\mathbb{R}$ and $\mathbb{R}^2$	76
Conclusion	83
Appendix A: A Lebesgue Nonmeasurable Set in $\mathbb{R}$	86
Appendix B: Useful Facts Regarding Integration	89
Works Cited	91

Preface:

The Axiom of Choice

The beauty of mathematics is found in the elegance of its theories, which flow logically from axioms into theorems and build upon one another to form a wide array of interesting and amazing results. Its power is in the fact that these theories and results can be applied so accurately to the world in which we live to help us better understand our universe and our reality. Some would say that the entire point of studying mathematics is to understand and perhaps further its applications, while others would argue that the study of mathematics is, in itself, a noble endeavor. This is the means vs. ends debate between pure and applied mathematics. The ultimate goal of the mathematical community is the strengthening of the bond between these two paradigms, uniting the beauty and power of mathematics into one. In doing this, one hopes that every new mathematical discovery will teach us more, not only about the universe of mathematics, but also about our own world.

As we study to understand the works of those who have come before us and to possibly make our own discoveries, it is important to keep in mind what mathematics really is. It is simply a series of logical steps rooted in sets of axioms that we consider to be “true”. It is the choice of axioms that is all-important in the study of mathematics. If one were studying Euclidean geometry, he or she would use Euclid’s five postulates (axioms). If studying the whole numbers, one may use Peano’s axioms, and so forth.

There have been many different sets (or systems) of axioms put forth over the centuries. Some are used in studying different areas of mathematics while others are refinements of systems proposed earlier. Some of these axiomatic systems (like those of non-Euclidean geometries) lend themselves to studies of a more ephemeral and abstract nature whose applications to the real world are not necessarily apparent. On the other hand, some systems of axioms have been developed specifically to describe those areas in which math is often used in the practice of applications. One characteristic of all axioms, however, is that they are not able to be proven, but are considered to be “true”. This notion of truth is a more complicated one than that of the usual meaning of the word. In theoretical mathematics, the choice of axioms may be a function of how interesting the results are that can be derived from them. In applied mathematics, axioms tend to be chosen based on how well they match our intuition or represent the ideas we wish to study. For instance, Euclid’s five postulates govern the relationships that we encounter between shapes and objects all around us, and Peano’s axioms lead to the rules that we use in counting every day. In any situation, if there is any argument over the validity of any axiom, then the debate is not over whether or not the axiom is true, but whether or not it should be an axiom in the first place.

One area of great discussion and study over the past century has been that of set theory. The axioms of set theory were formulated intuitively by Cantor in the late nineteenth century and then treated more rigorously by Zermelo and others in the early twentieth century. Now known as Zermelo-Fraenkel set theory, this system of axioms was meant to be a unifying theory that would bring together and serve as a foundation for

all areas of mathematics (Burton 615-616). However, it has brought with it a certain amount of controversy. This controversy has been centered mainly on a Trojan horse of sorts known as the Axiom of Choice.

First mentioned formally in 1902 by Beppo Levi (Burton 615), the Axiom of Choice states that given a collection of nonempty sets, one element (called a choice element) can be chosen from each set (Schechter). The axiom got its name two years later from Ernst Zermelo who used it to prove Cantor's earlier statement that every nonempty subset of a well-defined set has a first element with respect to some ordering. This claim, now known as the Well-Ordering Principle, has been shown to be equivalent to the Axiom of Choice (Burton 615).

Why has there been such controversy over this statement? It seems intuitive that if every set in the collection is nonempty, i.e. contains at least one element, then we should be able to choose an element from each one. The debate centered on the ability to back this claim up by actually defining the choice function being used. It seemed to some, the constructivists, that unless one actually specified how the elements were to be chosen, then he or she was just waving their hands, so to speak. Bertrand Russell exemplified the difference between these two points of view in his statement, "To choose one sock from each of infinitely many pairs of socks requires the Axiom of Choice, but for shoes the Axiom is not needed (Schechter)." The difference being that, since the shoes are distinguishable from each other, one can define that the left or the right shoe is to be chosen from each pair. However, since the socks are not distinguishable, an explicit algorithm cannot be defined, the socks must be chosen arbitrarily.

The debate concerning the Axiom of Choice, however, was not merely over semantics. Although the Axiom simplified a lot of mathematics involving infinite collections of sets, it also opened the door to some completely counter-intuitive notions, namely, the Banach-Tarski Paradox. In 1924, Stefan Banach and Alfred Tarski, mathematicians from the area that is now Poland, published a paper indicating that a solid ball can be broken up in such a way that one exact copy of the original sphere can be reassembled from some of the pieces, and another exact copy can be reassembled from those remaining (O'Connor, Alfred Tarski, Stefan Banach). Another consequence of this theorem is that any two bounded sets with nonempty interior in Euclidean three (or higher)-dimensional space are equivalent by finite decompositions (Wagon xi). What this means is that one set can be disassemble into finitely many pieces which can then be rearranged via translations and rotations and reassembled to form the other set with no gaps or overlapping. What's amazing about this (and the reason why it is called a paradox), is that it holds for any two such sets, regardless of their relative sizes. Thus, since our world is roughly considered to be a three-dimensional Euclidean space¹, we can imagine taking a tiny pebble, cutting it up, and reassembling the pieces into a rock the size of the earth.

Certainly these findings are astounding, indeed. If this is to be true, then we have found a solution to most all problems involving shortages ever facing us. Whether it be riches, food, land, etc., all that needs to be done is for the material to be duplicated vis-à-

---

¹ Further study by Einstein and others have shown that our universe is more accurately categorized as a Minkowski four-dimensional space (Einstein 61). However, the properties of Euclidean three-space accurately describe our world well enough for most mathematical applications. It is not until one operates in certain areas of mathematical physics that the validity of the model comes into question.

vis the Banach-Tarski Paradox. Of course, we know that it is not that simple. The hidden difficulty is that the proof of the Banach-Tarski Paradox relies on the use of the Axiom of Choice. This alludes to the fact that what Banach and Tarski have done is prove that such duplications are theoretically possible, although an explicit algorithm for exactly how to perform them cannot be specified. Much the same as saying that one member from every one of infinitely many sets can be chosen without giving an explicit algorithm for how to choose them.

As can be imagined, the existence of the Banach-Tarski Paradox sent shudders throughout the mathematical community. Especially in such applied areas as mathematical physics. The very notion of enlarging the size of an object through simple translations and rotations seems to contradict the Law of Conservation of Mass. At the very least, it goes against all intuition. The fact that it is allowed by the Axiom of Choice caused some applied mathematicians to join the constructivists in denouncing the use of the Axiom.

However, the fact that a theory leads to a contradiction of our intuition or common experience is not necessarily cause to abandon it. It may simply mean that our intuition and common experience do not reflect reality as well as previously thought. An example is Einstein's Special Theory of Relativity. According to this theory, if a plane were to fly around the world, the amount of time that would pass on the plane would be slightly less than the time that it takes the plane to complete the trip as measured on the ground (Einstein 40-42). (Of course, for this time difference to be noticeable, the plane would have to fly a much greater distance and at a speed near the speed of light.) This

goes completely against all intuition, as well as the laws of classical mechanics, which have been used to describe motion in our universe for hundreds of years. However, as paradoxical as it seems (in fact, one formulation of this situation is often referred to as the Twin Paradox (Bunch 189)), experiment upon experiment have shown it to be accurate. As our abilities to reach higher speeds and make more accurate measurements have increased, it has been found that Einstein's Theory of Relativity more accurately describes the physical relationships in our universe than does classical mechanics (Einstein 55-57). Does this mean that classical mechanics is worthless? Not at all. It simply means that the laws of classical mechanics have their place in calculations that do not involve astronomical distances being covered at very high speeds. They still form a very suitable model for our world even though a slightly better model exists that is better used in extreme situations. In other words, our pre-relativity ideas about physics may have been slightly naive, but are still suitable for most all applications as long as we are mindful of their limitations.

As it is, we should accept that we might have a similar case with the Banach-Tarski Paradox. It appears to violate the Law of Conservation of Mass by doubling the volume of a sphere without changing the density (the pieces are not stretched in any way, they are only translated and rotated, so the density remains unaffected). Thus, it would seem that the mass is somehow doubled. However, this is not necessarily the case if we accept the fact that our notions of volume may also be somewhat naïve and not as accurate as we may think in extreme situations.

As it so happens at least one of the pieces in the Banach-Tarski decomposition must have the quality that it is not Lebesgue measurable. When sets are not Lebesgue measurable, their measures are not necessarily fixed under translations and rotations (i.e. their measures are not necessarily translation-rotation invariant), and the sum of their measures is not necessarily the measure of their union (i.e. their measures are not necessarily additive). Thus, the “size” and “shape” of a Lebesgue nonmeasurable set is pathological at best. Hence, when we start dealing with such sets, our ordinary ideas about volume may not apply as accurately as we are accustomed to expect. We must, therefore, keep in mind the possibility that the Banach-Tarski Paradox doesn’t contradict the Law of Conservation of Mass, but rather deals with mass in a way we do not yet fully understand (Schechter).

With this in mind, and with consideration that the Axiom of Choice seems intuitively acceptable in itself (even if some of its consequence do not) and greatly simplifies so much mathematics, most mathematicians are willing to accept and use it. Although some mathematicians and physicists are somewhat uncomfortable with the idea and prefer to use some weaker notion such as Countable Choice, the great debate over the Axiom of Choice has, for the most part, been put to rest in most mathematical circles. The Axiom is now used regularly by most mathematicians without questioning its acceptability. However, its questionable beginnings will always be remembered as a great mathematical debate.

Part One :

**The Banach-Tarski
Paradox In
Higher Dimensions**

Chapter 1:

Free Groups and Paradoxical Sets

Since much of the work to produce the Banach-Tarski Paradox involves free groups acting on sets and the idea of a set being paradoxical, the first order of business is in defining these notions.

Definition (Group action on a set)

1.1

Suppose that G is a group and S is a set. G acts on S if, for every $g \in G$, there exists a function $f_g : S \rightarrow S$ such that:

$$a) f_{gh}(s) = f_g(f_h(s)) \text{ for all } g, h \in G \text{ and all } s \in S,$$

$$b) f_i(s) = s \text{ for all } s \in S, \text{ where } i \text{ is the identity element of } G.$$

(Fraleigh 197)

For an example of a group action on a set, consider the special orthogonal group SO_3 which is the group of orthogonal 3×3 matrices with determinant 1 (Wagon 249) and represents the rotations of $\mathbb{R}^3$ about lines through the origin. It can be shown that this group acts on $\mathbb{R}^3$ where the function f_A is simply left multiplication by the matrix A . The fact that $f_{AB}(\mathbf{x}) = f_A(f_B(\mathbf{x}))$ for all $A, B \in SO_3$ and all $\mathbf{x} \in \mathbb{R}^3$ follows from the fact that matrix multiplication is associative. Furthermore, since the identity element of SO_3 is the 3×3 identity matrix I_3 , the fact that $f_{I_3}(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^3$ follows from the well-known multiplicative property of the identity matrix. Thus, it can be said that SO_3 acts on $\mathbb{R}^3$ as prescribed in the above definition.

That having been said, it is helpful to look at this example in a different light. The fact that $f_{AB}(\mathbf{x}) = f_A(f_B(\mathbf{x}))$ for all $A, B \in SO_3$ and all $\mathbf{x} \in \mathbb{R}^3$ reflects the fact that performing the rotation on $\mathbb{R}^3$ represented by the matrix B , followed by the rotation represented by the matrix A , produces the same result as performing the rotation that is the composition of these two, which would be represented by the matrix AB . Furthermore, since the identity rotation represented by the matrix I_3 is the rotation of $\mathbb{R}^3$ about any line through the origin through an angle of 0° , the fact that $f_{I_3}(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^3$ simply reflects the fact that rotating a vector in $\mathbb{R}^3$ through an angle of 0° produces no change.

Seen in this light, it is more clear why the situation defined in definition 1.1 is called *group action* on a set. In the example given here, SO_3 is actually performing an action on 3-dimensional Euclidean space. It is taking the space and rotating it.

Notice that a subgroup acts on its parent group in a very natural way where $f_{g_1}(g_2) := g_1 * g_2$ (with $*$ being the operation of the group). Condition a) is satisfied by the associativity of $*$, and condition b) is satisfied by the fact that the identity element of the subgroup is also the identity element of the group (Fraleigh 199).

If group G acts on set S , then for each $g \in G$, the mapping f_g is not only a function from S into itself, but a permutation of S (i.e. a bijection from S onto itself), as the following proposition shows.

Proposition

1.2

If G is a group that acts on set S , then for each $g \in G$, the function f_g is a

permutation of S (Fraleigh 198).

Proof Let $s \in S$. Then $f_{g^{-1}}(s) \in S$ and $f_g(f_{g^{-1}}(s)) = f_i(s) = s$. Hence, f_g maps S onto S .

Also, suppose $f_g(s) = f_g(t)$ for some $s, t \in S$. Then

$$s = f_i(s) = f_{g^{-1}g}(s) = f_{g^{-1}}(f_g(s)) = f_{g^{-1}}(f_g(t)) = f_{g^{-1}g}(t) = f_i(t) = t.$$

Hence, f_g is one-to-one.

Thus, f_g is a permutation of S .

□

(For ease of notation, $f_g(s)$ will be referred to simply as $g(s)$ from this point forward.)

The following definition describes what it means for a set to be paradoxical.

Definition (*Paradoxical set*)

1.3

Suppose G is a group that acts on set S , and let $X \subseteq S$. If there exist pairwise disjoint subsets $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ of X and some $g_1, \dots, g_p, h_1, \dots, h_q \in G$ (for some $p, q \in \mathbb{N}$) such that $X = \bigcup_{i=1}^p g_i(Y_i)$ and also $X = \bigcup_{j=1}^q h_j(Z_j)$, then X is said to be paradoxical with respect to G (or, simply, G -paradoxical) (Wagon 4).

The idea here is that if X is G -paradoxical, then there must be two disjoint subsets $Y = \bigcup_{i=1}^p Y_i$

and $Z = \bigcup_{j=1}^q Z_j$ of X which can each be rearranged piecewise according to the action of G on S so

that each one covers X . Note that, if necessary, the sets in $\{Y_i\}_{i=1}^p$ and $\{Z_j\}_{j=1}^q$ can be scaled

down so that each of $\{g_i(Y_i)\}_{i=1}^p$ and $\{h_j(Z_j)\}_{j=1}^q$ are pairwise disjoint. In general cases, it will

be assumed that this has already been performed.

A nice example of a paradoxical set will be discussed in the next chapter vis-à-vis the Hausdorff Paradox, which shows that the 3-dimensional unit sphere is, for the most part, paradoxical with respect to rotations about the origin. For now, however, notice how the concept of a paradoxical set applies to the Banach-Tarski Paradox as stated in the preface. In one form, the Paradox claims that a 3-dimensional unit sphere can be dissected in such a way so that some of the pieces can be reassembled into the entire sphere, as can the remaining pieces, using only rotations and translations. This is usually referred to as duplicating the sphere. However, to prove this claim, it must still be shown that the unit sphere is indeed paradoxical with respect to the set of isometries of $\mathbb{R}^3$, and that the sets $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ actually do form a partition of X . (Notice that the definition of paradoxical only states that these sets are pairwise disjoint subsets of X , which is not enough to conclude that they are a partition of X , for there may be elements of X which are not included in any of these sets.) Doing so does require a bit of work, which will be covered over the course of the next two chapters.

Let $\mathcal{L} = \{l_i\}_{i \in I}$ be a set of elements where I is some (possibly infinite) indexing set. Call the elements l_i “letters” and the set of letters, $\mathcal{L}$, an “alphabet”. Similarly, call any l_i^k where $i \in I, k \in \mathbb{Z}$ a “syllable”. Then any juxtaposition of finitely many syllables is a “word”. Note that this allows for a word made up of 0 syllables. This word is called the “empty word” and is represented by the number 1 (Fraleigh 238).

Certain words may be able to be reduced to shorter words. For example, suppose

$\mathcal{L} = \{l_1, l_2, l_3, l_4\}$, then $l_3^5 l_2^{-3} l_2^7 l_4^{-7} l_1^2$ is a word. However, using the reducing operations discussed below, the word $l_3^5 l_2^{-3} l_2^7 l_4^{-7} l_1^2$ can be reduced to $l_3^5 l_2^{-1} l_1^2$. The two types of reductions allowed are replacing $l_i^p l_i^q$ by l_i^{p+q} , and replacing l_i^0 by 1. One can see that the word $l_3^5 l_2^{-3} l_2^7 l_4^{-7} l_1^2$ was reduced to $l_3^5 l_2^{-1} l_1^2$ by means of the following steps:

$$l_3^5 l_2^{-3} l_2^7 l_4^{-7} l_1^2 \rightarrow l_3^5 l_2^{-3+2} l_4^{7+(-7)} l_1^2 \rightarrow l_3^5 l_2^{-1} l_4^0 l_1^2 \rightarrow l_3^5 l_2^{-1} 1 l_1^2 \rightarrow l_3^5 l_2^{-1} l_1^2.$$

If a word cannot be reduced further, then it is called a “reduced word” (Fraleigh 238).

Definition (Free group)

1.4

Let $\mathcal{L}$ be an alphabet and let F be the set of all reduced words from $\mathcal{L}$. Then the group formed by F , along with the operation of concatenation, is called the free group generated by $\mathcal{L}$, and $\mathcal{L}$ is called a free generating set of F (Fraleigh 239).

The number of elements in a free generating set for a free group F is called the *rank* of F (Fraleigh 240). This number is unique (Fraleigh 239) since it can be proved that if $A = \{a_i\}_{i \in I}$ and $B = \{b_j\}_{j \in J}$ are both free generating sets for the free group F , then $|A| = |B|$. Note that the cardinality of F will always be countably infinite (unless, of course, the generating set of F is uncountable).

For an example of a free group, let θ be the counterclockwise rotation of the plane $\mathbb{R}^2$ about the origin through an angle of $\pi\sqrt{2}$ radians. Then θ^{-1} is the clockwise rotation through an angle of the same magnitude. Let F be the group generated by θ . Then F is the set of all rotations of $\mathbb{R}^2$ through angles of $n\pi\sqrt{2}$ radians for any $n \in \mathbb{Z}$. To see that F is a free group, it must be

shown that it is countably infinite. In other words, one must show that no nontrivial rotation in F equals the identity rotation of 0 radians. To see this, suppose that θ_I is a rotation of $n\pi\sqrt{2}$ radians (for some $n \neq 0$) that is equivalent to the identity rotation. Then $n\pi\sqrt{2} = 2m\pi$ for some $m \in \mathbb{Z}$, which implies that $n = \frac{2m}{\sqrt{2}} = m\sqrt{2} \notin \mathbb{Z}$. Thus, $\theta_I \notin F$. Hence, every element of F is a distinct rotation and F is, therefore, countably infinite. Therefore, F is a free group of rank 1 generated by θ .

Of course, the concept of free groups can be extended to encompass groups other than those mentioned above. Suppose G is a group generated by the elements of some set $X = \{x_i\}_{i \in I}$ and F is the free group generated by X . If G is isomorphic to F under the unique homomorphism $\sigma : F \rightarrow G$ defined by $\sigma(x_i) := x_i$, then G is also called free. Thus the term “free group” may also be used for a group that is isomorphic to a free group (Fraleigh 239).

Suppose F is the free group generated by $\mathcal{L}$. Notice that the words in F are composed of more than just the letters in $\mathcal{L}$. Namely, some words in F contain the inverses of letters from $\mathcal{L}$, even though those inverses are not, necessarily, elements of $\mathcal{L}$ themselves. Consider, then, how the notion of a free group would change if one allowed only the letters of $\mathcal{L}$ to be used. The most notable difference is that F would no longer be a group due to the lack of inverses. However, one would still have 1 as the word composed of 0 letters from $\mathcal{L}$, and the operation of concatenation would still be associative. Thus, F would become a semigroup.

Definition (*Free semigroup*)

1.5

Let $\mathcal{L}$ be an alphabet and let $\bar{F}$ be the set of all words composed strictly of letters

from $\mathcal{L}$. Then the set F along with the operation of concatenation is called the free semigroup generated by $\mathcal{L}$, and $\mathcal{L}$ is called a free generating set for F .

An easy example of a free semigroup is to take the free group F generated by θ that was previously described and restrict it to include only rotations composed of repeated applications of θ , without use of θ^{-1} . Then F would be the set of rotations of $\mathbb{R}^2$ through angles of $n\pi\sqrt{2}$ radians for $n \in \mathbb{N}$, along with the identity rotation of 0 radians. It can be seen that for any rotation ρ in F , ρ^{-1} is not in F . To demonstrate this fact, let ρ be a counterclockwise rotation of $n\pi\sqrt{2}$ radians for some $n \in \mathbb{N}$, and let ρ^{-1} be a counterclockwise rotation of $m\pi\sqrt{2}$ radians for some $m \in \mathbb{N}$. It must be the case that $(n\sqrt{2} + m\sqrt{2})\pi = 2k\pi$ for some $k \in \mathbb{N}$ since $\rho\rho^{-1}$ is the identity rotation. However, $(n\sqrt{2} + m\sqrt{2})\pi = 2k\pi \Rightarrow m = \frac{2k - n\sqrt{2}}{\sqrt{2}} = k\sqrt{2} - n \notin \mathbb{N}$. Thus, $\rho^{-1} \notin F$. Since the rotations in F do not have inverses in F , but F does retain the other characteristics of a free group, then F is a free semigroup of rank 1 generated by θ .

It is natural to think of a semigroup acting on a set in much the same way that a group would. However, a notable difference is that if $\overline{G}$ is a semigroup “acting” on set S , then the function $f_{\overline{g}} : S \rightarrow S$ for some $\overline{g} \in \overline{G}$ may not be a permutation of S due to the possible absence of $\overline{g}^{-1}$ in $\overline{G}$ (Wagon 6).

Proposition

1.6

If F is a free group, then F acts on itself by concatenation (Wagon 5).

Proof Let $\omega_i, \omega_j, \omega_k$ be any three words in F . Notice that concatenating ω_i onto the

front of ω_j and then concatenating that word onto the front of ω_k would produce $\omega_i\omega_j\omega_k$. However, the same word can be produced by concatenating ω_j onto the front of ω_k and then concatenating ω_i onto the front of $\omega_j\omega_k$. Hence,

$$(\omega_i\omega_j)(\omega_k) = \omega_i(\omega_j\omega_k) \text{ for all } \omega_i, \omega_j, \omega_k \in F.$$

Furthermore $1\omega_i$ is the word formed by concatenating 0 syllables onto the front of ω_i , which is simply ω_i . Hence, $1\omega_i = \omega_i$ for all $\omega_i \in F$.

Thus, F acts on itself by concatenation.

□

Since free groups act on themselves, it is possible, then, for some free groups to be paradoxical with respect to themselves. One example of just such a group is given below.

Proposition

1.7

If F is a free group of rank 2, then F is paradoxical with respect to itself (Wagon 5).

Proof Let $\mathcal{X}$ be a free generating set of F . Let W_x represent the subset of F containing exactly those words in F that begin with x . Since F is of rank 2, $\mathcal{X}$ must contain exactly 2 elements, call them δ, φ . Notice that any nonempty word in F must begin with exactly one of $\delta, \delta^{-1}, \varphi, \varphi^{-1}$. Hence, the sets $W_\delta, W_{\delta^{-1}}, W_\varphi$, and $W_{\varphi^{-1}}$ are pairwise disjoint and $\{1\} \cup W_\delta \cup W_{\delta^{-1}} \cup W_\varphi \cup W_{\varphi^{-1}} = F$.

