SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Choice Principles in Mathematics

av

Simon Almerström Przybyl

2015 - No 6

Choice Principles in Mathematics

Simon Almerström Przybyl

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Henrik Forssell

2015

Choice Principles in Mathematics

Simon Almerström Przybyl

Abstract

In this thesis we illustrate how mathematics is affected by the Axiom of Choice (AC). We also investigate how other choice principles affect mathematics. Proofs of the following three major results are presented:

(1) AC, Zorn's Lemma and the Well-Ordering Theorem are equivalent. We prove this equivalence without using transfinite techniques.

(2) The Banach-Tarski Paradox (BTP) holds in ZFC but fails in ZF + AD + DC and is thus independent of ZF + DC. The latter results are proved under certain consistency assumptions using the connection between BTP and non-measurable sets.

(3) AC and Tychonoff's Theorem are equivalent.

Proofs of other minor results regarding choice principles are also presented.

Contents

Chapter 1

Introduction

Berries in Bowls

The Axiom of Choice (AC) is the statement that it is possible to choose precisely one element from each and every set in any family of non-empty sets (we will call a set a family when we want to emphasize that its elements are sets). In less technical terms, AC is the following claim:

> Given a table filled with bowls such that the bowls themselves contain berries and no bowl is left empty, it is always possible to choose precisely one berry from each and every bowl on the table.

For every table (family) with finitely many bowls (sets), it could be said to be obvious that it is possible to choose one berry (element) from each and every bowl: We can manually choose one berry from each bowl until we are done. However, if we allow ourselves to consider abstract tables such as *the table with one bowl for each real number*, such a manual process of choosing berries will never come to an end.

Both in the finite and infinite case, we could cheat and give a rule for how to choose the berries from the bowls instead of manually specifying every choice. We could for example specify that we choose the berry which weighs the least from each bowl. In order for this rule to work, there needs to be a lightest berry in each bowl and this might not be the case in our abstract setting: Consider a bowl in which there exists a berry with weight r for every real number r strictly between 0 and 1. Obviously, we can manually specify a choice for this specific bowl if it is the only one showing this strange behavior, however there might be infinitely many bowls which behave in this way.

AC is however the claim that it is always possible to choose one berry from each and every bowl, no matter how many bowls there are and what berries they contain, as long as the bowls are non-empty.

Purpose of the Thesis

In this thesis we investigate how different areas of mathematics are affected by the presence or absence of AC. Throughout the thesis, we work in Zermelo-Fraenkel set theory unless otherwise stated. We denote this theory by ZF and denote ZF + AC by ZFC (see [Jec2006] for a list of the axioms of ZF).

We mainly discuss how elementary set theory, measure theory and topology are affected by AC. We illustrate the highly counterintuitive consequences of AC by presenting a detailed, almost complete, proof of the famous Banach-Tarski Paradox (BTP). Moreover, we also present a proof showing that a choice principle of similar strength as AC is necessary to yield BTP by proving that BTP is independent of ZF + Principle of Dependent Choices (DC).

We also motivate the need of a choice principle by illustrating that many statements are unprovable without a choice principle being present. We illustrate the problem of rejecting AC by presenting various innocent and important statements which are unprovable without full choice (AC).

Thus our programme is as follows:

In the next section, we describe the historical origins of AC. The history of AC is closely related to the beginning of the modern attempt of trying to define a solid foundation of mathematics.

In chapter 2, we begin by discussing some logical and set theoretic preliminaries. We also discuss alternative characterizations of AC using Cartesian products and present the implication from AC to DC to CC. We then give a presentation of the equivalence between AC, Zorn's Lemma (ZL) and the Well-Ordering Theorem (WOT). This equivalence is fundamental since the three different characterizations are seemingly unrelated yet of the same strength in ZF. We also prove that Hausdorff's Maximal Principle (HMP) can be added to the equivalence.

In chapter 3, a proof showing that ZFC implies BTP is presented in the first section. BTP illustrates the counterintuitive consequences of AC, the usual way of informally stating BTP is:

> A three-dimensional ball can be split up into finitely many pieces such that by only moving the individual pieces and rotating them, the pieces can be put together into two balls identical to the initial one.

As we will see, BTP can even be stated in a seemingly more general but equivalent form.

We begin the second section of this chapter by proving that BTP implies the existence of non-measurable subsets of $\mathbb{R}^3$. We then present a proof showing that BTP fails in ZF + Axiom of Determinateness (AD) + DC by proving that all subsets of any Euclidean space $\mathbb{R}^n$ are measurable in ZF + AD + DC. Given the consistency of ZF + AD + DC, it follows that BTP is independent of ZF

+ DC as well as the weaker theory ZF + Axiom of Countable Choice (CC, AC restricted to countable families). Regarding the consistency of ZF + AD + DC, we use a theorem stating the relative consistency of ZF + AD and ZF + AD + DC which we do not prove. We also assume the relative consistency of ZF and ZF + AD, whether this relation holds or not is still unknown.

In chapter 4, we present a few examples to illustrate that a lot of analysis can be developed in ZF + CC while some almost trivial set theoretic statements cannot be proved in ZF. We also present a proof of the equivalence between AC and Tychonoff's Theorem (TT).

In chapter 5, we finish the thesis by contemplating our results.

The results presented in this thesis are obviously well-know. My work has essentially been to find interesting theorems and then understand these theorems and express their proofs with my own words. In this process I have hopefully clarified some parts of the proofs which have been either omitted or unclear in the original material. For each proof in this thesis which has been inspired by another author, there is a footnote referring to the source. The main sources which have been used are [Coh2013] for the proof of BTP in ZFC, [Jec2006] for the measurability result in ZF + AD and [Her2006] for general results regarding AC and other choice principles.

Moreover, I have strived to present the material in a as self-contained way as possible: Section 2.1 requires some understanding of logic and section 3.2 is probably more easily read with some knowledge of measure theory. Otherwise only fundamental analysis and algebra is needed to understand this thesis.

History of AC

The sources of the historical statements made in this subsection are [Moo82] and to a smaller extent [Her2006].

The Well-Ordering Theorem (WOT), i.e. the statement that the every set X can be arranged in such a manner that every non-empty subset of X has a least element, is closely connected to the historical origins of AC. When Cantor developed the foundation of set theory at the end of the 19th century, he considered WOT to be a law of thought which was beyond the need of a proof.

Cantor's major innovation was his quantification of the infinite and the consequences it yields. The concept of cardinality is due to Cantor: A set X is *countable* if there exists a bijection f between X and a subset of $\mathbb{N}$. If a set is not countable, then it is *uncountable*. More generally, two sets X and Y are said to have the same *cardinality* if there exists a bijection $f : X \to Y$. We denote the cardinality of a set X by $|X|$, thus we define $|X| = |Y|$ to mean that there exists a bijection between X and Y. Moreover, we say that *the cardinality of Y is greater than the cardinality of X* if there exists an injection $h : X \to Y$. We denote this relation by $|X| \preceq |Y|$. The Schröder-Bernstein Theorem (see Theorem 6.1 in [Gol96]) allows us to conclude that $|X| \preceq |Y|$ and $|Y| \preceq |X|$ hold if and only if $|X| = |Y|$ holds.

Furthermore, it seems intuitive to provide an alternative definition of cardinality in terms of surjections. Thus we let $|X| \preceq^* |Y|$ denote the existence of a surjection $h : Y \to X$.

Proposition 1.0.0.1. *If X is a non-empty set and $|X| \preceq |Y|$, then $|X| \preceq^* |Y|$.*

Proof. Assume there exists an injection $f : X \to Y$. Then there exists a corresponding inverse $f^{-1} : f(X) \to X$. Note that f^{-1} is surjective: Each $x \in X$ has a unique image $y \in Y$, thus $f^{-1}(y) = x$. If f is surjective, then $f(X) = Y$ and the proof is finished. If f is not surjective, then we can extend f^{-1} to $g : Y \to X$ by defining $g(y) = \begin{cases} f^{-1}(y), \text{ if } x \in f(X) \\ x \text{ if } y \in Y \setminus f(X) \end{cases}$ for some arbitrary $x \in X$ since X is non-empty. $\square$

The reverse implication is related to AC and will be discussed in section 4.1.

