Russell Sets, Topology, and Cardinals

Ethan Thomas

Undergraduate Honors Thesis
Professor Marcia Groszek, Advisor

Department of Mathematics
Dartmouth College
May, 2014

Abstract

The Axiom of Choice is assumed in much of mathematics, and many important theorems in studies such as topology, logic, and algebra are known to be equivalent to the axiom. If we construct a model without the Axiom of Choice in which the failure of the axiom is witnessed by a particular set (such as the classic example of "Russell's Socks"), we can explicitly demonstrate the failure of these theorems and observe facts about the cardinality of sets without choice functions that vary greatly from familiar models. In this paper, we construct models of Zermelo-Fraenkel Set Theory with ur-elements without choice functions, and we show several implications and equivalences that result. Furthermore, we consider cardinals of infinite, choiceless sets in this model – which may not be comparable to any of the $\aleph$-cardinals – and what properties we can show of them.

Acknowledgements

First and foremost, I wish to thank my advisor, Professor Marcia Groszek, for spending several terms guiding this research and advising the authorship of this paper. This incredibly fulfilling process would not have been possible without her input and continued guidance.

I would also like to thank Professor Eleftherios Tachtsis of University of the Aegean – whose work I base much of the latter portion of this paper on – for his correspondence regarding my proofs on Russell cardinals. Additionally, I would like to acknowledge the Kaminsky Family Fund Award for providing funding which enabled my initial research on this topic.

Finally, I thank my family and friends for support throughout this process. In particular, I would like to thank Dr. David Cramer for inspiring my pursuit of study in mathematics and in large part motivating the decision to write this honors thesis.

Contents

1. Introduction

The Axiom of Choice is an axiom of set theory that, while now generally assumed true, has been a controversial subject in mathematics since its formulation. Basically, the Axiom of Choice states that from any collection of non-empty sets, an element can be chosen from every set in the collection, even if the elements have no distinguishing property and there are infinitely many sets. The classic example involves choosing shoes and socks. Given a finite set of pairs of shoes or socks, we need to invoke no principle or assumption to choose one from each pair, as we can simply make a finite list of what we will select (specifically, we achieve this through induction on the number of sets). Take infinitely many pairs of shoes, and the ability remains – since the shoes in each pair have some distinguishing property, we can for instance just take the right shoe from each pair. But if we wanted a sock out of each pair, assuming that the socks are not distinguishable, we do not necessarily have this ability.

The existence of such a choice function – a means of choosing an element from a collection of sets – on all sets was implicitly assumed to exist for centuries of mathematical study. Logician and mathematician Ernst Zermelo, in his quest to axiomatize mathematics and develop a rigorous foundation for the work done in the generations before him, was the first to explicitly formulate something close to the Axiom of Choice in 1908. At first, many were reluctant to accept this axiom, especially because it would imply the truth of Georg Cantor's Well-Ordering Principle which had been debated for almost three decades at the time. But as it was revealed that mathematicians, including Cantor, had implicitly taken the axiom to be true in fields of analysis, algebra, and set theory and that their results depended on it, the axiom became more controversial to reject – this discovery threatened to undermine massive amounts of work in several fields because they used the Axiom of Choice in some way or another as the foundation, and though some proofs were shown to have workarounds to avoid making such assumptions, others were realized to rely completely on the ability to make arbitrary choices from infinitely many sets.[1]

Zermelo indeed wished to include the Axiom of Choice in his set theory axioms, because he desired to prove Cantor's Well-Ordering Principle from an axiomatic system. Eventually, though, it was realized that the Axiom of Choice in fact was independent from the other eight axioms, meaning that a model of these original axioms could be consistent whether it assumed the Axiom of Choice or not. [1]

The non-intuitive nature of the Axiom of Choice is well summarized by Jerry Bona:

> "The Axiom of Choice is obviously true; the Well Ordering Principle is obviously false; and who can tell about Zorn's Lemma?" [2]

The three principles, although equivalent, seem to be intuitively difficult to collectively accept or reject. Today, most mathematicians accept the Axiom of Choice and its equivalents, and they recognize that many valuable theorems require the axiom. Still, a model of mathematics without the axiom is equally "valid", and studying how mathematical concepts differ in these models is an interesting endeavor.

We will first show that we can construct a "Weak Russell sequence" – a countable set of pairwise-disjoint pairs (of ur-elements) that has no choice function, and that this is consistent with Zermelo-Fraenkel Set Theory with ur-elements. In particular, we will show an interesting example that specifies exactly which subsets have choice functions to demonstrate that lacking the Axiom of Choice still allows specificity on where choice functions exist. We will then show the implications that such a set has in logic and topology by demonstrating explicitly the failures of important theorems, and we will show what additional conditions are necessary for equivalence theorems. Finally, we examine the cardinality of the unions of these collections of infinite sets without choice functions and see how they differ from the familiar infinite cardinals.

2. Preliminary Definitions

Remark. In this paper, the designation "Fact" will precede a theorem or result generally known to be true, along with the axiom system in which it is true if applicable; these results are not original and will often be assumed without proof. The designation "Theorem" will precede results that we are proving originally in this paper.

Definition 1 (Choice Function)**.** Given a set of sets S, a choice function is a function f such that $f(X) \in X$ for all $X \in S$.

Axiom 1 (Axiom of Choice)**.** *There exists a choice function for every collection of non-empty sets.*

Recall that for a partially-ordered $(S, \leq)$ and a subset $A \subseteq S$, an *upper bound* u of A is an element such that for all $a \in A, a \leq u$ A *maximal element* m of A is $m \in A$ such that there exist no $a \in A$ such that $m < a$.

Axiom 2 (Zorn's Lemma)**.** *If every chain (subset of comparable elements) in a partially-ordered set $(S, \leq)$ has an upper bound, then S has a maximal element.*

For example, if a set S was well-ordered by $\subseteq$, a maximal element would be a set in S not properly contained in any set of S. Zorn's Lemma guarantees such an element under certain conditions.

Axiom 3 (Well-Ordering Principle). *Every set can be well ordered, i.e., every set has a linear ordering $<$ under which each non-empty subset has a least element.*

Fact 1. *The Axiom of Choice, Zorn's Lemma, and the Well-Ordering Principle are equivalent.*

Zermelo-Fraenkel Set Theory (ZF) is a theory consisting of 8 axioms (two of which are actually axiom schema)[3][4]:

(1) *Axiom of Extensionality*
If X and Y have exactly the same elements, then $X = Y$.
(2) *Axiom of Pair*
For any a, b, there exists a set $\{a, b\}$.
(3) *Axiom of Union*
For any X, there exists a set $\bigcup X$, the union of all elements of X.
(4) *Axiom of Power Set*
For any X, there exists a set $\mathcal{P}(X)$, the set of all subsets of X.
(5) *Axiom of Foundation (or Axiom of Regularity)*
Every nonempty set contains a $\in$-minimal element (equivalently, each non-empty set, contains a set with which it is disjoint).
(6) *Axiom of Infinity*
A set I exists such that $\emptyset \in I$ and $\forall x \in I[(x \cup \{x\}) \in I]$ (I is called an inductive set).
(7) *Axiom Schema of Comprehension*
If $\varphi(x, p)$ is a formula in the language of set theory where p is a parameter, then for any X there exists a set $\{u \in X | \varphi(u, p)\}$ that contains all $u \in X$ with the property φ.
(8) *Axiom Schema of Replacement*
If $\varphi(x, p)$ is a formula in the language of set theory, where p is a parameter, which defines a function f on a set X, then the range of f, $\{f(x)|x \in X\}$, is a set.

ZF does not allow for *ur-elements* (elements which are not themselves sets) but we may alter ZF to include them by changing the language of the axioms to apply only to sets. We denote ZF with ur-elements by ZF$^\circ$

Fact 2. *The Axiom of Choice is independent of ZF. Therefore, the Axiom of Choice and its negation are both consistent with ZF and ZF$^\circ$.*

The Axiom of Choice is not assumed by ZF (when it is, it is denoted ZFC), so the theory does not preclude the possibility of models with choice functions on only some sets.

Definition 2 (Filter)**.** A filter $\mathcal{F}$ on S is a subset of $\mathcal{P}(S)$ such that:

(1) $S \in \mathcal{F}$ and $\emptyset \notin \mathcal{F}$
(2) If $X \in \mathcal{F}$ and $Y \in \mathcal{F}$ then $X \cap Y \in \mathcal{F}$
(3) If $X \in \mathcal{F}$ and $X \subseteq Y$ then $Y \in \mathcal{F}$

Definition 3 (Non-principal Filter)**.** A filter $\mathcal{F}$ is non-principal if it contains no finite sets.

Definition 4 (Ultrafilter)**.** A filter $\mathcal{F}$ on S is an ultrafilter if for every $X \subseteq S$, either $X \in \mathcal{F}$ or $S - X \in \mathcal{F}$.

Definition 5 (Measure)**.** A measure on a set S is a function $m : \mathcal{P}(S) \to \mathbb{R}$ such that:

(1) $m(\emptyset) = 0, m(S) > 0$.
(2) $A \subseteq B \Rightarrow m(A) \leq m(B)$.
(3) $A \cap B = \emptyset \Rightarrow m(A \cup B) = m(A) + m(B)$.

Fact 3. *If U is an ultrafilter on S, then the function m on $\mathcal{P}(S)$ defined by*

$$m(A) = \begin{cases} 1 & \textit{if } A \in U \\ 0 & \textit{if } A \notin U \end{cases}$$

is a two-valued measure on S. [4]

Remark. We will denote such a two-valued measure defined by an ultrafilter U as m_U, and use the terms "measure-zero set" and "measure-one set" accordingly when it is obvious which filter is being referenced.

3. A Non-Principal Ultrafilter

If a non-principal ultrafilter were to exist on a particular set, it would provide a very useful way to intuitively consider certain sets "large" (in the filter) and others "small" (not in the filter), where the large subsets contain "almost everything" in the original set. We would have properties such that all finite sets are small, if a set contains a large set it must be large itself, if a set is contained in a small set it is small itself, etc.

First, notice that an ultrafilter need not be complex:

Example 1. $\mathcal{F} = \{X \subseteq \mathbb{N} | 1 \in X\}$ is an ultrafilter on $\mathbb{N}$.

Proof. We can see immediately that the set $\mathcal{F}$ excludes the empty set and includes $\mathbb{N}$, that F is closed under intersection, and that a set's membership in F guarantees the membership of all of its supersets, so F is by definition a filter. It is an ultrafilter because for every set X, either X or its complement contains 1 (but not both). $\square$

Now consider some non-principal filters:

Example 2. $\mathcal{F} = \{X \subseteq \mathbb{N} | X \text{ is cofinite}\}$ is a non-principal filter (where a cofinite set in $\mathbb{N}$ is a set whose complement in $\mathbb{N}$ is infinite).

This follows immediately from the definition, but some non-principal filters are more complex:

Example 3. If $\mathcal{F}$ is a nonprincipal filter on $\mathbb{N}$, and X is a set such that for every $Y \in \mathcal{F}$, $X \cap Y$ is infinite, then

$$\mathcal{G} = \{Z \subseteq \mathbb{N} | \exists Y \in \mathcal{F}(X \cap Y \subseteq Z)\}$$

is a nonprincipal filter on $\mathbb{N}$ such that $\mathcal{F} \subseteq \mathcal{G}$ and $X \in \mathcal{G}$.

Proof. We must show that (1) $\mathcal{G}$ is a filter, (2) $\mathcal{G}$ is non-principal, (3) $\mathcal{F} \subseteq \mathcal{G}$, and (4) $X \in \mathcal{G}$.

Fix some filter $\mathcal{F}$ and set X, and define $\mathcal{G}$ as above.

(1) The three properties of filters follow from the definition:
  (a) $\mathbb{N} \in \mathcal{G}$ since for any $X, Y \subseteq \mathbb{N}$, their intersection is also contained in $\mathbb{N}$. $\emptyset \notin \mathcal{G}$ since that would require $X \cap Y = \emptyset$ for some $Y \in \mathcal{F}$, but $X \cap Y$ is infinite by assumption.
  (b) If $A, B \in \mathcal{G}$, then by definition there is a $Y_A \in \mathcal{F}$ such that $X \cap Y_A \subseteq A$, and there is a $Y_B \in \mathcal{F}$ such that $X \cap Y_B \subseteq B$. But let $Y = Y_A \cap Y_B \in \mathcal{F}$, so $X \cap Y \subseteq A \cap B$, and then $A \cap B \in \mathcal{G}$.
  (c) Consider $A \in \mathcal{G}$ and $A \subseteq B$. There exists then a $Y_A \in \mathcal{F}$ such that $X \cap Y_A \subseteq A$. But $A \subseteq B$, $X \cap Y_A \subseteq B$, so $B \in \mathcal{G}$.

  $\mathcal{G}$ is thus a filter.

(2) For $Z \in \mathcal{G}$, there exists $Y \in \mathcal{F}$ such that $X \cap Y \subseteq Z$. But by definition, the set X has an infinite intersection with every set in $\mathcal{F}$. So $X \cap Y$ is infinite, and so is its superset Z. Thus $\mathcal{G}$ contains only infinite sets and is therefore non-principal.
(3) Assume $Y \in \mathcal{F}$. Clearly $X \cap Y \subseteq Y$, so then $Y \in \mathcal{G}$.
(4) $X \cap Y \subseteq X$ for any Y, in particular some $Y \in \mathcal{F}$. So $X \in \mathcal{G}$.

$\square$

A few attempts to create a non-principal ultrafilter, however, will quickly demonstrate that it is not so intuitively possible. In fact, we need Zorn's Lemma (and thus the Axiom of Choice) to do so.

Fact 4. *Assuming the Axiom of Choice, there exists a non-principal ultrafilter on* $\mathbb{N}$.

Proof. Consider $\mathcal{F} = \{Z \subseteq \mathbb{N} | Z \text{ is cofinite}\}$. This is a non-principal filter (Example 2). Consider a set of filters $F = \{\mathcal{G} \subseteq \mathcal{P}(\mathbb{N}) | \mathcal{F} \subseteq \mathcal{G}, \mathcal{G} \text{ is non-principal}\}$. All filters in F are non-principal and contain $\mathcal{F}$, and F is a partially-ordered set by inclusion.

Consider $\mathcal{C} = \{\mathcal{G}_i | i \in I\}$, a $\subseteq$-chain in F such that for all $i \in I$, $\mathcal{F} \subseteq \mathcal{G}_i$. $\cup\mathcal{C}$ is an upper-bound for the chain, and we can see that $\cup\mathcal{C} \in F$:

- $\cup\mathcal{C}$ is a filter:
  (1) $\mathbb{N} \in \cup\mathcal{C}$ since it is in all filters in $\mathcal{C}$, and $\emptyset \notin \cup\mathcal{C}$ since it is in none of the filters in $\mathcal{C}$.
  (2) If $X, Y \in \cup\mathcal{C}$, then for some i, j $X \in \mathcal{G}_i, Y \in \mathcal{G}_j$. So X and Y are both in $\mathcal{G}_{\max(i,j)}$, so $X \cap Y \in \mathcal{G}_{\max(i,j)}$.
  (3) If $X \in \cup\mathcal{C}$, then for some i, $X \in \mathcal{G}_i/$. Then if $X \subseteq Y, Y \in \mathcal{G}_i$, so $Y \in \cup\mathcal{C}$.
- $\cup\mathcal{C}$ is nonprincipal, since a finite element in $\cup\mathcal{C}$ would be in one of the filters in $\mathcal{C}$, but all filters in $\mathcal{C}$ are non-principal.
- $\cup\mathcal{C}$ contains $\mathcal{F}$, since it is the first filter in $\mathcal{C}$.