Furthermore, if $f \in F$ is a word in F that does not begin with δ , then $\delta^{-1}(f)$ begins with δ^{-1} . Hence, $\delta^{-1}(f) \in W_{\delta^{-1}}$. Then

$f = 1(f) = (\delta\delta^{-1})(f) = \delta(\delta^{-1}(f)) \in \delta(W_{\delta^{-1}})$. Thus, for all $f \in F, f \in W_\delta$ or $f \in \delta(W_{\delta^{-1}})$. Therefore, $F = W_\delta \cup \delta(W_{\delta^{-1}})$.

Similarly, $F = W_\varphi \cup \varphi(W_{\varphi^{-1}})$.

Thus, F is paradoxical with respect to itself.

□

Thus, some groups are, in fact, paradoxical with respect to themselves. This collection of groups will prove to be very important for the Banach-Tarski Paradox. The reason for this is because the Banach-Tarski Paradox follows rather elegantly from an earlier theorem known as the Hausdorff Paradox. As the reader will see in the next chapter, it is one of these self-paradoxical groups that is the key to the Hausdorff Paradox. However, in addition to the existence of such a group one will need the following result.

Proposition

1.8

Let G be a paradoxical group with respect to itself and suppose that G acts on the set S in such a way that for all $s \in S$ and for all nonidentity $g \in G$, we have $g(s) \neq s$. Then S is paradoxical with respect to G (Wagon 11).

Proof Suppose that the paradoxical decomposition of G is brought about by

$Y_1, \dots, Y_p, Z_1, \dots, Z_q \subseteq G$ and $g_1, \dots, g_p, h_1, \dots, h_q \in G$ such that

$$G = \bigcup_{i=1}^p g_i(Y_i) \text{ and } G = \bigcup_{j=1}^q h_j(Z_j).$$

For each $s \in S$, consider the orbit of s under G , which is the set

$\{x \in S : x = \omega(s) \text{ for some } \omega \in G\}$. Invoking the Axiom of Choice, we can create a

choice set C consisting of one element from each of these orbits.

Now consider the set $\{g(C) : g \in G\}$. Suppose $g(C) \cap h(C) \neq \emptyset$ for some $g, h \in G$. Then $g(c_1) = h(c_2)$ for some $c_1, c_2 \in C$. Thus, $c_1 = (g^{-1}h)(c_2)$, implying that c_1 is in the same orbit under G as c_2 . Since C only contains one element from

each orbit, then $c_1 = c_2$. Hence, $c_1 = (g^{-1}h)(c_1)$. But there are no nontrivial fixed points under G 's action on S , so $g^{-1}h$ is the identity element of G . Thus, $g = h$.

Therefore, $\{g(C) : g \in G\}$ is pairwise disjoint.

Furthermore, $\bigcup_{g \in G} g(C) = S$. To see this, note that $\bigcup_{g \in G} g(C) \subseteq S$ since $g : S \rightarrow S$ for every $g \in G$. Also, suppose there exists some $s \in S$ such that $s \notin \bigcup_{g \in G} g(C)$. Then s is not in the same orbit as any $c \in C$. But C contains one element from every orbit of S under G . Thus, for all $s \in S$, $s \in \{g(c) : g \in G\}$ for some $c \in C$. Hence, $S \subseteq \bigcup_{g \in G} g(C)$.

Therefore, $\{g(C) : g \in G\}$ is a partition of S .

Next, let $\overline{Y}_i = \bigcup_{g \in Y_i} g(C)$ and $\overline{Z}_i = \bigcup_{g \in Z_i} g(C)$. Since $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ are pairwise disjoint, as is $\{g(C) : g \in G\}$, then $\overline{Y}_1, \dots, \overline{Y}_p, \overline{Z}_1, \dots, \overline{Z}_q$ are also pairwise disjoint.

Furthermore, $\bigcup_{i=1}^p \overline{Y}_i = \bigcup_{i=1}^p (g_i Y_i)(C) = G(C) = S$ and

$$\bigcup_{j=1}^q \overline{Z}_j = \bigcup_{j=1}^q (h_j Z_j)(C) = G(C) = S.$$

Therefore, S is paradoxical with respect to G .

□

Notice the use of the Axiom of Choice in the proof. This is the first occurrence of the Axiom in this thesis and it alludes to the necessity of the Axiom in the Banach-Tarski Paradox. A weaker form of the proposition can be proved without relying on the Axiom of Choice (Wagon 11), but then one is denied the full power of the proposition. This is the case in many instances in mathematics and it has been central in the debate over whether or not to accept the

Axiom. Those who reject the Axiom and choose instead to use a weaker form of choice are often forced to accept weaker conclusions (Schechter). Sometimes, however, the conclusions that must be abandoned agree exactly with our intuition based on what is already known about mathematics. This is the main reason why most mathematicians have chosen to accept the Axiom of Choice, along with its theoretical baggage.

Chapter 2: The Hausdorff Paradox

In this chapter, the work performed in chapter 1 to build up the foundation of paradoxical sets will pay off, as one will see the first specific example of the concept being applied to a well-known set. Let S^2 denote the 3-dimensional unit sphere. This is the set of all points in $\mathbb{R}^3$ that are one unit from the origin. Recall that SO_3 is the set of all 3×3 orthogonal matrices with determinant 1, and that each element represents a rotation of $\mathbb{R}^3$ about a line through the origin. Hausdorff's Paradox, whose proof will be forthcoming, states that the unit sphere is, for the most part, paradoxical with respect to these rotations.

Theorem (*The Hausdorff Paradox*)

2.1

Let S^2 be the unit sphere in $\mathbb{R}^3$. There is a countable subset X of S^2 such that $S^2 \setminus X$ is paradoxical with respect to SO_3 (Wagon 18).

To establish this result, one must first construct a free subgroup of SO_3 of rank 2. The approach used here has been credited to Swierczkowski. Let δ be the counterclockwise rotation of $\mathbb{R}^3$ about the z -axis through an angle θ , where $\cos(\theta) = \frac{1}{3}$. Also, let φ be the counterclockwise rotation of $\mathbb{R}^3$ about the x -axis through the same angle θ . Notice that $\delta(1, 0, 0) = (\frac{1}{3}, \frac{2}{3}\sqrt{2}, 0)$, $\delta(0, 1, 0) = (\frac{-2}{3}\sqrt{2}, \frac{1}{3}, 0)$, and $\delta(0, 0, 1) = (0, 0, 1)$. Hence, δ is

represented by the matrix
$$\begin{bmatrix} \frac{1}{3} & \frac{-2}{3}\sqrt{2} & 0 \\ \frac{2}{3}\sqrt{2} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \in SO_3.$$
 Similarly, $\varphi(1, 0, 0) = (1, 0, 0)$,

$\varphi(0, 1, 0) = (0, \frac{1}{3}, \frac{2}{3}\sqrt{2})$, and $\varphi(0, 0, 1) = (0, -\frac{2}{3}\sqrt{2}, \frac{1}{3})$ implying that φ is represented by

the matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & -\frac{2}{3}\sqrt{2} \\ 0 & \frac{2}{3}\sqrt{2} & \frac{1}{3} \end{bmatrix} \in SO_3$. Hence, δ^{-1} and φ^{-1} are represented by the matrices

$$\begin{bmatrix} \frac{1}{3} & \frac{2}{3}\sqrt{2} & 0 \\ -\frac{2}{3}\sqrt{2} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3}\sqrt{2} \\ 0 & -\frac{2}{3}\sqrt{2} & \frac{1}{3} \end{bmatrix} \in SO_3, \text{ respectively (Wagon 15).}$$

Lemma

2.2

Let F be the subgroup of SO_3 generated by δ, φ . Then F is a free group of rank 2 (Wagon 15-16).

Proof F is, by definition, a group generated by δ, φ . What remains to be shown is that no nontrivial reduced word in F equals the identity rotation. This will be accomplished using strong induction.

Let ω be a nontrivial reduced word in F and let the notation ω_{-n} denote the word formed by removing the first n letters from ω . Consider the two possible cases according to whether ω ends in $\delta^{\pm 1}$ or $\varphi^{\pm 1}$.

Case 1: Suppose ω ends in $\delta^{\pm 1}$.

Notice that if the length of ω is 1, then $\omega = \delta^{\pm 1}$ and $\omega(1, 0, 0) = (\frac{1}{3}, \pm \frac{2}{3}\sqrt{2}, 0)$, respectively.

If, however, the length of ω is 2, then there are six possibilities for ω :

$\omega = \delta\delta, \delta^{-1}\delta^{-1}, \varphi\delta, \varphi\delta^{-1}, \varphi^{-1}\delta, \text{ or } \varphi^{-1}\delta^{-1}$. Notice that in these

cases, one obtains $\omega(1, 0, 0) = (-\frac{7}{9}, \frac{4}{9}\sqrt{2}, 0), (-\frac{7}{9}, -\frac{4}{9}\sqrt{2}, 0),$

$(\frac{3}{9}, \frac{2}{9}\sqrt{2}, \frac{8}{9}), (\frac{3}{9}, -\frac{2}{9}\sqrt{2}, -\frac{8}{9}), (\frac{3}{9}, \frac{2}{9}\sqrt{2}, -\frac{8}{9}), \text{ or } (\frac{3}{9}, -\frac{2}{9}\sqrt{2}, \frac{8}{9}),$

respectively.

Suppose that for any $k > 2$ and any $n < k$, if the length of ω is n , then $\omega(1, 0, 0) = \left(\frac{x}{3^n}, \frac{y}{3^n} \sqrt{2}, \frac{z}{3^n} \right)$ for some $x, y, z \in \mathbb{Z}$ where $3 \nmid y$.

If, then, the length of ω is k , four subcases occur depending on whether ω begins with δ , δ^{-1} , φ , or φ^{-1} .

case 1a: ω begins with δ .

Suppose $\omega = \delta \varphi^{\pm 1} \omega_{-2}$. Notice that $\varphi^{\pm 1} \omega_{-2}(1, 0, 0) = \left(\frac{3x}{3^{k-1}}, \frac{y \mp 2z}{3^{k-1}} \sqrt{2}, \frac{\pm 4y + z}{3^{k-1}} \right)$ for some $x, y, z \in \mathbb{Z}$ where $3 \nmid y \mp 2z$ by the inductive hypothesis. Hence,

$$\begin{aligned} \omega(1, 0, 0) &= \delta \left(\frac{3x}{3^{k-1}}, \frac{y \mp 2z}{3^{k-1}} \sqrt{2}, \frac{\pm 4y + z}{3^{k-1}} \right) \\ &= \left(\frac{3x - 4y \pm 8z}{3^k}, \frac{6x + y \mp 2z}{3^k} \sqrt{2}, \frac{\pm 12y + 3z}{3^k} \right) \end{aligned}$$

where $3 \nmid 6x + y \mp 2z$ since $3 \mid 6x$ and $3 \nmid y \mp 2z$.

Suppose, on the other hand, that $\omega = \delta \delta \omega_{-2}$. Notice that $\delta \omega_{-2}(1, 0, 0) = \left(\frac{x - 4y}{3^{k-1}}, \frac{2x + y}{3^{k-1}} \sqrt{2}, \frac{3z}{3^{k-1}} \right)$ where $3 \nmid 2x + y$. Hence,

$$\begin{aligned} \omega(1, 0, 0) &= \delta \left(\frac{x - 4y}{3^{k-1}}, \frac{2x + y}{3^{k-1}} \sqrt{2}, \frac{3z}{3^{k-1}} \right) \\ &= \left(\frac{-7x - 8y}{3^k}, \frac{2(2x + y) - 9y}{3^k} \sqrt{2}, \frac{9z}{3^k} \right) \end{aligned}$$

where $3 \nmid 2(2x + y) - 9y$ since $3 \nmid 2(2x + y)$ and $3 \mid -9y$.

case 1b: ω begins with δ^{-1}

Suppose $\omega = \delta^{-1} \varphi^{\pm 1} \omega_{-2}$. Recall from case 1a that

$$\varphi^{\pm 1} \omega_{-2}(1, 0, 0) = \left(\frac{3x}{3^{k-1}}, \frac{y \mp 2z}{3^{k-1}} \sqrt{2}, \frac{\pm 4y + z}{3^{k-1}} \right) \text{ where } 3 \nmid y \mp 2z.$$

Hence,

$$\omega(1, 0, 0) = \delta^{-1} \left(\frac{3x}{3^{k-1}}, \frac{y \mp 2z}{3^{k-1}} \sqrt{2}, \frac{\pm 4y + z}{3^{k-1}} \right)$$

$$= \left(\frac{3x+4y+8z}{3^k}, \frac{-6x+y+2z}{3^k} \sqrt{2}, \frac{\pm 12y+3z}{3^k} \right)$$

where $3 \nmid -6x + y + 2z$ since $3 \mid -6x$ and $3 \nmid y + 2z$.

Suppose, on the other hand, that $\omega = \delta^{-1} \delta^{-1} \omega_{-2}$. Notice that

$$\delta^{-1} \omega_{-2}(1, 0, 0) = \left(\frac{x+4y}{3^{k-1}}, \frac{-2x+y}{3^{k-1}} \sqrt{2}, \frac{3z}{3^{k-1}} \right) \text{ where } 3 \nmid -2x + y.$$

Hence,

$$\begin{aligned} \omega(1, 0, 0) &= \delta^{-1} \left(\frac{x+4y}{3^{k-1}}, \frac{-2x+y}{3^{k-1}} \sqrt{2}, \frac{3z}{3^{k-1}} \right) \\ &= \left(\frac{-7x+8y}{3^k}, \frac{2(-2x+y)-9z}{3^k} \sqrt{2}, \frac{9z}{3^k} \right) \end{aligned}$$

where $3 \nmid 2(-2x + y) - 9z$ since $3 \nmid 2(-2x + y)$ and $3 \mid -9z$.

case 1c: ω begins with φ .

Suppose $\omega = \varphi \delta^{\pm 1} \omega_{-2}$. Notice that $\delta^{\pm 1} \omega_{-2}(1, 0, 0) =$

$$\left(\frac{x+4y}{3^{k-1}}, \frac{\pm 2x+y}{3^{k-1}} \sqrt{2}, \frac{3z}{3^{k-1}} \right) \text{ where } 3 \nmid \pm 2x + y. \text{ Hence,}$$

$$\begin{aligned} \omega(1, 0, 0) &= \varphi \left(\frac{x+4y}{3^{k-1}}, \frac{\pm 2x+y}{3^{k-1}} \sqrt{2}, \frac{3z}{3^{k-1}} \right) \\ &= \left(\frac{3x+12y}{3^k}, \frac{\pm 2x+y-6z}{3^k} \sqrt{2}, \frac{\pm 8x+4y+3z}{3^k} \right) \end{aligned}$$

where $3 \nmid \pm 2x + y - 6z$ since $3 \nmid \pm 2x + y$ and $3 \mid -6z$.

Suppose, on the other hand, that $\omega = \varphi \varphi \omega_{-2}$. Notice that

$$\varphi \omega_{-2}(1, 0, 0) = \left(\frac{3x}{3^{k-1}}, \frac{y-2z}{3^{k-1}} \sqrt{2}, \frac{4y+z}{3^{k-1}} \right) \text{ where } 3 \nmid y - 2z. \text{ Hence,}$$

$$\begin{aligned} \omega(1, 0, 0) &= \varphi \left(\frac{3x}{3^{k-1}}, \frac{y-2z}{3^{k-1}} \sqrt{2}, \frac{4y+z}{3^{k-1}} \right) \\ &= \left(\frac{9x}{3^k}, \frac{-9y+2(y-2z)}{3^k} \sqrt{2}, \frac{8y-7z}{3^k} \right) \end{aligned}$$

where $3 \nmid -9y + 2(y - 2z)$ since $3 \mid -9y$ and $3 \nmid 2(y - 2z)$.

case 1d: ω begins with φ^{-1} .

Suppose $\omega = \varphi^{-1} \delta^{\pm 1} \omega_{-2}$. Recall from case 1c that

$$\delta^{\pm 1} \omega_{-2}(1, 0, 0) = \left(\frac{x \mp 4y}{3^{k-1}}, \frac{\pm 2x + y}{3^{k-1}} \sqrt{2}, \frac{3z}{3^{k-1}} \right) \text{ where } 3 \nmid \pm 2x + y.$$

Hence,

$$\begin{aligned} \omega(1, 0, 0) &= \varphi^{-1} \left(\frac{x \mp 4y}{3^{k-1}}, \frac{\pm 2x + y}{3^{k-1}} \sqrt{2}, \frac{3z}{3^{k-1}} \right) \\ &= \left(\frac{3x \mp 12y}{3^k}, \frac{\pm 2x + y + 6z}{3^k} \sqrt{2}, \frac{\mp 8x - 4y + 3z}{3^k} \right) \end{aligned}$$

where $3 \nmid \pm 2x + y + 6z$ since $3 \nmid \pm 2x + y$ and $3 \mid 6z$.

Suppose, on the other hand, that $\omega = \varphi^{-1} \varphi^{-1} \omega_{-2}$. Notice that

$$\varphi^{-1} \omega_{-2}(1, 0, 0) = \left(\frac{3x}{3^{k-1}}, \frac{y + 2z}{3^{k-1}} \sqrt{2}, \frac{-4y + z}{3^{k-1}} \right) \text{ where } 3 \nmid y + 2z.$$

Hence,

$$\begin{aligned} \omega(1, 0, 0) &= \varphi^{-1} \left(\frac{3x}{3^{k-1}}, \frac{y + 2z}{3^{k-1}} \sqrt{2}, \frac{-4y + z}{3^{k-1}} \right) \\ &= \left(\frac{9x}{3^k}, \frac{-9y + 2(y + 2z)}{3^k} \sqrt{2}, \frac{-8y + 7z}{3^k} \right) \end{aligned}$$

where $3 \nmid -9y + 2(y + 2z)$ since $3 \mid 9y$ and $3 \nmid 2(y + 2z)$.

Thus, for any nontrivial reduced word ω ending in δ or δ^{-1} , one

sees that $\omega(1, 0, 0) = \left(\frac{x'}{3^k}, \frac{y'}{3^k} \sqrt{2}, \frac{z'}{3^k} \right)$ where $x', y', z' \in \mathbb{Z}$ and

$3 \nmid y'$. Since $3 \nmid y'$, then $\frac{y'}{3^k} \sqrt{2} \neq 0$. Hence, $\omega(1, 0, 0) \neq (1, 0, 0)$.

Therefore, every such word fails to hold at least one element of $\mathbb{R}^3$

(namely, $(1, 0, 0)$) fixed.

Case 2: Suppose ω ends in $\varphi^{\pm 1}$.

One can show that any such word that fixes the point (x, y, z)

induces a word ending in $\delta^{\mp 1}$ that fixes (z, y, x) . To see this, let ω' be

the word formed by replacing $\delta, \delta^{-1}, \varphi, \varphi^{-1}$ in ω by $\varphi^{-1}, \varphi, \delta^{-1}, \delta$,

respectively.

Notice that $\varphi(x, y, z) = \left(x, \frac{y-2z\sqrt{2}}{3}, \frac{z+2y\sqrt{2}}{3}\right)$ while $\delta^{-1}(z, y, x) = \left(\frac{z+2y\sqrt{2}}{3}, \frac{y-2z\sqrt{2}}{3}, x\right)$, and $\varphi^{-1}(x, y, z) = \left(x, \frac{y+2z\sqrt{2}}{3}, \frac{z-2y\sqrt{2}}{3}\right)$ while $\delta(z, y, x) = \left(\frac{z-2y\sqrt{2}}{3}, \frac{y+2z\sqrt{2}}{3}, x\right)$. Thus, if $\omega = \varphi^{\pm 1}$, then the coordinates of $\omega'(z, y, x)$ agree with those of $\omega(x, y, z)$ with the exception that the first and third coordinates are interchanged.

Now suppose that this same relationship holds when ω is of length k . That is to say, if $\omega(x, y, z) = (x', y', z')$, then $\omega'(z, y, x) = (z', y', x')$. Then, if ω is of length $k + 1$, one sees that

$$\begin{aligned}\omega(x, y, z) &= \varphi^{\pm 1} \omega_{-1}(x, y, z) = \varphi^{\pm 1}(x', y', z') \\ &= \left(x', \frac{y' \mp 2z'\sqrt{2}}{3}, \frac{z' \pm 2y'\sqrt{2}}{3}\right).\end{aligned}$$

However, in this same case, one has

$$\begin{aligned}\omega'(z, y, x) &= \delta^{\mp 1} \omega'_{-1}(z, y, x) = \delta^{\mp 1}(z', y', x') \\ &= \left(\frac{z' \pm 2y'\sqrt{2}}{3}, \frac{y' \mp 2z'\sqrt{2}}{3}, x'\right).\end{aligned}$$

Therefore, regardless of the length of ω , one always finds that $\omega'(z, y, x) = (z', y', x')$, where $\omega(x, y, z) = (x', y', z')$. In other words, $\omega'(z, y, x)$ is the reflection of $\omega(x, y, z)$ across the plane $x = z$.

Suppose, then, that ω fixes some point (x, y, z) . Then $\omega(x, y, z) = (x, y, z)$. Hence, ω' fixes the point (z, y, x) since $\omega'(z, y, x)$ would equal (z, y, x) . Therefore, by contraposition, no nontrivial reduced word ending in $\varphi^{\pm 1}$ can fix the point $(0, 0, 1)$ since it has been shown that no such word ending in $\delta^{\mp 1}$ fixes $(1, 0, 0)$.

Thus, every nontrivial reduced word in F fails to fix at least one element of $\mathbb{R}^3$.

Therefore, no such word equals the identity rotation, implying that F is a free group.

□

To see how this free subgroup of SO_3 applies to the Hausdorff Paradox, let $\rho \in F$ (where ρ is not the identity rotation). If l is the axis of rotation of ρ , then every point on l is fixed by ρ . Now, l intersects S^2 in exactly two points. Hence, ρ has two fixed points on S^2 . Since there is a countably infinite number of nontrivial rotations in F , then F has a countably infinite number of fixed points on S^2 (since $2\aleph_0 = \aleph_0$). Let X be the set of all such points (Wagon 17-18).

Furthermore, let $s \in S^2 \setminus X$. If $\rho(s) \notin S^2 \setminus X$, then $\rho(s) \in X$ and, hence, is a fixed point of some $\overline{\rho} \in F$. Then $\rho^{-1}\overline{\rho}\rho(s) = \rho^{-1}\rho(s) = s$. Hence, s is a fixed point of $\rho^{-1}\overline{\rho}\rho \in F$. However, since $s \in S^2 \setminus X$, this cannot be true. Thus, $\rho(s) \in S^2 \setminus X$. Therefore, if $s \in S^2 \setminus X$ and ρ is any rotation in F , then $\rho(s) \in S^2 \setminus X$ and, hence, F acts on $S^2 \setminus X$ with no nontrivially fixed points. Thus, by proposition 1.8 $S^2 \setminus X$ is paradoxical with respect to F . Furthermore, since $F \subset SO_3$, then $S^2 \setminus X$ is paradoxical with respect to SO_3 (Wagon 18).

It is important, now that the Hausdorff Paradox has been proven, to take a step back and translate the mathematics into more familiar terms so that one can see the overall picture. After all, the beauty of such a paradox is not just in the elegance of the mathematics, but also in what it implies about the assumptions and beliefs regarding what is “obviously” true or untrue.

Try to visualize $S^2 \setminus X$. One cannot completely imagine X and, therefore, $S^2 \setminus X$.

However, since the surface of the sphere S^2 is uncountable, then even though X is countably infinite, it would seem very small in comparison to the rest of the sphere. Therefore, a

reasonable representative of $S^2 \setminus X$ may be to imagine a sphere peppered with tiny pinholes.