It is worth noting that even though the concept of differently large infinities (such as the difference between the countable and uncountable) is not controversial today, it was controversial in Cantor's time. Even the concept of the existence of an actual infinite was doubted by distinguished mathematicians such as Poincaré.

At the turn of the century, Cantor started doubting the validity of WOT as a law of thought and sought to prove it. At the same time, Cantor was also trying to prove another statement:

Definition 1.0.0.2 (Continuum Hypothesis - CH). If A is an infinite subset of $\mathbb{R}$, then A is bijective either with $\mathbb{N}$ or with $\mathbb{R}$ itself.

CH is essentially the statement that the infinity represented by $\mathbb{R}$ is the next infinity after that represented by $\mathbb{N}$.

Hilbert was highly interested in Cantor's set theory. In a influential lecture in Paris in 1900, Hilbert presented a list of 23 problems which he considered to be the most important mathematical problems of the 20th century and the first of these problems was to prove if $\mathbb{R}$ could be well-ordered and to prove or disprove CH. Hilbert thought these two questions were connected.

In 1904, Zermelo explicitly defined AC and presented a proof of WOT from AC. The axiom (AC), and even more often CC, had been used by the mathematical community during the 19th century: The principle is present in proofs from that time, sometimes hidden and probably used without the author's knowing and sometimes used more consciously. We will see examples of hidden use of CC in section 4.1. However, it was first when Zermelo explicitly defined AC that the fundamental difference between making finitely many arbitrary choices and infinitely many such choices was noticed and debated by the mathematical community. The debate was heated and many notable mathematicians such as Borel and Lebesgue were skeptic towards AC, even though their own work preceding their criticism turned out to build on results motivated by AC or at least CC.

The controversy regarding AC and other contemporary events (such as the discoveries of Russell's Paradox and the Burali-Forti Paradox) essentially forced set theory to be axiomatized. Cantor and others had treated set theory as ordinary mathematics, reasoning in a way of common sense without caring too much about the underlying assumptions and without any clear syntax of how formal objects interfere. A few decades into the 20th century, the assumptions underlying Cantor's set theory were formulated in a formal language and became what we know as set theory.

The hesitation about the intuitive nature of AC proved to be justified. During the first decades of the 20th century, AC was used in the construction of several strange sets of real numbers, such as the Vitali sets. The Banach-Tarski Paradox was discovered around the 1920s and is probably one of the most remarkable consequences of AC.

However, in the 1930s, Gödel proved the relative consistency of ZF and ZFC. This was seen as a result in favor of the validity of AC: If adding AC to ZF would not ruin the assumed consistency of ZF, then AC seemed reasonable. It was to take another 30 years until Cohen in the 1960s used his newly invented method of forcing to prove the relative consistency of ZF + ¬AC and ZF, thus establishing the independence of AC from ZF. This method and result brought set theory into the modern era and finishes our historical voyage.

Notational Remarks

Note that even though I have written this thesis alone, I write using *we* throughout the thesis, referring to the mental collective the reader and the writer constitute.

Moreover, for every proof which is not done in ZF, a parentheses at the statement of the proposition is used to indicate that we are currently working in another variant of ZF.

Another notational remark is that $\bigcup X$ is used for the singleton union:

$$x \in \bigcup X \iff \exists Y \in X \ (x \in Y)$$

While $X \cup Y$ of course denotes the set of elements of X or Y:

$$x \in X \cup Y \iff (x \in X \vee x \in Y).$$

Moreover, $\bigcup\limits_{i \in I} X_i$ is used to denote the set of elements of all X_i together:

$$x \in \bigcup_{i \in I} X_i \iff \exists i \in I \ (x \in X_i).$$

Similar remarks apply to the intersection symbol.

Also note that an indexed family $\{X_i \mid i \in I\}$ sometimes will be written as $\{X_i\}$ to ease notation. The same remark applies to situations involving sequences.

Finally, we use $X \subseteq Y$ to denote that X is a subset of Y, possibly with $X = Y$. The symbol $\subset$ is used for the strict relation.

Acknowledgement

I would like to thank my supervisor Henrik Forssell for all the help and support I have received while writing this thesis. I would also like to thank Håkon Robbestad Gylterud for reviewing the thesis.

Chapter 2

Axiom of Choice

2.1 Definition

2.1.1 Logical & Set Theoretic Preliminairies

As noted, we work in ZF unless otherwise stated. In our case working in ZF simply means that we do ordinary mathematics only using the axioms of ZF, i.e. our reasoning is done in an arbitrary model of ZF. We naïvely think of a *model of ZF* as a universe of sets that satisfies the axioms of ZF. Thus all of our proofs can be formalized into derivations in natural deduction where a subset of the formalized versions of the ZF axioms are our only undischarged assumptions and the only non-logical symbol is the binary relation symbol $\in$ (see [Car2013] for a presentation of first-order logic).

A (well-formed) *formula* is simply a statement in first-order logic. In this subsection, by derivation we mean a derivation in natural deduction. Note that if we say that Γ derives φ, then we mean that there exists a derivation of φ where the undischarged assumptions constitute a subset of Γ.

Definition 2.1.1.1. A set Γ of formulas is *consistent* if there does not exist a derivation from Γ to $\perp$. If Γ is not consistent, i.e. if there exists a derivation from Γ to $\perp$, then Γ is *inconsistent.*

We may thus wonder whether ZF is consistent or inconsistent. However, this discussion turns out to be quite unfulfilling since Gödel's Second Incompleteness Theorem implies that ZF cannot prove its own consistency (given that ZF actually is consistent). Thus, we assume the following to be true for the rest of the thesis:

Assumption 2.1.1.2. ZF *is consistent.*

By the Model Existence Lemma, this assumption ensures the existence of a model of ZF and we may thus reason in an arbitrary model of the theory. Even though we are unable to prove the consistency of ZF, we can obtain positive results by discussing the concept of relative consistency:

Definition 2.1.1.3. Two sets of formulas Γ and Λ are *relatively consistent* if Γ is consistent if and only if Λ is consistent.

If $\Lambda \subseteq \Gamma$, then Λ is clearly consistent if Γ is. Thus only the reverse direction is non-trivial when speaking about relative consistency when one set of formulas is a subset of the other.

ZF is expressed in classical logic, thus *tertium non datur* is assumed to be valid, that is:

$$\varphi \vee \neg\varphi \tag{tertium}$$

is assumed to be valid for any formula φ. However, φ may have different truth values in different models of ZF, thus we define the following concept:

Definition 2.1.1.4. Let Γ be a set of formulas. A formula φ is *provable* in Γ if there exists a derivation from Γ to φ and is *unprovable* if there does not. If both φ and $\neg\varphi$ are unprovable in Γ, then φ is *independent* of Γ.

By the soundness and completeness of first-order logic, note that a formula φ is provable in Γ if and only if φ is true in every model of Γ.

When we say that φ holds in Γ, φ is a theorem of Γ or state some similar assertion, then we mean that φ is provable in Γ. If we say that φ fails, then we simply mean that $\neg\varphi$ holds. Note that if we say that φ holds or fails in a certain model of Γ, then we are only describing the truth value of φ in the given model and are not discussing the provability of φ in Γ.

Assume that Γ is a consistent set of formulas. Note that $\Gamma + \varphi$ and $\Gamma + \neg\varphi$ are both consistent if and only if φ is independent of Γ. Also note that if φ holds in $\Gamma + \psi$, then $\neg\varphi$ cannot hold in Γ as this would contradict the consistency of $\Gamma + \psi$. However, φ might be unprovable in Γ, in this case φ is independent of Γ. Finally, if φ is independent of Γ and $\Gamma + \psi$ derives φ, then ψ is not provable in Γ. If ψ were provable, it would yield a contradiction against the independence of φ from Γ. Moreover, since ψ is not provable in Γ, there exists a model $\mathcal{M}$ of Γ such that ψ fails in $\mathcal{M}$.

2.1.2 Choice Principles

Definition 2.1.2.1. A *choice function* for a family X of non-empty sets is a function $f : X \to \bigcup X$ such that $f(x) \in x$ holds for every $x \in X$.