Because we assume the Axiom of Choice and thus Zorn's Lemma, F has a maximal element which we will call $\mathcal{F}^*$. Since this element is maximal in F, it is a non-principal filter containing $\mathcal{F}$. It remains to show that $\mathcal{F}^*$ is an ultrafilter.

Assume for sake of contradiction that $\mathcal{F}^*$ is not an ultrafilter, so there exists some $X \subseteq \mathbb{N}$ such that $\mathcal{F}^*$ contains neither X nor $\mathbb{N} - X$. We will contradict this assumption by finding a filter $\mathcal{H} \in F$ such that $\mathcal{F}^* \subsetneq \mathcal{H}$.

Either X or $\mathbb{N} - X$ has the property that for every $Y \in \mathcal{F}^*, X' \cap Y$ is infinite. We show this claim by contradiction: assume there exists $Y', Y'' \in \mathcal{F}$ such that $X \cap Y'$ and $(\mathbb{N} - X) \cap Y''$ are finite. Let $Z = Y' \cap Y''$; then $X \cap Z$ and $X \cap (\mathbb{N} - X)$ are finite as well. $(Z \cap X) \cup (Z \cap (\mathbb{N} - X))$ is finite as the union of two finite sets, but this is equal to $Z \cap (X \cup (\mathbb{N} - X)) = Z \cap \mathbb{N} = Z$. So Z is finite. But since Z is the intersection of two sets in $\mathcal{F}$, it is in $\mathcal{F}$ by the intersection property of filters. But then Z is a finite set in a non-principal filter, contradicting our assumption. Therefore (still assuming that $\mathcal{F}^*$ contains neither X nor $\mathbb{N} - X$) , either X or $\mathbb{N} - X$ has the above property.

Let X' be one of X and $\mathbb{N} - X$ which satisfies the property that for every $Y \in \mathcal{F}^*, X' \cap Y$ is infinite. Then consider $\mathcal{H} = \{Z \subseteq \mathbb{N} | \exists Y \in \mathcal{F}^* (X' \cap Y \subseteq Z)\}$. By Example 3, $\mathcal{H}$ is a non-principal filter on $\mathcal{P}(\mathbb{N})$ such that $\mathcal{F}^* \subseteq \mathcal{H}$ and $X' \in \mathcal{H}$. But $X' \notin \mathcal{F}^*$, so $\mathcal{F}^* \subsetneq \mathcal{H}$, and $\mathcal{F}^*$ is not maximal in F $\lightning$.

Therefore, there exists a non-principal ultrafilter. $\square$

The existence of such a filter will be a useful way to differentiate subsets, and in fact we will use this to create a special kind of set which has choice functions on only "small" subsets.

4. Models of ZF° with Limited Choice Functions

We will now explore sets without choice functions – we use a non-principal ultrafilter to derive a model of set theory in which a set of sets does not always have a choice function.

More specifically, a model $\mathcal{M}$ of ZF° with ur-elements $\{u_n | n \in \mathbb{N}\}$ and a non-principal ultrafilter on the natural numbers, $\mathcal{G} \in \mathcal{M}$, can be used to construct a model $\mathcal{M}'$ with a countable set of pairwise-disjoint sets of two-ur-elements, $\{X_n | n \in \mathbb{N}\}$, such that for all $A \subseteq \mathbb{N}$, $\{X_n | n \in A\}$ has a choice function iff $A \notin \mathcal{G}$.

The filter $\mathcal{G}$ is on the natural numbers, but we are really interested in which subsets of X are considered "large" or "small". Thus we will define a filter on X, $\mathcal{F} \subseteq X$, by $Y = \{X_n | n \in A\} \in \mathcal{F} \iff A \in \mathcal{G}$.

First, assume we do in fact have a model $\mathcal{M}$ of ZF° with countably many ur-elements and a non-principal ultrafilter $\mathcal{F}$. Now consider a set of pairwise disjoint two-element sets $X = \{X_n | n \in \mathbb{N}\}$, where $X_n = \{u_{2n}, u_{2n+1}\}$ for all $n \in \mathbb{N}$.

Remark. Many definitions in this section are from Kern's paper [5], which begins by creating a similar model. We expand on proofs where possible, but proof that are cited were originally provided or outlined in [5].

Definition 6 (Natural Extension)**.** [5, p. 5] Consider a permutation $\varphi : \bigcup X \to \bigcup X$. Let the natural extension of φ to $\mathcal{M}$ be

$$\varphi'(Y) = \begin{cases} \varphi(Y) & \text{if } Y \in \bigcup X \\ \{\varphi'(y) | y \in Y\} & \text{if } Y \text{ is a set} \end{cases}.$$

Remark. We will freely assume that a permutation φ refers to its natural extension, as we are interested in the permutation of the ur-elements contained in a set of sets. Note that the natural extension of φ is an automorphism of $\mathcal{M}$; $x \in y \iff \varphi(x) \in \varphi(y)$.

Definition 7. A permutation φ fixes a set Y if $\varphi(Y) = Y$

Note that a permutation may fix a set without being the identity function on that set. If we want to express the more strict condition that $\forall y \in Y, \varphi(y) = y$, we will say that φ *pointwise-fixes* Y.

Definition 8 (Choice Structure)**.** [5, p. 5] A choice structure Γ on X is a set of permutation groups on $\bigcup X$ such that:

(1) For finite $Y \subseteq X$, there is some $S \in \Gamma$ such that for every $\varphi \in S$ pointwise-fixes Y.
(2) For $S, T \in \Gamma$, there is some $U \in \Gamma$ such that $U \subseteq S, T$.

The following sets will be important in constructing such a choice structure:

Definition 9 ($G, G(Y), \Gamma(X, \mathcal{F})$)**.** [5, p. 6] For our given X, the group of permutations of $\bigcup X$ that permute the elements of X locally (meaning, roughly, that the permutations map all elements to elements in the same set):

$$G = \{\varphi : \bigcup X \to \bigcup X | \varphi(u) = u' \wedge u \in x \in X \Rightarrow u' \in x\}$$

The subgroup of G of permutations pointwise-fixing all elements of $Y \subseteq X$:

$$G(Y) = \{\varphi \in G | \forall u \in x \in Y : \varphi(u) = u\} \subseteq G$$

The collection of all such sets of permutations pointwise-fixing elements *not in* the filter $\mathcal{F}$ (i.e., all measure-zero sets are pointwise-fixed for the measure $m_{\mathcal{F}}$):

$$\Gamma(X, \mathcal{F}) = \{G(Y) | Y \notin \mathcal{F}\}$$

Lemma 1. *For our given X and a non-principal ultrafilter $\mathcal{F}$ on X, $\Gamma(X, \mathcal{F})$ is a choice structure.* [5, p. 6]

Proof. We must show that $\Gamma(X, \mathcal{F})$ satisfies the two properties of choice structures.

First, assume A is a finite subset of X. Since $\mathcal{F}$ is non-principal, $A \notin \mathcal{F}$ and thus $G(A) \in \Gamma(X, \mathcal{F})$ by definition. Since all permutations in $G(A)$ pointwise-fix all elements of the sets in A by construction, the first condition holds.

Second, say $G(S), G(T) \in \Gamma$. So $S, T \notin \mathcal{F} \Rightarrow X - S, X - T \in \mathcal{F} \Rightarrow (X - S) \cap (X - T) \in \mathcal{F} \Rightarrow X - (S \cup T) \in \mathcal{F} \Rightarrow S \cup T \notin \mathcal{F}$. So $G(S \cup T) \in \Gamma$. If $\varphi \in \mathcal{G}(S \cup T)$, then it pointwise-fixes all elements of both S and T, so $\varphi \in G(S), G(T)$. Thus $G(S \cup T)$ is a set satisfying the definition. ☐

Definition 10 (Γ-rooted)**.** [5, p. 6] Given a choice structure Γ, Y is Γ-rooted if for some $S \in \Gamma$, every $\varphi \in S$ fixes Y.

Definition 11 (Γ-grounded)**.** [5, p. 6] Given a choice structure Γ, Y is Γ-grounded if every element of its transitive closure is Γ-rooted.

Definition 12 ($\mathcal{M}^\Gamma$)**.** [5, p. 7] For a choice structure Γ, let

$$\mathcal{M}^\Gamma = \{T \in \mathcal{M} | T \text{ is } \Gamma(X, \mathcal{F})\text{-grounded}\}$$

Lemma 2. *For a choice structure Γ, $\mathcal{M}^\Gamma$ is a model of ZF$^\circ$.*

Proof. Proving that the axioms of ZF$^\circ$ are true in $\mathcal{M}^\Gamma$ suffices to show that it is a model of ZF$^\circ$.

(1) *Axiom of Extensionality*
Consider sets $A, B \in \mathcal{M}^\Gamma$ and assume $A \neq B$. Since extensionality holds in $\mathcal{M}$, and A, B are also sets in this model there is some $a \in A$ such that $a \notin B$ (without loss of generality). $a \in \mathcal{M}^\Gamma$ since it is in A, a Γ-grounded set so it is also Γ-grounded. So the axiom is inherited from $\mathcal{M}$.

(2) *Axiom of Pair*
Assume $A, B \in \mathcal{M}^\Gamma$ (so they are Γ-grounded). It will suffice to show that $C = \{A, B\}$ is Γ-rooted, thus in $\mathcal{M}^\Gamma$. So by definition there exist $S_A, S_B \in \Gamma$ which witness the Γ-rootedness of A, B respectively. By the second property of choice structures, there exists $S \subseteq S_A, S_B$ in Γ; this set witnesses the Γ-rootedness of C since it contains permutations which fix both A and B.

(3) *Axiom of Union*
Suppose $A = \{A_i\}_{i \in I} \in \mathcal{M}^\Gamma$. Let $\bigcup A = \bigcup\limits_{i \in I} A_i$. Since A is Γ-grounded, so are all A_i, thus it suffices to show that $\bigcup A$ is Γ-rooted. Assume S is the permutation group which witnesses the Γ-rootedness of A; then all $\varphi \in S$ fix A, meaning that they map elements of A (and only elements of A) into A. So for all $\varphi \in S$, $a \in A$ if and only if $a \in A_i$ for some $i \in I$, if and only if $\varphi(a) \in A_j$ for some $j \in I$, if and only if $\varphi(a) \in \bigcup A$. Thus φ fixes $\bigcup A$, as desired.

(4) *Axiom of Power Set*
Suppose $A \in \mathcal{M}^\Gamma$. Let $\mathcal{P}(A)^{\mathcal{M}^\Gamma} = \{B | B \subseteq A, B \in \mathcal{M}^\Gamma\}$. Since $A \in \mathcal{M}^\Gamma$, A is Γ-grounded. Since the elements of $\mathcal{P}(A)^{\mathcal{M}^\Gamma}$ are Γ-grounded subsets of A, it suffices to show that $\mathcal{P}(A)^{\mathcal{M}^\Gamma}$ is Γ-rooted. Let S be a permutation group which witnesses the Γ-rootedness of A – so every permutation in S fixes A. Thus every permutation in S maps a subset of A to another subset of A. We must show that every $\varphi \in S$ maps elements of $\mathcal{P}(A)^{\mathcal{M}^\Gamma}$ to other elements of $\mathcal{P}(A)^{\mathcal{M}^\Gamma}$, and non-elements to non-elements.

If $B \in \mathcal{P}(A)^{\mathcal{M}^\Gamma}$, then since φ fixes A (i.e., maps elements of A to other elements of A), $\varphi(Y)$ will be a subset of A and thus an element of $\mathcal{P}(A)^{\mathcal{M}^\Gamma}$, as desired. If $B \notin \mathcal{P}(A)^{\mathcal{M}^\Gamma}$, then B is a set which contains an element b not in A. Since φ fixes A, $\varphi(b) \notin A$. Thus $\varphi(B)$ cannot be a subset of A, and $\varphi(B) \notin \mathcal{P}(A)^{\mathcal{M}^\Gamma}$, as desired.

(5) *Axiom of Foundation*
If there exists a set which violates the axiom in $\mathcal{M}^\Gamma$ (a non-empty set in which has a non-empty intersection with each of its elements), it must also exist in $\mathcal{M}$, which is impossible by foundation in $\mathcal{M}$. Therefore the axiom is inherited from $\mathcal{M}$.

(6) *Axiom of Infinity*
An inductive set exists in $\mathcal{M}$, and it contains no ur-elements and is therefore fixed by local permutations. Therefore the axiom is inherited from $\mathcal{M}$, for the set is clearly Γ-grounded.

(7) *Axiom Schema of Comprehension*
Suppose $\psi(x,p)$ is a formula in $\mathcal{M}^\Gamma$, where $p \in \mathcal{M}^\Gamma$ is a parameter. Suppose A is a set in $\mathcal{M}^\Gamma$. We must show that $\mathcal{M}^\Gamma \models \exists Y(Y = \{x \in A | \psi(x,p)\})$. Equivalently, it suffices to show that $\mathcal{M} \models \exists Y(Y = \{x \in A | \psi^{\mathcal{M}^\Gamma}(x,p)\} \wedge Y \in \mathcal{M}^\Gamma)$, where $\psi^{\mathcal{M}^\Gamma}$ is the formula ψ with all quantifiers changed from "$\forall y$" to "$\forall\Gamma$-grounded y".

Define in $\mathcal{M}$ the set $B = \{x \in A | \mathcal{M}^\Gamma \models \psi(x,p)\}$ – such a set exists in $\mathcal{M}$ by comprehension, because $\mathcal{M}^\Gamma \models \psi(x,p) \iff \mathcal{M} \models \psi^{\mathcal{M}^\Gamma}(x,p)$. So we must show that B is Γ-grounded. $A \in \mathcal{M}^\Gamma \Rightarrow A \subseteq \mathcal{M}^\Gamma \Rightarrow B \subseteq \mathcal{M}^\Gamma$, so every element of B is Γ-grounded. Thus it remains to show that B is Γ-rooted.

Because A and p are in $\mathcal{M}^\Gamma$, there must be permutation groups S_A, S_p in which all permutations fix A and p, respectively. By the second property of choice structures, there exists some permutation group S which is contained in both of these groups, and thus also fixes both A and p. Because every permutation $\varphi \in S$ is an automorphism of $\mathcal{M}$, and automorphisms preserve structure (and therefore preserve the truth of formulas),

$$\forall x(\psi^{\mathcal{M}^\Gamma}(x,p) \wedge x \in A \iff \psi^{\mathcal{M}^\Gamma}(\varphi(x), \varphi(p)) \wedge \varphi(x) \in A).$$

But for $\varphi \in S$, we have $\varphi(a) = A$ and $\varphi(p) = p$, so it follows that

$$\forall x(\psi^{\mathcal{M}^\Gamma}(x,p) \wedge x \in A \iff \psi^{\mathcal{M}^\Gamma}(\varphi(x), p) \wedge \varphi(x) \in A).$$

Equivalently, we have that $\forall x(x \in B \iff \varphi(x) \in B)$, and thus $\varphi(B) = B$ for all $\varphi \in S$. Thus B is Γ-rooted.