What Hausdorff has shown, is that this tattered sphere can be decomposed into two collections of subsets, where all the subsets are pairwise disjoint, such that the sets in either one of the collections can be individually rotated and then “glued” back together to produce the same tattered sphere. The amazing part is that one would still have the other collection of subsets left over. This is like the mechanic who tells you that he put your car’s engine back together, but he still had some pieces left over afterward.

Furthermore, the other collection of subsets can also be individually rotated and “glued” back together to get yet another copy of the same tattered sphere. (Whereas it is doubtful that the mechanic would be able to build a whole new engine from the pieces left over.) This is truly amazing and it is understandable that one may only accept it with a certain amount of hesitation. However, keep in mind that since Hausdorff’s Paradox relies on Proposition 1.8, which in turn relies on the Axiom of Choice, this is a purely theoretical construction and no one is claiming to be able to perform such a feat in front of one’s eyes. This author proposes that this is not a weakness, but rather a tribute to the beauty of mathematics. It is as though the mathematics gives a glimpse of what could be done if one could just reach into the infinite.

It would, of course, be desirable to be able to extend Hausdorff’s Paradox to the entire sphere S^2 rather than just “the vast majority” of it. One should keep in mind that, due to the relative differences in size between an uncountable and a countable set, the result of theorem 2.1 does practically hold for the entire sphere as it is. In fact, in mathematical terms, one would

say that the Hausdorff Paradox holds *almost everywhere* on the sphere. Just the same, however, it would be nice if it held everywhere on the sphere instead of just almost everywhere.

One attempt at achieving this, although it is not completely successful, is to include the set X in one of the two collections of subsets of $S^2 \setminus X$. For instance, let A and B be the two collections and let $\bar{A} = A \cup X$. Since $X \cap S^2 \setminus X = \emptyset$, then $\bar{A} \cup B$ will still be pairwise disjoint. Now if one performs the same rotations on $\bar{A} \setminus X$ as those required for the Hausdorff Paradox, and performs the identity rotation on X itself, then the pieces could be “glued” back together and the result would be the entire sphere S^2 since X would fill in the gaps. The problem is that the subsets in B would still not be able to cover the entire sphere, which would be required in order to say that the sphere is paradoxical with respect to rotations. Thus, one does not have a legitimate paradoxical decomposition of the entire sphere. The only way to overcome this shortcoming of B would be to add X to B as well. The problem with that is that the decomposition subsets would no longer be pairwise disjoint since X would be used twice. Thus, one can only get a “one-sided” paradoxical decomposition of the entire sphere. However, as will be seen in the next chapter, the relatively small size of X will enable one to disregard it completely. In fact, $S^2 \setminus X$ will be shown to be, in a certain way, equivalent to all of S^2 and one will get a paradoxical decomposition of the entire sphere.

Chapter 3: The Banach-Tarski Paradox

It has been shown via the Hausdorff Paradox that, except for some countable subset X , the unit sphere in $\mathbb{R}^3$ is paradoxical with respect to rotations. The goal is to eliminate X from this argument, thus extending the paradoxicality to the entire sphere S^2 . Thus, S^2 would have the property that it could be decomposed into disjoint subsets which can be grouped into two families such that each family can be rotated to cover the entire sphere. Since the size of X is so small compared to the entire sphere, with X being countable and S^2 being uncountable, there is reason to believe that the effect of X may be negligible. To achieve this, one must show that S^2 is somehow equivalent to $S^2 \setminus X$, which is already known to be paradoxical with respect to SO_3 , and that the paradoxicality is preserved by this equivalence. The notion of equivalence to be used, called equidecomposability, is essentially analogous to saying that two jigsaw puzzles are equivalent (equidecomposable) if one puzzle can be obtained from the other by simply taking the pieces apart and rearranging them.

Definition (*Equidecomposable*)

3.1

Suppose that X and Y are subsets of S on which group G acts, and that there exist partitions of X and Y containing the same number of pieces ($\bigcup_{i=1}^n X_i$ and $\bigcup_{i=1}^n Y_i$) such that for each i there exists a $g_i \in G$ where $g_i(X_i) = Y_i$. Then X and Y are equidecomposable with respect to G , denoted $X \sim_G Y$ (Wagon 23).

For instance, let $X = \{1, 5, 7, 10, 11\}$, $Y = \{2, 4, 6, 8, 10\} \subset \mathbb{R}$ and let T be the group of

translations of $\mathbf{R}$. Then $X \sim_T Y$ since $X = \{1\} \cup \{5, 7, 11\} \cup \{10\}$ which, by translating $\{1\}$ one unit to the right, $\{5, 7, 11\}$ one unit to the left, and $\{10\}$ two units to the left, can be translated into $\{2\} \cup \{4, 6, 10\} \cup \{8\} = Y$. To verify, however, that this is in fact an equivalence relation, it must be shown that equidecomposability possesses the reflexive, symmetric, and transitive properties.

Lemma

3.2

The relation of equidecomposability is an equivalence relation.

Proof Let S be a set that is being acted upon by group G .

Reflexivity: Suppose $X \subseteq S$. Since $i(X) = X$, where i is the identity element of G , $X \sim_G X$.

Symmetry: Suppose $X, Y \subseteq S$ and $X \sim_G Y$. Then $X = \bigcup_{i=1}^n X_i$ and $Y = \bigcup_{i=1}^n Y_i$ such that for each i there exists a $g_i \in G$ where $g_i(X_i) = Y_i$. Then $g_i^{-1}(Y_i) = X_i$ for each i . Thus, $Y \sim_G X$.

Transitivity: Suppose $X, Y, Z \subseteq S$ and $X \sim_G Y \sim_G Z$. Then $X = \bigcup_{i=1}^n X_i$ and $Y = \bigcup_{i=1}^n Y_i$ such that for each i there exists a $g_i \in G$ where $g_i(X_i) = Y_i$, and $Y = \bigcup_{j=1}^m Y^j$ and $Z = \bigcup_{j=1}^m Z^j$ such that for each j there exists a $h_j \in G$ where $h_j(Y^j) = Z^j$. (Note that the use of the superscript j is not to imply that the sets are being raised, somehow, to the power j . Rather, it is used simply to avoid confusion between sets. For instance, Y_1 and Y^1 .)

Now, further partition X as follows: For each $1 \leq i \leq n$, let

$X_i^j = \{x \in X_i : g_i(x) \in Y^j\}$. This partitions X_i into at most m subsets and doing so for all $1 \leq i \leq n$ partitions X into p subsets for some $p \leq mn$. Now, for each X_i^j in this

partition, define $Z_i^j := (h_j g_i)(X_i^j)$. This gives p disjoint subsets of Z . To see that this is a partition of Z , one must show that it covers Z . Note for all $z \in Z$, $z \in Z^j$ for some $1 \leq j \leq m$. Then $h_j^{-1}(z) \in Y$ and, hence, it is an element of Y_i for some $1 \leq i \leq n$. Hence, $g_i^{-1}(h_j^{-1}(z)) \in X_i^j$. Thus, $z \in Z_i^j$ for some $1 \leq i \leq n$, $1 \leq j \leq m$.

Therefore, one has that $X = \bigcup X_i^j$ and $Z = \bigcup Z_i^j$ where these are partitions of X and Z , respectively, and $(h_j g_i)(X_i^j) = Z_i^j$. Hence $X \sim_G Z$.

□

It is desirable, at this point, to take a few moments to point out a couple of characteristics of the equidecomposability relation.

Lemma

3.3

If $\bar{X} \subseteq X \sim_G Y$, then there is a bijection $f: X \rightarrow Y$ such that $\bar{X} \sim_G f(\bar{X})$ (Wagon 25).

Proof Since $X \sim_G Y$, we know there are partitions $X = X_1 \cup \dots \cup X_n$ and

$Y = Y_1 \cup \dots \cup Y_n$ with the property that for all $1 \leq i \leq n$, there is a $g_i \in G$ such that $g_i(X_i) = Y_i$. Pick one such g_i for each X_i , call it g_i^* . Note that each of these g_i^* is a bijection from X_i onto Y_i by Proposition 1.2. Now, define the bijection $f: X \rightarrow Y$ as follows: Let $x \in X$. Then $x \in X_i$ for some $1 \leq i \leq n$, and one defines $f(x) = g_i^*(x)$.

One wishes to show that $\bar{X} \sim_G f(\bar{X})$. Form the partition $\bar{X} = \bar{X}_1 \cup \dots \cup \bar{X}_n$, where $\bar{X}_i = \bar{X} \cap X_i$. Then a partition of $f(\bar{X})$ is given by

$$f(\bar{X}) = \bigcup_{i=1}^n f(\bar{X}_i) = \bigcup_{i=1}^n g_i^*(\bar{X}_i), \text{ which witnesses the fact that } \bar{X} \sim_G f(\bar{X}).$$

□

Lemma Suppose $X \sim_G Y$ and $\bar{X} \sim_G \bar{Y}$. If $X \cap \bar{X} = \emptyset$ and $Y \cap \bar{Y} = \emptyset$, then

$X \cup \bar{X} \sim_G Y \cup \bar{Y}$ (Wagon 25).

Proof Let the partitions that witness $X \sim_G Y$ be $X = X_1 \cup \dots \cup X_n$ and $Y = Y_1 \cup \dots \cup Y_n$, and let the partitions that witness $\bar{X} \sim_G \bar{Y}$ be $\bar{X} = X_{n+1} \cup \dots \cup X_m$ and $\bar{Y} = Y_{n+1} \cup \dots \cup Y_m$. Then $X_1 \cup \dots \cup X_m$ and $Y_1 \cup \dots \cup Y_m$ are partitions of $X \cup \bar{X}$ and $Y \cup \bar{Y}$, respectively, since $X \cap \bar{X} = \emptyset$ and $Y \cap \bar{Y} = \emptyset$. Furthermore, since for every $1 \leq i \leq m$ there exists a $g_i \in G$ such that $g_i(X_i) = Y_i$, one has $X \cup \bar{X} \sim_G Y \cup \bar{Y}$.

□

Now that one has this new equivalence relation of equidecomposability, one can prove that the sphere S^2 is equivalent to $S^2 \setminus X$ via equidecomposability with respect to rotations of $\mathbb{R}^3$ (where X is the countable subset of S^2 from the Hausdorff Paradox).

Proposition

3.5

If X is a countable subset of S^2 , then $S^2 \sim S^2 \setminus X$ with respect to SO_3 (Wagon 27).

Proof Let $\mathcal{L}$ be a line through the origin such that $\mathcal{L} \cap X = \emptyset$ and let S be the set of θ for which there exists some $n \in \mathbb{N}$ and some $x \in X$ such that the image of x under a rotation about $\mathcal{L}$ through an angle of $n\theta$ radians is in X .

Since $\mathbb{N}$ and X are both countable, then S is countable, also. Hence, since $(0, 2\pi]$ is uncountable, there must be some $\bar{\theta}$ that is not in S . Choose any such $\bar{\theta}$ and let ρ be the rotation about $\mathcal{L}$ through an angle of $\bar{\theta}$ radians.

Notice that $\rho\left(\bigcup_{n=0}^{\infty} \rho^n(X)\right) = \bigcup_{n=1}^{\infty} \rho^n(X)$, and that, since $\bar{\theta} \notin S$, $\rho^n(X) \cap X = \emptyset$ for all $n \in \mathbb{N}$.

Note also that $S^2 = \left(\bigcup_{n=0}^{\infty} \rho^n(X)\right) \cup S^2 \setminus \left(\bigcup_{n=0}^{\infty} \rho^n(X)\right)$ and

$$\begin{aligned}
S^2 \setminus X &= \left(\bigcup_{n=1}^{\infty} \rho^n(X) \right) \cup S^2 \setminus \left[X \cup \left(\bigcup_{n=1}^{\infty} \rho^n(X) \right) \right] \\
&= \rho \left(\bigcup_{n=0}^{\infty} \rho^n(X) \right) \cup S^2 \setminus \left(\bigcup_{n=0}^{\infty} \rho^n(X) \right).
\end{aligned}$$

Thus, one has $S^2 = (A \cup B) \sim_{SO_3} [\rho(A) \cup B] = S^2 \setminus X$, where

$$A = \bigcup_{n=0}^{\infty} \rho^n(X) \text{ and } B = S^2 \setminus \left(\bigcup_{n=0}^{\infty} \rho^n(X) \right).$$

□

It would appear, then, that the work of showing that the sphere S^2 is paradoxical with respect to SO_3 has for the most part been completed. After all, it has been demonstrated that $S^2 \setminus X$ is paradoxical with respect to SO_3 , and that S^2 is equivalent to $S^2 \setminus X$ by means of equidecomposability. However, making such a claim about the paradoxical nature of S^2 at this point would still be somewhat premature. Although it has been shown that equidecomposability is an equivalence relation, it still remains to be shown that it preserves paradoxicality.

To see how equidecomposability preserves paradoxicality, begin by comparing their respective definitions. Notice that if a set X is paradoxical with respect to some group G , where Y and Z are the two collections of subsets that yield the paradoxical decomposition, then $Y \sim_G X$ and $Z \sim_G X$. In fact, it can be shown that these two conditions are equivalent.

Proposition

3.6

A set X is paradoxical with respect to group G if and only if X contains two disjoint sets, Y and Z , such that $Y \sim_G X$ and $Z \sim_G X$ (Wagon 24).

Proof ($\Rightarrow$) Suppose that X is paradoxical with respect to G . Then, by the definition of a paradoxical set, there exist pairwise disjoint subsets $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ of X and some $g_1, \dots, g_p, h_1, \dots, h_q \in G$ (for some $p, q \in \mathbb{N}$) such that $X = \bigcup_{i=1}^p g_i(Y_i)$

and also $X = \bigcup_{j=1}^q h_j(Z_j)$.

Let $Y = \bigcup_{i=1}^p Y_i$ and $Z = \bigcup_{j=1}^q Z_j$. By the pairwise disjointness of the sequence $\{Y_i\}_{i=1}^p \cup \{Z_j\}_{j=1}^q$, $Y \cap Z = \emptyset$ and $\{Y_i\}_{i=1}^p$ and $\{Z_j\}_{j=1}^q$ form partitions of Y and Z , respectively. Also, recall that $\{g_i(Y_i)\}_{i=1}^p$ and $\{h_j(Z_j)\}_{j=1}^q$ are each pairwise disjoint and, hence, form partitions of X . Therefore, $Y \sim_G X$ and $Z \sim_G X$.

($\Leftarrow$) Suppose that X contains two disjoint sets Y and Z such that $Y \sim_G X$ and $Z \sim_G X$. Then, regarding Y , there exist partitions of X and Y containing the same number of pieces ($\bigcup_{i=1}^p Y_i$ and $\bigcup_{i=1}^p X_i$) such that for each i there exists a $g_i \in G$ where $g_i(Y_i) = X_i$. From this, one obtains $X = \bigcup_{i=1}^p g_i(Y_i)$. Similarly, regarding Z , there exist partitions of X and Z containing the same number of pieces ($\bigcup_{j=1}^q Z_j$ and $\bigcup_{j=1}^q X^j$) such that for each j there exists a $h_j \in G$ where $h_j(Z_j) = X^j$, from which one obtains $X = \bigcup_{j=1}^q h_j(Z_j)$.

Now since the Y_i form a partition of Y , they must be pairwise disjoint. Likewise for the Z^j . Furthermore, since $Y \cap Z = \emptyset$, then all of the Y_i and Z^j must be pairwise disjoint.

Thus, X is paradoxical with respect to G .

□

Corollary

3.7

Suppose $\bar{X} \subset X$ and that $\bar{X} \sim_G X$ for some group G . If $\bar{X}$ is paradoxical with respect to G , then so is X .

Proof

Let $\bar{X}$ be a subset of X that is equidecomposable with X with respect to G , and

suppose that $\bar{X}$ is paradoxical with respect to the same group G . Then, by Proposition 3.6, $\bar{X}$ contains two disjoint sets Y and Z such that $Y \sim_G \bar{X}$ and $Z \sim_G \bar{X}$. But $Y, Z \subseteq \bar{X} \Rightarrow Y, Z \subset X$ since $\bar{X} \subset X$, and the transitivity of $\sim_G$ gives $Y \sim_G \bar{X} \sim_G X \Rightarrow Y \sim_G X$ and $Z \sim_G \bar{X} \sim_G X \Rightarrow Z \sim_G X$. Thus, by Proposition 3.6, X is paradoxical with respect to G .

□

Thus, equidecomposability preserves paradoxicality as one moves up a chain of subsets. Then, since $S^2 \setminus X$ is a subset of S^2 that is SO_3 -equidecomposable to S^2 (by proposition 3.5), and $S^2 \setminus X$ is paradoxical with respect to SO_3 (theorem 2.1), then S^2 is also SO_3 -paradoxical by corollary 3.7. It is this fact that the unit sphere is paradoxical with respect to rotations that allows one to accomplish the first major goal: the proof of the Banach-Tarski Paradox.

Theorem (*The Banach-Tarski Paradox*)

3.8

Any solid ball in $\mathbb{R}^3$ is paradoxical with respect to G_3 , where G_3 is the group of isometries of $\mathbb{R}^3$. In fact, all of $\mathbb{R}^3$ is paradoxical with respect to G_3 (Wagon 27-28).

Proof Let β be any ball of radius r centered at the origin. One can use the paradoxical decomposition of the unit sphere S^2 to induce a similar decomposition of β , by invoking the radial correspondence ζ that takes any $s \in S^2$ onto the set $\{as : 0 < a \leq r\}$. In this manner, the subsets of S^2 are replaced by wedge shaped sets emanating from the origin and extending out to the boundary of β . Thus, if $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ and $\delta_1, \dots, \delta_p, \varphi_1, \dots, \varphi_q$ are the sets and rotations that yield a paradoxical decomposition of S^2 , then $\zeta(Y_1), \dots, \zeta(Y_p), \zeta(Z_1), \dots, \zeta(Z_q)$ and

$\delta_1, \dots, \delta_p, \varphi_1, \dots, \varphi_q$ likewise yield a paradoxical decomposition of $\beta \setminus \mathbf{0}$. However, notice that in order to keep the wedge shaped sets pairwise disjoint, as demanded by Definition 1.3, α must be restricted from 0. Otherwise, the wedges would all overlap at the origin. Thus, these wedge shaped sets do not actually yield a paradoxical decomposition of β . To extend this decomposition of $\beta \setminus \mathbf{0}$ to a paradoxical decomposition of β , one can show $\beta \setminus \mathbf{0} \sim_{SO_3} \beta$ and invoke Corollary 3.7 once again.

Let $b = (\frac{r}{2}, 0, 0)$ and let $\mathcal{L}$ be any line through b that does not pass through the origin. Furthermore, let ρ be the counterclockwise rotation about $\mathcal{L}$ through an angle of $\pi\sqrt{2}$ radians. As shown in Chapter 1 in the example of a free group, ρ is of infinite order (i.e. ρ^n is not the identity rotation for any $n \in \mathbb{N}$). Hence, since $\mathbf{0}$ does not lie on $\mathcal{L}$, $\rho^n(\mathbf{0}) \neq \mathbf{0}$ for any $n \in \mathbb{N}$. Thus, similarly to the situation involving S^2 and $S^2 \setminus X$ in the proof of Proposition 3.5, one has

$$\begin{aligned}\beta &= \bigcup_{n=0}^{\infty} \rho^n(\mathbf{0}) \cup \beta \setminus \left(\bigcup_{n=0}^{\infty} \rho^n(\mathbf{0}) \right) \\ &\sim_{SO_3} \rho \left(\bigcup_{n=0}^{\infty} \rho^n(\mathbf{0}) \right) \cup \beta \setminus \left(\bigcup_{n=0}^{\infty} \rho^n(\mathbf{0}) \right) = \beta \setminus \{\mathbf{0}\}\end{aligned}$$

Therefore, since $\beta \setminus \{\mathbf{0}\}$ is a subset of β that is both paradoxical and also equidecomposable with β , both with respect to SO_3 , then β is, itself, paradoxical with respect to SO_3 by Corollary 3.7.

Thus, any ball centered at the origin is paradoxical with respect to SO_3 . To extend this to any ball $\bar{\beta}$ in $\mathbb{R}^3$, let τ be the translation that sends the center of $\bar{\beta}$ to the origin. Then, if $\beta = \bigcup_{i=1}^p \delta_i(Y_i) = \bigcup_{j=1}^q \varphi_j(Z_j)$ is a paradoxical decomposition of β , the ball of the same size centered at the origin, then

$$\bar{\beta} = \bigcup_{i=1}^p (\tau^{-1} \delta_i \tau)(Y_i) = \bigcup_{j=1}^q (\tau^{-1} \varphi_j \tau)(Z_j)$$

yields a paradoxical decomposition of $\bar{\beta}$ (where $\bar{Y}_i = \tau^{-1}(Y_i)$, $\bar{Z}_j = \tau^{-1}(Z_j)$).

Since this would require translations, it is G_3 that would be performing the action instead of SO_3 (since SO_3 contains only the rotations of $\mathbb{R}^3$ about the origin, and not translations). However, since $SO_3 \subset G_3$, the paradoxical actions that would take place at the origin would not be lost by switching to G_3 . In this manner, one sees that any ball in $\mathbb{R}^3$ is paradoxical with respect to G_3 .

Finally, to see that all of $\mathbb{R}^3$ is paradoxical with respect to G_3 , all that is required is to review the argument above and replace the radial correspondence of s onto the set $\{\alpha s \mid 0 < \alpha \leq r\}$ with that of s onto the set $\{\alpha s \mid 0 < \alpha\}$. In doing so, the paradoxicality of $\mathbb{R}^3$ is derived from that of S^2 in exactly the same manner that the paradoxicality of β was.

□

Notice that the Banach-Tarski Paradox is actually stronger than it gives itself credit for.

Examining the proof, one finds that reflections are not needed to prove the theorem. Thus, the paradoxicality in the Banach-Tarski Paradox is with respect to the set of orientation preserving isometries (those composed solely of translations and rotations), which is a subset of G_3 .

Stating that this smaller group is being used makes the assertion of the theorem mathematically stronger. However, this author does not wish to depart from the classical wording of the Banach-Tarski Paradox, which is why the slightly weaker but traditionally more recognizable wording has been used here. The same situation arises with the Strong Form of the Banach-Tarski Paradox, as one will soon find.

Thus, it has been shown that any ball in $\mathbb{R}^3$ can be disassembled into finitely many pairwise disjoint pieces which can be grouped into two collections with the outcome that, by using simple

isometries, the pieces in each of the collections can be rearranged and then reassembled to give a new copy of the origin ball. Thus getting two copies from the one original. Now, this does not address the necessary shape or characteristics of these pieces. However, it goes without saying that the shape of the pieces would almost certainly have to be pathological in some respect in order for such a result to occur.

This is truly an amazing result and worthy of a brief pause for reflection. However, if one recalls from the preface, there is another statement of the Banach-Tarski Paradox that seems possibly even more amazing. Not only can a ball be disassembled and reassembled to copy itself twice, it can be disassembled and reassembled to give a copy of any other ball in $\mathbb{R}^3$, regardless of how big or small, with no gaps or overlapping. In fact, not only is this true for balls, but for any bounded set with nonempty interior in $\mathbb{R}^3$. By now, one knows that this is to say that the two sets are equidecomposable. This equally astonishing result is known as the strong form of the Banach-Tarski Paradox. To prove this result requires establishing a partial ordering of the $\sim_G$ -equivalence classes in $\wp(\mathbb{R}^3)$, the power set of $\mathbb{R}^3$.

Theorem (Banach-Schröder-Bernstein Theorem)

3.9

Let S be a set that is being acted on by the group G . Then " $\leq_G$ " (where $X \leq_G Y$ is taken to mean " X is $\sim_G$ with some subset of Y ") is a partial ordering of the $\sim_G$ -equivalence classes in $\wp(S)$ (Wagon 25).