Definition 2.1.2.2 (Axiom of Choice - AC). Every family of non-empty sets has a choice function.

A definition like the one above is to be read as *The Axiom of Choice, hereafter AC, is the statement: Every family....*

A priori, AC could be a theorem of ZF. Then there would be little sense in distinguishing between ZF and ZFC. Alternatively, it could be the case that $\neg$AC were a theorem of ZF. The theory ZFC would then be inconsistent. As noted in the section about the history of AC, the following famous theorems due to Gödel (see [Göd38]) and Cohen (see Theorem 14.36 in [Jec2006]) respectively resolve these considerations:

Theorem 2.1.2.3. ZF *and* ZFC *are relatively consistent.*

Theorem 2.1.2.4. ZF *and* ZF + ¬AC *are relatively consistent.*

We will not prove these theorems as they require advanced methods from set theory but we note that they imply the following theorem:

Theorem 2.1.2.5. AC *is independent of* ZF*.*

We will now present the first equivalence of AC in ZF, first we need to define some concepts. Also note that by the reasoning of the previous subsection, as soon we prove that a statement is a sufficient condition for AC, then the statement is unprovable in ZF and there exists a model of ZF where its negation holds.

Definition 2.1.2.6. Let X be a set, then I is an *index set* of X if there exists a surjection $j : I \to X$. We call such a j an *index function* and say that X *is indexed by* I with respect to j. We denote $j(i)$ by x_i.

We often let the j be implicit and simply say that X is indexed by I and write $X = \{x_i \mid i \in I\}$ without discussing what j we are referring to. If we view X as a family, then we often write X_i instead of x_i.

Every set X can be indexed: Consider the *canonical indexing* defined by letting X index itself using the identity function $id_X : X \to X$ as the index function.

Definition 2.1.2.7. Let X be a family indexed by I, then the *Cartesian product* of X with respect to I is:

$$\prod_{i\in I} x_i = \{h : I \to \bigcup_{i\in I} x_i \mid \forall i \in I\ (h(i) \in X_i)\}$$

A Cartesian product as defined above generalizes the finite Cartesian product denoted by $\underbrace{x_1 \times ... \times x_n}_{n \text{ times}}$ for n not necessarily distinct sets x_i.

We call the Cartesian product X with respect to the canonical indexing *the Cartesian product* of X. By simplifying the notation from x_x to x and using $\bigcup X = \bigcup_{x\in X} x$, we can obviously write the Cartesian product of X as:

$$\prod X = \prod_{x\in X} x = \{h : X \to \bigcup X \mid \forall x \in X\ (h(x) \in x)\}$$

Proposition 2.1.2.8. *Let X be a family of non-empty sets. Then the Cartesian product of X is non-empty if and only if the Cartesian product of X with respect to an arbitrary indexing I is non-empty.*

Proof. The left direction follows trivially. For the right direction, let I be an index function of X and let j be the corresponding index function. Moreover, let $f \in \prod X$ and for all $x \in X$, define h by letting $h(i) = f(x)$ for all i such that $j(i) = x$. Then $h \in \prod_{i\in I} x_i$. □

Proposition 2.1.2.9. AC *holds if and only if the Cartesian product of a family of non-empty sets is non-empty.*

Proof. Let X be any family of non-empty sets. By AC, there exists a choice function f for X. Define $h : X \to \bigcup X$ by $h(x) = f(x)$, then $h \in \prod X$.

Conversely, let X be any family of non-empty sets. By assumption, there exists $h \in \prod X$. Define $f(x) = h(x)$, this yields a choice function f for X. $\square$

Moreover, the following is an axiom of ZF (we state it as in [Jec2006]):

Definition 2.1.2.10 (Axiom Schema of Replacement - ASR). If a class F is a function, then for any X there exists a set $Y = F(X) = \{F(x) \mid x \in X\}$.

ASR essentially says that the image of any definable function is a set. If $(X_i)_{i\in I}$ is a generalized sequence of sets X_i, i.e. if there exists a function such that $f(i) = X_i$ holds for all $i \in I$, then ASR implies that $\{X_i \mid i \in I\}$ is a set. Thus AC holds if and only if $\prod\limits_{i\in I} X_i$ is non-empty for any generalized sequence $(X_i)_{i\in I}$ of non-empty sets X_i.

We now present some weaker choice principles:

Definition 2.1.2.11 (Principle of Dependent Choices - DC). Let X be a non-empty set and let R be a relation on X such that for each $x \in X$, there exists $y \in X$ satisfying xRy. Then there exists a sequence $(x_n)_{n=0}^{\infty}$ with $x_n \in X$ such that $x_n R x_{n+1}$ holds for each $n \in \mathbb{N}$.

Proposition 2.1.2.12. [1] AC $\Rightarrow$ DC.

Proof. Define $S_x = \{y \in X \mid xRy\}$. By assumption, S_x is non-empty for each $x \in X$, thus $\mathscr{S} = \{S_x \mid x \in X\}$ has a choice function f by AC. Let x_0 be an arbitrary element of X and define $x_{n+1} = f(S_{x_n})$ for all $n \in \mathbb{N}$, this recursively defines a sequence $(x_n)_{n=0}^{\infty}$ such that $x_n R x_{n+1}$ holds for all $n \in \mathbb{N}$. $\square$

Definition 2.1.2.13 (Axiom of Countable Choice - CC). Every countable family of non-empty sets has a choice function.

By the previous discussion, we see that CC holds if and only if $\prod\limits_{n\in\mathbb{N}} X_n$ is non-empty for any sequence $(X_n)_{n=0}^{\infty}$ of non-empty sets X_n.

Proposition 2.1.2.14. [2] DC $\Rightarrow$ CC.

Proof. Let $\{X_n \mid n \in \mathbb{N}\}$ be a countable family of non-empty sets X_n. Define $Y_n = \prod\limits_{m\leq n} X_m$ for all $n \in \mathbb{N}$ and let $Y = \bigcup\limits_{n\in\mathbb{N}} Y_n$. Define a relation R on Y by:

$$(\alpha^0, ..., \alpha^m) R (\beta^0, ..., \beta^n) \text{ if and only if,}$$
$$m + 1 = n \text{ and } x^i = y^i \text{ holds for all } 0 \leq i \leq m.$$

[1] Corresponds to part one of Theorem 2.12 in [Her2006].
[2] Corresponds to part two of Theorem 2.12 in [Her2006].

Since each X_n is non-empty, Y is clearly non-empty and every $\alpha = (\alpha^0, ..., \alpha^m) \in Y$ relates to some $\beta = (\beta^0, ..., \beta^{m+1}) \in Y$. Thus by DC, there exists a sequence $(\gamma_n)_{n=0}^{\infty}$ such that $\gamma_n R \gamma_{n+1}$. Moreover, by the proof of DC from AC, we see that we are free to choose the first element of the sequence which DC claims exists. We choose it to be some arbitrary element $\gamma_0 = (\gamma_0^0)$ in the singleton product $\prod\limits_{n=0} X_n$. Using $\gamma_n = (\gamma_n^0, ..., \gamma_n^n)$, we define $f(n) = \gamma_n^n$ for all $n \in \mathbb{N}$, then $f \in \prod\limits_{n \in \mathbb{N}} X_n$ holds. $\square$

2.2 Zorn's Lemma & Well-Ordering Theorem

Given a family X, it seems as though AC knows a rule for choosing an element from each $X_i \in X$ even when no definable rule seems to exist. This vaguely suggest that there exists some rule for choosing elements from each X_i which we are unable to see. In this section, we will formalize these thoughts by proving that AC implies the Well-Ordering Theorem (WOT) which says that the every set X can be arranged in such a manner that every non-empty subset of X has a least element. Note that WOT easily implies AC as WOT gives us a rule for choosing elements: For any family X of non-empty sets X_i, arrange $\bigcup X$ so every non-empty subset has a least element and specify the choice from X_i to be the least element of X_i under the arrangement (this reasoning is formalized in Theorem 2.2.3.1). Thus we will in this section prove that AC and WOT are equivalent.