(8) *Axiom Schema of Replacement*
Suppose $\varphi(x,y,p)$ is a formula in $\mathcal{M}^\Gamma$ defining a function on a set A, where $p \in \mathcal{M}^\Gamma$ is a parameter. So by definition, $\mathcal{M}^\Gamma \models \forall x \in A \exists! y(\varphi(x,y,p))$. Then we have equivalently that $\mathcal{M} \models \forall x \in A \exists! y(\psi(x,y,p))$,

where $\psi(x,y,p) = y \in \mathcal{M}^\Gamma \wedge \varphi^{\mathcal{M}^\Gamma}(x,y,p)$, and where $\varphi^{\mathcal{M}^\Gamma}$ is the formula φ with all quantifiers changed from "$\forall y$" to "$\forall\Gamma$-grounded y". So in $\mathcal{M}$, ψ defines a function on A.

Say B is the range of this function defined by ψ in $\mathcal{M}$ – in particular, $B = \{y \in \mathcal{M}^\Gamma | \exists x \in A(\psi^{\mathcal{M}^\Gamma}(x,y,p))\}$. By replacement in $\mathcal{M}$, B is a set in $\mathcal{M}$. But in ψ we specify that all range elements are elements of $\mathcal{M}^\Gamma$, and thus Γ-grounded. So, to show $B \in \mathcal{M}^\Gamma$, it remains to show that B is Γ-rooted. Since A and p are elements of $\mathcal{M}^\Gamma$, they are Γ-grounded, and so there are permutation groups S_A, S_p in which all permutations fix A and p, respectively. Then there exists some permutation group S which is contained in both of these groups, and thus also fixes both A and p. Thus for every permutation $\phi \in S$,

$$\forall y(\psi^{\mathcal{M}^\Gamma}(x,y,p) \wedge x \in A \iff \psi^{\mathcal{M}^\Gamma}(\phi(x),\phi(y),p) \wedge \phi(x) \in A).$$

Equivalently, we have that $\forall y(y \in B \iff \phi(y) \in B)$, and thus $\phi(B) = B$ for all $\phi \in S$. Thus B is Γ-rooted, so $B \in \mathcal{M}^\Gamma$. B is the range of the function defined by $\psi(x,y,p)$ in $\mathcal{M}^\Gamma$, therefore replacement is satisfied.

Thus $\mathcal{M}^\Gamma$ is a model of ZF°. □

Now we prove the desired conclusions for our set of ur-element pairs, X, and its non-principal ultrafilter $\mathcal{F}$.

Lemma 3. $X \in \mathcal{M}^{\Gamma(X,\mathcal{F})}$. [5, p. 8]

Proof. Each $u_n \in \bigcup X$, as a ur-element, is clearly $\Gamma(X,\mathcal{F})$-rooted (if $u_n \in X_m$, consider $G(\{X_m\})$). Each $X_n \in X$, a set of two ur-elements, is also $\Gamma(X,\mathcal{F})$-rooted: $\{X_n\}$ is finite, so not in $\mathcal{F}$, and thus the set is pointwise-fixed by $G(\{X_n\}) \in \Gamma(X,\mathcal{F})$.

X itself is fixed by any $\varphi \in G$: each such φ permutes elements within each X_n, so it fixes these sets and thus also fixes X. Therefore X is $\Gamma(X,\mathcal{F})$-rooted, and we have thus shown that X is $\Gamma(X,\mathcal{F})$-grounded. □

Lemma 4. *$Y \subseteq X$ has a choice function defined on it in $\mathcal{M}^{\Gamma(X,\mathcal{F})} \iff Y \notin \mathcal{F}$.* [5, p. 8]

Proof. It is useful to recall that a function f on a set S is a set of ordered pairs $\{(s,f(s)) | s \in S\}$.

First, suppose $Y \notin \mathcal{F}$ and $f \in \mathcal{M}$ is a choice function on $Y \subseteq X$. By definition, $G(Y) \in \Gamma(X,\mathcal{F})$, so this is a set in $\Gamma(X,\mathcal{F})$ containing all permutations pointwise-fixing Y. But all of these permutations must also pointwise-fix f (in other words, $f = \varphi \circ f$ for all permutations in $G(Y)$), since the sets contain exactly the same ur-elements ($\bigcup Y$). So f is $\Gamma(X,\mathcal{F})$-rooted. Each set in f is clearly $\Gamma(X,\mathcal{F})$-grounded; a permutation fixes an ordered pair if it fixes both elements, so we know the set $\{\{u,v\},u\} \in f$ – for example – is $\Gamma(X,\mathcal{F})$-grounded because $G(\{\{u,v\}\}) \in \Gamma(X,\mathcal{F})$.

Thus f is $\Gamma(X,\mathcal{F})$-grounded, so there exists a choice function on Y in $\mathcal{M}^{\Gamma(X,\mathcal{F})}$, namely f.

Now, suppose $Y \in \mathcal{F}$. Assume that $f \in \mathcal{M}^{\Gamma(X,\mathcal{F})}$ is a choice function on $Y \subseteq X$ for sake of contradiction. f then is $\Gamma(X,\mathcal{F})$-grounded, and in particular $\Gamma(X,\mathcal{F})$-rooted. So we can choose a $Z \subseteq X$ such that $Z \notin \mathcal{F}$ and all $\varphi \in G(Z)$ fix f. $Y \nsubseteq Z$ (since then we would have $Z \in \mathcal{F}$ by properties of filters), so there is some $y = \{u,v\} \in Y$ not in Z. Say without loss of generality that $f(y) = u$.

Consider a permutation $\psi : \bigcup X \to \bigcup X$ such that

$$\psi(x) = \begin{cases} v & \text{if } x = u \\ u & \text{if } x = v \\ x & \text{otherwise} \end{cases}$$

Since $y \notin Z$, ψ pointwise-fixes Z and thus $\psi \in G(Z)$; clearly it does not fix f though. But Z was chosen specifically such that all $\varphi \in G(Z)$ fixed f, thus we have a contradiction and conclude that there cannot exist a choice function for Y in $M^{\Gamma(X,\mathcal{F})}$. $\square$

Thus we have shown that a non-principal ultrafilter allows us to construct a model of ZF° in which we have only choice functions on "small" sets. This is somewhere between the classic example (only finite choice functions) and what the Axiom of Choice guarantees (choice functions on all collections of nonempty sets) – measure-zero sets (of nonempty sets), as defined by the ultrafilter, are exactly the sets with choice functions.

The existence of this set is a theoretically interesting result. Recall the idea of a set of socks used to explain choice functions. We have constructed a set of socks with an interesting modification to this classic example – the ultrafilter allows us to differentiate between "small infinity" and "large infinity" for pairs of socks, and we can indeed make choices on finite and small infinite sets. The existence of this set thus allows us to have choice functions on *some* infinite sets, where in the classic model all infinite sets of non-distinct pairs lack choice functions, and the existence of this set is consistent with ZF°.

We will call such a sequence a Pseudo-Russell sequence (for reasons made more clear in Chapter 9 when we discuss Russell sequences), which we define formally as follows:

Definition 13 (Pseudo-Russell Sequence)**.** A sequence $(X_n)_{n\in\mathbb{N}}$, where the X_n are pairwise-disjoint sets of two ur-elements, is a Pseudo-Russell sequence if for some non-principal ultrafilter $\mathcal{F}$ on $\mathbb{N}$, for all $M \subseteq \mathbb{N}$:

$$M \in \mathcal{F} \iff \prod_{m\in M} X_m \text{ is empty.}$$

The elements of this product are precisely choice functions, as a point in the product set contains exactly one element from each set in X; thus the product being empty is equivalent to there existing no choice functions on the sequence.

We will also use the term "Weak Russell sequence" to refer to the more general sequence which has no choice function over the whole sequence (though there may be choice functions on infinite subsequences).

Definition 14 (Weak Russell Sequence)**.** A sequence $(X_n)_{n\in\mathbb{N}}$, where the X_n are pairwise-disjoint sets of two ur-elements, is a Weak Russell sequence if $\prod\limits_{n\in\mathbb{N}} X_n = \emptyset$.

Definition 15 (Weak Russell's Socks Set, Pseudo-Russell's Socks Set)**.** If $(X_n)_{n\in\mathbb{N}}$ is a Weak (Pseudo) Russell sequence, we will call $\{X_n | n \in \mathbb{N}\}$ a Weak (Pseudo) Russell's Socks set.

Note that a Pseudo-Russell's Socks set is more generally a Weak Russell's Socks set, so we have also in fact shown that there exists a model where a Weak Russell's Socks set exists. Now that we have such a model, we can explore the implications of the existence of such sets.

5. Implications in Logic

A fundamental theorem of sentential logic is the following:

Fact 5 (Compactness Theorem (ZFC))**.** *A set of sentences is satisfiable if and only if it is finitely satisfiable.*

The Compactness Theorem, however, follows from the Axiom of Choice; in fact, the existence of a Pseudo-Russell's Socks set can be used to construct an explicit violation of the Compactness Theorem in our model.

Fact 6. *If a model contains a Pseudo-Russell's Socks set $X = \{X_n | n \in \mathbb{N}\}$, then the Compactness Theorem is false in that model.*

Proof. For all $n \in \mathbb{N}$, say $X_n = \{a_n, b_n\}$, and let these ur-elements represent sentence symbols. Consider for every n a set of sentences $\sigma_n = \{a_n \iff \neg b_n, b_n \iff \neg a_n\}$. Note that these sentences are logically equivalent, but by including both we avoid creating a selection rule between the symbols. For a set $A \subseteq \mathbb{N}$, let $\Sigma_A = \bigcup\{\sigma_n | n \in A\}$.

For Σ_A to be satisfiable, there must be a model of the set of sentences – this means simply that we must be able to pick out either a_n or b_n for each $n \in A$ and designate it to be true in the model, while the other is false (we can just assign "true" to the rest of the symbols in the language, since symbols not in Σ_A are irrelevant to whether the structure satisfies Σ_A). So a Pseudo-Russell's Socks set is equivalent to a set of pairs of indistinguishable sentence symbols. Therefore Σ_A is satisfiable if and only if $m(A) = 0$, since then and only then are we able to choose a sock from each pair.

Consider now $\Sigma = \Sigma_{\mathbb{N}}$. Since all finite sets are measure-zero, all finite $\Sigma_0 \subseteq \Sigma$ are satisfiable. Σ, however, is not itself satisfiable, for a model of Σ is exactly a choice function over a Pseudo-Russell's Socks set. Therefore Σ is finitely satisfiable but not satisfiable, and the Compactness Theorem is false. $\square$

This result already indicates that the existence of a Pseudo-Russell's Socks set is going to be very upsetting to logic – we have already shown that one of the fundamental theorems for logic is in fact provably false in our model, and as one would expect, this issue propagates rather quickly to other important pillars of logic.

Definition 16 (Completeness Theorem)**.** Every consistent set of formulas is satisfiable.

Fact 7. *The Completeness Theorem implies the Compactness Theorem.*

Proof. In any model, a satisfiable set of sentences is clearly finitely satisfiable, so we need to show the reverse holds. Assume the Completeness Theorem holds in a particular model. Consider a finitely satisfiable set of sentences Σ in the same model, and assume for sake of contradiction that Σ is not satisfiable. Then by the Completeness Theorem, Σ is inconsistent, so a contradiction can be derived in finitely many steps. But since Σ is finitely satisfiable, a finite-step proof of a contradiction from Σ is not possible, and therefore Compactness holds. $\square$

The following theorem thus follows immediately by contrapositive.

Fact 8. *The existence of a Pseudo-Russell's Socks set implies the failure of the Completeness Theorem.*

So the Completeness Theorem is false in our model $\mathcal{M}^{\Gamma(X,\mathcal{F})}$, and there exist consistent sets of formulas which are not satisfiable in this model.

6. Implications in Topology

Tychonoff's Theorem, equivalent to the Axiom of Choice, suggests that the existence of a Pseudo-Russell's Socks set (or some variation thereof) should allow us to find explicit failures of topological theorems as well. Recall the following:

Definition 17 (Compactness)**.** A topological space is compact if every open cover of the space contains a finite subset which also covers the space.

Definition 18 (Product Topology)**.** Let $X = \prod_{i \in I} X_i$, where for all $i \in I$ X_i has the topology τ_i. Consider

$$\mathcal{B} = \{\prod_{i \in I} U_i | \forall i \in I : U_i \in \tau_i, \text{ and only finitely many } U_i \neq X_i\}.$$

The product topology on X is the topology generated by the basis $\mathcal{B}$.

Fact 9 (Tychonoff's Theorem (ZFC))**.** *The product of compact topological spaces is compact in the product topology.*

We encounter a difficulty here, because the Axiom of Choice is required to ensure that infinite products are non-empty, and an empty product prevents us from achieving any interesting results. An element of Cartesian product is in fact a choice function over all of the component sets – but what if we only had to make a choice on measure-zero-many sets?

Let's consider a modification of a Pseudo-Russell's Socks set. Instead of each set containing a pair of socks, let's say each set has two white socks and a black one – the black sock is distinct from the white ones, but we cannot distinguish between the white socks. More formally:

$$Z = \{\{a_n, b_n, ab_n\} | n \in \mathbb{N}\}$$

where a_n, b_n are indistinguishable within our model – we can say that the other element is distinguishable from these two by first defining a set $AB = \{ab_1, ab_2, \ldots\}$ and using it to extend the Pseudo-Russell's Socks set. The product of the sets in Z then is not empty, but it also does not contain all of the points it would in a model of ZFC. It instead contains only points z where $\pi_i(z) = ab_i$ for measure-one-many $i \in \mathbb{N}$ (where $\pi_n(z)$ is the i^{th} coordinate of z).

Lemma 5. *For any $I \subseteq \mathbb{N}$, there exists a $z \in Z$ with $I = \{i \in \mathbb{N} | \pi_i(z) = ab_i\}$ if and only if $m(I) = 1$.*

Proof. Fix some $I \subseteq \mathbb{N}$.

To prove the forward direction, suppose there exists a $z \in Z$ such that $I = \{i \in \mathbb{N} | \pi_i(z) = ab_i\}$. Suppose for sake of contradiction that $m(I) = 0$. Then we are choosing a_i or b_i for all $i \in (\mathbb{N} - I)$, a measure-one set. By construction, there is no such choice function on these sets, contradicting that I is measure-zero.

To prove the reverse direction, let $z = \prod_{n \in \mathbb{N}} z_n$, where $z_n = ab_n$ if $n \in I$ and let $z_n = a_n$ if $n \notin I$. This requires choice on only a measure-zero set $(\mathbb{N} - I)$ which is possible by construction. $\square$

Now we have a non-empty product set which extends the Pseudo-Russell's Socks set without using the Axiom of Choice, as desired.

Assign the discrete topology to each Z_n. Since there are finitely many open sets in each space, each space is clearly compact. Consider $P = \prod_{n \in \mathbb{N}} Z_n$, a product of compact spaces.