Proof To show that " $\leq_G$ " is a partial ordering, it must be shown that it is reflexive, transitive, and antisymmetric.

Reflexivity: Let X be any subset of S . Since $X = i(X)$, where i is the identity element of G , then $X \sim_G X$. Furthermore, $X \subseteq X$. Hence, $X \leq_G X$.

Transitivity: Suppose $X, Y, Z \subseteq S$ with $X \leq_G Y \leq_G Z$. Then there is some $\bar{Y} \subseteq Y$ and $\bar{Z} \subseteq Z$ such that $X \sim_G \bar{Y}$ and $Y \sim_G \bar{Z}$. Then there are partitions $X = X_1 \cup \dots \cup X_n$, $\bar{Y} = \bar{Y}_1 \cup \dots \cup \bar{Y}_n$, and $Y = Y_1 \cup \dots \cup Y_m$, $\bar{Z} = \bar{Z}_1 \cup \dots \cup \bar{Z}_m$ with the property that for all $1 \leq i \leq n$ there is some $g_i \in G$ such that $g_i(X_i) = \bar{Y}_i$, and for all $1 \leq j \leq m$ there is some $h_j \in G$ such that $h_j(Y_j) = \bar{Z}_j$.

Now partition each $\bar{Y}_i$ into its intersections with the sets in the partition of Y to get $\bar{Y} = \bar{Y}_{11} \cup \dots \cup \bar{Y}_{1m} \cup \dots \cup \bar{Y}_{n1} \cup \dots \cup \bar{Y}_{nm}$ (keeping in mind that some of these subsets might be empty). Also partition each X_i into the separate pre-images of each $\bar{Y}_{ij}$ under g_i . This gives $X = X_{11} \cup \dots \cup X_{1m} \cup \dots \cup X_{n1} \cup \dots \cup X_{nm}$, where $g_i(X_{ij}) = \bar{Y}_{ij}$ for each $1 \leq i \leq n$ and $1 \leq j \leq m$.

Thus, one has $h_j g_i(X_{ij}) = h_j(\bar{Y}_{ij}) \subseteq \bar{Z}_j$ (since $\bar{Y}_{ij} \subseteq \bar{Y}_j$). Hence, $X \sim_G \bigcup_{j=1}^m \bigcup_{i=1}^n h_j g_i(X_{ij}) \subseteq Z$ and, therefore, $X \leq_G Z$.

Antisymmetry: Suppose that $X \leq_G Y$ and $Y \leq_G X$. Then $X \sim_G \bar{Y}$ and $\bar{X} \sim_G Y$ for some $\bar{X} \subseteq X$, $\bar{Y} \subseteq Y$. Let $f_1 : X \rightarrow \bar{Y}$ and $f_2 : \bar{X} \rightarrow Y$ be the bijections described in Lemma 3.3. Furthermore, let $Z = \bigcup_{n=0}^{\infty} Z_n$ where $Z_0 = X \setminus \bar{X}$ and, inductively, $Z_{n+1} = f_2^{-1} f_1(Z_n)$.

Consider the set $X \setminus Z$. Notice that

$$\begin{aligned} X \setminus Z &= \bigcap_{n=0}^{\infty} X \setminus Z_n = X \setminus Z_0 \cap \left(\bigcap_{n=1}^{\infty} X \setminus Z_n \right) = \bar{X} \cap \bigcap_{n=0}^{\infty} [X \setminus f_2^{-1} f_1(Z_n)] \\ &= \bar{X} \cap X \setminus \bigcup_{n=0}^{\infty} (f_2^{-1} f_1(Z_n)) = \bar{X} \cap X \setminus [f_2^{-1} f_1(\bigcup_{n=0}^{\infty} Z_n)] = \bar{X} \cap X \setminus f_2^{-1} \end{aligned}$$

Thus, $X \setminus Z$ consists of exactly those points that are in X , but are not in $f_2^{-1}f_1(Z)$.

Hence, $f_2(X \setminus Z) = Y \setminus f_1(Z)$. However, by definition of f_2 , $(X \setminus Z) \sim_G f_2(X \setminus Z)$ since

$X \setminus Z \subseteq X$ (See Lemma 3.3). Therefore, $(X \setminus Z) \sim_G (Y \setminus f_1(Z))$. Similarly,

$Z \sim_G f_1(Z)$. Thus, by Lemma 3.4, one has $[(X \setminus Z) \cup Z] \sim_G [(Y \setminus f_1(Z)) \cup f_1(Z)]$.

But $(X \setminus Z) \cup Z = X$ and $[(Y \setminus f_1(Z)) \cup f_1(Z)] = Y$. Hence, $X \sim_G Y$.

Therefore, $X \preceq_G Y$ and $Y \preceq_G X$ together imply that $X \sim_G Y$. Hence, " $\preceq_G$ " is antisymmetric.

□

The immediate uses of the Banach-Schröder-Bernstein Theorem are two-fold. First of all, it allows one to make the definition of paradoxical much more user friendly. Recall that a set X is paradoxical with respect to G if and only if there exist pairwise disjoint subsets

$Y_1, \dots, Y_p, Z_1, \dots, Z_q \subseteq X$ such that $X = \bigcup_{i=1}^p g_i(Y_i) = \bigcup_{j=1}^q h_j(Z_j)$ for some

$\{g_i\}_{i=1}^p, \{h_j\}_{j=1}^q \subseteq G$, which one now knows is equivalent to $Y \sim_G X \sim_G Z$ where $Y = \bigcup_{i=1}^p Y_i$

and $Z = \bigcup_{j=1}^q Z_j$. It can be shown, using the Banach-Schröder-Bernstein Theorem, that the Y_i

and Z_j can always be chosen in such a way that they partition X (Wagon 25).

Suppose that X is paradoxical with respect to G and that the subsets that witness this are

$Y_1, \dots, Y_p, W_1, \dots, W_n$. Let $Y = \{Y_1, \dots, Y_p\}$ and $W = \{W_1, \dots, W_n\}$. Although the sets

$Y_1, \dots, Y_p, W_1, \dots, W_n$ must be pairwise disjoint, it is not necessarily true that

$(\bigcup_{i=1}^p Y_i) \cup (\bigcup_{j=1}^n W_j) = X$. Hence, these sets do not necessarily form a partition of X .

However, since $X \sim_G W$ and $W \subseteq X \setminus Y$, then $X \preceq_G X \setminus Y$. Also, since $X \setminus Y \subseteq X$, then, trivially,

$X \setminus Y \preceq_G X$. Hence, $X \setminus Y \sim_G X$. Therefore, there is a partition, call it $\{Z_j\}_{j=1}^q$, of $X \setminus Y$ such

that $X = \bigcup_{j=1}^q h_j(Z_j)$ for some $\{h_j\}_{j=1}^q \subseteq G$. Thus, the sets $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ also give a paradoxical decomposition of X with respect to G .

Furthermore, since $(\bigcup_{i=1}^p Y_i) \cup (\bigcup_{j=1}^q Z_j) = Y \cup (X \setminus Y) = X$, then $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ form a partition of X . Thus, from this point forward, when dealing with a set that is paradoxical, one can automatically refer to a partition of the set that witnesses such, allowing the reader to take advantage of the useful qualities that partitions hold.

The second immediate use of the Banach-Schröder-Bernstein Theorem is in proving the strongest and most interesting result thus far:

Theorem (*Strong Form of the Banach-Tarski Paradox*)

3.10

Any two bounded sets with nonempty interior in $\mathbb{R}^3$ are equidecomposable with respect to G_3 (Wagon 29).

Proof Let X and Y be any two bounded sets with nonempty interior in $\mathbb{R}^3$, and let β_1 and β_2 be two solid balls in $\mathbb{R}^3$ such that $X \subseteq \beta_1$ and $\beta_2 \subseteq Y$. Take $n \in \mathbb{N}$ to be large enough so that β_1 may be covered by the union of n copies of β_2 (not necessarily disjoint).

Now, if S is a set of n disjoint copies of β_2 , then β_1 is equidecomposable (with respect to isometries) with a subset of S . The explanation for this is as follows: One can take the n copies of β_2 in S and translate them to cover β_1 . Let B be the union of these balls which covers β_1 . Then $\beta_1 \subseteq B$ and, hence, it is trivial that β_1 is equidecomposable with a subset of B . Furthermore, B is equidecomposable with a

subset of S since it was formed by translating and (perhaps) overlapping the pieces of S . However, if the balls do overlap when forming B , then B is not equidecomposable with all of S since, when partitioning B , the areas that overlap may only be counted once. Nonetheless, one has $\beta_1 \preceq B \preceq S$ and, hence, $\beta_1 \preceq S$.

Using the Banach-Tarski Paradox, one can also show that $S \preceq \beta_2$. It is already known that β_2 is paradoxical with respect to isometries. Thus, β_2 can be self-duplicated, with one of the copies translated to cover a ball in S and the other copy set aside to be self-duplicated again. After $n - 1$ iterations, one will have obtained S through repeated applications of the Banach-Tarski Paradox on β_2 , showing that $\beta_2 \sim_{G_1} S$.

Thus, $X \subseteq \beta_1 \preceq S \sim_{G_1} \beta_2 \subseteq Y$. However, Lemma 3.3 implies that if $X \subseteq \beta_1$ and $\beta_1 \preceq S$, then X is equidecomposable with some subset of S , resulting in $X \preceq S$. Hence, $X \preceq S \preceq Y$. Therefore $X \preceq Y$ by the transitivity of “ $\preceq$ ”.

A similar argument to the one above also shows that $Y \preceq X$.

Therefore, since “ $\preceq$ ” is antisymmetric, $X \sim_{G_1} Y$.

□

Chapter 4: The Banach-Tarski Paradox In Euclidean Spaces Higher Than $\mathbb{R}^3$

After proving the Banach-Tarski paradox in three dimensional Euclidean space, the next logical question is whether or not the paradox exists in Euclidean spaces of higher and lower dimensions. Of course, the notions of spheres and balls are somewhat different in these other spaces. For instance, a “sphere” in $\mathbb{R}^2$ is actually a circle and a “ball” is a disk. Similarly, in higher dimensions one finds hyperspheres and hyperballs, respectively. However, these are really just restrictions (in lower dimensional spaces) and extensions (in higher dimensional spaces) of the familiar notions of spheres and balls in three dimensional space. Thus, a logical plan of attack is to see if the concepts and methods used to prove the Banach-Tarski Paradox in $\mathbb{R}^3$ can, likewise, be restricted or extended to apply to other dimensional spaces.

Contrary to what may be a reasonable assumption that lower dimensions are easier to work in than higher dimensions, it is actually more straightforward to approach the existence of the Banach-Tarski Paradox in higher dimensions (greater than three) than it is in $\mathbb{R}$ and $\mathbb{R}^2$. This is because one can use mathematical induction to extend the argument for three dimensional Euclidean space into Euclidean spaces of dimensions greater than three. It is in order to take advantage of this directness that this author approaches the higher dimensions first. The following proposition bridges the gap between three dimensional and higher spaces, and allows one to extend the Banach-Tarski Paradox to $\mathbb{R}^n$ for $n \geq 3$.

Proposition**4.1**

Any unit sphere in $\mathbb{R}^n$ (with $n \geq 3$) centered at the origin is paradoxical with respect to SO_n (Wagon 53).

Proof It has already been proven in the previous chapter that S^2 (the unit sphere in $\mathbb{R}^3$ centered at the origin) is paradoxical with respect to SO_3 , which is the group of rotations of $\mathbb{R}^3$ (See Theorem 3.8). Suppose that this same result is true for S^{n-1} in $\mathbb{R}^n$. Then, recalling the definition of paradoxical, there is a partition $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ of S^{n-1} and elements $\delta_1, \dots, \delta_p, \varphi_1, \dots, \varphi_q$ of SO_n such that $S^{n-1} = \bigcup_{i=1}^p \delta_i Y_i = \bigcup_{j=1}^q \varphi_j Z_j$. (One should take note that when $n > 3$, SO_n is not the group of rotations of $\mathbb{R}^n$, but rather the group generated by the set of rotations of $\mathbb{R}^n$.)

Now consider S^n (the unit sphere in $\mathbb{R}^{n+1}$ centered at the origin). Let p^+ and p^- be the two poles, $(0, \dots, 0, 1)$ and $(0, \dots, 0, -1)$, respectively, of S^n . One begins by forming a partition of $S^n \setminus \{p^+, p^-\}$ as follows. For any point $(x_1, \dots, x_n, x_{n+1}) \in S^n \setminus \{p^+, p^-\}$, taking the n -tuple formed by the first n coordinates and normalizing it (by dividing the n -tuple by its norm) yields a point on S^{n-1} . Namely, $\frac{(x_1, \dots, x_n)}{|(x_1, \dots, x_n)|}$. Since this point is an element of S^{n-1} , it must belong to exactly one of

$Y_1, \dots, Y_p, Z_1, \dots, Z_q$, which partition S^{n-1} . Thus, define the sets

$\bar{Y}_1, \dots, \bar{Y}_p, \bar{Z}_1, \dots, \bar{Z}_q$ as follows:

$$\bar{Y}_i = \left\{ (x_1, \dots, x_n, x_{n+1}) \in S^n \setminus \{p^+, p^-\} : \frac{(x_1, \dots, x_n)}{|(x_1, \dots, x_n)|} \in Y_i \right\}$$

$$\bar{Z}_j = \left\{ (x_1, \dots, x_n, x_{n+1}) \in S^n \setminus \{p^+, p^-\} : \frac{(x_1, \dots, x_n)}{|(x_1, \dots, x_n)|} \in Z_j \right\}.$$

To verify that this is, in fact, a partition of $S^n \setminus \{p^+, p^-\}$, one needs only to call upon the facts that $\bar{Y}_1, \dots, \bar{Y}_p, \bar{Z}_1, \dots, \bar{Z}_q$ cover $S^n \setminus \{p^+, p^-\}$ since $\frac{(x_1, \dots, x_n)}{|(x_1, \dots, x_n)|}$ belongs

to some Y_i or Z_j for every $(x_1, \dots, x_n, x_{n+1}) \in S^n \setminus \{p^+, p^-\}$, and that

$Y_1, \dots, Y_p, Z_1, \dots, Z_q$ are pairwise disjoint since $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ are. As a

sidebar, it should be easily noted that the reason why one is restricted to $S^n \setminus \{p^+, p^-\}$,

and not able to partition all of S^n in this way, is because the first n -tuples of both p^+

and p^- are $(0, \dots, 0)$, and $\frac{(0, \dots, 0)}{|(0, \dots, 0)|}$ is undefined since $|(0, \dots, 0)| = 0$.

Now one turns to $\delta_1, \dots, \delta_p, \varphi_1, \dots, \varphi_q \in SO_n$ mentioned earlier. Recall that each one can be represented as an $n \times n$ orthogonal matrix with a determinant of 1.

Extend each of their representative matrices as follows. Let

$$\bar{\delta}_i = \begin{bmatrix} & & & 0 \\ & \delta_i & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \text{ and } \bar{\varphi}_i = \begin{bmatrix} & & & 0 \\ & \varphi_i & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \text{ giving}$$

$\bar{\delta}_1, \dots, \bar{\delta}_p, \bar{\varphi}_1, \dots, \bar{\varphi}_q \in SO_{n+1}$. Each of these extended matrices will act upon the

first n coordinates of $(x_1, \dots, x_n, x_{n+1})$ in exactly the same manner as their

predecessors, but will leave the $n+1^{\text{st}}$ coordinate fixed. Notice that each of these

extended matrices fixes the line through the two poles p^+ and p^- since

$$\bar{\delta}_i(0, \dots, 0, x_{n+1}) = (\delta_i(0, \dots, 0), x_{n+1}) = (0, \dots, 0, x_{n+1}) \text{ and}$$

$$\bar{\varphi}_i(0, \dots, 0, x_{n+1}) = (\varphi_i(0, \dots, 0), x_{n+1}) = (0, \dots, 0, x_{n+1}).$$

The claim, now, is that these extended matrices, along with the partition

$Y_1, \dots, Y_p, Z_1, \dots, Z_q$ form a paradoxical decomposition of $S^n \setminus \{p^+, p^-\}$ with respect

to SO_{n+1} . To see this, consider $\bar{\delta}_i(x_1, \dots, x_n, x_{n+1})$ for some $(x_1, \dots, x_n, x_{n+1}) \in Y_i$.

Since $\bar{\delta}_i$ fixes the last coordinate, x_{n+1} , and acts on the first n coordinates as δ_i acts on

$\frac{(x_1, \dots, x_n)}{|(x_1, \dots, x_n)|}$, the resulting vector is as follows: $\bar{\delta}_i(x_1, \dots, x_n, x_{n+1}) = (\bar{x}_1, \dots, \bar{x}_n, x_{n+1})$,

where $(\mathfrak{X}_1, \dots, \mathfrak{X}_n)$ is the vector obtained by normalizing $(x_1, \dots, x_n)$ to get $\frac{(x_1, \dots, x_n)}{|(x_1, \dots, x_n)|}$, applying the δ_i , and then multiplying by $|(x_1, \dots, x_n)|$ to reestablish the length of $(x_1, \dots, x_n)$. Since $|(\mathfrak{X}_1, \dots, \mathfrak{X}_n)| = |(x_1, \dots, x_n)| = \sqrt{1 - x_{n+1}^2}$, then $|(\mathfrak{X}_1, \dots, \mathfrak{X}_n, x_{n+1})| = 1$. Furthermore, since $\bigcup_{i=1}^p \delta_i Y_i$ gives all possible n -tuples of length 1 (since $\bigcup_{i=1}^p \delta_i Y_i = S^{n-1}$), then applying $\overline{\delta}_i$ to all $(x_1, \dots, x_n, x_{n+1})$ in Y_i , for all $1 \leq i \leq p$, gives all possible $n+1$ -tuples of length 1 ending in x_{n+1} . Since, according to the definition of Y_i , x_{n+1} can be any element of $\mathbb{R} \setminus \pm 1$, this yields all unit vectors in $\mathbb{R}^{n+1}$ except $(0, \dots, 0, \pm 1)$. In other words, $\bigcup_{i=1}^p \overline{\delta}_i Y_i = S^n \setminus \{p^+, p^-\}$. Furthermore, similarly defining sets Z_j and matrices $\overline{\varphi}_j$ yields $\bigcup_{j=1}^q \overline{\varphi}_j Z_j = S^n \setminus \{p^+, p^-\}$.

The above argument shows that $S^n \setminus \{p^+, p^-\}$ is paradoxical with respect to SO_{n+1} , but what affect does this have on the entirety of S^n ? Begin by letting

$$\rho = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & \cos(\pi\sqrt{2}) & -\sin(\pi\sqrt{2}) \\ 0 & \dots & \dots & \sin(\pi\sqrt{2}) & \cos(\pi\sqrt{2}) \end{bmatrix} \in SO_{n+1}.$$

One can see that ρ is a rotation of $\mathbb{R}^{n+1}$ which effects only the last two coordinates of any point, and leaves the first $n-1$ coordinates fixed. Furthermore, the discussion in chapter 1 verifies that ρ is of infinite order. Therefore,

$$\rho^n(\{p^+, p^-\}) \cap \{p^+, p^-\} = \emptyset \text{ for all } n \in \mathbb{N}.$$

Thus, if one lets $P = \bigcup_{n=0}^{\infty} \rho^n(\{p^+, p^-\})$, then one gets $S^n = P \cup S^n \setminus P$ and $S^n \setminus \{p^+, p^-\} = (\bigcup_{n=1}^{\infty} \rho^n(\{p^+, p^-\})) \cup S^n \setminus P = \rho(P) \cup S^n \setminus P$.

Therefore, S^n and $S^n \setminus \{p^+, p^-\}$ are equidecomposable with respect to SO_{n+1} and, hence, S^n is paradoxical with respect to SO_{n+1} by Corollary 3.7.

Thus, it has been shown that S^2 is paradoxical with respect to SO_3 , and that if S^{n-1} is paradoxical with respect to SO_n , then S^n is paradoxical with respect to SO_{n+1} .

Therefore, by mathematical induction, then unit sphere centered at the origin in $\mathbb{R}^n$, for any $n \geq 3$, is paradoxical with respect to SO_n .

□

Of course, the above argument could very easily be extended to encompass spheres of varying size located elsewhere in $\mathbb{R}^n$. All that would be required is to temporarily translate the center of the sphere to the origin, and to replace the sets in the paradoxical decomposition of the unit sphere with their projections onto the alternate sphere. Naturally, since translations are required, the sphere would not be paradoxical with respect to SO_n alone, but rather with respect to G_n , the set of isometries of $\mathbb{R}^n$. This argument, however, will not be formalized here since the main goal is to speak of balls rather than spheres. However, if one is interested in the formalization of the argument for arbitrary spheres, one need only to replace the radial correspondence ζ in the following argument with one that takes any $s \in S^{n-1}$ onto the point rs , rather than onto the set $\{\alpha s : 0 < \alpha \leq r\}$.

The following two theorems are of utmost importance because they expand the Banach-Tarski Paradox to higher dimensions. However, they follow from Proposition 4.1 in exactly the same manner in which Theorems 3.8 and 3.10 follow from the fact that S^2 is G_3 -paradoxical. Thus, if the proofs of Theorem 3.8 and 3.10 have been well understood, then a

quick overview should be sufficient enough to dispense with the following proofs. All steps, however, have been included in case a more careful perusal is desired.

Theorem (*Banach-Tarski Paradox for Higher Dimensions*)

4.2

Any solid ball in $\mathbb{R}^n$, with $n \geq 3$, is paradoxical with respect to G_n . In fact, this property holds for all of $\mathbb{R}^n$ (Wagon 53).

Proof Since it has already been established that S^{n-1} is paradoxical with respect to SO_n and, therefore, with respect to G_n , proving the first part of this statement is a relatively easy task. The second part will then be almost automatic.

Let, β be any ball of radius r centered at the origin. The unit sphere S^{n-1} can be taken onto $\beta \setminus 0$ by the radial correspondence ζ which takes the point $s \in S^{n-1}$ onto the set $\{\alpha s \mid 0 < \alpha \leq r\}$. Thus, if $S^{n-1} = \bigcup_{i=1}^p \delta_i Y_i = \bigcup_{j=1}^q \phi_j Z_j$ is a paradoxical decomposition of S^{n-1} , then $\beta \setminus 0 = \bigcup_{i=1}^p \delta_i(\zeta(Y_i)) = \bigcup_{j=1}^q \phi_j(\zeta(Z_j))$ is, likewise, a paradoxical decomposition of $\beta \setminus 0$.

Now, take ρ to be any rotation of $\mathbb{R}^n$ such that $\rho^n(0) \in \beta \setminus 0$ for all $n \in \mathbb{N}$. This is done by choosing ρ such that $\rho^n(0) \neq 0$ for any $n \in \mathbb{N}$, and such that $(0, \dots, 0, \frac{r}{2})$ belongs to the fixed set of ρ . (This fixed set is analogous to the axis of rotation if one were working in $\mathbb{R}^3$. However, in $\mathbb{R}^4$ it is a plane and in $\mathbb{R}^n$, $n > 4$, it is a $(n-2)$ -dimensional hyperplane. Choosing ρ so that this fixed set passes within $\frac{r}{2}$ of 0 ensures that $\rho^n(0) \in \beta$ for all $n \in \mathbb{N}$.) If ρ is chosen in this manner, then one has $\beta = P \cup \beta \setminus P$ and $\beta \setminus 0 = \bigcup_{n=1}^{\infty} \rho^n(0) \cup \beta \setminus P = \rho(\beta) \cup \beta \setminus P$, where $P = \bigcup_{n=0}^{\infty} \rho^n(0)$. Hence, β and $\beta \setminus 0$ are equidecomposable with respect to SO_n and β

is, therefore, paradoxical with respect to SO_n by Corollary 3.7.