The equivalence of AC, ZL and WOT is often proved using techniques related to Cantor's quantification of the infinite, such as transfinite induction (induction generalized to other sets than $\mathbb{N}$). However, we will prove the equivalence without using these techniques, essentially to prove that the equivalence holds independently of the concept of ordinals and cardinals. Note that the proof of the equivalence becomes much shorter when the transfinite techniques are employed.

2.2.1 AC $\Rightarrow$ ZL

We use the standard definitions of partial and total (i.e. linear) orders: Partial orders are binary relations satisfying reflexivity, antisymmetry and transitivity while total orders also satisfy totality (all elements are comparable). A poset is an ordered pair (X, P) of a set X and a partial order P on X. We will often use $\leq$ to denote the partial order and we use $a < b$ as shorthand for $(a \leq b) \wedge (a \neq b)$.

We continue with some more standard definitions. Note that in some of the definitions below, it would be more correct to say *in* $(X, \leq)$ than *in* X. We use the latter to ease notation.

Definition 2.2.1.1. Let $(X, \leq)$ be a poset and $S \subseteq X$:

- $u \in X$ is an *upper bound* for S if $s \in S \Rightarrow s \leq u$. Moreover, if $u \in S$ we say that u is a *greatest element* of S.
- $m \in S$ is a *maximal element* of S if $\forall s \in S\ (m \leq s \Rightarrow m = s)$.
- The *initial segment* of x in X is $\downarrow x = \{y \in X \mid y \leq x\}$. The *proper initial segment* $\downarrow^* x$ is defined similarly with strict inequality. We denote $\downarrow x \bigcap S = \{s \in S \mid s \leq x\}$ by $\downarrow_S x$ and similarly define $\downarrow_S^* x$ with strict inequality.
- If for any two elements s_1 and s_2 of S, either $s_1 \leq s_2$ or $s_1 \geq s_2$ holds, then S is called a *chain* in X.
- If Y and Z are chains in X and $Y \subseteq Z$ and $\forall y \in Y\ (\downarrow_Z y \subseteq Y)$ hold, then Y is an *initial chain* of Z in X. We denote it by $Y \sqsubseteq Z$.

We also define *lower bound*, *least element*, *minimal element*, *(proper) terminal segment* and *terminal chain* in a dual way. Note that the greatest or least element of a subset $S \subseteq X$ is unique since if two exist, then they will be equal by antisymmetry.

Definition 2.2.1.2 (Zorn's Lemma - ZL). Let $(X, \leq)$ be a poset such that X is non-empty and every chain in X has an upper bound in X. Then X has at least one maximal element.

Theorem 2.2.1.3. [3] AC $\Rightarrow$ ZL.

Proof. Let $(X, \leq)$ be a poset satisfying the preconditions of ZL. For every chain C in X, let C^* be the set of upper bounds u of C not in C, i.e. $C^* = \{u \in X \setminus C \mid \forall c \in C\ (c < u)\}$. By AC, there exists a function f which chooses one element from each non-empty C^*. We define a non-empty chain C in X to be an f-chain if the following implication holds:

$$\Big(S \subset C \wedge S^* \cap C \neq \emptyset\Big) \Rightarrow f(S^*) \text{ is a minimal element of } S^* \cap C. \qquad (\phi)$$

The existence of an f-chain is proved as follows: Since X is non-empty, there exists an $x \in X$. The set $\{x\}$ is a chain and contains no non-empty strict subsets, thus ϕ is vacuously true for $\{x\}$.

We will now prove some properties for chains and f-chains. We will use ■ to denote that the proof of a specific property is finished:

Lemma (a). Let C be a chain in X. If $S \subseteq C$ and $S^* \cap C = \emptyset$, then $S^* = C^*$.

Proof of Lemma (a). Obviously $C^* \subseteq S^*$ holds since every upper bound u of C is an upper bound of S, and if $u \notin C$ then $u \notin S$. To prove $S^* \subseteq C^*$, assume

[3] This proof corresponds *closely* to Theorem 4.19 in [RuRu85].

there exists $u \in S^* \setminus C^*$. Then for all $s \in S$, $s < u$ holds. Since also $S^* \cap C = \emptyset$ holds by assumption, we obtain $u \notin C^* \cup C$. Thus there exists $c \in C$ such that either $u < c$ or u is not comparable with c. If $u < c$ holds, then by transitivity $s < c$ holds for every $s \in S$. Then c is an upper bound of S, contradicting $S^* \cap C = \emptyset$. Instead assume u and c are not comparable. Since $S^* \cap C = \emptyset$, $c \notin S^*$ holds and since C is a chain with $S \subseteq C$ there exists $s \in S$ such that $c \leq s$. Since $s < u$ holds, by transitivity we reach a contradiction again. ■

Lemma (b). If C is an f-chain and $C^* \neq \emptyset$, then $D = C \cup \{f(C^*)\}$ is also an f-chain.

Proof of Lemma (b). Assume $S \subset D$ and $S^* \cap D \neq \emptyset$, we divide our analysis into three cases depending on the relation between S and C:

(i) $S \subset C$ and, $S^* \cap C = \emptyset$ or $S^* \cap C \neq \emptyset$.

(ii) $S = C$, if this holds then $S^* \cap C = \emptyset$ by definition of S^*.

(iii) $S \nsubseteq C$ and, $S^* \cap C = \emptyset$ or $S^* \cap C \neq \emptyset$.

These cases exhaust all possible forms of S. We may rewrite conditions (i) and (ii) as:

(i') $S \subset C$ and $S^* \cap C \neq \emptyset$.

(ii') $S \subseteq C$ and $S^* \cap C = \emptyset$.

Since C is an f-chain by assumption, (i') implies that $f(S^*)$ is a minimal element of $S^* \cap C$. Since $f(C^*)$ is an upper bound of C and thus of $S^* \cap C$, we obtain $f(S^*) \leq f(C^*)$. Thus $f(S^*)$ is also a minimal element of $S^* \cap (C \cup \{f(C^*)\}) = S^* \cap D$.

Assume (ii') holds. Then as noted in the proof of lemma (a), $C^* \subseteq S^*$ so $f(C^*) \in S^*$ and thus $S^* \cap D = \{f(C^*)\}$ since $S^* \cap C = \emptyset$ by assumption. Applying lemma (a) to condition (ii'), we obtain $S^* = C^*$. Thus $f(S^*) = f(C^*)$ and then $f(S^*)$ is definitely a minimal element of $S^* \cap D = \{f(S^*)\}$.

Finally, we note that given our assumption $S \subset D$ and $S^* \cap D \neq \emptyset$, (iii) cannot hold: $S \nsubseteq C$ and $S \subset D$ implies $f(C^*) \in S$. Thus $S^* \cap C = \emptyset$ since the elements of S^* has to be upper bounds of $f(C^*)$ and no element of C satisfy this since $f(C^*) \in X \setminus C$ is an upper bound of C. Now $S^* \cap C = \emptyset$ and $f(C^*) \in S$ (so $f(C^*) \notin S^*$) gives $S^* \cap D = \emptyset$, contradicting our initial assumption. ■

Lemma (c). Given any two f-chains Y and Z in X, one is an initial chain of the other.

Proof of Lemma (c). Assume w.l.o.g that there exists $t \in Z \setminus Y$ as if $Z \setminus Y$ is empty, then we just let $t \in Y \setminus Z$ and if this set is empty too, then $Z = Y$. Define:

$$U_t = \{s \in Z \cap Y \mid s \leq t\} = \downarrow_{Z \bigcap Y} t$$

Note that we could define U_t with strict inequality since $t \notin U_t$.

Clearly $U_t \subseteq Z$ and $U_t \subseteq Y$ hold. We will prove $Y = U_t$, note that this yields $Y \subseteq Z$. Also note that since $t \in Z \setminus Y$ was chosen arbitrarily, $Y = U_t$

holds for all such t. We now prove that $Y = U_t$ for all concerned t implies that also the second property of $Y \sqsubseteq Z$ is satisfied and thus the whole relation: Assume $Y = U_t$ holds for every $t \in Z \setminus Y$ and $Y \not\sqsubseteq Z$. Then there exists some $y \in Y$ such that $\downarrow_Z y \not\subseteq Y$, implying that there exists $s \in \downarrow_Z y \setminus Y \subseteq Z \setminus Y$. Note that $s < y$ holds. Since $s \in Z \setminus Y$, the equality $Y = U_s$ holds. However, $y \in Y$ holds but $y \notin U_s$ does not hold since this would imply that both $s < y$ and $y \leq s$ hold. Having reached a contradicting, we conclude that $Y \sqsubseteq Z$ holds.