For $i \in \mathbb{N}$, let $U_i \in P$ be the set where

$$\pi_j(U_i) = \begin{cases} Z_i & \text{if } j \neq i \\ \{ab_j\} & \text{if } j = i \end{cases}.$$

Each U_i is open in the product topology – U_i is essentially the set of all points in P where we choose the black sock as the i^{th} coordinate. Consider a finite collection of these sets, $\{U_{i_1}, \ldots, U_{i_m}\}$. This set does not cover P – a point in which we select white socks on the coordinates $i_1, \ldots, i_m$ but black socks on all remaining coordinates is a point in the product (since we choose only finitely many white socks), but it is not in the cover. We then assume for sake of contradiction that P is compact, and therefore by the definition of compactness, $\{U_n\}_{n \in \mathbb{N}}$ does not cover P (since none of its finite subcovers can).

Since $\{U_n\}_{n \in \mathbb{N}}$ does not cover P, there is some $p \in P$ such that $p \notin \bigcup U_n$. In other words, there is a p that corresponds to choosing no black socks, thus choosing a white sock out of every set (requiring a choice function on the original Pseudo-Russell's Socks set), a contradiction. Therefore, P cannot be compact.

Therefore, this extended version of the Pseudo-Russell's Socks set Z witnesses an interesting violation of Tychonoff's Theorem - it is a noncompact product of compact subspaces which is not empty.

Theorem 1. *If Z is an extension of a Pseudo-Russell's Socks set as constructed above and it is given the discrete topology, then Z witnesses the failure of Tychonoff's Theorem.*

Proof. We showed above that the existence of the set Z proved the existence of a non-compact set (P) which was the product of countably many compact spaces. P witnesses a failure of Tychonoff's Theorem. But since a Pseudo-Russell's Socks set requires only a simple extension to become Z (we need only to add a single element to each component), its existence is sufficient to prove the failure of Tychonoff's Theorem. $\square$

7. An Equivalence Theorem in Logic

Now that we have seen some implications of the existence of the existence of a Pseudo-Russell Socks set, what can we show in the reverse direction? While it is well known the Axiom of Choice implies the Compactness Theorem, what additional conditions do we need to place on the failure of the Compactness Theorem to show the existence of a Weak Russell Socks set?

We create the following definition for the next theorem:

Definition 19 (Component-Countable)**.** Let Σ be a set of sentences of the form $A \leftrightarrow \neg B$ where A, B are sentence symbols. Let G be the graph with

- $V = \{A | A \text{ is a sentence symbol occurring in } \Sigma\}$
- $E = \{(A, B) | \Sigma \vdash (A \leftrightarrow B) \text{ or } \Sigma \vdash (A \leftrightarrow \neg B)\}$

If G has countably many components, say Σ is component-countable.

Theorem 2 (Weak Russell's Socks Set Equivalence in Logic)**.** *The existence of a Weak Russell's Socks set is equivalent to the existence of a component-countable set of sentences of the form $A \leftrightarrow \neg B$ (where A and B are sentence symbols) which is not satisfiable but finitely satisfiable.*

Remark. The use of provability ($\vdash$), rather than implication ($\models$), is important in this proof, as the failure of the Compactness Theorem on Σ means that Σ must also be unsatisfiable, but an inconsistency cannot be derived from Σ since doing so would require a finite proof.

Proof. Suppose such a set of sentences, Σ, exists. Let S be the set of sentence symbols occurring in the sentences of Σ. Define an relation on S where $A \sim B \iff \Sigma \vdash (A \leftrightarrow B)$ – we can see that this is an equivalence relation (reflexive, symmetric, and transitive), and we will denote equivalence classes by $[E] = \{D | E \sim D\}$. We define a Weak Russell's Socks set whose socks are equivalence classes, and we will show that each component of the graph G (defined in Definition 19) represents a pair of socks, so there are finitely many pairwise-disjoint pairs of socks with no choice function. We define:

$$X = \{\{[A], [B]\} | \Sigma \vdash (A \leftrightarrow \neg B\}.$$

First, we note that the pair $\{[A], [A]\}$ cannot occur in X for any $A \in S$, since that would imply that $\Sigma \vdash (A \leftrightarrow \neg A)$, but Σ is finitely satisfiable and thus we cannot prove contradictions.

Next, we show that each $[A]$ occurs in exactly one pair in X. If $A \in S$, then there is a B such that $A \leftrightarrow \neg B$ or $B \leftrightarrow \neg A$ by construction. So each $[A]$ is certainly in at least one pair in X since $(A \leftrightarrow \neg B) \in \Sigma \Rightarrow \Sigma \vdash (A \leftrightarrow \neg B) \wedge \Sigma \vdash (B \leftrightarrow \neg A)$. Now say $\{[A], [B]\}$ and $\{[A], [C]\}$ are in X. Then $\Sigma \vdash (A \leftrightarrow \neg B)$ and $\Sigma \vdash (A \leftrightarrow \neg C)$. Therefore $\Sigma \vdash (B \leftrightarrow C)$ and thus $[B] = [C]$.

So X is a set of pairwise-disjoint sets, each containing two different elements. We can show that X is countable since Σ is component-countable. If A and B are in the same component of G, then they are connected by a finite number of edges. So either $A \leftrightarrow \neg B$ or $A \leftrightarrow B$ is provable from Σ, thus either $[A] = [B]$ or $\{[A], [B]\} \in X$. Inversely, if A and B are not in the same component, then they cannot be equivalent or in the same pair, since either of those statements would require a proof from Σ, but there is no finite set of sentences connecting A and B in Σ. Thus:

Claim. *A and B are in the same component of G $\iff$ either $[A] = [B]$ or $\{[A], [B]\} \in X$.*

Therefore, X has the same cardinality as the number of components in G. But since G has countably many components (not finite, otherwise Σ would be finite and thus satisfiable), $|X| = \aleph_0$.

Finally, assume for sake of contradiction that a choice function f exists on X. We wish to construct a model $\mathfrak{A}$ satisfying Σ. Index X by $X = \{X_n | n \in \mathbb{N}\}$ and let $\mathfrak{A}$ assign "true" to all elements in the equivalence class $f(X_n)$ for each $n \in \mathbb{N}$ and "false" for the remaining symbols in S. Then $\mathfrak{A} \models \Sigma$; if $(A \leftrightarrow \neg B) \in \Sigma$, $\mathfrak{A}$ assigns "true" to exactly one of A or B, and thus $\mathfrak{A} \models (A \leftrightarrow \neg B)$. Thus Σ is satisfiable, a contradiction.

Therefore the set X is a countable set of pairwise-disjoint sets, each containing two distinct elements, with no choice function; X is a Weak Russell's Socks set.

The proof of the forward direction can be proved by a simple modification to the proof of Fact 6; as long as the entire set has no choice function, the proof is the same. □

8. An Equivalence Theorem in Topology

Now, we seek an equivalence in topology. We saw earlier that the existence of a Pseudo-Russell's Socks set implied the failure of Tychonoff's Theorem, so what particular failure of Tychonoff's Theorem must we witness for a Weak Russell's Socks set to exist?

First, recall the definition of a basis for a topology:

Definition 20 (Basis)**.** $\mathcal{B} \subseteq \mathcal{P}(X)$ is a basis for a given topology on a set X if it satisfies the following:

(1) If $x \in Y \subseteq X$ and Y is open, then $x \in B \subseteq Y$ for some $B \in \mathcal{B}$
(2) Every $B \in \mathcal{B}$ is open

We will use a tree to represent topological product spaces as follows:

Definition 21 (Tree, Branch, Cone)**.** We will define a tree to be a rooted graph, i.e., a collection of discrete nodes connected by directed edges with a designated root node and no cycles. Say a node is at *level* m if its distance from the root node is m.

Let a branch be an infinite path originating at the root node (a branch will pass through a node at each level).

Let a cone originating at node σ be the collection of branches through σ.

Theorem 3 (Weak Russell's Socks Set Equivalence in Topology)**.** *The existence of a Weak Russell's Socks set is equivalent to the existence of a non-compact space that is the product of countably many compact 3-element spaces with the discrete topology.*

Proof. Let X be a non-compact space that is the product of countably many 3-element compact spaces with the discrete topology. So, say for all $n \in \mathbb{N}$, $X_n = \{a_n, b_n, c_n\}$ with the discrete topology and $X = \prod_{n\in\mathbb{N}} X_n$.

Since X is not compact, it must be non-empty. So say $p \in X$ and say without loss of generality that $\pi_n(p) = c_n$ for all $n \in \mathbb{N}$ (where $\pi_n(p)$ is the n^{th} coordinate of p). Let $Y = \prod_{n\in\mathbb{N}} Y_n$, where $Y_n = \{a_n, b_n\}$ for all n. Y is a countable product of pairwise-disjoint two-element sets. Therefore, $Y = \emptyset$ if and only if $\{Y_n | n \in \mathbb{N}\}$ is a Weak Russell's Socks set.

Let T be an infinite tree with one root node σ_0 and each node having three children. We will say the nodes of the tree are $\{\sigma_i | i \in I\}$ for an index set I including 0. For each node σ_i, denote its three children $\sigma_{i,a}$, $\sigma_{i,b}$, and $\sigma_{i,c}$ (this is not to imply a selection rule - given a particular node, we can arbitrarily label its three children in this way).

Observe that a branch in T represents an element of the product space X, since we make one of three choices at each level $n \in \mathbb{N}$. Observe also that a node at level m of T represents a finite sequence of choices on $X_1, \ldots, X_m$.

We will show that the cones of T represent a basis for X.

Consider a cone originating from a node at level m of T. To arrive at this node, we have chosen a specific node at each previous level, but we allow all choices at further levels. So, if we let c be a choice function on the lower levels, the cone can be represented as $\{c(S_1)\} \times \ldots \times \{c(S_m)\} \times S_{m+1} \times S_{m+2} \times \ldots$.

The collection of all cones of this tree (when represented as sets as described above) is a basis for X:

(1) It will suffice to show that all basic open sets of X are the unions of cones. Suppose $A \subseteq X$ is a basic open set (i.e., an element of the basis for the product topology defined in Definition 18). So $A = \prod_{n\in\mathbb{N}} A_n$, where for some finite $I \subseteq \mathbb{N}$, $i \notin I \iff A_i = X_i$. If $m = \max(I)$, then A is simply the union of cones originating at the nodes in $\prod_{i\in\{1,\ldots,m\}} A_i$.

(2) By definition, a cone is the product of finitely many singletons (selected by a choice function) followed by infinitely many repetitions of X_n. Therefore, a cone is the product of open sets, only finitely many of which are proper subsets of their space. Thus each cone is an open set in the product topology.

Therefore, the cones of T represent a basis for X. Note also that every set in the usual basis of the product topology can be expressed as unions and finite intersections of cones.

For each node $\sigma_i \in T$, define the sets:

$$\begin{aligned} B_{\sigma_i} &= \{b|b \text{ is a branch, } \sigma_i \subseteq b\} \\ C_{\sigma_i} &= \{B_\tau|\tau \text{ is an initial segment of } \sigma_i\} \\ [T] &= \{b|b \text{ is a branch of } T\} \end{aligned}$$

So B_{σ_i} is the cone originating at σ_i, C_{σ_i} is the set of cones containing B_{σ_i}, and $[T]$ represents exactly the elements of X.

Let S be an open cover of $[T]$ with no finite subcover, and let S' be the set of basis elements (cones) which are subsets of elements in S. Then since $\bigcup S = \bigcup S' = X$, if S' has a finite subcover, S must have one as well. Thus S' has no finite subset which covers X, and therefore any finite collection of cones in S' fails to cover $[T]$.

Let $T' = \{\tau|\forall\sigma(B_\sigma \in S' \Rightarrow \sigma \text{ is not an initial segment of } \tau)\}$; this is a subtree of T consisting of all nodes $\sigma \in T$ such that the cone originating at σ is not contained in any single element of S'.

T' is infinite: otherwise, for some minimal n, every element of T' is at a level less than or equal to n. So for every node σ on level n, there is a cone $B_{\tau(\sigma)} \in S'$ such that $\tau(\sigma)$ is an initial segment of σ (that is, $B_\sigma \subseteq B_{\tau(\sigma)}$). But X is the union of the finitely many cones B_σ for σ on level n, so X is the union of the finitely many cones $B_{\tau(\sigma)}$ for σ on level n. These finitely many cones form a finite subcover of S', a contradiction.

T' has no branch: otherwise, suppose b is a branch of T'. Since S' is a cover of X, there is some cone $B_\sigma \in S'$ such that $b \in B_\sigma$, that is, such that $\sigma \in b$. But since $B_\sigma \in S'$, $\sigma \notin T'$ by definition of T', contradicting that b is a branch of T'.

Assume now for sake of contradiction that there exists a choice function f on $\{Y_n|n \in \mathbb{N}\}$. We will show that X is isomorphic to $\{0,1,2\}^\omega$ by mapping the black sock in each set to 0, the white sock defined by the choice function to 1, and the other white sock to 2. Define an isomorphism $j : X \to \{0,1,2\}^\omega$

by $j(x) = \prod_{n\in\mathbb{N}} g(\pi_n(x))$, where

$$g(x' \in X_n) = \begin{cases} 0 & \text{if } x' = c_n \\ 1 & \text{if } x' = f(X_n) \\ 2 & \text{otherwise} \end{cases} .$$

Thus $X \cong \{0,1,2\}^\omega$.

We can inductively define a branch through any infinite subtree of T: suppose $\sigma_i = f|n = (f(0),\ldots,f(n-1))$ has been defined so that $T' \cup B_{\sigma_i}$ is infinite. Then at least one of $T' \cup B_{\sigma_i,j}$ for $j \in \{0,1,2\}$ must be infinite, so we can let $f(n)$ be the least such j. Thus in T, every infinite subtree must have a branch, and therefore T' has a branch. $\lightning$

Thus there is no choice function on $\{Y_n|n \in \mathbb{N}\}$, and this set is by construction a Weak Russell's Socks set. Therefore a Weak Russell's Socks set exists.

The proof of Theorem 1 proves the converse with only slight modification – the proof still holds after dropping the assumption that some infinite sets have choice functions.

$\square$

9. Russell Sets

The sets we have been discussing which have no choice function are particularly interesting in terms of cardinality. If we take the union of a Weak Russell set, it is clear that it will not be countable (as an ordering can be used to make a choice function), but then what can we say about its cardinality? Though intuitively it seems "larger" than $\aleph_0$, since it is infinite and not countable, is such a cardinal even comparable to $\aleph_0$?

First, recall the following definitions:

Definition 22 (Cardinality)**.** $|A|$ is the cardinality of the set A. c is said to be a cardinal if for some set C, $c = |C|$.

Definition 23. For sets A and B, we say $|A| \leq |B|$ if there exists an injective function $j : A \to B$.

Definition 24. For sets A and B, we say $|A| = |B|$ if $|A| \leq |B|$ and $|B| \leq |A|$, or equivalently if there exists an bijective function $j : A \to B$.

Definition 25. For sets A and B, we say $|A| < |B|$ if $|A| \leq |B|$ and $|A| \neq |B|$.

Definition 26 (Disjoint Union)**.** Suppose A and B are sets. The disjoint union of A and B is $A \uplus B = \{(A,a)|a \in A\} \cup \{(B,b)|b \in B\}$. If A and B are disjoint, $A \cup B \cong A \uplus B$.