If, on the other hand, β is any arbitrary ball in $\mathbb{R}^n$, let τ be the required translation to shift the center of β to the origin and let $\bar{\beta} = \tau\beta$. Since $\bar{\beta}$ is centered at the origin, it must have a paradoxical decomposition with respect to SO_n . Let $\bar{\beta} = \bigcup_{i=1}^p \delta_i Y_i = \bigcup_{j=1}^q \phi_j Z_j$ represent this decomposition. Then $\beta = \bigcup_{i=1}^p (\tau^{-1} \delta_i \tau)(Y_i) = \bigcup_{j=1}^q (\tau^{-1} \phi_j \tau)(Z_j)$, where $Y_i = \tau^{-1} Y_i$ and $Z_j = \tau^{-1} Z_j$, gives a paradoxical decomposition of β with respect to G_n .

Finally, if one were to define $\zeta(s)$ above to be the set $\{\alpha s \mid 0 < \alpha\}$, as opposed to $\{\alpha s \mid 0 < \alpha \leq r\}$, then the above argument would have proven that $\mathbb{R}^n$ is paradoxical with respect to SO_n since the sphere would correspond to the entire space rather than just to β .

□

Theorem (*Strong Form of the Banach-Tarski Paradox for Higher Dimensions*)

4.3

Any two bounded sets in $\mathbb{R}^n$ ($n \geq 3$) with nonempty interior are equidecomposable with respect to G_n (Wagon 53).

Proof Let X and Y be any two bounded sets with nonempty interior in $\mathbb{R}^n$, and let β_1 and β_2 be two solid balls in $\mathbb{R}^n$ such that $X \subseteq \beta_1$ and $\beta_2 \subseteq Y$. Take $k \in \mathbb{N}$ to be large enough so that β_1 may be covered by the union of k (not necessarily disjoint) copies of β_2 , and let S be a set of k disjoint copies of β_2 . These copies can be translated to cover β_1 . Let B be this covering. Keep in mind that B is not necessarily k copies of β_2 since

these copies may have overlapped to form the covering of β_1 . Note also that since $\beta_1 \subseteq B$ and every set is (trivially) equidecomposable to itself, then $\beta_1 \preceq B$.

If the k copies of β_2 in S do not overlap when forming B , then B is equidecomposable to S . If, however, the sets of S do overlap when forming B , then the overlapping regions can only be counted once in any partition of B . Thus, the most that can be said is that B is equidecomposable with a subset of S . Either way, one has $B \preceq S$. Combining this with $\beta_1 \preceq B$, and the transitivity of “ $\preceq$ ”, yields $\beta_1 \preceq S$. Furthermore, since $X \subseteq \beta_1$, then $X \preceq S$ by Lemma 3.3.

Now suppose that $A_1, \dots, A_p, B_1, \dots, B_q$ is a partition of β_2 that gives a paradoxical decomposition of β_2 with respect to G_n . One knows such a partition exists by Theorem 4.2. Then $\beta_2 = \bigcup_{i=1}^p g_i A_i$ for some collection of $g_i \in G_n$. Suppose this replication of β_2 by the sets $A = \{A_1, \dots, A_p\}$ were performed $k - 1$ times to achieve a set, call it Γ , of k distinct replicas of β_2 . Since these replicas could be translated to cover S , then S is equidecomposable with Γ . Furthermore, Γ is equidecomposable with A due to the way it was formed by successively repeating the Banach-Tarski replicating process $k - 1$ times. Therefore, one has $S \sim_{G_n} \Gamma \sim_{G_n} A \subset \beta_2 \subseteq Y$. Hence, $S \preceq Y$.

Combining the above results, $X \preceq S$ and $S \preceq Y$, yields $X \preceq Y$. A similar argument also yields $Y \preceq X$. Therefore, by the antisymmetry inherent in “ $\preceq$ ”, one has $X \sim_{G_n} Y$.

□

Part Two :

**The Banach-Tarski
Paradox In
Lower Dimensions**

Chapter 5: Lebesgue Measure

Thus far, it has been the goal of this thesis to show the existence of the Banach-Tarski Paradox in all Euclidean spaces of dimension three or higher. The attention will now be shifted to examining how the Paradox fares specifically in $\mathbb{R}$ and $\mathbb{R}^2$. To the astute reader, it may seem a logical assumption to surmise that the Paradox does not fare well in these lower dimensions, since they are being segregated from the other spaces in which the Paradox fares quite well. As it happens, this would be a very accurate observation. In fact, the Banach-Tarski Paradox does not exist at all in one and two dimensions, a fact which can be proven using the theories of measure and integration. (A brief review of integration has been included in appendix B.) To see this, one will need to make use of some notion of the relative “size” of a set. Notice that when a set is duplicated by a collection of subsets via the Banach-Tarski Paradox, the subsets seem to become somehow enlarged during the process so as to cover the space that was once occupied by the original set. If one can show that the relative “size” of a set in $\mathbb{R}$ or $\mathbb{R}^2$ remains unchanged under rigid motions, then there would seem to be no way to perform the Banach-Tarski Paradox in these spaces.

This notion of “size” that is being referred to is what mathematicians call the *measure* of the set. Specifically, a measure is a nonnegative countably additive (for pairwise disjoint sets) set function (Mattila 8) which maps certain subsets of a space into the set of extended real numbers, specifically mapping $\emptyset$ to 0 (Royden 253). In general terms, this simply means that

to each set a unique number, called the measure of the set, is assigned which indicates the relative size of the set. Relatively small sets may have measure zero, relatively large sets may have infinite measure, and other sets of varying size may have any measure in between. The countable additivity simply means that if the union of several mutually disjoint sets (up to countably many) is taken, then its measure will be the sums of the measures of the sets.

Now, there are two points of importance to remember when working with a measure. First, notice that the words “relative size” have been used here. This is deliberate as a set that we would normally view as being very large may actually have measure zero, meaning that it is actually quite small in relation to most other sets. For instance, the set of rational numbers, which has cardinality $\aleph_0$, has Lebesgue measure 0, whereas the interval $(0, 1)$ has Lebesgue measure 1. This is explained by the fact that even though there are infinitely many rational numbers, this is a countable infinity, which is relatively small to the uncountably many real numbers between 0 and 1. The second point to take note of is that a measure does not necessarily apply to all sets in a given space. In fact, there are certain sets, referred to as measurable sets, which make up the domain of the measure function. This will be explained in more detail as the particular measure that will be used here is developed.

First, begin in the space $\mathbb{R}$. Since the notion of a measure may not necessarily apply to all sets in $\mathbb{R}$, one can begin with a weaker notion known as outer measure which does apply to all of $\mathbb{R}$.

Definition (Outer Measure)**5.1**

An outer measure, m^* , on $\wp(\mathbb{R})$ is a function $m^* : \wp(\mathbb{R}) \rightarrow [0, \infty]$ such that:

- i) $m^*(\emptyset) = 0$
- ii) $X \subseteq Y \Rightarrow m^*(X) \leq m^*(Y)$
- iii) $m^*\left(\bigcup_{i=1}^{\infty} X_i\right) \leq \sum_{i=1}^{\infty} m^*(X_i)$

These three characteristics are known as (i and ii together) nonnegativity, (ii) monotonicity, and (iii) countable subadditivity. Notice that if the inequality in condition (iii) were changed to equality (with the condition that the X_i be pairwise disjoint), one would have a measure on $\wp(\mathbb{R})$ (Mattila 8, Royden 288).

Since any set function meeting the three conditions above is called an outer measure, one can imagine that several different outer measures can be defined on the same space. In fact, a wide range of work has been completed on the study of measures and outer measures in general, regardless of specifying a certain measure function. Many of the findings in this area of mathematics, known as measure theory, will be useful to us throughout this chapter. Whereas the specific findings pertinent to the discussion of the Banach-Tarski Paradox will be fully developed, the depth of general measure and integration theory will dictate that some more general facts be cited without proof. For the reader who is interested in further study of such topics, this author highly recommends chapters 3, 4, and 11-15 of Royden.

One of the most useful and intuitive forms of an outer measure on $\wp(\mathbb{R})$, which is the form

to be used here, is that known as Lebesgue outer measure in honor of its founder Henri Lebesgue.

Definition (*Lebesgue outer measure*)

5.2

Let $X \subseteq \mathbb{R}$ and consider all possible coverings of X by a collection of open intervals.

(i.e. Consider all collections $\{\mathcal{I}_i\}$ such that $X \subseteq \bigcup_i \mathcal{I}_i$.) Then the Lebesgue outer measure, λ^* , of X is:

$$\lambda^*(X) = \inf_{X \subseteq \bigcup_i \mathcal{I}_i} \sum_i l(\mathcal{I}_i)$$

where $l(\mathcal{I}_i)$ is the length of the interval $\mathcal{I}_i$ (Royden 56).

It is a worthwhile exercise to spend a few moments verifying that Lebesgue outer measure is, in fact, an outer measure as defined in Definition 5.1. The fact that $\lambda^*(\emptyset) = 0$ is fairly trivial and should not cause any difficulty. Consider, then, two sets X and Y such that $X \subseteq Y$. Any covering of Y must also be a covering of X since Y contains X . Hence, the collection of coverings of Y by open intervals is a subset of the collection of such coverings of X . In other words, $\{\{\mathcal{J}_j\} : Y \subseteq \bigcup \mathcal{J}_j\} \subseteq \{\{\mathcal{I}_i\} : X \subseteq \bigcup \mathcal{I}_i\}$. Thus, by a well-known property of the infimum, one has

$$\lambda^*(X) = \inf_{X \subseteq \bigcup \mathcal{I}_i} \sum_i l(\mathcal{I}_i) \leq \inf_{Y \subseteq \bigcup \mathcal{J}_j} \sum_j l(\mathcal{J}_j) = \lambda^*(Y).$$

Finally, consider a countable collection $\{X_k\}_{k=1}^{\infty}$ of subsets of $\mathbb{R}$ and let the notation

$X_k \subseteq \bigcup_i \mathcal{I}_{k,i}$ denote a covering of X_k by a collection open intervals. Suppose that $\lambda^*(X_k) < \infty$

for each k (since, if $\lambda^*(X_k) = \infty$ for any k , then condition (iii) would hold trivially). Since

$\lambda^*(X_k) = \inf_{X_k \subseteq \bigcup_i \mathcal{I}_{k,i}} \sum_i l(\mathcal{I}_{k,i})$, then given any $\varepsilon > 0$, there exists a covering $X_k \subseteq \bigcup_i \mathcal{I}_{k,i}$ such

that $\sum_i l(\mathcal{I}_{k,i}) < \lambda^*(X_k) + \frac{\varepsilon}{2^k}$. Consider such a covering for each X_k , $1 \leq k \leq \infty$. Then

$\bigcup_k \bigcup_i \mathcal{I}_{k,i}$ is a countable collection of open intervals that covers $\bigcup_k X_k$. Thus, one has

$$\lambda^*\left(\bigcup_k X_k\right) \leq \sum_k \sum_i l(\mathcal{I}_{k,i}) < \sum_k \left(\lambda^*(X_k) + \frac{\varepsilon}{2^k}\right) = \sum_k \lambda^*(X_k) + \varepsilon.$$

Hence, since ε can be any arbitrary nonnegative number, $\lambda^*\left(\bigcup_k X_k\right) \leq \sum_k \lambda^*(X_k)$ (Royden 57).

Thus, one sees that Lebesgue outer measure is indeed an outer measure as defined in

Definition 5.1. Now, one aspect of Lebesgue outer measure that is quite useful and should be fairly easily verified is the fact that the Lebesgue outer measure of any interval (open or closed) is simply the length of the interval.

It is easy to see from the definition of outer measure that Lebesgue outer measure is nonnegative. However, it is not countably additive but merely subadditive. It is this fact that withholds Lebesgue outer measure from being a measure. If, however, one restricts themselves to the sets in $\mathbb{R}$ for which Lebesgue outer measure is additive, then one would indeed have a measure on this subset of $\wp(\mathbb{R})$. The question, then, that emerges is that of deciding for which sets Lebesgue outer measure is countably additive. Let $\mathfrak{M}$ be the collection of sets X such that for any set Y , one has $\lambda^*(Y) = \lambda^*(Y \cap X) + \lambda^*(Y \setminus X)$. It can be shown that Lebesgue outer

measure is countably additive when applied to pairwise disjoint sets in $\mathfrak{M}$.

Proposition

5.3

If $\{X_i\}$ is a countable collection of pairwise disjoint sets from $\mathfrak{M}$, then

$$\lambda^*\left(\bigcup_i X_i\right) = \sum_i \lambda^*(X_i) \text{ (Royden 62).}$$

Proof Let X_1, X_2 be two disjoint sets from $\mathfrak{M}$. Then $\lambda^*(X_1 \cup X_2) =$

$$\lambda^*[(X_1 \cup X_2) \cap X_2] + \lambda^*[(X_1 \cup X_2) \setminus X_2] = \lambda^*(X_2) + \lambda^*(X_1) \text{ since } X_2 \in \mathfrak{M}.$$

Now suppose that if $\{X_i\}_{i=1}^n$ is a collection of pairwise disjoint sets from $\mathfrak{M}$, then $\lambda^*\left(\bigcup_{i=1}^n X_i\right) = \sum_{i=1}^n \lambda^*(X_i)$, and let $X_{n+1} \in \mathfrak{M}$ be disjoint with X_i for all

$1 \leq i \leq n$. Then

$$\begin{aligned} \lambda^*\left(\bigcup_{i=1}^{n+1} X_i\right) &= \lambda^*\left(\left(\bigcup_{i=1}^{n+1} X_i\right) \cap X_{n+1}\right) + \lambda^*\left(\left(\bigcup_{i=1}^{n+1} X_i\right) \setminus X_{n+1}\right) \\ &= \lambda^*(X_{n+1}) + \lambda^*\left(\bigcup_{i=1}^n X_i\right) = \sum_{i=1}^{n+1} \lambda^*(X_i). \end{aligned}$$

Thus, by mathematical induction, $\lambda^*\left(\bigcup_{i=1}^n X_i\right) = \sum_{i=1}^n \lambda^*(X_i)$ for all $n \in \mathbb{N}$ provided that the X_i are pairwise disjoint sets found in $\mathfrak{M}$.

The last case to verify, then, is that in which the collection $\{X_i\}$ is countably infinite. Note that for any $n \in \mathbb{N}$, $\bigcup_{i=1}^n X_i \subset \bigcup_{i=1}^{\infty} X_i$. Thus, $\sum_{i=1}^n \lambda^*(X_i) = \lambda^*\left(\bigcup_{i=1}^n X_i\right) \leq \lambda^*\left(\bigcup_{i=1}^{\infty} X_i\right)$ by the monotonicity of λ^* . Since this is true for all $n \in \mathbb{N}$, then $\sum_{i=1}^{\infty} \lambda^*(X_i) \leq \lambda^*\left(\bigcup_{i=1}^{\infty} X_i\right)$. Furthermore, since $\bigcup_{i=1}^{\infty} X_i$ is (trivially) a subset of itself, then invoking the subadditivity of λ^* yields

$$\lambda^*\left(\bigcup_{i=1}^{\infty} X_i\right) \leq \sum_{i=1}^{\infty} \lambda^*(X_i). \text{ Therefore, } \lambda^*\left(\bigcup_{i=1}^{\infty} X_i\right) = \sum_{i=1}^{\infty} \lambda^*(X_i).$$

Thus, for any countable collection $\{X_i\}$ of pairwise disjoint sets from $\mathfrak{M}$, $\lambda^*\left(\bigcup_i X_i\right) = \sum_i \lambda^*(X_i)$.

□

Thus, Lebesgue outer measure is a measure on the subset $\mathfrak{M}$ of $\mathcal{P}(\mathbb{R})$. Accordingly, the sets in $\mathfrak{M}$ are referred to as the Lebesgue measurable sets in $\mathbb{R}$ (Mattila 8, Royden 58). Thus, when dealing with a set in $\mathfrak{M}$, one refers to its Lebesgue measure, λ , as opposed to its Lebesgue outer measure (noting that these two are equivalent for such sets since Lebesgue measure is simply a restriction of Lebesgue outer measure to the collection $\mathfrak{M}$).

To get an idea of the scope of $\mathfrak{M}$, a fair amount of time will be spent discussing what sets in $\mathbb{R}$ are Lebesgue measurable. First of all, notice that the empty set is measurable since for any set $Y \subseteq \mathbb{R}$ one has $\lambda^*(Y \cap \emptyset) + \lambda^*(Y \setminus \emptyset) = \lambda^*(\emptyset) + \lambda^*(Y) = \lambda^*(Y)$, yielding $\emptyset \in \mathfrak{M}$.

In addition, for every measurable set X , the complement of X (denoted by X^c) must also be measurable (Royden 58) since $\lambda^*(Y) = \lambda^*(Y \cap X) + \lambda^*(Y \setminus X) = \lambda^*(Y \setminus X^c) + \lambda^*(Y \cap X^c)$. Now combining this with the previous fact that $\emptyset$ is measurable yields the fact that $\mathbb{R}$ is, itself, measurable also since it is the complement of $\emptyset$.

Furthermore, any set consisting of a single point must also be Lebesgue measurable, as witnessed by the following argument. Notice, first of all, that for any two sets X and Y , one has $Y = (Y \cap X) \cup (Y \setminus X)$. Hence, the subadditivity of λ^* will always yield $\lambda^*(Y) \leq \lambda^*(Y \cap X) + \lambda^*(Y \setminus X)$. Therefore, to show that $X \in \mathfrak{M}$, it is only necessary to show that $\lambda^*(Y) \geq \lambda^*(Y \cap X) + \lambda^*(Y \setminus X)$. That having been said, let $x \in \mathbb{R}$ and $Y \subset \mathbb{R}$. If $x \notin Y$, then $Y \cap x = \emptyset$ and $Y \setminus x = Y$. Hence,

$$\lambda^*(Y) = 0 + \lambda^*(Y) = \lambda^*(\emptyset) + \lambda^*(Y) = \lambda^*(Y \cap x) + \lambda^*(Y \setminus x).$$

Suppose, on the other hand, that $x \in Y$. Since $(Y \setminus x) \subset Y$, then using the monotonicity of

λ^* followed by the fact that $\lambda^*(x) = 0$, one obtains the following chain:

$$\lambda^*(Y) \geq \lambda^*(Y \setminus x) = 0 + \lambda^*(Y \setminus x) = \lambda^*(x) + \lambda^*(Y \setminus x) = \lambda^*(Y \cap x) + \lambda^*(Y \setminus x).$$

Therefore, $\{x\} \in \mathfrak{M}$.

This gives a reasonable starting point for understanding what sets in $\mathbb{R}$ are Lebesgue measurable, but they are merely the tip of the proverbial iceberg. In truth, the set $\mathfrak{M}$ is what is known as a σ -algebra, meaning that it is not only closed under complementation, but also under countable unions. (Notice that these two together ensure that $\mathfrak{M}$ is also closed under countable intersections since, by DeMorgan's Law, an intersection of sets is simply the complement of the union of their complements.)

Proposition

5.4

The collection of Lebesgue measurable sets in $\mathbb{R}$ is a σ -algebra (Mattila 8, Royden 60).

Proof It has already been shown that $\mathfrak{M}$ is closed under complementation, so begin by focusing attention on the case of $\bigcup_{i=1}^n X_i$ where $X_i \in \mathfrak{M}$ for each $1 \leq i \leq n$. One wishes to show first that $\bigcup_{i=1}^n X_i \in \mathfrak{M}$.

The case where $n = 1$ is trivial based on the conditions placed on the sets X_i .

Therefore, begin with the case $n = 2$. Let $Y \subseteq \mathbb{R}$. By the measurability of X_1 followed by that of X_2 , one has

$$\begin{aligned} \lambda^*(Y) &= \lambda^*(Y \cap X_1) + \lambda^*(Y \setminus X_1) \\ &= \lambda^*(Y \cap X_1) + \lambda^*[(Y \setminus X_1) \cap X_2] + \lambda^*[(Y \setminus X_1) \setminus X_2]. \end{aligned} \tag{1}$$

Notice, however, that

$$\lambda^*[Y \cap (X_1 \cup X_2)] \leq \lambda^*(Y \cap X_1) + \lambda^*[(Y \setminus X_1) \cap X_2]$$

since $Y \cap (X_1 \cup X_2) = (Y \cap X_1) \cup [(Y \setminus X_1) \cap X_2]$.

Thus, eqn. [1] becomes

$$\lambda^*(Y) \geq \lambda^*[Y \cap (X_1 \cup X_2)] + \lambda^*[(Y \setminus X_1) \setminus X_2].$$

Also, $(Y \setminus X_1) \setminus X_2 = Y \setminus (X_1 \cup X_2)$. Therefore,

$$\lambda^*(Y) \geq \lambda^*[Y \cap (X_1 \cup X_2)] + \lambda^*[Y \setminus (X_1 \cup X_2)]$$

and, hence, $X_1 \cup X_2 \in \mathfrak{M}$.

Suppose, now, that $\bigcup_{i=1}^n X_i \in \mathfrak{M}$, and consider $\bigcup_{i=1}^{n+1} X_i$ where X_{n+1} is also measurable. A similar argument as that above (replacing X_1 with $\bigcup_{i=1}^n X_i$ and X_2 with X_{n+1}) yields the following:

$$\begin{aligned} \lambda^*(Y) &\geq \lambda^*[Y \cap (\bigcup_{i=1}^n X_i \cup X_{n+1})] + \lambda^*[Y \setminus (\bigcup_{i=1}^n X_i \cup X_{n+1})] \\ &= \lambda^*[Y \cap (\bigcup_{i=1}^{n+1} X_i)] + \lambda^*[Y \setminus (\bigcup_{i=1}^{n+1} X_i)] \end{aligned}$$

Therefore, $\bigcup_{i=1}^{n+1} X_i \in \mathfrak{M}$, showing that $\mathfrak{M}$ is closed under finite unions.

For countable unions, take $\{X_i\}_{i=1}^{\infty}$ to be any sequence sets from $\mathfrak{M}$ having been partitioned to be pairwise disjoint. Since $\mathfrak{M}$ is closed under finite unions,

$\bigcup_{i=1}^n X_i \in \mathfrak{M}$ for any $n \in \mathbb{N}$. Thus, for any set Y ,

$$\lambda^*(Y) = \lambda^*(Y \cap \bigcup_{i=1}^n X_i) + \lambda^*(Y \setminus \bigcup_{i=1}^n X_i). \quad [2]$$

Also, $\bigcup_{i=1}^n X_i \subseteq \bigcup_{i=1}^{\infty} X_i$ implies $Y \setminus \bigcup_{i=1}^{\infty} X_i \subseteq Y \setminus \bigcup_{i=1}^n X_i$, yielding

$$\lambda^*(Y \setminus \bigcup_{i=1}^{\infty} X_i) \leq \lambda^*(Y \setminus \bigcup_{i=1}^n X_i) \quad [3]$$

Combining eqn. [2] and ineq. [3], one obtains

$$\lambda^*(Y) \geq \lambda^*(Y \cap \bigcup_{i=1}^n X_i) + \lambda^*(Y \setminus \bigcup_{i=1}^{\infty} X_i). \quad [4]$$

Now, examine the term $\lambda^*(Y \cap \bigcup_{i=1}^n X_i)$ in ineq. [4] and notice that since

$X_n \in \mathfrak{M}$, this term becomes

$$\begin{aligned}\lambda^*(Y \cap \bigcup_{i=1}^n X_i) &= \lambda^*[(Y \cap \bigcup_{i=1}^n X_i) \cap X_n] + \lambda^*[(Y \cap \bigcup_{i=1}^n X_i) \setminus X_n] \\ &= \lambda^*(Y \cap X_n) + \lambda^*(Y \cap \bigcup_{i=1}^{n-1} X_i)\end{aligned}$$

Since this is true for any $n \in \mathbb{N}$, then one can use induction to repetitively break down the final term as follows:

$$\begin{aligned}\lambda^*(Y \cap \bigcup_{i=1}^n X_i) &= \lambda^*(Y \cap X_n) + \lambda^*(Y \cap \bigcup_{i=1}^{n-1} X_i) \\ &= \lambda^*(Y \cap X_n) + \lambda^*(Y \cap X_{n-1}) + \lambda^*(Y \cap \bigcup_{i=1}^{n-2} X_i) \\ &\quad \vdots \\ &= \sum_{i=1}^n \lambda^*(Y \cap X_i)\end{aligned}$$

Thus, **ineq. [4]** becomes $\lambda^*(Y) \geq \sum_{i=1}^n \mu^*(Y \cap X_i) + \lambda^*(Y \setminus \bigcup_{i=1}^n X_i)$. Since this is true for all n , then $\lambda^*(Y) \geq \sum_{i=1}^\infty \mu^*(Y \cap X_i) + \lambda^*(Y \setminus \bigcup_{i=1}^\infty X_i)$. Furthermore, since $Y \cap \bigcup_{i=1}^\infty X_i = \bigcup_{i=1}^\infty (Y \cap X_i)$, then

$$\lambda^*(Y \cap \bigcup_{i=1}^\infty X_i) \leq \sum_{i=1}^\infty \lambda^*(Y \cap X_i)$$

by the subadditivity of λ^* .