Thus we will prove $Y = U_t$. For the rest of the proof we simply denote U_t by U, thus the t which defines this set is now considered to be fixed.

First, assume:

(i) $U \subset Y$

Since $t \in Z \setminus U$ and $U \subseteq Z$, the following holds:

(ii) $U \subset Z$

Clearly $t \in U^*$ holds, thus the following also holds:

(iii) $U^* \cap Z \neq \emptyset$

By assumption, Z is an f-chain so (ii) and (iii) implies that $f(U^*)$ is a minimal element of $U^* \cap Z$. Also, Z is a chain so all of its elements are comparable, thus:

(iv) $f(U^*) \leq t$

Now assume $f(U^*) \in Y$. Since $f(U^*) \in Z$, by (iv) and the definition of U we then have $f(U^*) \in U$. However, this contradicts $f(U^*) \in U^*$ so we conclude:

(v) $f(U^*) \notin Y$

Assume $U^* \cap Y \neq \emptyset$. By (i) and since Y is an f-chain, in particular this yields $f(U^*) \in U^* \cap Y \subseteq Y$, contradicting (v). Thus:

(vi) $U^* \cap Y = \emptyset$

By (i) and (vi), lemma (a) implies:

(vii) $U^* = Y^*$

By (vii), there does not exist $r \in Y \setminus U$ such that $r \in U^*$. Thus two alternatives can hold: Either $Y \setminus U$ is empty. Then $Y \subseteq U$ so $U = Y$ holds, finishing the proof. Otherwise, for any $r \in Y \setminus U$ there exists $u \in U$ such that either $r \leq u$ or r is incomparable with u. Since Y is a chain and $U \subseteq Y$, r must be comparable with u, thus $r \leq u$ holds. By definition of U, $u \leq t$ holds and by transitivity, $r \leq t$ thus holds. Since we have $r \in Y$, $r \leq t$ and $r \notin U$, we obtain $r \notin Z$. Define:

$$V = \{s \in Z \cap Y \mid s \leq r\}$$

Completely analogous with how we deduced (vi), we obtain:

(viii) $V^* \cap Z = \emptyset$

Since $t \in Z \setminus Y$ holds, we have $t \notin V$. Since also $r \leq t$ holds, $t \in V^* \cap Z$ holds, contradicting (viii). We may thus conclude $\neg(U \subset Y)$ from (i) and use $U \subseteq Y$ to conclude $U = Y$. ■

By lemma (c), given any two f-chains one is a subset of the other. Thus the set $\{C \mid C \text{ is } f\text{-chain in } X\}$ is a $\subseteq$-chain in $\mathcal{P}(X)$ and therefore its union $\bigcup\{C \mid C \text{ is } f\text{-chain in } X\} = \mathscr{C}$ is a chain in X: Let $x, y \in \mathscr{C}$. Then $x \in C_x$ and $y \in C_y$ for some C_x and C_y with either $C_x \subseteq C_y$ or vice versa. Since C_x and C_y are chains, it follows that x and y are comparable. We will prove that $\mathscr{C}$ is an f-chain:

Let $S \subset \mathscr{C}$ such that $S^* \cap \mathscr{C} \neq \emptyset$, then there exists an f-chain Z such that $S^* \cap Z \neq \emptyset$ since all elements of $\mathscr{C}$ are elements of f-chains. We will first prove $S \subseteq Z$: Let $c \in S^* \cap Z$, note that $S \subseteq {\downarrow_{\mathscr{C}}}\, c$ holds since c is an upper bound of S and S is contained in $\mathscr{C}$. Thus it is sufficient to prove ${\downarrow_{\mathscr{C}}}\, c \subseteq Z$: Assume ${\downarrow_{\mathscr{C}}}\, c \nsubseteq Z$. Then there exists $x \in \mathscr{C}$ such that $x < c$ and $x \notin Z$ (the inequality is strict since $c \in Z$). However, since $x \in \mathscr{C}$, there exists an f-chain A such that $x \in A$. By lemma (c), either A is an initial chain of Z or vice versa. $A \sqsubseteq Z$ is impossible since $A \subseteq Z$ cannot hold since $x \in A \setminus Z$. Thus $Z \sqsubseteq A$ must hold. This implies that ${\downarrow_A}\, c \subseteq Z$ holds (remember that $c \in Z$). However, $x < c$ and $x \in A$ hold so $x \in {\downarrow_A}\, c$ holds but $x \notin Z$. Having reached a contradiction, the proof of $S \subseteq Z$ is finished.

Moreover, $S = Z$ violates $S^* \cap Z \neq \emptyset$ so $S \subset Z$ holds. Thus Z is an f-chain such that $S \subset Z$ and $S^* \cap Z \neq \emptyset$, implying that $f(S^*)$ is a minimal element of $S^* \cap Z \subseteq S^* \cap \mathscr{C}$. We will now prove that $f(S^*)$ is a minimal element of $S^* \cap \mathscr{C}$: Suppose it is not. Then there exists $t \in S^* \cap \mathscr{C}$ such that $t < f(S^*)$ and $t \notin Z$ (the inequality couldn't hold if $t \in Z$ since then we would have $t \in S^* \cap Z$ which $f(S^*)$ is a minimal element of). Since $t \in \mathscr{C}$, there is an f-chain V such that $t \in V$. Since $t \in V \setminus Z$, lemma (c) implies $Z \sqsubseteq V$. However, $t \in {\downarrow_V}\, f(S^*)$ and thus ${\downarrow_V}\, f(S^*) \nsubseteq Z$, contradicting $Z \sqsubseteq V$. Thus $f(S^*)$ is a minimal element of $S^* \cap \mathscr{C}$ so $\mathscr{C}$ is an f-chain.

Furthermore, $\mathscr{C}^* = \emptyset$ since otherwise $f(\mathscr{C}^*) \in \mathscr{C}^*$ and then $\mathscr{C} \cup \{f(\mathscr{C}^*)\} = C$ is an f-chain by lemma (b). However, by construction of $\mathscr{C}$ we then have $C \subseteq \mathscr{C}$ so $f(\mathscr{C}^*) \in \mathscr{C}$ which is impossible. By the preconditions of ZL, each chain in X has an upper bound in X. Thus $\mathscr{C}$ has an upper bound m which by the previous considerations necessarily is a greatest element of $\mathscr{C}$. Assuming that m is not a maximal element of X directly leads to a contradiction against $\mathscr{C}^* = \emptyset$. □

2.2.2 ZL $\Rightarrow$ WOT

Definition 2.2.2.1. Let $(X, \leq)$ be a totally ordered set. Then X is *well-ordered* by $\leq$ if every non-empty subset of X has a least element under $\leq$.

Definition 2.2.2.2 (Well-Ordering Theorem - WOT). Every set can be well-ordered.

Theorem 2.2.2.3. [4] ZL $\Rightarrow$ WOT.

Proof. If X is empty the theorem follows trivially. Thus let X be a non-empty set and define $\mathscr{X} = \{(s, \leq) \mid s \subseteq X \text{ and } \leq \text{ well-orders } S\}$. Note that $\mathscr{X}$ is non-empty since X is non-empty: For $x \in X$, $(\{x\}, \{(x,x)\}) \in \mathscr{X}$. For $S_i = (s_i, \leq_i)$, $S_j = (s_j, \leq_j) \in \mathscr{X}$, we define the partial order $S_i \leq_{\mathscr{X}} S_j$ on $\mathscr{X}$ by the following three conditions:

$$s_i \subseteq s_j \tag{2.1}$$

$$\leq_i = \leq_j \restriction s_i \tag{2.2}$$

$$\forall x \in s_i \left(\downarrow_j^* (x) \subseteq s_i \right) \tag{2.3}$$

In (2), $\leq_j \restriction s_i = \{(x, y) \in s_i \times s_i \mid x \leq_j y\}$. In (3), $\downarrow_j^* (x) = \{y \in s_j \mid y <_j x\}$.