Definition 27 (Cardinal Arithmetic)**.** Suppose X and Y are sets and $a = |X|, b = |Y|$.

- $a + b = |X \uplus Y|$
- $ab = |X \times Y|$
- $a^b = |\{f | f : Y \to X\}|$

In models where the Axiom of Choice is true, familiar rules about cardinality hold: adding a finite number of elements to an infinite set does not alter its cardinality, all cardinals are comparable to $\aleph_0$, etc. However, as we will show, several of these rules do not hold without the Axiom of Choice.

Definition 28 (Dedekind Finite/Dedekind Set)**.** [6, p. 2-3] A set X is Dedekind finite if it satisfies the following equivalent conditions:

- $\aleph_0 \not\leq |X|$
- $|X| \neq |X| + 1$
- $A \subsetneq X \Rightarrow |A| < |X|$

A cardinal is Dedekind finite if it is the cardinal of a Dedekind finite set
A set (or cardinal) is Dedekind infinite if it is not Dedekind finite. A Dedekind set is a set which is infinite and Dedekind finite.

No Dedekind sets exist in models where the Axiom of Choice is true, as seems intuitive. However, we will see that this is not the case where AC fails.

We now use definitions from Herrlich and Tachtsis[6] to specifically identify the kind of sets and cardinals we are interested in:

Definition 29 (Russell sequence)**.** [6, p. 2] A Russell sequence is a sequence $(X_n)_{n\in\mathbb{N}}$ of pairwise-disjoint 2-element sets such that for each infinite subset M of $\mathbb{N}$, the product $\prod_{m\in M} X_m$ is empty. (Equivalently, there is no choice function on $\{X_m | m \in \mathbb{M}\}$.)

Definition 30 (Russell set)**.** [6, p. 2] A Russell set is the union $X = \bigcup_{n\in\mathbb{N}} X_n$ of some Russell sequence $(X_n)_{n\in\mathbb{N}}$.

Definition 31 (Russell cardinal)**.** [6, p. 2] A Russell cardinal is the cardinal $|X|$ of some Russell set X.

A Russell Set is then a particular kind of Weak Russell set, in that no infinite subsets have choice functions.

Herrlich and Tachtsis provide several propositions about Russell Sets. Here we will state some of these interesting properties and provide select proofs. (Propositions 1-14 are from [6]; proofs are original unless cited).

The first proposition will show that every Russell set is a Dedekind set. This is a key observation for many results in this section, but it is an interesting conclusion itself. By adding or removing a single sock or a pair from a Russell set, the cardinality of that set changes. This is already a fundamental difference between Russell sets and any infinite set that exists in ZFC (where there are no Dedekind sets).

Proposition 1. *Every Russell set is a Dedekind set.*

Proof. Let $(X_n)_{n\in\mathbb{N}}$ be a Russell sequence and let X be the corresponding Russell set. Suppose for sake of contradiction that X is not Dedekind finite. In particular, for $y \notin X$, $|X| = |X \cup \{y\}|$. So there exists a bijection $j : X \cup \{y\} \to X$, and $j|_X : X \to X - \{j(y)\}$ is also a bijection.

Consider the sequence S: $j(y), j(j(y)), \ldots$. This is an infinite sequence of distinct elements of X (since j is bijective).

Now consider $X_m = \{m \in \mathbb{N} | X_m$ contains an element of $S\}$. X_m is an infinite sequence, but it has a selection rule: choose the element of X_i which occurs in S (or the one that occurs first if they are both in S). Thus $\prod\limits_{m\in M} X_m$ is not empty. ↯ □

Remark. Cardinal subtraction often is not well-defined; however, subtraction by a finite cardinal is well-defined. If $a = |A|$ and n is a finite cardinal, then $a - n = |A - \{A_1, \ldots, A_n\}|$, where $A_1, \ldots, A_n$ are distinct elements of A. We can see that a is a Dedekind infinite cardinal if and only if $a = a - 1$.

Proposition 2. *Every Russell cardinal a has an immediate predecessor $a - 1$ and an immediate successor $a + 1$.*

Proof. Suppose a is the cardinality of a Russell set X. Then by Proposition 1, X is Dedekind finite, so $a < a + 1$. $a + 1$ must then be the direct successor of a, because it is a strictly larger cardinal obtained by the addition of a single element to X.
A similar proof to that of Proposition 1 shows that for a Russell Cardinal a, $a - 1 < a$, so we have that $a - 1$ is the direct predecessor of a. □

By definition, no infinite subsequence of a Russell sequence has a choice function. We now show that furthermore, if we take infinitely many pairs of socks from a Russell sequence, those pairs themselves will form a Russell sequence. The Russell cardinal, however, is strictly smaller.

Proposition 3. *If (X_n) is a Russell sequence and M is an infinite proper subset of $\mathbb{N}$, then $\bigcup\limits_{m\in M} X_m$ is a Russell set and $|\bigcup\limits_{m\in M} X_m| < |\bigcup\limits_{n\in\mathbb{N}} X_n|$.*

Proof. Since $\bigcup_{m\in M} X_m \subsetneq \bigcup_{n\in\mathbb{N}} X_n$, and both are Russell sets, it follows from the definition of Dedekind finite that $|\bigcup_{m\in M} X_m| < |\bigcup_{n\in\mathbb{N}} X_n|$. $\square$

Proposition 4. *If a is a Russell cardinal, then so are $a+2$ and $a-2$, and $a-2 < a < a+2$.*

Proof. Suppose $(X_n)_{n\in\mathbb{N}}$ is a Russell sequence, X is its Russell set, and $a = |X|$. Consider two elements not in X, u and v. Consider a sequence (Y_n) defined by

$$Y_i = \begin{cases} \{u, v\} & \text{if } i = 1 \\ X_{i-1} & \text{otherwise} \end{cases}.$$

Consider some infinite $M \subseteq \mathbb{N}$; there can be no choice function on $(Y_m)_{m\in M}$, for if there were, it would also be a choice function on $(X_{m-1})_{m\in M}$. Thus Y is a Russell set, and $|Y| = |X \uplus \{u, v\}| = |X| + 2 = a + 2$, so $a + 2$ is a Russell Cardinal.

Now consider $(Z_n)_{n\in\mathbb{N}}$, a sequence defined by $Z_n = X_{n+1}$. So $Z = X - X_1$ and $|Z| = |X| - 2 = a - 2$. For infinite $M \subseteq \mathbb{N}$, $(Z_m)_{m\in\mathbb{M}}$ has no choice function; if it did, then we would be able to construct a choice function on $(X_m)_{m\in M}$ by making at most one additional choice (for X_1). Thus Z is a Russell set and $a - 2$ is a Russell Cardinal.

Since $a < a + 1$ and $a + 1 < a + 2$ by Proposition 2 (with a and $a + 2$ being Russell cardinals), it follows by transitivity that $a < a + 2$. Since $a - 2$ is also a Russell cardinal, a similar argument shows that $a - 2 < a$. $\square$

We see then as we would hope, adding or removing a single pair of socks will still leave us with a Russell set, in particular a strictly larger or smaller Russell set. The next three propositions show what kind of arithmetic we can perform on Russell cardinals and the relationships they have with the original cardinal.

Proposition 5. *If a and b are Russell cardinals, then so is $a + b$, and there exists a family $(a_{(r,n)})_{(r,n)\in\mathbb{R}\times\mathbb{Z}}$ of Russell cardinals such that*

$$a < a_{(r,n)} < a + b \textit{ for each } (r, n) \in \mathbb{R} \times \mathbb{Z}$$

and

$$a_{(r,n)} < a_{(s,m)} \textit{ iff } r < s \textit{ or } (r = s \textit{ and } n < m).$$

This proposition is proven in [6, p. 4], but we will outline it here. We consider a Russell cardinal a and a Russell cardinal $b = |\bigcup_{n\in N} X_n|$, a bijection $\rho : \mathbb{N} \to \mathbb{Q}$, and a function $m : \mathbb{R} \to \mathbb{N}$, $m(r) = \min\{n \in \mathbb{N} \mid |r - \rho(n)| < \frac{1}{2}\}$.

Thus the family of Russell cardinals desired can be obtained by:

$$a_{(r,n)} = a + \begin{cases} |\bigcup\limits_{\rho(k)<r} X_k \cup \bigcup\limits_{k=1}^{n} X_{m(r+k)}| & \text{if } n \geq 0 \\ |\bigcup\limits_{\rho(k)<r} X_k \setminus \bigcup\limits_{k=1}^{n} X_{m(r-k)}| & \text{if } n < 0 \end{cases}.$$

Proposition 6. *If a is a Russell cardinal, then so is $2a$, and $a < 2a$.*

Proof. Suppose $(X_n)_{n\in\mathbb{N}}$ is a Russell sequence, X is its Russell set, and $a = |X|$. Let $Y = \{0,1\} \times X$ and observe that $|Y| = 2a$. Define now sequence $(Y_n)_{n\in\mathbb{N}}$ such that

$$Y_n = \begin{cases} 0 \times X_{\frac{n}{2}} & \text{if } n \text{ is even} \\ 1 \times X_{\frac{n+1}{2}} & \text{if } n \text{ is odd} \end{cases}.$$

Then $Y = \bigcup\limits_{n\in\mathbb{N}} Y_n$, and the sequence is a Russell sequence – if there is a choice function on an infinite subsequence, then there is either a choice function on infinitely many even sets in the sequence or infinitely many odd sets, and either way this is an infinite choice function on X. Thus $2a$ is a Russell Cardinal.

Since $a < a+1$ by Proposition 3 and $a+1 \leq 2a$, by transitivity $a < 2a$. $\square$

Proposition 7. *If a is a Russell cardinal, then so are all cardinals $n \cdot a$ with $n \in \mathbb{N}^+$, and*

$$a < a+1 < a+2 < \ldots < a+n < \ldots < 2a < 3a < \ldots < \aleph_0 \cdot a \leq 2^a.$$

Proof. Consider a Russell cardinal a. For all $n \in \mathbb{N}$, $n < a$, so for every $n \in \mathbb{N}$ we have $a + n < a + a = 2a$. By Proposition 3, for any Russell Cardinal a, $a < a+1 < a+2$ and $a+2$ is also a Russell Cardinal. Then we have that $a < a+2 < a+4 < \ldots$ is a sequence of Russell cardinals, and by adding one to each, we see $a < a+1 < a+2 < \ldots < a+n < \ldots < 2a$.

Consider $n \in \mathbb{N}$. By a similar argument to the proof of Proposition 6, $n \cdot a$ is a Russell cardinal. $n\cdot a < n\cdot a+1$ by Proposition 3 and $n\cdot a+1 \leq n\cdot a+n = (n+1)a$, so $n \cdot a < (n+1)a$. Thus we have $a < 2a < 3a < \ldots$, a sequence of Russell cardinals. We can also see that $n\cdot a < \aleph_0\cdot a$, since $n\cdot a < (n+1)a \leq \aleph_0\cdot a$.

It remains to show that $\aleph_0\cdot a \leq 2^a$. We will prove this by showing an injection from $\mathbb{N} \times X$ (where X is a Russell set $\bigcup\limits_{n\in\mathbb{N}} X_n$) into $\{f | f : X \to \{0,1\}\}$. For $n \in \mathbb{N}$ and $x \in X_i$, let $j(n,x) = f_{n,x}$, where (for $y \in X_j \in X$):

$$f_{n,x}(y) = \begin{cases} 1 & \text{if } x = y \text{ or } (i \neq j \text{ and } j \leq n+i) \\ 0 & \text{if } (x \neq y \text{ and } i = j) \text{ or } (i \neq j \text{ and } n+i < j). \end{cases}$$

In less technical terms, j maps a natural number n and an element x of the Russell set to a function which takes an element y of the Russell set. If y is in the same set as x, the function returns 1 if and only if $y = x$. If the elements are in different sets, the function returns 1 if and only if the index of y's pair is greater than $n + i$. We can see that j maps each pair (n, x) to a distinct function, so it is injective. □

Though we saw before that no Russell set is countable or even comparable with $\aleph_0$, the next two propositions show that Russell cardinals are in fact incomparable with any of the infinite cardinals from ZFC.

Proposition 8. *No Russell cardinal is comparable with any $\aleph$.*

Proof. Suppose a is a Russell cardinal for a Russell set $X = \bigcup_{n \in \mathbb{N}} X_n$ and α is an ordinal.

Assume for sake of contradiction that $a \leq \aleph_\alpha$. Then there is an injection from X into some infinite ordinal β. Since β would be well-orderable, we can well-order X by each element's image in this injection, and thus there would be a choice function on X, a contradiction.

If $\aleph_\alpha \leq a$, then by transitivity $\aleph_0 \leq a$, contradicting that X is Dedekind finite.

Thus a is incomparable with any $\aleph$ cardinal. □

Proposition 9. *No Russell cardinal is comparable with a cardinal of the form* $2^{\aleph}$

This proposition is proven in [6, p. 6].

We now discuss even and odd cardinals, to ultimately show in Proposition 13 that if we have two infinite cardinals, both less than or equal to a Russell cardinal and one the immediate successor of the other, then exactly one is a Russell cardinal. Since Proposition 4 showed that Russell cardinals alternate with non-Russell cardinals when adding and subtracting 1, this result confirms an intuition that in a sense "every other (infinite) cardinal" less than or equal to a Russell cardinal is a Russell cardinal.

Definition 32 (Even and Odd Cardinals)**.** A cardinal a is even if for some cardinal b, $a = 2b$. A cardinal a is odd if for some cardinal b, $a = 2b + 1$.

Definition 33 (Almost Even)**.** [6, p. 6] A cardinal $a = |A|$ is almost even if it satisfies the following equivalent conditions:

- There exists a fixpoint-free map $\sigma : A \to A$ with $\sigma^2 = \mathrm{id}_A$
- A can be expressed as the disjoint union of a family of 2-element sets

Clearly, all Russell cardinals and even cardinals are almost even.

Proposition 10. *If a and $a + 1$ are both almost even, then a is Dedekind infinite.*

Proof. Let X be a set such that $a = |X|$, and suppose that a and $a+1$ are almost even. Say $Y = X \uplus \{0\}$. Then there exists a fixpoint-free map $\sigma : X \to X$ with $\sigma^2 = \mathrm{id}_X$, and there exists a fixpoint-free map $\tau : Y \to Y$ with $\tau^2 = \mathrm{id}_Y$. Observe that both functions are bijective, which follows from $\sigma^2 = \mathrm{id}$. Recursively define a sequence by $x_0 = \tau(0)$, $x_{n+1} = \tau(\sigma(x_n))$.

We claim all elements of this sequence are distinct. Suppose for sake of contradiction that $x_i = x_j$ for $i < j$. Then:

$(\tau \circ \sigma)^i(x_0) = (\tau \circ \sigma)^j(x_0)$

$(\sigma \circ \tau)^i(\tau \circ \sigma)^i(x_0) = (\sigma \circ \tau)^i(\tau \circ \sigma)^j(x_0)$

$x_0 = (\tau \circ \sigma)^{j-i}(x_0)$, by repeatedly applying the identity rules for σ and τ

$\tau(0) = (\tau \circ \sigma)^{j-i}(x_0)$, by definition

$\tau(\tau(0)) = \tau((\tau \circ \sigma)^{j-i}(x_0))$

$0 = \tau((\tau \circ \sigma)^{j-i}(x_0))$, by the identity rule for τ

$0 = \tau((\tau \circ \sigma)((\tau \circ \sigma)^{j-i-1}(x_0)))$, by separating out one $\tau \circ \sigma$

$0 = \sigma((\tau \circ \sigma)^{j-i-1}(x_0))$, by the identity rule for τ.