Hence, $\lambda^*(Y) \geq \lambda^*(Y \cap \bigcup_{i=1}^\infty X_i) + \lambda^*(Y \setminus \bigcup_{i=1}^\infty X_i)$ and $\bigcup_{i=1}^\infty X_i \in \mathfrak{M}$.

Therefore, $\mathfrak{M}$ is closed under complementation and countable unions (and, therefore, countable intersections by DeMorgan's Law) and, hence, is a σ -algebra.

□

This fact that $\mathfrak{M}$ is a σ -algebra gives a fairly large degree of freedom in formulating sets that are Lebesgue measurable beyond just the empty set, $\{x\} : x \in \mathbb{R}$, and $\mathbb{R}$ itself. For instance, consider the open interval (x, ∞) for some $x \in \mathbb{R}$. For any set $Y \subseteq \mathbb{R}$ and any $\varepsilon > 0$,

there is a countable collection $\{I_n\}$ of open intervals such that $Y \subseteq \bigcup I_n$ and

$\sum l(I_n) \leq \lambda^*(Y) + \varepsilon$. If, for each n , one defines the sets I'_n and I''_n to be

$$I'_n = I_n \cap (x, \infty)$$

$$I''_n = I_n \cap (-\infty, x],$$

then $[Y \cap (x, \infty)] \subseteq \bigcup I'_n$ and $[Y \cap (-\infty, x]] \subseteq \bigcup I''_n$. Furthermore, since each I'_n and I''_n is

either empty or is an interval, its outer Lebesgue measure is its length. Thus, one has

$$\lambda^*[Y \cap (x, \infty)] + \lambda^*[Y \setminus (x, \infty)] = \lambda^*[Y \cap (x, \infty)] + \lambda^*[Y \cap (-\infty, x]]$$

$$\leq \lambda^*(\bigcup I'_n) + \lambda^*(\bigcup I''_n)$$

$$\leq \sum \lambda^*(I'_n) + \sum \lambda^*(I''_n)$$

$$= \sum [\lambda^*(I'_n) + \lambda^*(I''_n)]$$

$$= \sum [l(I'_n) + l(I''_n)] = \sum l(I_n) \leq \lambda^*(Y) + \varepsilon$$

Thus, since ε is any arbitrary positive number, $\lambda^*[Y \cap (x, \infty)] + \lambda^*[Y \setminus (x, \infty)] \leq \lambda^*(Y)$.

Therefore, $\lambda^*(Y) = \lambda^*[Y \cap (x, \infty)] + \lambda^*[Y \setminus (x, \infty)]$ and $(x, \infty) \in \mathfrak{M}$ (Royden 60).

It is at this point that the fact of $\mathfrak{M}$ being a σ -algebra can be called upon to produce numerous further results. Since $(x, \infty) \in \mathfrak{M}$, then $[x, \infty) = \{x\} \cup (x, \infty) \in \mathfrak{M}$. Furthermore, $(-\infty, x) = [x, \infty)^c \in \mathfrak{M}$ and $(-\infty, x] = (x, \infty)^c \in \mathfrak{M}$. Hence, for any $x, y \in \mathbb{R}$, the intervals $(x, y) \in \mathfrak{M}$ and $[x, y] \in \mathfrak{M}$ since $(x, y) = (-\infty, y) \cap (x, \infty)$ and $[x, y] = (-\infty, y] \cap [x, \infty)$ (Royden 60).

Therefore, every interval is Lebesgue measurable. Furthermore, since every open set is a countable union of open intervals, then every open set (and, hence, every closed set) is

Lebesgue measurable, as are their unions and intersections. The correct terminology for this is to say that every Borel set is Lebesgue measurable (Royden 61), where the family $\mathfrak{B}$ of Borel sets is the smallest σ -algebra containing the collection of open sets in $\mathbb{R}$ (Royden 52).

(However, it is not true that every Lebesgue measurable set is a Borel set.) Thus, a great many sets that one deals with in common mathematical work on $\mathbb{R}$ are Lebesgue measurable.

However, one should not rush to the assumption that every set in $\mathbb{R}$ is Lebesgue measurable, for this is not true. Based on what is known about the Lebesgue measurable sets, any set that is not measurable must be somewhat pathological and difficult to visualize in the mind, for any set that is easy to visualize is most likely a Borel set. These sets do, however, exist just as measurable sets do. For a specific example of a Lebesgue nonmeasurable set in $\mathbb{R}$, one should see appendix A included at the end of this thesis.

The final point to make about Lebesgue measure in $\mathbb{R}$ is about its invariant properties.

Notice that Lebesgue measure is based upon the length of an interval, which does not change when the interval is translated along the real line or reflected about some point on the real line.

Hence, Lebesgue measure is invariant with respect to translations and reflections. Since all

isometries of $\mathbb{R}$ consist of compositions of these two types of transformations, then the

Lebesgue measure of a set will remain invariant under any transformation in G_1 , the group of isometries of $\mathbb{R}$. This property is known as isometry invariance and guarantees that the

Lebesgue measure of a set in $\mathbb{R}$ will remain the same no matter where it is found on the real line.

Chapter 6: Lebesgue Measure In $\mathbb{R}^2$

Hitherto, the discussion of Lebesgue measure has been limited to the space $\mathbb{R}$. It is, then, a reasonable question to ask what Lebesgue measure would look like in higher dimensional spaces. The focus will now shift defining Lebesgue measure in $\mathbb{R}^2$, which is known as two dimensional Lebesgue measure.

Suppose that $X, Y \subseteq \mathbb{R}$. Then $X \times Y$ is a subset of $\mathbb{R}^2$ and is referred to as a rectangle. Consider the collection of rectangles formed by Lebesgue measurable sets from $\mathbb{R}$. This collection is $\mathfrak{M} \times \mathfrak{M}$, but will be denoted here by R . Notice that R is not an algebra since it is not closed under finite unions. For example, $[0, 1]^2 \cup [1, 2]^2 \notin R$ (where $[0, 1]^2 = [0, 1] \times [0, 1]$ and $[1, 2]^2 = [1, 2] \times [1, 2]$). Rather, R is what is referred to as a semi-algebra, meaning that it is closed under finite intersections and the complement of any measurable rectangle is the finite union of disjoint measurable rectangles (Royden 297). To see this, let $R_1 = X_1 \times Y_1$ and $R_2 = X_2 \times Y_2$ be measurable rectangles. Then

$$R_1 \cap R_2 = (X_1 \times Y_1) \cap (X_2 \times Y_2) = (X_1 \cap X_2) \times (Y_1 \cap Y_2)$$

$$R_1^c = (X_1 \times Y_1)^c = (X_1^c \times Y_1) \cup (X_1 \times Y_1^c) \cup (X_1^c \times Y_1^c) \text{ (Royden 303).}$$

It is desirable to define a measure on R . Notice, however, that measures can only be defined on σ -algebras since, in order to have $\mu\left(\bigcup_{i=1}^{\infty} X_i\right) = \sum_{i=1}^{\infty} \mu(X_i)$ for all disjoint sequences $\{X_i\}_{i=1}^{\infty}$, one needs $\bigcup_{i=1}^{\infty} X_i$ to be in the domain of μ . However, since a countable union of disjoint measurable rectangles is not necessarily a measurable rectangle, one would not be able

to achieve countable additivity for any function defined solely on R . Thus, what is required is to expand R to a σ -algebra, and try to build a measure on this expanded collection of sets.

First, begin by expanding R to include all finite unions of disjoint measurable rectangles.

Since this expanded collection is an algebra (called the algebra generated by R), it will be denoted by $\mathcal{A}_R$. Now, define the function $\alpha : \mathcal{A}_R \rightarrow [0, \infty]$ in the following way. Let $S \in \mathcal{A}_R$. If $S \in R$, then there exist two Lebesgue measurable sets, X and Y , from $\mathbb{R}$ such that $S = X \times Y$. In this case, let $\alpha(S) = \lambda(X) \cdot \lambda(Y)$ (Royden 303). If, however, S is a finite union of disjoint measurable rectangles $\bigcup_{i=1}^n R_i$ that is not found in R , then let $\alpha(S) = \sum_{i=1}^n \alpha(R_i)$.

Notice that α is nonnegative and $\alpha(\emptyset) = 0$. The fact, however, that α is also finitely additive is more difficult to see. Suppose $\{R_i\}_{i=1}^n$ is a finite sequence of pairwise disjoint measurable rectangles. If $\bigcup_{i=1}^n R_i \notin R$, then $\alpha(\bigcup_{i=1}^n R_i) = \sum_{i=1}^n \alpha(R_i)$ by the definition of α . Suppose, however, that $\bigcup_{i=1}^n R_i \in R$. Then $\bigcup_{i=1}^n R_i$ is a measurable rectangle and, hence, equals $X \times Y$ for some $X, Y \in \mathcal{M}$. Since each $R_i \in R$, let $X_i \times Y_i = R_i$ for each $1 \leq i \leq n$. It is true that $X = \bigcup_{i=1}^n X_i$ and $Y = \bigcup_{i=1}^n Y_i$. However, it is not true that $\{X_i\}_{i=1}^n$ and $\{Y_i\}_{i=1}^n$ are each disjoint sequences (for if they were, $\bigcup_{i=1}^n R_i$ would not be found in R). See, for instance, the example on the following page involving $\bigcup_{i=1}^4 R_i \in R$. Notice that $\bigcup_{i=1}^4 R_i = X \times Y$ where $X = \bigcup_{i=1}^4 X_i$ and $Y = \bigcup_{i=1}^4 Y_i$, but neither of these two unions are disjoint.

It is true in the above example that $\alpha\left(\bigcup_{i=1}^4 R_i\right) = \lambda(X) \cdot \lambda(Y)$. However, this is not simply $\sum_{i=1}^4 \lambda(X_i) \cdot \sum_{i=1}^4 \lambda(Y_i)$ since the additivity of λ applies only to disjoint unions. Thus, to find $\lambda(Y)$, begin by fixing one point x in X . Since the rectangles R_i are disjoint, then for each $y \in Y$, the point (x, y) can be found in exactly one R_i . Therefore, Y is the disjoint union $\bigcup_{i \in I} Y_i$, where $I = \{i : x \in X_i\}$ (Royden 304). For instance, if x is fixed as shown in the example below, then $I = \{1, 3\}$ and $Y = Y_1 \cup Y_3$ which is a union of disjoint sets.

This yields the equation $\lambda(Y) \cdot \chi_X(x) = \sum_{i=1}^n \lambda(Y_i) \cdot \chi_{X_i}(x)$ which can be used in the following argument to show the finite additivity of α .

$$\begin{aligned}
 \alpha\left(\bigcup_{i=1}^n R_i\right) &= \lambda(X) \cdot \lambda(Y) = \lambda(Y) \int_X \chi_X d\lambda \\
 &= \int_X \lambda(Y) \cdot \chi_X d\lambda = \int_X \left(\sum_{i=1}^n \lambda(Y_i) \cdot \chi_{X_i}\right) d\lambda \\
 &= \sum_{i=1}^n \left(\int_X \lambda(Y_i) \cdot \chi_{X_i} d\lambda\right) = \sum_{i=1}^n \lambda(Y_i) \int_X \chi_{X_i} d\lambda \\
 &= \sum_{i=1}^n \lambda(Y_i) \cdot \lambda(X_i) = \sum_{i=1}^n \alpha(R_i) \quad (\text{Royden 304})
 \end{aligned}$$

Thus, α has many of the important characteristics of a measure except for the fact that its domain is still not closed under countable unions since $\mathcal{A}_R$ is not a σ -algebra. It is impossible, then, for α to achieve countable additivity. Therefore, α is referred to as a measure on the algebra $\mathcal{A}_R$. A *measure on an algebra* is a nonnegative finitely additive function that maps $\emptyset$ to 0 (Royden 291), and is not to be confused with a measure in general. Thankfully, however, α can be extended to a measure using a process credited to Carathéodory. Although the full text of Carathéodory's Extension Theorem goes beyond that which is stated below, the results needed for this discussion are included here.

Theorem (*Carathéodory's Extension Theorem*)

6.1

If μ is a measure on the algebra $\mathcal{A}$, then its corresponding outer measure μ^* restricted to the collection of μ^* -measurable sets is a measure that extends μ to a σ -algebra containing $\mathcal{A}$ (Royden 295-296).

Proof   *This proof is somewhat tedious, so it is helpful to outline what must be shown*

before one begins. First, the outer measure μ^* needs to be defined. Then, it must be shown that the restriction of μ^* to M (the collection of μ^* -measurable sets) is a measure, which involves showing that M is a σ -algebra and that μ^* is a countably additive nonnegative function on M with $\mu^*(\emptyset) = 0$. Lastly, it needs to be shown that μ^* , when restricted to M , is an extension of μ . This is done by showing that $A \subset M$ and $\mu^*(X) = \mu(X)$ for all $X \in A$.

Let μ be a measure on the algebra A . The first order of business is to define μ^* and show that it is an outer measure. For any set X , consider all possible sequences $\{A_i\}$ of sets from A such that $X \subseteq \bigcup A_i$ and define

$$\mu^*(X) := \inf_{X \subseteq \bigcup A_i} \sum \mu(A_i).$$

Since μ is nonnegative with $\mu(\emptyset) = 0$, then the same is true of μ^* . Furthermore, μ^* , being a function based on the infimum, is monotonic. To see that μ^* is countably subadditive, let $X = \bigcup_{j=1}^{\infty} X_j$. If, for some j , $\mu^*(X_j) = \infty$, then it is given that $\mu^*(X) \leq \sum_{j=1}^{\infty} \mu^*(X_j)$. Suppose, then, that $\mu^*(X_j) < \infty$ for all j . Since $\mu^*(X) = \inf \sum \mu(A_i)$ over all sequences $\{A_i\}$ from A that cover X , then given any $\varepsilon > 0$, there is for each X_j a sequence $\{A_{ji}\}$ of sets from A such that $X_j \subseteq \bigcup_i A_{ji}$ and $\sum_i \mu(A_{ji}) < \mu^*(X_j) + \frac{\varepsilon}{2^j}$. Furthermore, since $X \subseteq \bigcup_{j=1}^{\infty} \bigcup_i A_{ji}$, then using the well known property of the infimum, one obtains

$$\mu^*(X) \leq \sum_{j=1}^{\infty} \sum_i \mu(A_{ji}) < \sum_{j=1}^{\infty} \left(\mu^*(X_j) + \frac{\varepsilon}{2^j} \right) = \sum_{j=1}^{\infty} \mu^*(X_j) + \varepsilon.$$

Thus, since ε is an arbitrary positive number, $\mu^*(X) \leq \sum_{j=1}^{\infty} \mu^*(X_j)$.

Therefore, μ^* is countably subadditive and, hence, an outer measure.

Now let M be the collection of μ^* -measurable sets. That is, the collection of

sets X such that, for every set Y , $\mu^*(Y) = \mu^*(Y \cap X) + \mu^*(Y \setminus X)$. Notice that $\emptyset \in M$ since $\mu^*(Y) = 0 + \mu^*(Y) = \mu^*(\emptyset) + \mu^*(Y) = \mu^*(Y \cap \emptyset) + \mu^*(Y \setminus \emptyset)$ for any Y .

Notice also that if $X \in M$, then its complement can also be found in M since, for any Y , $\mu^*(Y) = \mu^*(Y \cap X) + \mu^*(Y \setminus X) = \mu^*(Y \setminus X^c) + \mu^*(Y \cap X^c)$. Thus, M is closed under complementation.

To see that M is closed under countable unions, one must begin first by showing that the same is true for finite unions. To do this, let $X_1, X_2 \in M$ and let Y again be any set. Using the measurability of X_1 and X_2 , one obtains

$$\begin{aligned} \mu^*(Y) &= \mu^*(Y \cap X_1) + \mu^*(Y \setminus X_1) \\ &= \mu^*(Y \cap X_1) + \mu^*[(Y \setminus X_1) \cap X_2] + \mu^*[(Y \setminus X_1) \setminus X_2]. \end{aligned} \quad [1]$$

Also, since $Y \cap (X_1 \cup X_2) = (Y \cap X_1) \cup [(Y \setminus X_1) \cap X_2]$, the subadditivity of μ^* implies that $\mu^*[Y \cap (X_1 \cup X_2)] \leq \mu^*(Y \cap X_1) + \mu^*[(Y \setminus X_1) \cap X_2]$. Therefore, eqn. [1] becomes

$$\begin{aligned} \mu^*(Y) &\geq \mu^*[Y \cap (X_1 \cup X_2)] + \mu^*[(Y \setminus X_1) \setminus X_2] \\ &= \mu^*[Y \cap (X_1 \cup X_2)] + \mu^*[Y \setminus (X_1 \cup X_2)]. \end{aligned}$$

Since the subadditivity of μ^* automatically yields the same inequality in the opposite direction, $X_1 \cup X_2 \in M$. By induction, then, it is seen that M is closed under the union of any finite collection of sets.

To extend this closure to countably infinite unions, take $\{X_i\}_{i=1}^{\infty}$ to be any sequence of pairwise disjoint μ^* -measurable sets (keeping in mind that any sequence can be partitioned down so as to make it pairwise disjoint). Let $\bar{X}_n = \bigcup_{i=1}^n X_i$ and $\bar{X} = \bigcup_{i=1}^{\infty} X_i$. It has already been shown that for every $n \in \mathbb{N}$, $\bar{X}_n \in M$. Hence, $\mu^*(Y) = \mu^*(Y \cap \bar{X}_n) + \mu^*(Y \setminus \bar{X}_n)$ for any set Y . Furthermore, $Y \setminus \bar{X} \subseteq Y \setminus \bar{X}_n$

since $X_n \subseteq X$, yielding the inequality $\mu^*(Y \cap X) \leq \mu^*(Y \cap X_n)$. Therefore,

$$\mu^*(Y) \geq \mu^*(Y \cap X_n) + \mu^*(Y \setminus X). \quad [2]$$

Consider the term $\mu^*(Y \cap X_n)$ in the above inequality. Since $X_n \in M$, this term is equal to $\mu^*[(Y \cap X_n) \cap X_n] + \mu^*[(Y \cap X_n) \setminus X_n]$. This, however, is equal to $\mu^*(Y \cap X_n) + \mu^*(Y \cap X_{n-1})$ since $X_n \cap X_n = X_n$ and $X_n \setminus X_n = X_{n-1}$. Since this is true for all $n \in \mathbb{N}$, then by induction, one finds:

$$\begin{aligned} \mu^*(Y \cap X_n) &= \mu^*(Y \cap X_n) + \mu^*(Y \cap X_{n-1}) \\ &= \mu^*(Y \cap X_n) + \mu^*(Y \cap X_{n-1}) + \mu^*(Y \cap X_{n-2}) \\ &\vdots \\ &= \sum_{i=1}^n \mu^*(Y \cap X_i) \end{aligned}$$

Combining this with ineq. [2] yields $\mu^*(Y) \geq \sum_{i=1}^n \mu^*(Y \cap X_i) + \mu^*(Y \setminus X)$.

Since this holds for all $n \in \mathbb{N}$, then $\mu^*(Y) \geq \sum_{i=1}^{\infty} \mu^*(Y \cap X_i) + \mu^*(Y \setminus X)$.

Furthermore, $Y \cap X = \bigcup_{i=1}^{\infty} Y \cap X_i$, so $\mu^*(Y \cap X) \leq \sum_{i=1}^{\infty} \mu^*(Y \cap X_i)$ by the subadditivity of μ^* . Therefore, $\mu^*(Y) \geq \mu^*(Y \cap X) + \mu^*(Y \setminus X)$ (with the opposite direction guaranteed by subadditivity), implying $X \in M$.

Therefore, M is closed under countable unions (and countable intersections since, by DeMorgan's Law, $\bigcap X_i = (\bigcup X_i^c)^c$) and, hence, is a σ -algebra.

Now, since M is a σ -algebra, it is possible to define a full measure on M (meaning a function conforming to the actual definition of a measure, without having restrictions imposed upon it). To see that μ^* is just such a function, note that it has already been shown that μ^* is nonnegative with $\mu^*(\emptyset) = 0$. Thus, it remains only to be shown that μ^* is countably additive on M . To see this, let $X_1, X_2 \in M$ with

$X_1 \cap X_2 = \emptyset$. Then, since $X_1 \in M$, one obtains the following:

$$\mu^*(X_1 \cup X_2) = \mu^*[(X_1 \cup X_2) \cap X_1] + \mu^*[(X_1 \cup X_2) \setminus X_1] = \mu^*(X_1) + \mu^*(X_2).$$

Hence, μ^* is finitely additive on M .

Now let $X = \bigcup_{i=1}^{\infty} X_i$ for some disjoint sequence $\{X_i\}_{i=1}^{\infty}$ of μ^* -measurable sets. Then, for any $n \in \mathbb{N}$, $\sum_{i=1}^n \mu^*(X_i) = \mu^*(\bigcup_{i=1}^n X_i)$ by the finite additivity of μ^* on M . Furthermore, since $\bigcup_{i=1}^n X_i \subseteq X$, the subadditivity of μ^* yields $\mu^*(\bigcup_{i=1}^n X_i) \leq \mu^*(X)$. Thus, $\sum_{i=1}^n \mu^*(X_i) \leq \mu^*(X)$. Since this is true for any $n \in \mathbb{N}$, then $\sum_{i=1}^{\infty} \mu^*(X_i) \leq \mu^*(X)$. However, calling upon the subadditivity of μ^* once again yields $\mu^*(X) = \mu^*(\bigcup_{i=1}^{\infty} X_i) \leq \sum_{i=1}^{\infty} \mu^*(X_i)$. Therefore, $\mu^*(X) = \sum_{i=1}^{\infty} \mu^*(X_i)$, indicating that μ^* is countably additive on M .

Therefore, the function μ_M^* (where μ_M^* is used to denote the restriction of μ^* to M) is, indeed, a measure.

The only portion of Carathéodory's theorem that remains to be justified is that μ_M^* is, in fact, an extension of μ . This involves showing that $A \subset M$ and that $\mu_M^*(X) = \mu(X)$ for all $X \in A$.