$\leq_{\mathscr{X}}$ is a partial order because it inherits the needed properties from the subset partial order on $\mathcal{P}(X)$. We only prove the transitivity of $\leq_{\mathscr{X}}$. Thus, let $S_i, S_j, S_k \in \mathscr{X}$ and assume $S_i \leq_{\mathscr{X}} S_j$ and $S_j \leq_{\mathscr{X}} S_k$. Then clearly conditions (1) and (2) of $S_i \leq_{\mathscr{X}} S_k$ hold so we only have to prove (3): We want to prove $\downarrow_k^* (x) \subseteq s_i$ for all $x \in s_i$. Since $S_i \leq_{\mathscr{X}} S_j$ holds by assumption, specifically $s_i \subseteq s_j$ holds so for every $x \in s_i$ we have $x \in s_j$. Since $\downarrow_k^* (x) \subseteq s_j$ and $\leq_j = \leq_k \restriction s_j$ hold by assumption, $\downarrow_k^* (x) = \downarrow_j^* (x)$ holds. By assumption, $\downarrow_j^* (x) \subseteq s_i$ which finishes the proof of the transitivity of $\leq_{\mathscr{X}}$.

Now let C be any chain in $\mathscr{X}$. Let C be indexed by K so $C = \{S_k = (s_k, \leq_k) \mid k \in K\}$. Note that by (2) and the antisymmetry of $\leq_{\mathscr{X}}$, there is at most one element in C which has a given subset $s_k \subseteq X$ as its domain.

Define $C^* = (\bigcup\limits_{k \in K} s_k, \bigcup\limits_{k \in K} \leq_k) = (s^*, \leq^*)$. Note that $s^* \subseteq X$. Moreover, for all $i, j \in K$, either $\leq_i \subseteq \leq_j$ or $\leq_j \subseteq \leq_i$ holds since C is a chain. Thus $\leq^* \restriction s_i = \leq_i$. Therfore if $s \subseteq s_i \subseteq s^*$ and m is the least element of s under $\leq_i$, then m is the least element of s under $\leq^*$.

We will now prove some properties of C^*:

- $\leq^*$ is a total order on s^*: We only prove the transitivity of $\leq^*$, the other properties are proved similarly. Let $x, y, z \in s^*$ and assume $x \leq^* y$ and $y \leq^* z$. Then for some i and j we have $x \leq_i y$ and $y \leq_j z$. Since C is a chain in $\mathscr{X}$, either $\leq_i$ is a restriction of $\leq_j$ or vice versa, assume the first case w.l.o.g. Thus $x \leq_j y$ and $y \leq_j z$ holds, implying $x \leq_j z$ and since $\leq_j \subseteq \leq^*$ it follows that $x \leq^* z$.

- $\leq^*$ well-orders s^*: Let s be a non-empty subset of s^*. Define $\mathscr{S} = \{S_k \in C \mid \exists x \in s \ (x \in s_k)\}$. Let S_i be an element of $\mathscr{S}$. Obviously $s \cap s_i \subseteq s_i$ so

[4] The proof of this theorem is formalized version of the sketch of proof available at [I-Wiki1].

the non-empty set $s \cap s_i$ has a least element m under $\leq_i$ and thus m is also the least element of $s \cap s_i$ under $\leq^*$. We will prove that this m is the least element of s under $\leq^*$:

Let $S_k \in \mathscr{S}$ and assume $S_k <_{\mathscr{X}} S_i$. Then $s_k \subset s_i$ so $s \cap s_k \subseteq s \cap s_i$. Since m is the $\leq^*$-least element of s_i, it surely is the $\leq^*$-least element of its subsets as well and thus specifically of $s \cap s_k$.

Instead assume $S_i <_{\mathscr{X}} S_k$. Then similarly we have $s \cap s_i \subseteq s \cap s_k$. Assume the least element n of $s \cap s_k$ under $\leq_k$ satisfies $n <_k m$. By (3) in the definition of $\leq_{\mathscr{X}}$, this implies $n \in s_i$ (and thus $n \in s \cap s_i$) and then (2) implies $n <_i m$. The last inequality contradicts that m is the least element in $s \cap s_i$ under $\leq_i$. Thus there exists no $n \in s \cap s_k$ such that $n <_k m$. Since $\leq_k$ well-orders $s \cap s_k$, it follows that m is the least element of $s \cap s_k$ under $\leq_k$ and thus under $\leq^*$.

Now, $\mathscr{S}$ is a chain in $\mathscr{X}$ since $\mathscr{S} \subseteq C$. Thus the above reasoning gives that for all $S_k \in \mathscr{S}$, m is the least element of $s \cap s_k$ under $\leq^*$. Since $s \subseteq s^*$, we have $s \subseteq \bigcup\limits_{S_k \in \mathscr{S}} s_k$. Thus $s = (\bigcup\limits_{S_k \in \mathscr{S}} s_k) \cap s = \bigcup\limits_{S_k \in \mathscr{S}} (s_k \cap s)$. This implies that m is the least element of s under $\leq^*$ since assuming otherwise directly leads to a contradiction against m being the least element of each $s \cap s_k$.

Since $\leq^*$ well-orders $s^* \subseteq X$, we have $C^* \in \mathscr{X}$.

- C^* is an upper bound of C in $\mathscr{X}$: Let $S_i \in C$, then $s_i \subseteq \bigcup\limits_{k \in K} s_k = s^*$

holds. Moreover, as earlier noted, $\leq_k = \leq^* \upharpoonright s_k$ also holds. To prove that (3) is satisfied, assume $x \in s_i$ and let $y \in s^*$ be such that $y <^* x$. Then $y \in s_k$ for some k such that either $S_k <_{\mathscr{X}} S_i$ or $S_i \leq_{\mathscr{X}} S_k$. If the first inequality holds, then $s_k \subset s_i$ implying $y \in s_i$. Assuming the second inequality holds, then $\downarrow_k^*(x) \subseteq s_i$ holds for every $x \in s_i$ and thus $y \in s_i$.

Having proved that an arbitrary chain C in $\mathscr{X}$ has an upper bound C^* in $\mathscr{X}$, we apply ZL. Thus $\mathscr{X}$ has maximal element $M = (s, \leq)$. Assume there exists $x \in X$ such that $x \notin s$, then construct $s_* = s \cup \{x\}$ and $\leq_* = \leq_* \cup \{(y, x) \mid y \in s\}$ so $y <_* x$ for all $y \in s$ and define $M_* = (s_*, o_*)$. Obviously $s \subset s_*$ holds. Also $\leq = \leq^* \upharpoonright s$ holds by construction. Moreover, for $y \in s$, $\downarrow_*^*(x) = \{y \in s_* \mid y <_* x\} \subseteq s$. Thus $M <_{\mathscr{X}} M_*$ which contradicts M being a maximal element of $\mathscr{X}$.

Thus $s = X$, giving $M = (X, \leq)$ so $\leq$ well-orders X. □

2.2.3 WOT ⇒ AC & ZL $\iff$ HMP

Theorem 2.2.3.1. WOT ⇒ AC.

Proof. Let X be a family of non-empty sets indexed by I, moreover let $\mathscr{W} = \{\leq \mid \leq \text{ well-orders } \bigcup X\}$ and let $\leq$ be an arbitrary element of $\mathscr{W}$ (WOT implies that $\mathscr{W}$ is non-empty). Denote the least element of $X_i \subseteq \bigcup X$ under $\leq$ by u_i and define $f : X \to \bigcup X$ by $f(X_i) = u_i$. This f is a choice function for X, finishing the proof. □

The results of this section prove the equivalence:

Theorem 2.2.3.2. AC $\iff$ ZL $\iff$ WOT.

Moreover, we can also easily add the following principle to the list:

Definition 2.2.3.3 (**Hausdorff's Maximal Principle - HMP**)**.** Every partially ordered set contains a maximal chain.