So σ maps some element to 0, contradicting that $x_i = x_j$. Thus we have an infinite countable sequence of distinct elements in X, so $\aleph_0 \leq X$ and thus X is Dedekind infinite, contradicting Proposition 1. Therefore, a and $a+1$ cannot both be almost even. $\square$

Proposition 11. *No Russell cardinal is odd.*

Proof. If a Russell cardinal a were odd, there would be some cardinal b such that $a = 2b+1$. Since even cardinals and Russell cardinals are almost even, then we have that $2b+1$ and $2b+2$ (which is just $2(b+1)$) are almost even. So by Proposition 10, $a = 2b+1$ is Dedekind infinite and thus not a Russell cardinal. ↯So no Russell cardinal is odd. $\square$

Proposition 12. *If a is a Russell cardinal, then $a+1$ is not.*

Proof. For a Russell cardinal a, suppose for sake of contradiction that $a+1$ is a Russell cardinal. Then a and $a+1$ are almost even and $a+1$ is Dedekind infinite by Proposition 10. ↯ $\square$

Proposition 13. *If b is an infinite cardinal less than or equal to a Russell cardinal, then exactly one of b and $b+1$ is a Russell cardinal.*

Proof. (Proof outlined in [6, p. 6])
Suppose $(X_n)_{n\in\mathbb{N}}$ is a Russell sequence, X is its Russell set, and $a = |X|$. Suppose $b \leq a$ and for some set Y', $b = |Y'|$. Then there exists an injection from $Y' \to X$; let Y be the image of that injection (so $Y \subseteq X$ and $|Y| = b$).

Consider $M = \{n \in \mathbb{N} | |X_n \cap Y| = 1\}$, the set of indices of sets in X which have exactly one element in Y, and let $m = |M|$. M is a finite set, since if it were infinite we could have a choice function (choose the element in Y) on $X' = \{X_n | n \in M\}$. So we then have m elements of X which are in Y but are paired with an element not in Y. Y thus consists of infinitely many X_n as well as m "unpaired" socks.

If m is even, we can pair all of these unpaired socks into disjoint two-element sets $S_1, \ldots, S_{\frac{m}{2}}$. Consider $Z = \{X_n | n \in M\} \cup S_1 \cup \ldots \cup S_{\frac{m}{2}}$. Then $\bigcup Z = Y$, $|Z| = b$, and Z has no choice functions on infinite subsets, for it it did, we would have a choice function on infinitely many X_n. Thus $\bigcup Z$ is a Russell set and b is a Russell cardinal.

If m is odd, add one element to the unpaired socks. Then we can pair the extended set of unpaired socks into disjoint two-element sets $S_1, \ldots, S_{\frac{m}{2}+1}$. By the same argument, $\bigcup Z \uplus \{0\}$ is a Russell set and $b + 1$ is a Russell cardinal. $\square$

Finally, we show an interesting result that any permutation of a Russell set "separates" only finitely many pairs. Herrlich and Tachtsis[6] use the analogy that if you took Russell's socks to a laundromat and randomly re-paired all of the socks, almost all of them would be paired with their original "partner".

Definition 34 (Separation)**.** [6, p. 10] If $(X_n)_{n\in\mathbb{N}}$ is a Russell sequence with $X = \bigcup\limits_{n\in\mathbb{N}} X_n$, then a map $f : X \to X$ is said to separate some X_n if it maps the two elements of X_n into two different sets.

Proposition 14. *If $(X_n)_{n\in\mathbb{N}}$ is a Russell sequence, then each permutation of $X = \bigcup\limits_{n\in\mathbb{N}} X_n$ separates only finitely many X_n.*

Proof. Suppose for sake of contradiction that $(X_n)_{n\in\mathbb{N}}$ is a Russell sequence, X is its Russell set, and φ is a permutation of X which separates infinitely many X_n. Let $M = \{m \in \mathbb{N} | \varphi \text{ separates } X_m\}$. Then we can define a choice function on $\{X_m | m \in M\}$, an infinite subset of a Russell sequence, by choosing the sock in each X_m which is mapped to a lower index set by φ. ↯ $\square$

10. Weak Russell Sets and Pseudo-Russell Sets

While we define Russell sequences to have no choice function on any infinite subsequence, we can define this more loosely in two ways. Recall the definitions of Weak Russell sequences and Pseudo-Russell sequences :

Definition 14 (Pseudo-Russell Sequence)**.** A sequence $(X_n)_{n\in\mathbb{N}}$, where the X_n are pairwise-disjoint sets of two ur-elements, is a Pseudo-Russell sequence if for some non-principal ultrafilter $\mathcal{F}$ on $\mathbb{N}$, for all $M \subseteq \mathbb{N}$:

$$M \in F \iff \prod_{m\in M} X_m \text{ is empty.}$$

Definition 15 (Weak Russell Sequence)**.** A sequence $(X_n)_{n\in\mathbb{N}}$, where the X_n are pairwise-disjoint sets of two ur-elements, is a Weak Russell sequence if $\prod\limits_{n\in\mathbb{N}} X_n = \emptyset$.

We define sets and cardinals as we did for Russell sequences. Revisiting the propositions we showed for Russell Sets, we can see how these types of sets differ.

Because the definition of Weak Russell set is very general, some results are interesting when we consider specifically those Weak Russell sets which are not Russell sets. Thus the following definition will be useful:

Definition 35 (Strictly-Weak Russell Set)**.** A Strictly-Weak Russell Set (cardinal, sequence) is a Weak Russell set (cardinal, sequence) which is not a Russell set (cardinal, sequence).

The following fact is obvious from the definitions, but it is useful to state explicitly as we consider the variations of Russell sets.

Fact 10. *Every Pseudo-Russell set is a Strictly-Weak Russell set, and every Strictly-Weak Russell Set is a Weak Russell set. Every Weak Russell set is either a Russell set or s Strictly-Weak Russell set. The same relationship holds for sequences and cardinals.*

Fact 10 implies that every proposition that holds for a Weak Russell set (or a Strictly-Weak Russell Set) will be true for a Pseudo-Russell set. However, we will provide proofs for both propositions in some cases if the explicit proof for the more specific set requires a different and interesting approch.

First, we see that Strictly-Weak Russell sets are not Dedekind sets. Because many of the proofs in the previous section depended on this fact, we can predict that many of the facts about Russell sets will not be true for Strictly-Weak Russell sets.

Proposition 15. *No Strictly-Weak Russell set is a Dedekind set.*

Proof. Consider $(X_n)_{n\in\mathbb{N}}$, a Strictly-Weak Russell Sequence; let $M \subseteq N$ be an infinite set such that there exists a choice function, f, on $\{X_m | m \in M\}$.

$(f(X_m))_{m\in M}$ is thus a countably infinite sequence of distinct elements of $X = \bigcup\limits_{n\in\mathbb{N}}$. Thus $\aleph_0 \leq X$, and X is Dedekind infinite. □

Proposition 16. *No Pseudo-Russell set is a Dedekind set.*

Proof. Note that this follows directly from Proposition 15, since every Pseudo-Russell set is a Strictly-Weak Russell set. We provide an independent proof below.

Let $(X_n)_{n\in\mathbb{N}}$ be a Pseudo-Russell sequence (so $X = \bigcup_{n\in\mathbb{N}}$ is a Pseudo-Russell set), on which the non-principal ultrafilter $\mathcal{F}$ defines which subsequences have choice functions. If we partition the natural numbers into even and odd numbers, by definition of ultrafilter exactly one of these parts is in $\mathcal{F}$. Let M be the partition not in $\mathcal{F}$; then there is a choice function in $\{X_m | m \in M\}$. Call this choice function f.

Consider $j : \mathbb{N} \to X$, which maps $n \to f(X_{2n})$ if M is the even numbers or $n \to f(X_{2n-1})$ if M is the odd numbers. j is an injection from the natural numbers to X, thus $\aleph_0 \leq |X|$, and X is by definition not Dedekind finite. Thus X is not a Dedekind Set. □

Proposition 17. *For a Strictly-Weak Russell cardinal a, $a-1$ is not a direct predecessor and $a+1$ is not a direct successor. In particular, $a-1 = a = a+1$.*

Proof. This follows directly from Proposition 15 and the definition of Dedekind infinite. □

Proposition 18. *For a Pseudo-Russell cardinal a, $a-1$ is not a direct predecessor and $a+1$ is not a direct successor. In particular, $a-1 = a = a+1$.*

Proof. This follows directly from Proposition 16 and the definition of Dedekind infinite. This also follows from Proposition 17, since every Pseudo-Russell set is a Strictly-Weak Russell set. □

Proposition 19. *If (X_n) is a Weak Russell sequence and M is an infinite proper subset of $\mathbb{N}$, then $\bigcup_{m\in M} X_m$ is a possibly, but not necessarily, a Weak Russell set. If not, $|\bigcup_{m\in M} X_m| < |\bigcup_{n\in\mathbb{N}} X_n|$. Otherwise, the cardinalities may or may not be equal.*

Proof. Suppose $X = \bigcup_{n\in\mathbb{N}} X_n$ is a Weak Russell set, and $a = |X|$. Suppose that $M \subseteq N$ is infinite. Since X is a Weak Russell set, there may be choice functions on some infinite subsequences of $(X_n)_{n\in\mathbb{N}}$, but in general we do not know which – if any – subsequences have choice functions. If $Y = \bigcup_{m\in M} X_m$ and $a = |X|$, then Y is a Weak Russell set if and only if there is not a choice function on $(X_m)_{m\in M}$.

If Y is not a Weak Russell set, then Y is countable. Thus $b = \aleph_0 < a$.

If Y is a Weak Russell set, then it could be the case that either $b = a$ or $b < a$. As an example, suppose X is a Russell set (which is more generally a Weak Russell set). Then by Proposition 3, $|Y| < |X|$ for any Russell set $Y \subsetneq X$.

But suppose we partition $\mathbb{N}$ into infinite sets A, B, C, and there is a choice function on $(X_a)_{a\in A}$ and on $(X_b)_{b\in B}$.
Then $|\bigcup_{n\in\mathbb{N}} X_n|$
$= |\bigcup_{a\in A} X_a| + |\bigcup_{b\in B} X_b| + |\bigcup_{c\in C} X_c|$
$= \aleph_0 + \aleph_0 + |\bigcup_{c\in C} X_c|$
$= \aleph_0 + |\bigcup_{c\in C} X_c|$
$= |\bigcup_{b\in B} X_b| + |\bigcup_{c\in C} X_c|$
$= |\bigcup_{i\in B\cup C} X_i|$.
So $\bigcup_{i\in B\cup C} X_i \subsetneq \bigcup_{n\in\mathbb{N}} X_n$, and these Weak Russell Sets have the same cardinality.

$\square$

Proposition 20. *If (X_n) is a Pseudo-Russell sequence (using the non-principal ultrafilter $\mathcal{F}$) and M is an infinite proper subset of $\mathbb{N}$, then $\bigcup_{m\in M} X_m$ is a Pseudo-Russell set $\iff M \in \mathcal{F} \iff |\bigcup_{m\in M} X_m| = |\bigcup_{n\in\mathbb{N}} X_n|$.*

Proof. First, we show that $M \in \mathcal{F} \iff |\bigcup_{n\in\mathbb{N}} X_n| = |\bigcup_{m\in M} X_m|$.

If $M \notin \mathcal{F}$, then $|\bigcup_{m\in M} X_m| = \aleph_0$. Thus $|\bigcup_{n\in\mathbb{N}} X_n| \neq |\bigcup_{m\in M} X_m|$.

If $M \in \mathcal{F}$, then there is no choice function on $(X_m)_{m\in M}$ by definition, and there is a choice function on $(X_i)_{i\in\mathbb{N}-M}$ since $\mathcal{F}$ is an ultrafilter. Then $\bigcup_{i\in\mathbb{N}-M} X_i$ is countable and has cardinality $\aleph_0$. Furthermore, if we partition M into two infinite sets A, B, then there will be a choice function on either (but not both) $(X_a)_{a\in A}$ or $(X_b)_{b\in B}$ – to see this, consider a partition of $\mathbb{N}$ into A and $(\mathbb{N} - M) \cup B$, keeping in mind that $\mathbb{N} - M$ is not in the filter. Say without loss of generality that there is a choice function on $(X_a)_{a\in A}$.
So $|\bigcup_{n\in\mathbb{N}} X_n| = |\bigcup_{m\in M} X_m| + |\bigcup_{i\in\mathbb{N}-M} X_i|$
$= |\bigcup_{a\in A} X_a| + |\bigcup_{b\in B} X_b| + |\bigcup_{i\in\mathbb{N}-M} X_i|$
$= \aleph_0 + |\bigcup_{b\in B} X_b| + \aleph_0$
$= \aleph_0 + |\bigcup_{b\in B} X_b|$
$= |\bigcup_{a\in A} X_a| + |\bigcup_{b\in B} X_b|$
$= |\bigcup_{m\in M} X_m|$, as desired.

Now, we show that $M \in \mathcal{F} \iff \bigcup_{m\in M} X_m$ is a Pseudo-Russell set.

If $M \in \mathcal{F}$, then by the previous solution, $|\bigcup_{m \in M} X_m| = |\bigcup_{n \in \mathbb{N}} X_n|$. These sets are then isomorphic, and thus $|\bigcup_{m \in M} X_m|$ is also a Pseudo-Russell set.

If $M \notin \mathcal{F}$, then by definition there exists a choice function on $(X_m)_{m \in M}$, so $\bigcup_{m \in M} X_m$ is not a Pseudo-Russell set. $\square$

Proposition 15 showed that adding a sock to a Strictly-Weak Russell set did not change its cardinality. The following lemma shows that we can add any finite number of socks, or even a countably infinite number of socks, and still not affect the cardinality. Removing a finite number of socks will also not affect the cardinality.

Lemma 6. *Suppose a is a Strictly-Weak Russell cardinal. Then for all $n \in \mathbb{N}$, $a - n = a = a + n$. Furthermore, $a = a + \aleph_0$.*

Proof. By Proposition 15, Strictly-Weak Russell cardinals are Dedekind infinite. By definition of Dedekind infinite, $a - n = a - n + 1$. Applying this property recursively, we have $a - n = a = a + n$.

Suppose $X = \bigcup_{n \in \mathbb{N}} X_n$ is a Strictly-Weak Russell set with cardinality a. By definition, there exists some infinite M such that $(X_m)_{m \in M}$ has a choice function. Let $Y = \bigcup_{m \in M} X_m$ and observe that $|Y| = \aleph_0$. Since $\aleph_0 = 2 \cdot \aleph_0$, we have that $a = |X| = |X - Y| + \aleph_0 = |X - Y| + 2 \cdot \aleph_0 = a + \aleph_0$. $\square$

We see now that adding or removing a pair of socks to a Strictly-Weak/Pseudo-Russell set will not affect the cardinality, but the new set will still be a Strictly-Weak/Pseudo-Russell set.