To see that M contains A , let $X \in A$ and $\varepsilon > 0$. Take Y to be any set such that $\mu^*(Y) < \infty$. Since $\mu^*(Y) = \inf \sum \mu(A_i)$ over all sequences $\{A_i\}$ from A that cover Y , there exists some such sequence such that $\sum \mu(A_i) < \mu^*(Y) + \varepsilon$.

Furthermore, since μ is additive on A , then

$$\sum \mu(A_i) = \sum \mu(A_i \cap X) + \sum \mu(A_i \setminus X)$$

for each i . Hence, $\sum \mu(A_i \cap X) + \sum \mu(A_i \setminus X) < \mu^*(Y) + \varepsilon$.

Finally, $Y \cap X \subseteq \bigcup (A_i \cap X)$ implies $\mu^*(Y \cap X) \leq \sum \mu(A_i \cap X)$ and $Y \setminus X \subseteq \bigcup (A_i \setminus X)$ implies $\mu^*(Y \setminus X) \leq \sum \mu(A_i \setminus X)$. Therefore,

$\mu^*(Y \cap X) + \mu^*(Y \setminus X) < \mu^*(Y) + \varepsilon$. Since this is true for any $\varepsilon > 0$, then

$\mu^*(Y) \geq \mu^*(Y \cap X) + \mu^*(Y \setminus X)$, and since the subadditivity of μ^* yields the same inequality in the opposite direction, then $\mu^*(Y) = \mu^*(Y \cap X) + \mu^*(Y \setminus X)$. Hence,

$X \in M$ and, therefore, $A \subset M$.

To see that $\mu(X) = \mu^*(X)$, suppose now that $\{A_i\}$ is any sequence of sets from A that covers X and let $X_i = X \cap (A_i \setminus \bigcup_{k=1}^{i-1} A_k)$. Then, for all i , $X_i \subseteq A_i$ and, hence, $\mu(X_i) \leq \mu(A_i)$. Furthermore, the sequence $\{X_i\}$ is pairwise disjoint with $\bigcup X_i = X$. Thus, $\mu(X) = \sum \mu(X_i) \leq \sum \mu(A_i)$. Since this is true for all such sequences $\{A_i\}$, then $\mu(X) \leq \inf_{X \subseteq \bigcup A_i} \sum \mu(A_i) = \mu^*(X)$. However, since $\{X_i\}$ is a disjoint sequence of sets from A whose union is X , then $\mu(X) = \sum \mu(X_i) \geq \inf_{X \subseteq \bigcup A_i} \sum \mu(A_i) = \mu^*(X)$.

Therefore, $\mu(X) = \mu^*(X)$ for all $X \in A$, thus showing that μ^* (and, hence, μ_M^*) is an extension of μ .

□

Therefore, since α is a measure on the algebra A_R , it can be extended to a true measure using its outer measure $\alpha^*(S) := \inf_{S \subseteq \bigcup R_i} \sum \alpha(R_i)$. This is known as the Carathéodory process.

If $\mathfrak{M}^2$ is used to denote the collection of α^* -measurable sets (which is not to be confused with $\mathfrak{M} \times \mathfrak{M} = R$ but, rather, is a σ -algebra containing $\mathfrak{M} \times \mathfrak{M}$), then the restriction of α^* to $\mathfrak{M}^2$ is a measure that extends α .

The outer measure α^* is known as two dimensional Lebesgue outer measure. Hence, one will usually find it denoted by λ^* . Since the restriction of λ^* to $\mathfrak{M}^2$ is a measure, this restriction

will be denoted from this point forward simply by λ . This measure, λ , is two dimensional Lebesgue measure (Royden 304), and $\mathcal{M}^2$ is the collection of Lebesgue measurable sets in $\mathbb{R}^2$. (Notice that the same notation is used for Lebesgue measure regardless of the dimension. Thus one is to keep track of which measure is being denoted based on the space from which the sets originate.)

As a point of interest referring back to chapter 5, note that one dimensional Lebesgue measure may very well have been developed using the Carathéodory process as well. Let L denote the collection of intervals of the real line. Since the intersection of any two intervals is a (possibly empty) interval, and the complement of any interval is the finite union of disjoint intervals, then L is a semi-algebra. Hence, by including all finite unions of disjoint intervals, L can be extended to the algebra $\mathcal{A}_L$. If one defines the measure on $\mathcal{A}_L$

$$l(X) := \begin{cases} \text{the length of } X \text{ if } X \text{ is an interval} \\ \text{the sum of the lengths of } \mathcal{I}_i \text{ if } X \text{ is the union of the disjoint intervals } \mathcal{I}_i \end{cases},$$

then the corresponding outer measure is the usual Lebesgue outer measure for the real line.

Extending this outer measure to the collection λ^* -measurable sets yields a measure, λ , on $\mathcal{M}$ which is the same Lebesgue measure that was defined earlier in chapter 5.

Now, the argument that two dimensional Lebesgue measure is isometry invariant is, unfortunately, not as straightforward as that for one dimensional Lebesgue measure. It is clear that two dimensional Lebesgue measure is translation invariant since it is constructed directly from one dimensional Lebesgue measure which already possesses this property. However,

reflections in $\mathbb{R}^2$ are not always as simple as those in $\mathbb{R}$. In addition to this, G_2 also contains rotations which are not found on the real line at all. Thus, to show that two dimensional Lebesgue measure is isometry invariant, one uses a clever argument utilizing the characteristics of a certain class of measures known as the Baire measures. Although the proof is quite lengthy and tedious (it is understandable, but beyond the scope of this thesis), it can be shown that there is a unique Baire measure that is invariant with respect to G_2 . (In $\mathbb{R}^n$, a Baire measure is a measure that is defined for all Borel sets and is finite for all compact Borel sets (Royden 331-332), where the collection of Borel sets is the smallest σ -algebra containing the collection of closed sets.) Call this G_2 -invariant Baire measure μ . Since μ is invariant with respect to the isometries of $\mathbb{R}^2$, then it must be invariant with respect to T_2 , the group of translations of $\mathbb{R}^2$ (which is a subgroup of G_2). However, recall that it is already known that Lebesgue measure is invariant with respect to T_2 . Furthermore, it is readily verified that Lebesgue measure is a Baire measure. Hence, μ must be equivalent to (i.e. a constant multiple of) Lebesgue measure. Therefore, two dimensional Lebesgue measure must be isometry invariant (Royden 383-384).

On a final note, let it be understood that the process followed here to construct Lebesgue measure in two dimensions can be continued to any dimension desired. Notice that, at its very basic level, two dimensional Lebesgue measure is a product of one dimensional Lebesgue measure with itself, and that it operates in the space that is the cross product of $\mathbb{R}$ with itself. The Carathéodory process is then used to extend this basic level to a full measure. A measure that is created in this manner is known as a product measure (Royden 304), and one can create

Lebesgue measure of any dimension, n , by inductively constructing the product measure of one dimensional Lebesgue measure with $n - 1$ dimensional Lebesgue measure. In doing so, the same argument concerning isometry invariance holds. Therefore, Lebesgue measure provides an element of consistency when changing the positions of sets in $\mathbb{R}^n$, which is why it is such a useful tool for the analyst working in Euclidean spaces.

Chapter 7: The Solvability of G_1 and G_2

In this chapter, the discussion will focus on the solvability of the isometry groups for the one and two dimensional Euclidean spaces. Recall that a group, G (with identity element e), is solvable if and only if there exists a finite chain of subgroups $\{e\} = H_0 < \cdots < H_n = G$ where, for every $0 \leq i \leq n-1$, H_i is a normal subgroup of H_{i+1} (making the chain a subnormal series) and the factor group H_{i+1}/H_i is abelian (Beachy). Remember also that if K is the kernel of a group homomorphism $\phi : A \rightarrow B$, then K is a normal subgroup of A (Fraleigh 175) and the factor group A/K is isomorphic to B (Fraleigh 173).

Consider now the group of isometries of $\mathbb{R}$, which is denoted by G_1 . The isometries in this group are composed of translations and reflections. Needless to say, not all of these transformations of the real line preserve inner products. In fact, the only two isometries of $\mathbb{R}$ which do preserve $\|x\|$ for all $x \in \mathbb{R}$ are the identity transformation, I , and the reflection about the origin, $r : r(x) = -x$. These two isometries are called the orthogonal transformations of $\mathbb{R}$ because of their ability to preserve inner products. Indeed, their matrix representations are given by $I(x) = [1]x$ and $r(x) = [-1]x$, which we know to be orthogonal matrices. Thus, these two isometries comprise a group called the orthogonal group of $\mathbb{R}$, which is denoted by O_1 . It is easy to see that this group $O_1 = \{I, r\}$ is abelian. In fact, since it is a cyclic group of order 2 generated by r , it is isomorphic to Z_2 (Wagon 223, 224).

Now, let T_1 be the group of translations of $\mathbb{R}$. One wishes to show that G_1 is homomorphic

to O_1 with T_1 as the kernel of the homomorphism. To see this, note that every isometry of the real line can be uniquely represented by a translation, or a reflection across the origin followed by a translation. In other words, for any $g \in G_1$, $g = \tau$ ($= \tau I$) or $g = \tau r$ for some $\tau \in T_1$.

Define the function $\phi : G_1 \rightarrow O_1$ by

$$\phi(g) := \begin{cases} I & \text{if } g = \tau \\ r & \text{if } g = \tau r \end{cases} \quad (\text{Wagon 223}).$$

To see that this is a homomorphism, let $g_1, g_2 \in G$. If $g_1 = \tau_1$ and $g_2 = \tau_2$ for some $\tau_1, \tau_2 \in T_1$, then $g_1 g_2 = \tau_1 \tau_2 \in T_1$. Hence, $\phi(g_1 g_2) = I = I^2 = \phi(g_1) \phi(g_2)$. Similarly, if $g_1 = \tau_1 r$ and $g_2 = \tau_2 r$, then $g_1 g_2 = (\tau_1 r)(\tau_2 r) = \tau_1 r \tau_2 r = \tau_1 \tau_2^{-1} r^2 = \tau_1 \tau_2^{-1}$ and, hence, $\phi(g_1 g_2) = I = r^2 = \phi(g_1) \phi(g_2)$. Suppose, however, that $g_1 = \tau_1 r$ and $g_2 = \tau_2$. Then $g_1 g_2 = (\tau_1 r) \tau_2 = \tau_1 r \tau_2 = \tau_1 \tau_2^{-1} r = (\tau_1 \tau_2^{-1}) r$. Hence, $\phi(g_1 g_2) = r = r I = \phi(g_1) \phi(g_2)$, with similar results occurring when $g_1 = \tau_1$ and $g_2 = \tau_2 r$. Thus, $\phi(g_1 g_2) = \phi(g_1) \phi(g_2)$ for all $g_1, g_2 \in G_1$ and, therefore, ϕ is a homomorphism.

Notice above that the kernel of the homomorphism from G_1 to O_1 is T_1 . Hence, T_1 is a normal subgroup of G_1 and G_1/T_1 is isomorphic to O_1 and, hence, is abelian. Furthermore, $\{I\}$ is a normal subgroup of T_1 and $T_1/\{I\}$ is abelian since T_1 is, itself, abelian. Thus, the sequence of subgroups $\{I\} < T_1 < G_1$ (Wagon 224) is a subnormal series with $T_1/\{I\}$ and G_1/T_1 abelian. Therefore, G_1 is solvable.

The discussion of the solvability of G_2 is, naturally, somewhat more involved since one has also to account for rotations along with translations and reflections. First, take a few moments

to consider the notion of the determinant. The reader is undoubtedly very familiar with the idea of the determinant of a matrix. In fact, when the term determinant is mentioned, one's normal reaction is to think of it as a characteristic of matrices. It is true, indeed, that determinants and matrices tend to go hand in hand. This is not, however, the entire picture. One can, in fact, discuss the determinant of a transformation that has no matrix representation at all. In fact, the early study of the determinant predates the study of matrices altogether (Goldberg 60). It was only as the study of matrices became increasingly more popular that the determinant became linked with them so intrinsically.

That having been said, consider how the determinant applies to G_2 . Let A_2 be the group of affine transformations of $\mathbb{R}^2$ (i.e. the transformations generated by translations and linear operators). The determinant is a homomorphic function on A_2 . Thus, every affine transformation, f , of the real plane, whether or not it has a matrix representation, has a determinant associated with it. This determinant is denoted by $\det(f)$ and provides a measure of the effect of f on the areas of sets in $\mathbb{R}^2$. To be more specific, $|\det(f)|$ is the scale factor by which the area of a set is changed under the transformation f (Wagon 224).

Now, since G_2 consists of affine transformations that preserve area, we know that $|\det(g)| = 1$ for every $g \in G_2$. Some of these isometries have matrix representations, while others do not. In the case of an isometry that is representable by a matrix (in other words, an isometry that is linear transformation), since the determinant is ± 1 , then the matrix that represents this transformation must be an orthogonal matrix. Thus, the group O_2 is exactly the

Now that one has a better understanding of how G_2 breaks down, one can go about proving that it is solvable. Let I be the identity transformation and T_2 be the group of translations of $\mathbb{R}^2$. We know that $\{I\}$ is a normal subgroup of T_2 and that $T_2/\{I\}$ is abelian since T_2 is, itself, abelian.

Next, one wishes to show that SG_2 is homomorphic to SO_2 with T_2 as the kernel of the homomorphism (Wagon 224). Recall that the orientation preserving isometries of $\mathbb{R}^2$, which are the elements of SG_2 are those that are composed of translations and rotations about the origin. Thus, every element of SG_2 can be uniquely represented by a rotation about the origin followed by a translation. In other words, for every $g \in SG_2$, there exists some $\tau \in T_2$ and $\rho \in SO_2$ such that $g = \tau\rho$. Define the function $\varphi : SG_2 \rightarrow SO_2$ by $\varphi(g) := \rho$. Then, for any $g_1 = \tau_1\rho_1, g_2 = \tau_2\rho_2 \in SG_2$, one has $g_1g_2 = (\tau_1\rho_1)(\tau_2\rho_2) = \tau_1\rho_1\tau_2\rho_2 = \tau_1\tau_3\rho_1\rho_2 = (\tau_1\tau_3)(\rho_1\rho_2)$ where τ_3 is the translation of $\mathbb{R}^2$ by a vector of $\rho_1(\mathbf{v})$ (with $\mathbf{v}$ being the vector of translation of τ_2). $\varphi(g_1g_2) = \varphi[(\tau_1\tau_3)(\rho_1\rho_2)] = \rho_1\rho_2 = \varphi(g_1)\varphi(g_2)$. Thus, φ is a homomorphism. Furthermore, $\varphi(g) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ if and only if $g \in T_2$. Hence, T_2 is the kernel of φ . Therefore, T_2 is a normal subgroup of SG_2 and SG_2/T_2 is isomorphic to SO_2 , which is abelian.

Finally, examine the connection between SG_2 and G_2 . Start by defining the function $\psi : G_2 \rightarrow \mathbb{Z}_2$ by

$$\psi(g) := \begin{cases} 0 & \text{if } \det(g) = 1 \\ 1 & \text{if } \det(g) = -1 \end{cases},$$

and let $g_1, g_2 \in G_2$. Examine the possible cases:

Case 1: $\det(g_1) = \det(g_2) = 1$

If $\det(g_1) = \det(g_2) = 1$, then $\det(g_1 g_2) = \det(g_1) \cdot \det(g_2) = (1)(1) = 1$ and, hence,
 $\psi(g_1 g_2) = 0 = 0 + 0 = \psi(g_1) + \psi(g_2)$.

Case 2: $\det(g_1) = \det(g_2) = -1$

If $\det(g_1) = \det(g_2) = -1$, then $\det(g_1 g_2) = \det(g_1) \cdot \det(g_2) = (-1)(-1) = 1$ and,
 hence, $\psi(g_1 g_2) = 0 \equiv (1 + 1)(\text{mod } 2)$. Thus, $\psi(g_1 g_2) = \psi(g_1) + \psi(g_2)$.

Case 3: $\det(g_1) = 1, \det(g_2) = -1$

If $\det(g_1) = 1$ and $\det(g_2) = -1$, then $\det(g_1 g_2) = \det(g_1) \cdot \det(g_2) = (1)(-1) = -1$ and,
 hence, $\psi(g_1 g_2) = 1 = 0 + 1 = \psi(g_1) + \psi(g_2)$.

Case 4: $\det(g_1) = -1, \det(g_2) = 1$

If $\det(g_1) = -1$ and $\det(g_2) = 1$, then $\det(g_1 g_2) = \det(g_1) \cdot \det(g_2) = (-1)(1) = -1$ and,
 hence, $\psi(g_1 g_2) = 1 = 1 + 0 = \psi(g_1) + \psi(g_2)$.

Thus, $\psi(g_1 g_2) = \psi(g_1) + \psi(g_2)$ for all $g_1, g_2 \in G_2$ and, therefore, ψ is a homomorphism.

Furthermore, since SG_2 is the kernel of this homomorphism, then SG_2 is a normal subgroup of G_2 and G_2/SG_2 is isomorphic to $\mathbb{Z}_2$ (and, hence, is abelian).

Thus, the sequence of subgroups $\{I\} < T_2 < SG_2 < G_2$ (Wagon 224) is a subnormal series with $T_2/\{I\}$, SG_2/T_2 , and G_2/SG_2 abelian and G_2 is, therefore, solvable.

This exercise in solvability, although a departure from the previous work, will prove to be invaluable in the closing chapter of this thesis. Beside giving a better idea of how the isometry

groups of $\mathbb{R}$ and $\mathbb{R}^2$ are structured, the solvability of G_1 and G_2 will allow for the construction of an invariant extension of Lebesgue measure in these two spaces, precluding any chance of paradoxical decompositions with respect to rigid motions.

Chapter 8: The Banach-Tarski Paradox In $\mathbb{R}$ and $\mathbb{R}^2$

The work of the second part of this thesis will now culminate in the examination of one final topic pertinent to the study of the Banach-Tarski Paradox, the failure of the paradox to hold in one and two dimensional Euclidean spaces. Proving this fact relies heavily on measure theory. Special emphasis will be placed on the facts about Lebesgue measure developed in the previous chapters, and of particular interest will be the following class of groups known as the amenable groups.

Definition (*Amenable Group*)

8.1

A group is known as amenable if on $\wp(G)$ there exists a (finitely additive) measure, μ , such that $\mu(G) = 1$ and $\mu(gX) = \mu(X)$ for all $g \in G$ and $X \subset G$ (Wagon 146).

From this point forward, the function μ described above will be referred to simply as an invariant measure on the amenable group G . However, it is important to keep in mind that this function is only finitely additive, unlike the true measures described earlier. Thus, it is somewhat of a “psuedo-measure”, much like a measure on an algebra. Just the same, it is commonly accepted to use the general term of measure for all of these functions provided that any specific limitations have been initially pointed out.

The reason this class of groups is of particular interest is because it contains the class of

solvable groups (Wagon 149) and, hence, G_1 and G_2 . Thus, by the definition of amenable group, one is assured that an invariant measure (as described above) exists for each of G_1 and G_2 .

What one would like to do at this point is to extend the one and two dimensional Lebesgue measures discussed in earlier to apply to all subsets of $\mathbb{R}$ and $\mathbb{R}^2$, respectively. Recall that (one dimensional) Lebesgue measure, λ , is not a function on $\wp(\mathbb{R})$, but rather on a particular subset of $\wp(\mathbb{R})$ known as the collection of Lebesgue measurable sets and denoted by $\mathfrak{M}$. Likewise, two dimensional Lebesgue measure is a function on $\mathfrak{M}^2 \subset \wp(\mathbb{R}^2)$, not on $\wp(\mathbb{R}^2)$, itself. Thus, Lebesgue measure cannot be applied to all sets in $\mathbb{R}$ (or $\mathbb{R}^2$). The next best option is to construct a function that is similar to Lebesgue measure, and which extends Lebesgue measure to $\wp(\mathbb{R})$ (or $\wp(\mathbb{R}^2)$). In other words, a function on $\wp(\mathbb{R})$ (or $\wp(\mathbb{R}^2)$) that retains as many of the important qualities of Lebesgue measure as possible, and is equivalent to Lebesgue measure when applied to Lebesgue measurable sets.

There is, in measure theory, a very useful tool for this particular task known as the Measure Extension Theorem. Although the proof of this theorem is outside the realm of this thesis, we may, if the reader is willing, still use it to our advantage. The theorem reads as follows:

Theorem (*Measure Extension Theorem*)

8.2

If μ is a measure on S , a subring of the algebra A , then there exists a finitely additive nonnegative function ν that is an extension of μ to A (Wagon 153).

Since $\mathfrak{M}$ is a subalgebra of the algebra $\wp(\mathbb{R})$, then it is automatically a subring of $\wp(\mathbb{R})$ (meaning that it is closed under unions and set differences) (Wagon 153). Thus, by the Measure Extension Theorem, one knows that there is a finitely additive nonnegative function, ν , on $\wp(\mathbb{R})$ such that $\nu(X) = \lambda(X)$ for all $X \in \mathfrak{M}$. In other words, ν agrees with Lebesgue measure on the Lebesgue measurable sets. This function ν is very useful in that it can be applied to any set in $\mathbb{R}$ and preserves the nonnegativity and additivity of Lebesgue measure (although the additivity is only preserved for a finite sequence of disjoint sets). However, it does not necessarily preserve one of the qualities of Lebesgue measure that will prove to be one of the most useful, and that is isometry invariance. For that, a stronger theorem is needed.

Theorem (*Invariant Extension Theorem*)

8.3

Let G be an amenable group of automorphism of the algebra A , and let S be a G -invariant subring of A . If μ is a measure on S that is invariant with respect to G , then there exists a finitely additive nonnegative function ν that is an extension of μ to A and is also invariant with respect to G (Wagon 155).

Proof First of all, note that an automorphism of A is any bijection $g : A \rightarrow A$ such that $g(X_1 \cup X_2) = g(X_1) \cup g(X_2)$, $g(X_1 \cap X_2) = g(X_1) \cap g(X_2)$, and $g[(X_1)^c] = [g(X_1)]^c$ for all $X_1, X_2 \in A$, and that a G -invariant subring of A is any subring, S , such that $g(Y) \in S$ for all $g \in G$ and $Y \in S$.

Now, we know by the Measure Extension Theorem that there exists a finitely additive nonnegative function that extends μ to A . Call this function ν . Then for every $X \in A$, define the functions $f_X : G \rightarrow \mathbb{R}$, $\nu : A \rightarrow [0, \infty]$ by $f_X(g) := \nu[g^{-1}(X)]$ and

$$\nu(X) := \begin{cases} \int_G f_X d\gamma & \text{if } f_X \text{ is bounded} \\ \infty & \text{if } f_X \text{ is unbounded} \end{cases}$$

where γ is the invariant measure on the amenable group G (See Definition 8.1).

The claim is that ν is a finitely additive nonnegative extension of μ to A that is invariant with respect to G .

The nonnegativity of ν follows from the fact that the integral of a nonnegative function is nonnegative (See Appendix B).