Theorem 2.2.3.4. ZL $\iff$ HMP.

Proof. Let $(X, \leq)$ be a poset satisfying the preconditions of ZL. By HMP, there exists a maximal chain C in X which (by the preconditions of ZL) is bounded from above by $u \in X$. Thus $u \in C$ holds (otherwise C is not a maximal chain) and assuming that u is not a maximal element of X directly leads to a contradiction: If some $x \in X$ satisfies $u < x$, then x is an upper bound of C. By the previous reasoning, x is then included in C. However, this contradicts u being an upper bound of C.

Conversely, let $(X, \leq)$ be a poset and consider the set $\mathscr{P} = \{C \in \mathcal{P}(X) \mid C$ is a chain in $X\}$ partially ordered by $\subseteq$. Let $\mathscr{C}$ be a chain in $\mathscr{P}$. Then $\bigcup \mathscr{C} \in \mathcal{P}(X)$ holds and also $C \subseteq \bigcup \mathscr{C}$ holds for any $C \in \mathscr{C}$. Moreover, if $x, y \in \bigcup \mathscr{C}$ then there exists $C_x, C_y \in \mathscr{C}$ containing x and y respectively, with one being a subset of the other. Assume w.l.o.g that $C_x \subseteq C_y$ holds, then x and y are $\leq$-comparable since C_y is a chain. Thus $\bigcup \mathscr{C}$ is a chain in X, so $\bigcup \mathscr{C} \in \mathscr{P}$ holds. These considerations establish that every chain $\mathscr{C}$ in $\mathscr{P}$ has an upper bound $\bigcup \mathscr{C}$ in $\mathscr{P}$. Thus by ZL, $\mathscr{P}$ has a maximal element i.e. X has a maximal chain. $\square$

Chapter 3

Banach-Tarski Paradox

The Banach-Tarski Paradox (BTP) can as noted be phrased even more generally than the description we gave in the introduction: BTP is the statement that any bounded subset of $\mathbb{R}^3$ with non-empty interior can be partitioned into a finite number of subsets in such a way that by only moving and rotating these subsets, any bounded subset of $\mathbb{R}^3$ with non-empty interior can be obtained. As will be proved in the first section of this chapter, this description of BTP is equivalent with the one given in the introduction.

If we apply BTP to our three-dimensional physical surrounding, BTP says that we can for example split up a straw of grass into some finite number of pieces, and then by just moving and rotating these pieces construct the whole Earth or even all of the planets in our galaxy. Needless to say, BTP is a result which contradicts our fundamental intuition of how volume behaves.

In the first section of this chapter, we prove that the two forms of BTP we have stated are equivalent and that they hold in ZFC. In the second section we prove that BTP fails in ZF + Axiom of Determinateness (AD) + DC (AD will be defined in section 3.2). Thus, given the consistency of ZF + AD + DC, BTP is independent of ZF + DC. This result also establishes that ZFC + AD is inconsistent.

3.1 ZFC $\Rightarrow$ BTP

We begin with a summary of the proof of BTP in ZFC. Even though the summary lies before the formal proof, it is recommended to read it both before and after having read the actual proof.

3.1.1 Summary of the Proof

We begin by formalizing the notion of deconstructing an object into pieces and obtaining a new object by rotating and moving these pieces individually.

Thus we consider the *rotation group* $SO(3)$ which in matrix form represents all possible rotations of $\mathbb{R}^3$ about lines passing through the origin. By combining an element M of $SO(3)$ with a translation of $\mathbb{R}^3$, wee are able to rotate and move an object as we please. The *group of rigid motions* G_3 is thus defined as the set of functions $Mx + b$ with $M \in SO(3)$ and $b \in \mathbb{R}^3$. As these functions are defined on $\mathbb{R}^3$, we say that G_3 defines a *group action* on $\mathbb{R}^3$.

We now define A and B, subsets of $\mathbb{R}^3$, to be G_3-*equidecomposable* if it is possible to finitely partition A and B into equally many sets A_i and B_i for which there exist elements g_i of G_3 such that $g_i \cdot A_i = B_i$ holds for all i. We denote this property by $A \sim B$. We note that equidecomposability is clearly reflexive and symmetric, it is also transitive since two successive relations $A \sim B$ and $B \sim C$, each involving partitions containing n and m sets respectively, yield $A \sim C$ with partitions containing nm sets (see Proposition 3.1.3.3).

If a set A is equidecomposable with a subset of B and B is equidecomposable with a subset of A, then $A \sim B$ holds. This is the content of Proposition 3.1.3.4, which we for now label (a).

Moreover, we define A to be G_3-*paradoxical* if there exists a partition $\{A'_1, A'_2\}$ of A such that A is G_3-equidecomposable with both A'_1 and A'_1. Using (a), we see (in Proposition 3.1.3.7) that a sufficient condition for A being paradoxical is that A is equidecomposable with two disjoint subsets of itself. We note that if A is paradoxical and both $A \subseteq B$ and $A \sim B$ hold, then by this sufficient condition for paradoxicality, B is also paradoxical.

The statement (b) that *the unit ball of* $\mathbb{R}^3$ *is* G_3*-paradoxical* is the formal version of the informal description of BTP we gave in the introduction of this thesis. The seemingly more general description of BTP given in the beginning of this chapter is the statement (c) that *any two bounded subsets* A *and* B *of* $\mathbb{R}^3$ *with non-empty interiors are* G_3*-equidecomposable.*

The fact that (c) implies (b) is clear. Conversely, assuming (b) and letting A and B be as in (c), we let $B(x, r)$ be a ball contained in A (remember that the interior of A is non-empty). By seeing that the geometry of $B(x, r)$ is the same as that of the unit ball and using (a) repeatedly, we may duplicate $B(x, r)$ until we have obtained so many balls that it is possible to cover B by translating the balls individually (remember that B is bounded). By the transitivity of equidecomposability, we see that A is equidecomposable with a set containing B and we can now restrict the equidecomposability so that a smaller subset of A is equidecomposable with B. Similarly, we can establish that a subset of B is equidecomposable with a A. Now applying (a) yields (c).

The concepts of equidecomposability and paradoxicality of course generalizes to other groups than just G_3.

To prove that BTP holds in ZFC, we prove (b) as follows:

We say that a group G is *generated* by two elements $\{\sigma, \tau\}$ if any element g of G can be obtained by a finite group product of σ, τ and their respective inverses. We call such a finite product a *word* over $\{\sigma, \tau\}$. We consider e to be described by the empty word not having any elements. Moreover, a word such that no element stands besides its inverse is called a *reduced word*. If any g of G has a unique representation by a reduced word over some $\{\sigma, \tau\} \subseteq G$, then G is said to be *free on two generators*.

The underlying group set of any group G free on two generators is G-paradoxical, this is the content of Proposition 3.1.4.6. We are then given two elements of $SO(3)$ which are claimed to generate a free subgroup $\mathscr{G}$ of $SO(3)$. In Proposition 3.1.4.7, we describe how to verify this statement.

We continue by proving that given a certain condition, it is possible to transfer the paradoxicality of a subgroup G to images of group actions of G (see Proposition 3.1.5.3). The condition is that the group action does not allow any elements to be to be sent to themselves by any other element than e under the group action. This is the only stage of the proof where we invoke the (seemingly) full AC.

Using the above result, we prove (Proposition 3.1.5.7) that the unit sphere S (the points with absolute value 1) of $\mathbb{R}^3$ has a countable subset D such that $S \setminus D$ is $SO(3)$-paradoxical: Seeing that each non-trivial rotation corresponds to exactly two fixed points on the sphere and that $\mathscr{G}$ only has countably many elements, it follows that the set obtained by removing the fixed points D of S under $\mathscr{G}$ contains no non-trivial fixed points. Letting $\mathscr{G}$ act on $S \setminus D$ gives that the latter set is $\mathscr{G}$-paradoxical and thus $SO(3)$-paradoxical.