Proposition 21. *If a is a Weak Russell cardinal, then so are $a + 2$ and $a - 2$. However, $a - 2 = a = a + 2$ iff a is a Strictly-Weak Russell cardinal.*

Proof. Let X be a Weak Russell set with cardinality a. Because Strictly-Weak Russell sets are Dedekind infinite (Proposition 15), $\aleph_0 \leq a$, and we can partition X into $Y \cup Z$, where $|Y| = \aleph_0$. So then $X \uplus \{u, v\} = (Y \uplus \{u, v\}) \cup Z$, but since Y is a countable set, Y is isomorphic to $Y \uplus \{u, v\}$, and thus X is isomorphic to $X \uplus \{u, v\}$. So the latter must also be a Strictly-Weak Russell set, and thus $a + 2$ is a Strictly-Weak Russell cardinal. A similar argument shows that $a - 2$ is a Strictly-Weak Russell cardinal.

If a is a Strictly-Weak Russell cardinal, then by Lemma 6, $a - 2 = a = a + 2$. Else, a is a Russell cardinal, and $a - 2 < a < a + 2$ by Proposition 4. $\square$

Proposition 22. *If a is a Pseudo-Russell cardinal, then so are $a + 2$ and $a - 2$. However, $a - 2 = a = a + 2$.*

Proof. The result follows directly from Proposition 21. Note that a direct proof of this theorem would be identical to Proposition 21, since the proof only requires a to be Dedekind infinite. $\square$

Proposition 23. *If a is a Weak Russell cardinal, then so is $2a$. It is possible that either $a < 2a$ or $a = 2a$.*

Proof. Suppose $X = \bigcup_{n\in\mathbb{N}} X_n$ is a Weak Russell set, and $a = |X|$. Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence where

$$Y_n = \begin{cases} \{0\} \times X_{\frac{n+1}{2}} & \text{if } n \text{ is odd} \\ \{1\} \times X_{\frac{n}{2}} & \text{if } n \text{ is even} \end{cases}.$$

Let $Y = \bigcup_{n\in\mathbb{N}} Y_n$, and observe $|Y| = 2a$. If there is a choice function on $(Y_n)_{n\in\mathbb{N}}$, then there is a choice function on the subsequence with odd indices, which is the same sequence as $(\{0\} \times X_n)_{n\in\mathbb{N}}$, contradicting that a is a Weak Russell cardinal. So there is no such choice function, and $2a$ must be a Weak Russell cardinal.

Clearly $a \leq 2a$. Proposition 19 shows (and Proposition 24 will show) it is possible that $2a \neq a$. It is also possible that $2a = a$: consider a Weak Russell cardinal $b = |Y|$ and let $a = \aleph_0 \cdot b$. Then a is also a Weak Russell cardinal, since if there is no choice function on Y, there can be no choice function on $\mathbb{N} \times Y$. But then $2a = 2 \cdot (\aleph_0 \cdot b) = (2 \cdot \aleph_0) \cdot b = \aleph_0 \cdot b = a$. $\square$

Proposition 24. *If a is a Pseudo-Russell cardinal, then $2a$ is not, and $a < 2a$.*

Proof. Suppose a is a Pseudo-Russell cardinal and, for sake of contradiction, $2a$ is a Pseudo-Russell cardinal. Then X and $X \times \{0, 1\}$ are both Pseudo-Russell sets. By definition then, there must be a choice function on either $\{0\} \times \{X_n | n \in \mathbb{N}\}$ or $\{1\} \times \{X_n | n \in \mathbb{N}\}$, but this would induce a choice function on all of $\{X_n | n \in \mathbb{N}\}$. ↯

Clearly $a \leq 2a$, and since a is a Pseudo-Russell cardinal and $2a$ is not, $a \neq 2a$. Thus $a < 2a$. $\square$

Because Strictly-Weak Russell cardinals are Dedekind infinite, we know that they are comparable to $\aleph_0$ unlike Russell cardinals. However, we show that they are not comparable to any other $\aleph$ cardinal.

Proposition 25. *Suppose a is a Strictly-Weak Russell cardinal. Then a is comparable with $\aleph_\alpha$ if and only if $\alpha = 0$.*

Proof. Suppose X is a Strictly-Weak Russell set with $X = \bigcup_{n\in\mathbb{N}} X_n$ and $a = |X|$. Because X is Dedekind infinite, $\aleph_0 \leq a$.

If it were the case that $a \leq \aleph_\alpha$ for $\alpha \geq 0$, there would be an injection from X into some infinite ordinal β. Since β would be well-orderable, we can well-order X (by ordering the elements in the image of the injection and taking their inverses), and thus there would be a choice function on X. Thus for all $\alpha > 0$, $a \not\leq \aleph_\alpha$.

If it were the case that $\aleph_\alpha \leq a$ for some $\alpha \geq 0$, then there would exist an injection $f : \beta \to X$ for some ordinal β with cardinality $\aleph_\alpha$. Then we could construct an injection $g : \beta \to \aleph_0 \times \{0, 1\}$ by:

$$g(b) = \begin{cases} (n, 0) & \text{if } f(b) \in X_n \text{ and } \forall c < b : f(c) \notin X_n \\ (n, 1) & \text{if } f(b) \in X_n \text{ and } \exists c < b : f(c) \in X_n \end{cases}.$$

Thus we would have a contradiction, $\aleph_\alpha \leq \aleph_0 \times 2 = \aleph_0$.

Thus $\aleph_\alpha$ is comparable with a if and only if $\alpha = 0$. □

Proposition 26. *Suppose a is a Pseudo-Russell cardinal. Then a is comparable with $\aleph_\alpha$ if and only if $\alpha = 0$.*

Proof. The proof for Proposition 25 suffices. □

Proposition 14 states that no permutation on any Russell set will separate infinitely many sets. This is not true for Weak/Pseudo-Russell sets. In fact, we will prove that every Strictly-Weak Russell set and Pseudo-Russell set has such a permutation.

Proposition 27. *There exists a permutation on every Strictly-Weak Russell set which separates infinitely many sets.*

Proof. Let $X = \bigcup_{n \in \mathbb{N}}$ be a Weak Russell set. Let M be an infinite subset of $\mathbb{N}$ such that $(X_m)_{m \in M}$ has a choice function, f. Let ψ be some derangement of M, and consider the following permutation on X:

$$\varphi(x \in X_n) = \begin{cases} x & \text{if } n \notin M \\ x & \text{if } n \in M \text{ and } f(X_n) \neq x \\ f(X_{\psi(n)}) & \text{if } n \in M \text{ and } f(X_n) = x \end{cases}$$

This permutation separates every X_m for $m \in M$ by deranging the socks in the image of the choice function – since they are already in the image of the choice function, φ does not induce any choice functions not already present on X. □

Proposition 28. *There exists a permutation on every strictly Pseudo-Russell set which separates infinitely many sets.*

Proof. Let $X = \bigcup_{n\in\mathbb{N}}$ be a Pseudo-Russell set. By definition, there exist infinite $M \subseteq \mathbb{N}$ such that $(X_m)_{m\in M}$ has a choice function. Using this M, a similar proof to the proof of Proposition 27 suffices. □

11. Russell Cardinal Theorems

The following theorems are posed as open questions at the end of Herrlich and Tachtsis's paper on Russell cardinals [6, p. 11]. We prove them in the affirmative here.

Theorem 4. *If a is a Russell cardinal, $a^2 < 2^a$.*

A more generalized version of Theorem 4 was proven by Herrlich, Howard, and Tachtsis[7] (in particular, Lemma 1 and Corollary 1). Thomas [8, p. 160] proves that for every cardinal $\alpha \geq 5$, $2^\alpha \nleq \alpha^2$. The proof below, however, is an original and direct proof of the theorem.

Proof. Let $X = \bigcup_{n\in\mathbb{N}} X_n$ be a Russell Set such that $a = |X|$. Note that by definition, $a^2 = |X \times X|$ and $2^a = |\mathcal{P}(X)| = |\{f | f : X \to \{0,1\}\}|$.

First we show that $a^2 \leq 2^a$ by showing that there exists an injection $j : X \times X \to \mathcal{P}(X)$. j will be defined without appeal to the Axiom of Choice, since it is defined from the Russell Set X. We will code into $j(x, y)$ both the elements x, y and the ordering of the pair (x, y). We consider the following cases for elements of $X \times X$:

(1) (x, x)
(2) (x, y), where $\{x, y\} = X_n$
(3) (x, y), where $x \in X_n, y \in X_m, n < m$
(4) (x, y), where $x \in X_n, y \in X_m, m < n$

We will define j as follows, by cases:

(1) $j(x, x) = \{x\}$
(2) $j(x, y) = \{x\} \cup X_{n+1}$
(3) $j(x, y) = \{x, y\} \cup X_{m+1}$
(4) $j(x, y) = \{x, y\} \cup X_{n+1} \cup X_{n+2}$

Since none of these cases overlap, j is a well-defined function. To show that j is injective, it will suffice to show that every element in its image has a unique inverse. For each element of the image of j, we can tell immediately which case we are in by the size of the set; consider $S \in Im(j) \subset \mathcal{P}(X)$, and the following cases which are disjoint and completely describe elements of $Im(j)$:

(1) ($|S|$=1) $S = \{x\}$, and clearly by construction of j, $j^{-1}(S) = (x, x)$.

(2) ($|S|$=3) $S = \{u, v, x\}$. By construction, two of these elements must be in the same set X_n (say without loss of generality that these are u and v). Then if $x \in X_m = \{x, y\}$, $j^{-1}(S) = (x, y)$.
(3) ($|S|$=4) $S = \{u, v, x, y\}$, where (without loss of generality) $u, v \in X_k$, $x \in X_n$, $y \in X_m$, and $n < m$. We must have then that $k = m + 1$ by construction of j, so we can identify y as the second element of the pair. Thus $j^{-1}(S) = (x, y)$.
(4) ($|S|$=6) $S = \{u, v, f, g, x, y\}$, where (without loss of generality) $u, v \in X_k$, $f, g \in X_{k+1}$, $x \in X_m$, $y \in X_n$, and $m < n$. We must have that $k = n + 1$, so we can identify x as the first element of the pair. Thus $j^{-1}(S) = (x, y)$.

The rules we defined for j allow us to determine a unique inverse for every element of the image of j, so j is injective. Therefore $a^2 \leq 2^a$.

Now, to prove $a^2 < 2^a$, we must show that $a^2 \neq 2^a$. Assume otherwise; in particular, assume $2^a \leq a^2$. Then there exists an injection $j : F \to X \times X$, where $F = \{f | f : X \to \{0, 1\}\}$. For $n \in \mathbb{N}$, define $f_n \in F$ to be the function such that

$$f_n(x) = \begin{cases} 1 & \text{if } x \in X_n \\ 0 & \text{otherwise} \end{cases}.$$

So there exists a countable sequence of distinct elements of F: $f_1, f_2, \ldots$

Now consider the countable sequence $j(f_1), j(f_2), \ldots$; these are all distinct elements of $X \times X$ since j is injective. So we have a countable sequence of distinct ordered pairs $(a_1, b_1), (a_2, b_2), \ldots$. If the set $\{a_i | i \in \mathbb{N}\}$ is infinite, then we have found a countable subset of X. If not, then for some $k \in \mathbb{N}$, there must be infinitely many elements of the sequence such that the first of the pair equals a_k; in this case, $\{b_i | a_i = a_k\}$ must be infinite as well, and we have found a countable subset of X.

Therefore, X contains a countable subset. So $\aleph_0 \leq |X|$, X is Dedekind infinite, and thus X is not a Russell Set $\lightning$.

Therefore, $a^2 < 2^a$. $\square$

Theorem 5. *Let B be the theory that the class of Russell Cardinals is bounded from above, and let R be the theory that there is a Russell set of atoms (urelements). If ZFC$^\circ$ is consistent, then so is ZF$^\circ$+R+B.*

Herrlich and Tachtsis present this as an open problem at the end of [6]. It has been shown previously that it is consistent that the class of Dedekind cardinals is bounded from above by Blass[9] and Herrlich, Howard, and Tachtsis[10]. We provide below an original proof explicitly for Russell cardinals.

Proof. [1] We will define a model, $L(A)[f]$, by a construction hierarchy based on the atoms in a Russell set, and show that in that model there exists a Russell cardinal and the class of Russell cardinals is bounded. We will show this by defining some set and creating an injection from an arbitrary Russell Set into that set. For each element of the Russell Set, we will consider it as an element of the constructible universe relative to the set of atoms A, and we will assign to it a finite sequence of atoms and an ungrounded construction tree.

Let $(A_n)_{n\in\mathbb{N}}$ be a Russell sequence such that the Russell set $\bigcup\limits_{n\in\mathbb{N}} A_n$ us a set of atoms. We will define the following construction hierarchy in accordance with the axioms of the constructible universe.[11, p. 128]

First, we include in our language $\in$ and a function f (which for our given Russell Set, $f(a)$ returns $n \in \mathbb{N}$ such that $a \in A_n$). Now we define the steps of construction:

- $L_0(A)[f] = A$.
- $L_{\alpha+1}(A)[f] = A \cup \{x | x = \{y \in L_\alpha | L_\alpha \models \varphi(y, p_1, \ldots, p_n\}\}$, where $p_1, \ldots, p_n \in L_\alpha(A)$ are parameters and $\varphi(y, x_1, \ldots, x_n)$ is a well-formed formula (here $x_1, \ldots, x_n$ are variables for the formula φ).
- $L_\lambda(A)[f] = \bigcup\limits_{\alpha<\lambda} L_\alpha(A)$, where λ denotes a limit cardinal.
- $L(A)[f] = \bigcup\limits_{\alpha} L_\alpha(A)$.

Note that $L(A)[f]$ is a model of ZF$^\circ$ and it contains the Russell sequence $(A_n)_{n\in\mathbb{N}}$.

Now for all $x \in L(A)[f]$, we define a *construction tree of* x as follows (nodes will be triples containing a natural number, an ordinal, and either an atom or a formula along with variables). Note that a construction tree uniquely determines a set/atom, but a set/atom could have many different construction trees.

- If $x \in A$, then the construction tree of x is the node $(0, 0, x)$.
- If $x \in \{y \in L_\alpha | L_\alpha \models \varphi(y)\}$, then a construction tree of x is the node $(1, \alpha + 1, \varphi(y))$, for the smallest such α.
- If $x \in \{y \in L_\alpha | L_\alpha \models \varphi(y, p_1, \ldots, p_n)\}$, where φ takes n parameters $p_1, \ldots, p_n$, then a construction tree of x is a tree with root node $(2, \alpha + 1, \varphi(y, x_1, \ldots, x_n))$, for the smallest such α and n children: the construction trees for $p_1, \ldots, p_n$, in this order.

Claim. *Construction trees as defined above are finite.*

The second value of the labeling-triple of each node (the smallest step at which the element is constructed) strictly decreases at each level, the node is

[1] Credit for this proof belongs jointly to Marcia Groszek and Ethan Thomas.

terminal when this value is zero, and each node has only finitely many children if any. This is proved by a simple induction argument, which we will omit.