To see the finite additivity of ν , let $X_1, X_2 \in A$ such that $X_1 \cap X_2 = \emptyset$. If f_{X_1} is unbounded, then $\nu[g^{-1}(X_1)] = \infty$ for some $g \in G$. Then

$$\nu[g^{-1}(X_1 \cup X_2)] = \nu[g^{-1}(X_1) \cup g^{-1}(X_2)] = \nu[g^{-1}(X_1)] + \nu[g^{-1}(X_2)] = \infty.$$

Hence, $f_{X_1 \cup X_2}$ is unbounded. Thus, $f_{X_1 \cup X_2} = \infty = \infty + f_{X_2} = f_{X_1} + f_{X_2}$. A similar situation holds if f_{X_2} is unbounded. Suppose, then, that both f_{X_1} and f_{X_2} are bounded.

$$\begin{aligned} \text{Then } \nu(X_1 \cup X_2) &= \int_G f_{X_1 \cup X_2} d\gamma = \int_G \nu[g^{-1}(X_1 \cup X_2)] d\gamma \\ &= \int_G \nu[g^{-1}(X_1)] + \nu[g^{-1}(X_2)] d\gamma \\ &= \int_G \nu[g^{-1}(X_1)] d\gamma + \int_G \nu[g^{-1}(X_2)] d\gamma \\ &= \int_G f_{X_1} d\gamma + \int_G f_{X_2} d\gamma = \nu(X_1) + \nu(X_2) \end{aligned}$$

Now let $Y \in S$. To show that ν is an extension of μ , one must show that

$$\begin{aligned} \nu(Y) &= \mu(Y). \text{ If } f_Y \text{ is bounded, then } \nu(Y) = \int_G f_Y d\gamma = \int_G \nu[g^{-1}(Y)] d\gamma \\ &= \int_G \mu[g^{-1}(Y)] d\gamma \text{ since } S \text{ is } G\text{-invariant and } \nu(Y) = \mu(Y) \text{ for all } Y \in S \end{aligned}$$

$$\begin{aligned}
&= \int_G \mu(Y) d\gamma \text{ since } \mu \text{ is invariant with respect to } G \\
&= \int_G \mu(Y) \cdot \chi_G d\gamma \\
&= \mu(Y) \text{ since } \gamma(G) = 1.
\end{aligned}$$

Suppose, on the other hand, that f_Y is unbounded. Then $\nu(Y) = \infty$. Also, $\nu[g^{-1}(Y)] = \infty$ for some $g \in G$. But $\nu[g^{-1}(Y)] = \mu[g^{-1}(Y)]$ since S is G -invariant and $\nu(Y) = \mu(Y)$ for all $Y \in S$. Furthermore, $\mu(Y) = \mu[g^{-1}(Y)]$ for all $g \in G$. Thus, $\nu(Y) = \infty = \mu(Y)$. Therefore, ν is an extension of μ .

Finally, one wishes to show that ν is invariant with respect to G . In other words, $\nu[h(X)] = \nu(X)$ for any $h \in G$ and $X \in A$. To see this, let $h \in G$ and $X \in A$.

If $f_{h(X)}$ is bounded, then

$$\nu[h(X)] = \int_G f_{h(X)} d\gamma = \int_G \nu(g^{-1}[h(X)]) d\gamma = \int_G \nu[(g^{-1}h)(X)] d\gamma.$$

However, notice that if $h \in G$, then $\{g^{-1}h : g \in G\} = \{g^{-1} : g \in G\}$ since, for any $g \in G$, $g^{-1}h = j^{-1}$ for some $j \in G$ and $g^{-1} = (hg)^{-1}h$ (by the familiar group inverse and closure properties). Hence,

$$\{\nu[(g^{-1}h)(X)] : g \in G\} = \{\nu[g^{-1}(X)] : g \in G\}.$$

Thus, one finds that

$$\int_G \nu[(g^{-1}h)(X)] d\gamma = \int_G \nu[g^{-1}(X)] d\gamma = \int_G f_X d\gamma = \nu(X),$$

yielding $\nu[h(X)] = \nu(X)$.

If, on the other hand, $f_{h(X)}$ is unbounded, then $\nu(g^{-1}[h(X)]) = \infty$ for some $g \in G$. However, $\nu[(g^{-1}h)(X)] = \nu[g^{-1}[h(X)]] = \infty$. Hence, f_X is unbounded and, therefore, $\nu(X) = \infty = \nu[h(X)]$.

□

The importance of this result as it pertains to the Banach-Tarski Paradox in $\mathbb{R}$ and $\mathbb{R}^2$ cannot be emphasized enough. It is specifically this result which dictates that a Banach-Tarski type paradoxical decomposition simply cannot take place in either of these two spaces. Note that G_1 is an amenable group of automorphisms of $\mathcal{P}(\mathbb{R})$, and $\mathcal{M}$ is an isometry invariant subalgebra of $\mathcal{P}(\mathbb{R})$. Thus, by the Invariant Extension Theorem, there is a finitely additive isometry invariant extension of Lebesgue measure to $\mathcal{P}(\mathbb{R})$. Therefore, even though one may not be able to apply Lebesgue measure to every set in $\mathbb{R}$, one now has a new function (which shall be denoted by $\vec{\lambda}$) that agrees exactly with Lebesgue measure when applied to sets that are Lebesgue measurable, and still retains the important properties of invariance and (finite) additivity even when applied to sets that are not. Furthermore, a similar extension of Lebesgue measure exists in $\mathbb{R}^2$ since G_2 is also amenable.

To see why this precludes the Banach-Tarski Paradox, suppose X is any bounded set in $\mathbb{R}$ or $\mathbb{R}^2$ with nonempty interior. Since X is bounded, it can be enclosed by a bounded interval (if in $\mathbb{R}$) or rectangle (if in $\mathbb{R}^2$), call it $\bar{X}$. Since $\bar{X}$ is a bounded Lebesgue measurable set, it has finite Lebesgue measure $\lambda(\bar{X}) < \infty$. Then, by the nonnegativity and finite additivity of $\vec{\lambda}$, one has $\vec{\lambda}(X) \leq \vec{\lambda}(X) + \vec{\lambda}(X \setminus X) = \vec{\lambda}(\bar{X}) = \lambda(\bar{X}) < \infty$. Furthermore, since X has nonempty interior, then it must contain a nonempty interval (if in $\mathbb{R}$) or rectangle with nonempty interior (if in $\mathbb{R}^2$), call it $\underline{X}$. Similarly to above, then, one has $0 < \lambda(\underline{X}) = \vec{\lambda}(\underline{X}) \leq \vec{\lambda}(X) + \vec{\lambda}(X \setminus \underline{X}) = \vec{\lambda}(X)$. Thus, for any bounded set, X , with nonempty interior, one finds that $0 < \vec{\lambda}(X) < \infty$ (Wagon 156).

Now suppose there existed some partition of $\{Y_1, \dots, Y_p, Z_1, \dots, Z_q\}$ of X , and some set of isometries $\{g_1, \dots, g_p, h_1, \dots, h_q\}$ such that $X = \bigcup_{i=1}^p g_i(Y_i)$ and $X = \bigcup_{j=1}^q h_j(Z_j)$. Then, by the finite additivity of $\vec{\lambda}$, we have $\sum_{i=1}^p \vec{\lambda}[g_i(Y_i)] = \vec{\lambda}(\bigcup_{i=1}^p g_i(Y_i)) = \vec{\lambda}(X)$ and $\sum_{j=1}^q \vec{\lambda}[h_j(Z_j)] = \vec{\lambda}(\bigcup_{j=1}^q h_j(Z_j)) = \vec{\lambda}(X)$. But $\{Y_1, \dots, Y_p, Z_1, \dots, Z_q\}$ is a partition of X , so the same properties of $\vec{\lambda}$ also yield:

$$\begin{aligned} \vec{\lambda}(X) &= \vec{\lambda}(\left[\bigcup_{i=1}^p g_i(Y_i)\right] \cup \left[\bigcup_{j=1}^q h_j(Z_j)\right]) \\ &= \sum_{i=1}^p \vec{\lambda}[g_i(Y_i)] + \sum_{j=1}^q \vec{\lambda}[h_j(Z_j)] \\ &= \vec{\lambda}(X) + \vec{\lambda}(X). \end{aligned}$$

Thus, $\vec{\lambda}(X)$ must equal $2 \cdot \vec{\lambda}(X)$, implying that either $\vec{\lambda}(X) = 0$ or $\vec{\lambda}(X) = \infty$. Since this contradicts what we already know to be true about bounded sets in $\mathbb{R}$ and $\mathbb{R}^2$ with nonempty interior, then no such sets are paradoxical with respect to isometries.

Thus, the Banach-Tarski Paradox is nonexistent in $\mathbb{R}$ and $\mathbb{R}^2$. A result which stems from the fact that the isometry groups of these spaces are solvable.

Conclusion

There are two basic reasons for studying the Banach-Tarski Paradox, one philosophical and the other practical. The philosophical reason for studying the paradox is for the simple reason of witnessing the sheer beauty of such an incredibly counter-intuitive, and yet easily justifiable mathematical construction. The very notion of taking, for instance, a ball (or, for that matter, any bounded object with nonempty interior) and dissecting it in such a way as to reassemble two separate facsimiles of it is enough to make one's mind turn to flights of fancy. Indeed, the idea is dismissed as magic by those who are not privy to the mathematical details. Those who have studied the details of the paradox see it as another type of magic, the magic of mathematics. It is a perfect example of the reason why many mathematicians have been drawn into the field in the first place: the wonder of being able to take something, albeit humanly impossible, and show that it can be reduced to a number of mathematically indisputable steps. Of course, remember that the community of mathematicians has, itself, not been without dissension. The term "mathematically indisputable" is not exactly black and white when it comes to such topics as the Axiom of Choice, one of the key elements to proving the Banach-Tarski Paradox. In fact, this paradox is a good example of why the Axiom of Choice had such a turbulent beginning. It is by now, however, generally accepted, which makes the paradox herein, at the very least, a fairly indisputable mathematical curiosity.

Although one would hope that its philosophical implications and mathematical beauty would be enough reason for anyone to study the Banach-Tarski Paradox, it does have more practical

applications. For instance, the nonexistence of the paradox in $\mathbb{R}$ and $\mathbb{R}^2$ relies on the fact that G_1 and G_2 are solvable. Hence, the fact that the paradox exists in all Euclidean spaces of higher dimensions implies the fact that the isometry groups of these spaces are not solvable. In addition to this, the paradox is of particular interest to the measure theorist because of its implications about the nonexistence of certain measures and the existence of certain nonmeasurable sets. It has been shown in chapter 8 that the Banach-Tarski Paradox is inconsistent with Lebesgue measure (or, rather, any isometry invariant extension thereof) over all sets in the space. In fact, it is the very fact that Lebesgue measure has such an extension in $\mathbb{R}$ and $\mathbb{R}^2$ that preclude the paradox in these two spaces. Hence, the existence of the Banach-Tarski Paradox in dimensions higher than two mandates that Lebesgue measure may not be applied to all sets in these spaces. Hence, one is assured a priori of the existence of Lebesgue nonmeasurable sets. Indeed, if $Y_1, \dots, Y_p, Z_1, \dots, Z_q$ is a paradoxical partition of X , then at least one each of the Y_i and Z_j must be Lebesgue nonmeasurable, since the isometry invariance and additivity of Lebesgue measure make it impossible to have both $\sum_{i=1}^p \lambda(Y_i) + \sum_{j=1}^q \lambda(Z_j) = \lambda(X)$ and $\sum_{i=1}^p \lambda(g_i Y_i) = \sum_{j=1}^q \lambda(h_j Z_j) = \lambda(X)$. Furthermore, at least one each of the Y_i and Z_j must have nonempty interior (and all are bounded by X). Therefore, it can be said that not only does $\mathbb{R}^n$ ($n \geq 3$) contain a Lebesgue nonmeasurable set, but every bounded set in the space with nonempty interior contains infinitely many Lebesgue nonmeasurable subsets.

This may be a treasure trove of information about Lebesgue measure, but it can actually be

extended further to apply to any finitely additive isometry-invariant measure that normalizes the unit cube (where μ normalizes the unit cube in $\mathbb{R}^n$ if $\mu([0, 1]^n) = 1$). Suppose that μ is just such a measure and X is any bounded set with nonempty interior. Since μ normalizes the unit cube, then $0 < \mu([0, \alpha]^n) < \infty$ for all $0 < \alpha < \infty$. (If $\alpha < 1$, then $[0, 1]^n$ can be covered with finitely many copies of $[0, \alpha]^n$, which implies that $0 < \mu([0, \alpha]^n)$. If $1 < \alpha$, then $[0, \alpha]^n$ can be covered with finitely many copies of $[0, 1]^n$, which implies that $\mu([0, \alpha]^n) < \infty$.) Furthermore, since X is bounded, it must be contained in a translation of $[0, \alpha]^n$ for some $\alpha > 0$, and since X has nonempty interior, it must contain a translation of $[0, \beta]^n$ for some $\beta < \infty$. Hence, $0 < \mu(X) < \infty$. Thus, one arrives at the the same contradiction between μ and any paradoxical decomposition of X that was seen with $\vec{\lambda}$ at the end of chapter 8. Therefore, the Banach-Tarski Paradox realizes the nonexistence of any finitely additive isometry-invariant measure on $\mathcal{P}(\mathbb{R}^n)$ ($n \geq 3$) that normalizes the unit cube. Equivalently, for any such measure in $\mathbb{R}^n$ ($n \geq 3$), the Banach-Tarski realizes the existence of nonmeasurable sets.

In closing, let it be said that whether one's interests lie in discovery for its own sake or for the sake of its applications, the Banach-Tarski Paradox is truly a remarkable and noteworthy mathematical discovery.

Appendix A: A Lebesgue Nonmeasurable Set in $\mathbb{R}$

The Following argument describes a specific set in $\mathbb{R}$ that is not Lebesgue measurable. A similar description of the same set can be found on pages 64 and 65 of Royden.

Define the function $\cup: \mathcal{P}([0, 1)) \times [0, 1) \rightarrow [0, 1)$ as follows. For any $X \subseteq [0, 1)$ and $y \in [0, 1)$, $X \cup y := \{z \in [0, 1) : x + y \equiv z \pmod{1} \text{ for some } x \in X\}$. The notation $\cup$ is used here because $X \cup y$ is equivalent to taking the set $\{2\pi x : x \in X\}$ on the unit circle and rotating it through an angle of $2\pi y$ radians.

The first point of order is to note that Lebesgue measure is $\cup$ invariant. That is, if X is a Lebesgue measurable subset of $[0, 1)$, then for any $y \in [0, 1)$, $X \cup y$ is Lebesgue measurable and $\lambda(X \cup y) = \lambda(X)$. To see that this is the case, let $X_1 = X \cap [0, 1 - y)$ and $X_2 = X \cap [1 - y, 1)$. Since, both X_1 and X_2 are the intersections of Lebesgue measurable sets (recall that all intervals are Lebesgue measurable), then they are, themselves, Lebesgue measurable. Furthermore, X_1 and X_2 are disjoint sets whose union is X . Thus, since Lebesgue measure is additive, one has $\lambda(X) = \lambda(X_1) + \lambda(X_2)$. Now consider $X_1 \cup y$ and $X_2 \cup y$. Notice that $X_1 \cup y = X_1 + y$ and $X_2 \cup y = X_2 + (y - 1)$. Hence, by the translation invariance of Lebesgue measure, both $X_1 \cup y$ and $X_2 \cup y$ are Lebesgue measurable, and one has $\lambda(X_1 \cup y) = \lambda(X_1)$ and $\lambda(X_2 \cup y) = \lambda(X_2)$. Finally, since $X \cup y = (X_1 \cup y) \cup (X_2 \cup y)$, which are both disjoint Lebesgue measurable sets, then $X \cup y$ is Lebesgue measurable and

$$\lambda(X \cup y) = \lambda(X_1 \cup y) + \lambda(X_2 \cup y) = \lambda(X_1) + \lambda(X_2) = \lambda(X).$$

Next define the equivalence relation “ $\sim$ ” on $[0, 1) \times [0, 1)$ as $x \sim y$ if and only if $x - y \in \mathbb{Q}$.

It is easy to see that “ $\sim$ ” is reflexive (for any $x \in [0, 1)$, $x - x = 0 \in \mathbb{Q}$), symmetric (if $x - y \in \mathbb{Q}$, then $y - x = -(x - y) \in \mathbb{Q}$), and transitive (if $x - y \in \mathbb{Q}$ and $y - z \in \mathbb{Q}$, then $x - z = (x - y) + (y - z) \in \mathbb{Q}$) and, therefore, is indeed an equivalence relation. Therefore, one can partition the interval $[0, 1)$ into its $\sim$ -equivalence classes. Note that each equivalence class will be composed of numbers that differ from each other by a rational number. Any two numbers that differ from one another by an irrational number will be in different equivalence classes.

Invoking the axiom of choice, one can establish the existence of a set Ω that is composed of exactly one element from each of the $\sim$ -equivalence classes. Define the sets $\Omega_i := \Omega \cup r_i$ where $\{r_i\}_{i=0}^{\infty}$ is the sequence of rational numbers from 0 to 1 with $r_0 = 0$. The claim is that the sequence of sets $\{\Omega_i\}$ is a partition of the interval $[0, 1)$. To see this, we must show that it is a sequence of pairwise disjoint sets that covers $[0, 1)$. (Note that $\{\Omega_i\} \subseteq [0, 1)$ since, by the definitions of Ω_i and $\cup$, $\Omega_i = \{z \in [0, 1) : z + r_i \equiv x \pmod{1} \text{ for some } x \in \Omega\}$ for each $0 \leq i < \infty$.)

Suppose that $\Omega_m \cap \Omega_n \neq \emptyset$. Then there exists some $x \in \Omega_m \cap \Omega_n$. In other words, $x \in \Omega \cup r_m$ and $x \in \Omega \cup r_n$, where r_m and r_n are the m^{th} and n^{th} , respectively, rational numbers in $(0, 1)$. Hence, there exist some $\gamma_m, \gamma_n \in \Omega$ such that $\gamma_m + r_m = x = \gamma_n + r_n$. From this one sees that $\gamma_m - \gamma_n = r_n - r_m \in \mathbb{Q}$. Hence, $\gamma_m \sim \gamma_n$. But Ω contains only one element from each $\sim$ -equivalence class and, hence, $\Omega_m = \Omega_n$. Therefore, if $m \neq n$, then $\Omega_m \cap \Omega_n = \emptyset$.

Thus, the sequence $\{\Omega_i\}$ is pairwise disjoint. To see that this sequence covers $[0, 1)$, start by taking any $x \in [0, 1)$. This x must be in one of the $\sim$ -equivalence classes and must, therefore, be equivalent to some $\gamma \in \Omega$. Hence, $x - \gamma = r_n$ for some n since the elements in any one equivalence class differ from each other by rational numbers. Thus, $x \in \Omega_n$. Therefore, the sequence $\{\Omega_i\}$ is a partition of the interval $[0, 1)$.

Now suppose that Ω is Lebesgue measurable. Since $\Omega_i := \Omega \cup r_i$ and Lebesgue measure is $\cup$ invariant, then for each i , Ω_i is also Lebesgue measurable with $\lambda(\Omega_i) = \lambda(\Omega)$. Furthermore, $\{\Omega_i\}$ is a partition of $[0, 1)$, so $\lambda[0, 1) = \sum_{i=0}^{\infty} \lambda(\Omega_i)$. Thus, one has $\lambda[0, 1) = \sum_{i=0}^{\infty} \lambda(\Omega_i) = \sum_{i=0}^{\infty} \lambda(\Omega)$. However, this yields $\lambda[0, 1) = 0$ if $\lambda(\Omega) = 0$ and $\lambda[0, 1) = \infty$ if $\lambda(\Omega) \neq 0$. Both of which are incorrect since $\lambda[0, 1) = 1$. Therefore, the set Ω is not Lebesgue measurable.

Appendix B: Useful Facts Regarding Integration

This section is not meant to be a thorough discussion of integration, but rather a brief review of how an integral is to be defined, along with a short list of important properties of an integral which may be useful to the reader in studying this thesis.

For the entirety of this discussion, let μ be a measure, M be the collection of μ -measurable sets, and $X \in M$.

A *simple function* is defined to be a finite linear combination of the characteristics functions of measurable sets (Royden 77). Thus, every simple function has the form:

$$\varphi(x) = \sum_{i=1}^n a_i \cdot \chi_{X_i}(x)$$

For example, step function is a simple function where each of the X_i are adjacent intervals.

A *measurable function* is defined to be any extended real valued function, f , such that, for any $y \in \mathbb{R}$, one of the following holds (Royden 259):

$$\text{i) } \{x : f(x) < y\} \in M$$

$$\text{ii) } \{x : f(x) \leq y\} \in M$$

$$\text{iii) } \{x : f(x) \geq y\} \in M$$

$$\text{iv) } \{x : f(x) > y\} \in M$$

The integral over X with respect to μ is a function on the the class of measurable functions on X . It is first defined for simple functions, then for nonnegative measurable functions, and finally for the remaining measurable functions. (Since this thesis only requires integration of

nonnegative functions, this discussion will not involve functions that assume negative values).

The integral over X of the simple function ϕ with respect to μ is (Royden 263):

$$\int_X \phi \, d\mu := \sum_{i=1}^n a_i \cdot \mu(X_i \cap X)$$

The integral over X of the nonnegative measurable function f with respect to μ is (Royden 264):

$$\int_X f \, d\mu := \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu$$

(i.e. The integral of f over X is the supremum of the integrals of all nonnegative simple functions that do not surpass f .)

Below is a list of results concerning integration that may prove useful to the reader in chapter 6 and chapter 8 (ii and iv can be found on page 266 of Royden):

i) if f is a nonnegative measurable function, then $\int_X f \, d\mu$ is nonnegative.

ii) if f, g are nonnegative measurable functions, and α, β are nonnegative

real numbers, then $\int_X \alpha f + \beta g \, d\mu = \alpha \cdot \int_X f \, d\mu + \beta \cdot \int_X g \, d\mu$

iii) if $Y \subseteq X$, then $\int_X \chi_Y \, d\mu = \mu(Y)$

iv) if $\{f_i\}$ is a sequence of nonnegative measurable functions, then

$$\int_X (\sum f_i) \, d\mu = \sum \left(\int_X f_i \, d\mu \right)$$

Works Cited

- Beachy, John. Abstract Algebra On Line. Northern Illinois University. 9 Apr. 2001
<<http://www.math.niu.edu/~beachy/aaol/>>.
- Bunch, Bryan. Mathematical Fallacies and Paradoxes. Mineola, New York: Dover Publications, Inc., 1982.
- Burton, David M. The History of Mathematics. Ed. Karen M. Minette and John M. Morris. 3rd ed. San Francisco, California: McGraw-Hill, 1991.
- Einstein, Albert. "Relativity: The Special and General Theory." Albert Einstein: In His Own Words. By Einstein. New York, New York: Portland House, 2000.
- Fraleigh, John B. A First Course in Abstract Algebra. 6th ed. Menlo Park, California: Addison-Wesley, 1999.
- Goldberg, Jack L. Matrix Theory with Applications. International Series in Pure and Applied Mathematics. New York: McGraw-Hill, Inc., 1991.
- Mattila, Pertti. Geometry of Sets and Measures in Euclidean Spaces: Fractals and rectifiability. Cambridge studies in advanced mathematics 44. New York, New York: Cambridge University Press, 1995.
- O'Connor, J. J., and E. F. Robertson. Alfred Tarski. School of Mathematics and Statistics, University of St. Andrews, Scotland. 12 Sep. 2000
<<http://www-history.mcs.st-andrews.ac.uk/history/Mathematicians/Tarski.html>>.
- O'Connor, J. J., and E. F. Robertson. Stefan Banach. School of Mathematics and Statistics, University of St. Andrews, Scotland. 12 Sep. 2000
<<http://www-history.mcs.st-andrews.ac.uk/history/Mathematicians/Banach.html>>.
- Royden, H. L. Real Analysis. Ed. Carolyn J. Lee-Davis. 3rd ed. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1988.
- Schechter, Eric. A Home Page for the Axiom of Choice. Vanderbilt University. 12 Sep. 2000 <<http://www.math.vanderbilt.edu/~schectex/ccc/choice.html>>.
- Strang, Gilbert. Linear Algebra and Its Applications. 3rd ed. Fort Worth, Texas: Saunders College Publishing, Harcourt Brace Jovanovich, Inc., 1988.
- Wagon, Stan. The Banach-Tarski Paradox. New York, New York: Cambridge University Press, 1994.