Now, we define an *ungrounded construction tree* as a tree resulting from the removal of atoms from a construction tree, i.e., replacing every node $(0, 0, x)$ with $(0, 0, 0)$. This transformation causes the resulting tree to correspond only to a method of constructing an element, and not the atoms used to do so.

We now aim to well-order the collection of all ungrounded construction trees, to show that all such methods can be well-ordered. We provide the following method (by induction on the second entry of the root node) for comparing two trees to show that the collection has a total ordering:

Consider T_1 and T_2, ungrounded construction trees for x_1 and x_2, respectively, with root nodes (n_1, k_1, ψ_1) and (n_2, k_2, ψ_2), respectively (where ψ_1 is a placeholder for either 0, $\varphi(y_1)$, or $\varphi(y_1, x_1^1, \ldots, x_n^1)$ and similarly for ψ_2).

(1) If $n_1 < n_2$, then $T_1 < T_2$.
(2) If $n_1 = n_2$ and $k_1 < k_2$, then $T_1 < T_2$.
(3) If $n_1 = n_2$ and $k_1 = k_2$, then if $\psi_1 < \psi_2$ (recall that well-formed formulas are well-ordered), then $T_1 < T_2$.
(4) If $n_1 = n_2$, $k_1 = k_2 = 2$, and $\psi_1 = \psi_2$, then ψ_1 and ψ_2 both have n subtrees corresponding to parameters. Denote these sequences of subtrees $x_1^1, \ldots, x_n^1$ and $x_1^2, \ldots, x_n^2$ for x_1 and x_2, respectively. Let i be the smallest index such that $x_i^1 \neq x_i^2$. If $x_i^1 < x_i^2$, then $T_1 < T_2$.
(5) As desired, if none of the steps above produce a result for the ordering of T_1 and T_2, then $T_1 = T_2$; the trees have the exact same root node and thus are identical by construction.

Consider a descending $<$ chain of ungrounded construction trees under this total-order. Since root nodes consist of a natural number, and ordinal, and formulas with finitely many parameters (which are themselves ungrounded construction trees), all of which are decreasing and bounded from below, the chain cannot be infinite. Otherwise, there would be an infinite descending sequence of natural numbers, ordinals, or formulas.

Thus the collection of ungrounded construction trees is well-ordered.

Suppose $X \in L(A)[f]$ is a Russell set and $x \in X$. Say (l, κ, φ) is the root node for the least ungrounded construction tree T of x, and A_m is the set of finite sequences of atoms that, when placed in the tree T, give a construction tree for x; we place atoms from a finite sequence into an ungrounded construction tree by iterating through the tree and replacing the i^{th} node of the form $(0, 0, 0)$ with $(0, 0, a_i)$, a_i being the i^{th} atom of the sequence. We define the order of iterating through a tree recursively as follows: if tree T has root node x and immediate subtrees $x_1, \ldots, x_n$, then compare nodes in the same subtree using the ordering on that subtree, nodes in different subtrees by placing nodes in x_i before nodes in x_j for $i < j$, and making the root node the largest element.

Define $j : X \to \mathbb{N} \times \{0, 1, 2\} \times A^{<\omega}$ (where $A^{<\omega}$ is the set of all finite sequences of elements of A) by $j(x) = (n, i, A_m)$, where if $X_n = \{x, \bar{x}\}$ is the set in X containing x and $\bar{T}$ is the least ungrounded construction tree for $\bar{x}$, then

$$i = \begin{cases} 0 & \text{if } T \leq \bar{T} \\ 1 & \text{if } T > \bar{T} \end{cases};$$

define j similarly for all other elements of X.

Suppose that for x and y in a Russell set X, $j(x) = j(y) = (n, i, A_m)$ with j defined as above. Then $x, y \in X_n$ and $T = \bar{T}$ (since we have $T \leq \bar{T}$ and $\bar{T} \leq T$). So x and y are in the same set and have the same least ungrounded construction tree. So we are placing the same set of finite sequences of atoms, A_m, into the same ungrounded construction tree. Since this creates the same construction tree, and a construction tree uniquely defines an element, we must have that $x = y$. Thus j is injective.

So, an arbitrary Russell set can be injected into $\mathbb{N} \times \{0, 1\} \times A^{<\omega}$, so the class of Russell Cardinals is bounded from above by $\aleph_0 \cdot 2 \cdot 2^{|A|}$. □

Finally, we will show a stronger version of Theorem 4. This proof uses the same approach, but the number of cases to consider is much higher, so we use a slightly different technique in "coding" the products of elements from a Russell set.

Theorem 6. *If a is a Russell cardinal, $a^3 < 2^a$.*

Proof. Suppose for this proof that $(X_n)_{n\in\mathbb{N}}$ is a Russell sequence, that $X = \bigcup_{n\in\mathbb{N}} X_n$ is a Russell set, and $a = |X|$.

First, we show that $a^3 \leq 2^a$. We prove this in a similar manner to the proof of Theorem 4 – by constructing an injection to code the elements of a Russell set – however we have many more cases to consider.

Define a function $I : X \to \mathbb{N}$ to identify the index of the set to which a particular element in X belongs, so if $x \in X_n$, then $I(x) = n$. As we show below, we will have 22 cases for the elements of $X \times X \times X$, so it will be useful to have a way to code which case we are in as well as what elements we need. This function will code the case number and a nautral number (which will ensure that this case-coding set does not intersect with any other coded elements) into an element of $\mathcal{P}(X)$. To do this, we will define $C : \{1, 2, \ldots, 22\} \times \mathbb{N} \to \mathcal{P}(X)$ by $C(c, m) = X_{m+1} \cup X_{m+2} \cup \ldots \cup X_{m+c}$.

Now, we construct an injection $j : (X \times X \times X) \to \mathcal{P}(X)$ by

$$j(u, v, w) = Q(u, v, w) \cup C(c, \max\{I(u), I(v), I(w)\}),$$

where c and $Q(u, v, w)$ are dependent on case as follows:

Case number (c)	Conditions	$Q(u,v,w)$
1	$u = v = w$	$\{u\}$
2	$u = v, I(u) < I(w)$	$\{u, w\}$
3	$u = v, I(u) > I(w)$	$\{u, w\}$
4	$u = v, I(u) = I(w)$	$\{u\}$
5	$v = w, I(u) < I(v)$	$\{u, v\}$
6	$v = w, I(u) > I(v)$	$\{u, v\}$
7	$v = w, I(u) = I(v)$	$\{u\}$
8	$u = w, I(u) < I(v)$	$\{u, v\}$
9	$u = w, I(u) > I(v)$	$\{u, v\}$
10	$u = w, I(u) = I(v)$	$\{u\}$
For cases 11-22, u, v, w are distinct		
11	$I(u) < I(v) < I(w)$	$\{u, v, w\}$
12	$I(u) < I(w) < I(v)$	$\{u, v, w\}$
13	$I(v) < I(u) < I(w)$	$\{u, v, w\}$
14	$I(v) < I(w) < I(u)$	$\{u, v, w\}$
15	$I(w) < I(u) < I(v)$	$\{u, v, w\}$
16	$I(w) < I(v) < I(u)$	$\{u, v, w\}$
17	$I(u) = I(v) < I(w)$	$\{u, w\}$
18	$I(u) = I(w) < I(v)$	$\{u, v\}$
19	$I(v) = I(w) < I(u)$	$\{u, v\}$
20	$I(u) < I(v) = I(w)$	$\{u, v\}$
21	$I(v) < I(u) = I(w)$	$\{u, v\}$
22	$I(w) < I(u) = I(v)$	$\{u, w\}$

No case above overlaps, so j is a well-defined function. Finally, it suffices to show that given any element in $\mathrm{Im}(j)$, we can uniquely determine its inverse. Consider an element of the image, $Z \in \mathcal{P}(X)$. Consider the greatest $c \in \{1, \ldots, 22\}$ such that for some $m \in \mathbb{N}$, $X_{m+1} \cup \ldots \cup X_{m+c} \subset Z$. This c is then our case number, and we will have 1, 2, or 3 elements in Z whose "partners" are not in Z (i.e., $u \in Z$, $u \in X_n$ and $X_n \not\subset Z$). Then we can use the table above to determine how to construct an element in $X \times X \times X$.

For example, suppose we have $Z = X_5 \cup X_6 \cup X_7 \cup \{j, k\}$, where $j \in X_J, k \in X_K, J < K$. Then we see that we are in case 3, so we let the element in the higher-indexed set (k) take the place of u and v, and we let the other element (j) take the place of w. So $j^{-1}(Z) = (k, k, j)$. As another example, let $Z = X_7 \cup \ldots \cup X_{25} \cup \{j, k\}$, where $j \in X_J, k \in X_K, J < K$. We are then in case 19, so we let k take the place of u, j take the place of v, and j's "partner" (the other element in X_J; call it l) take the place of w. So $j^{-1}(Z) = (k, j, l)$.

There is thus a unique inverse for every element in the image of j, and therefore j is an injection and $a^3 \leq 2^a$.

Now, we show $2^a \not\leq a^3$. Assume for sake of contradiction that there exists an injection $j : F \to (X \times X \times X)$, where $F = \{f | f : X \to \{0, 1\}\}$. For all $n \in \mathbb{N}$, define $f_n \in F$ by

$$f_n(x) = \begin{cases} 1 & \text{if } x \in X_n \\ 0 & \text{otherwise} \end{cases}.$$

So there exists a countable sequence of distinct elements of F: $f_1, f_2, \ldots$

Now consider the countable sequence $j(f_1), j(f_2), \ldots$; these are all distinct elements of $X \times X \times X$ since j is injective. So we have a countable sequence of distinct ordered triples $(a_1, b_1, c_1), (a_2, b_2, c_2), \ldots$. If any of the sets $\{a_i | i \in \mathbb{N}\}$, $\{b_i | i \in \mathbb{N}\}$, or $\{c_i | i \in \mathbb{N}\}$ are infinite, then we have found a countable subset of X. But this must be the case, for we can only have finitely many triples if we can choose from only finitely many elements for each coordinate. Therefore, X contains a countable subset. So $\aleph_0 \leq |X|$, X is Dedekind infinite, and thus X is not a Russell Set ↯.

Therefore, $a^3 < 2^a$. □

Can we generalize this result to show that for all n, $a^n < 2^a$? Neither of the methods used for either direction of Theorem 6 make particular use of the fact that $n = 3$. Below we show that these methods can in fact be generalized.

Theorem 7. *If a is a Russell cardinal, then for all $n \in \mathbb{N}$, $a^n < 2^a$.*

Proof. Fix some $n \in \mathbb{N}$. Suppose for this proof that $(X_k)_{k \in \mathbb{N}}$ is a Russell sequence, that $X = \bigcup_{k \in \mathbb{N}} X_k$ is a Russell set, and $a = |X|$.

First, we show that $a^n \leq 2^a$. We will essentially generalize the method used in the proof of Theorem 6.

Define a function $I : X \to \mathbb{N}$ to identify the index of the set to which a particular element in X belongs, so if $x \in X_m$, then $I(x) = m$. Define also a function $P : X \to X$ to identify the "partner" of an element in X; so if $x \in X_m$ and $X_m = \{x, \bar{x}\}$, then $P(x) = \bar{x}$.

We will now define "patterns" among the elements of X^n to be sets which partition the elements of X^n based on which coordinates of the n-tuple are equal, which coordinates belong to the same set, and the order of the indices of the sets to which the coordinates belong. We explicitly identified these patterns in the proofs of Theorem 4 (4 patterns) and Theorem 6 (22 patterns). There must be only finitely many distinct patterns, because the criteria on which we define patterns depends on comparing n items for equality and order of the set to which they belong. Suppose that there are p distinct patterns on the elements of X^n.

Define $C : \{1, \ldots, p\} \times \mathbb{N} \to \mathcal{P}(X)$ by $C(c, i) = X_{i+1} \cup \ldots \cup X_{i+c}$. Define $Q : X^n \to \mathcal{P}(X)$ by

$$a_i \in Q(a_1, \ldots, a_n) \iff a_i \in \{a_1, \ldots, a_n\} \wedge \forall j < i(a_j \neq p(a_i)).$$

Then, $Q(a_1, \ldots, a_n)$ includes all coordinates of the element in X^n except those whose "partner" appears as an earlier coordinate – this avoids including a complete set X_m where we cannot differentiate the elements, but the pattern lets us know that the unlisted element still belongs as a coordinate (see case 17 in the proof of Theorem 6 for an example of this).

Now, define $j : X^n \to \mathcal{P}(X)$ by

$$j(a_1, \ldots, a_n) = C(c, \max\{I(a_1), \ldots, I(a_n)\}) \cup Q(a_1, \ldots, a_n),$$

where c is the number identifying the pattern to which $(a_1, \ldots, a_n)$ belongs. By construction, this function will be injective – if $j(x) = j(y)$, it must be the case that x and y have the same pattern and the same constructing elements, so $x = y$. Therefore j is an injection and $a^n \leq 2^a$.

It remains to show $2^a \not\leq a^n$. This is proven in a nearly identical manner to for the $n = 3$ case, except that we observe that at least one of the n sequences generated by taking the i^{th} coordinate of each n-tuple must be infinite, and thus the Russell set is Dedekind infinite, deriving a contradiction.

Therefore, for all $n \in \mathbb{N}$ and for every Russell cardinal a, $a^n < 2^a$. $\square$

References

[1] Gregory H. Moore. *Zermelo's Axiom of Choice: Its Origins, Development & Influence*. Springer-Verlag, 1982.

[2] Eric Schechter. *Handbook of Analysis and its Foundations*. Elsevier Science and Technology Books, 1st edition, 1996.

[3] Thomas J. Jech. *Set Theory*. Springer, 3rd edition, 2002.

[4] Karel Hrbacek and Thomas Jech. *Introduction to Set Theory, Revised and Expanded*. Crc Press, 1999.

[5] Tom Kern. Equivalents of failures of the axiom of choice. Undergraduate honors thesis, Dartmouth College, 2009.

[6] Horst Herrlich and Eleftherios Tachtsis. On the number of russell's socks or 2+ 2+ 2+...=? *Commentationes Mathematicae Universitatis Carolinae*, 47(4):707–718, 2006.

[7] Horst Herrlich, Paul Howard, and Eleftherios Tachtsis. The cardinal inequality $\alpha 2$¡ 2α. *Quaestiones Mathematicae*, 34(1):35–66, 2011.

[8] Jech Thomas. The axiom of choice. *Studies in Logic and the Foundations of Mathematics*, 75, 1973.

[9] Andreas Blass. Injectivity, projectivity, and the axiom of choice. *Transactions of the American Mathematical Society*, 255:31–59, 1979.

[10] Horst Herrlich, Paul Howard, and Eleftherios Tachtsis. Finiteness classes and small violations of choice. 2011.

[11] Frank Robert Drake. *Set theory: An Introduction to Large Cardinals*. North-Holland Publishing Company Amsterdam, 1974.

[12] Herbert B. Enderton. *A Mathematical Introduction to Logic*. Access Online via Elsevier, 2001.

[13] James R. Munkres. *Topology*. Prentice Hall Upper Saddle River, 2nd edition, 2000.

[14] Herman Rubin and Jean E Rubin. *Equivalents of the Axiom of Choice, II*. Elsevier Science Ltd, 1st edition, 1985.