Since D is countable and thus contains very few points compared to the uncountable set S, we can construct a line passing through the origin but through none of the points of D. Moreover, again using the countability of D, we can define a rotation $\rho_0 \in SO(3)$ such that the images of D under successive applications of ρ_0 are pairwise disjoint. To prove that $S \sim S \setminus D$ holds, and that S thus is $SO(3)$-paradoxical by our earlier remarks, we see that:

$$e \cdot S \setminus \bigcup_{i=0}^{\infty} \rho_0^i(D) = S \setminus \bigcup_{i=0}^{\infty} \rho_0^i(D)$$
$$\rho_0 \cdot \bigcup_{i=0}^{\infty} \rho_0^i(D) = \bigcup_{i=1}^{\infty} \rho_0^i(D)$$

Since the union of the sets involved on the lefthand side is S and the union of the sets involved on the righthand side is $S \setminus D$, we have proved $S \sim S \setminus D$. See Proposition 3.1.5.8 for details regarding the preceding paragraph.

Seeing that S is $SO(3)$-paradoxical, it is intuitively clear that every sphere of arbitrary radius is paradoxical (as earlier noted, the geometry of the sphere is not affected by changes in scale) and have corresponding scaled down paradoxical partitions. Using this, we can (Proposition 3.1.5.9) choose one paradoxical partition of S and scale it down towards zero to obtain paradoxical partitions for every sphere of smaller radius. Taking the union of all of these paradoxical partitions yields a paradoxical partition of the unit ball without the origin, B'.

Proving that B' is G_3-equidecomposable with the unit ball B finishes the proof of BTP: We use a similar technique as above and define a rotation ρ of G_3 such that the images of the origin under successive applications ρ are pairwise disjoint yet contained B. We then define an equidecomposition between B and B':

$$e \cdot B \setminus \bigcup_{i=0}^{\infty} \rho^i(0) = B \setminus \bigcup_{i=0}^{\infty} \rho^i(0)$$
$$\rho \cdot \bigcup_{i=0}^{\infty} \rho^i(0) = \bigcup_{i=1}^{\infty} \rho^i(0)$$

Since $SO(3)$ is a subgroup of G_3, and B' is $SO(3)$-paradoxical, it follows that B is G_3-paradoxical.

3.1.2 Basic Definitions

We will now go through most of the details of the above summary. This subsection as well as the following are highly inspired by Appendix G of [Coh2013] and several definitions and formulations of theorems and propositions are taken more or less literally from it.

Definition 3.1.2.1. [1] A *group* is a set G with an associative binary operation $\cdot : G \times G \to G$ such that G contains an identity element e for the operation and each element x of G has an inverse x^{-1} in G.

We will write $g_1 \cdot g_2$ as $g_1 g_2$. The identity element e of a group is unique. A subgroup is a subset S of G which is itself a group under the restriction of the binary operation of G (note that S thus has to be closed under the group operation).

Definition 3.1.2.2. A real valued $n \times n$-matrix is *orthogonal* if its column vectors are orthonormal under the Euclidean scalar product. $SO(3)$ is the set of orthogonal 3×3-matrices M such that $\det(M) = 1$.

Note that a matrix M is orthogonal if and only if $M^T M = I$. Moreover, $M^T M = I$ holds if and only if $MM^T = I$ (this statement is valid for arbitrary square matrices, not only orthogonal, see any book on linear algebra). Since $(M^T)^T = M$, it follows that M is orthogonal if and only if M^T is orthogonal.

For a proof of the following theorem, see [PaPa2007]:

Theorem 3.1.2.3 (Euler's Rotation Theorem). *If $M \in SO(3)$, then there exists a non-zero vector v of $\mathbb{R}^3$ such that $Mv = v$.*

[1] This definition is taken essentially literally from [BeBl2006].

The above theorem essentially establishes that each M represents a rotation through the line μv with $\mu \in \mathbb{R}$. The given line obviously passes through the origin, thus does every element of $SO(3)$ represent a rotation about a line through the origin.

Conversely, one can prove that every rotation around a line through the origin can be expressed by an element of $SO(3)$: One essentially decomposes an arbitrary rotation φ into rotations about the x-, y- and z-axes. These rotations can be described by certain standard matrices M_x, M_y and M_z. Through some trigonometric manipulation of these standard matrices, one can prove that their product $M_xM_yM_z$ (which represents φ) is an element of $SO(3)$. Thus:

Corollary 3.1.2.4. *$SO(3)$ represents the set of rotations of $\mathbb{R}^3$ about a line through the origin.*

Proposition 3.1.2.5. *SO(3) is a group under matrix multiplication.*

Proof. Matrix multiplication is associative. The identity matrix is clearly in $SO(3)$, proving the non-emptiness of $SO(3)$ and the existence of an identity element.

Let $M \in SO(3)$, as noted above this implies that M^T is the inverse of M and that M^T is orthogonal. Moreover $\det(MM^T) = \det(I) = 1$ holds. Thus by the linearity of the determinant, $\det(M^T) = \frac{1}{\det(M)} = 1$ holds, establishing $M^T \in SO(3)$.

Lastly, if $M, N \in SO(3)$ then $\det(MN) = \det(M)\det(N) = 1$. Moreover, $(MN)^T = N^TM^T$ (this formula is valid for arbitrary square matrices), giving $MN(MN)^T = I$ so MN is orthogonal, establishing $MN \in SO(3)$. □

Definition 3.1.2.6. For $M \in SO(3)$ and $b \in \mathbb{R}^3$, let G_3 be the set of all functions $T : \mathbb{R}^3 \to \mathbb{R}^3$ of the form $T(x) = Mx + b$. G_3 is called the set of *rigid motions* in $\mathbb{R}^3$.

We want G_3 to represent precisely arbitrary rotations followed by arbitrary translations, yet it seems to only to be able to perform rotations about lines through the origin followed by translation. The following remark from [Whi88] (p. 4 in chapter 1) resolves these issues:

Theorem 3.1.2.7. *"[A] rotation about any axis is equivalent to a rotation through the same angle about any axis parallel to it, together with a simple translation in a direction perpendicular to the axis."*

For the following proposition, note that $I(x)$ denotes the identity function of $\mathbb{R}^3$ while Ix denotes matrix multiplication between the identity matrix of $\mathbb{R}^3$ and the column vector $x \in \mathbb{R}^3$.

Proposition 3.1.2.8. *G_3 is a group under function composition.*

Proof. $T(x) = Ix \in G_3$, which establishes non-emptiness and existence of an identity element. The associativity of the function composition follows from the associativity of matrix multiplication and vector addition.

G_3 is closed under function composition since for any $T_1, T_2 \in G_3$, the following holds: $T(x) = T_2 \circ T_1(x) = M_2(M_1x + b_1) + b_2 = M_2M_1x + M_2b_1 + b_2$. Since $M_1, M_2 \in SO(3)$ we have $M_2M_1 = M \in SO(3)$ and setting $M_2b_1 + b_2 = b \in \mathbb{R}^3$ gives $T(x) = Mx + b$.

Every $T \in G_3$ has an inverse $T^{-1} \in G_3$ since: For $T(x) = Mx + b = y$, define $T^{-1}(y) = M^{-1}(y - b)$. Then $(T^{-1} \circ T)(x) = M^{-1}((Mx + b) - b) = Ix$. $T^{-1}(y)$ is in G_3 since $M^{-1} \in SO(3)$ and $-M^{-1}b \in \mathbb{R}^3$, finishing the proof. □

Definition 3.1.2.9. Let G be a group. Then a *group action* of G on X is a function $h : G \times X \to X$ for which we use the notation $h(g, x) = g \cdot x$, such that the following two conditions are satisfied for all $g_1, g_2 \in G$ and all $x \in X$:

$$\begin{aligned} &g_1 \cdot (g_2 \cdot x) = (g_1g_2) \cdot x \\ &e \cdot x = x \end{aligned}$$

Here g_1g_2 denotes result of g_1 and g_2 under the group operation. When we are talking about a group action of G on X we will often say that G acts on X. Note that for any group G, the binary operation of G defines a group action on the group set of G. Finally, note that for the rest of the thesis we will only deal with group actions on non-empty sets X since the group action $\cdot$ otherwise becomes the trivial function from the empty set to the empty set.

Proposition 3.1.2.10. *G_3 defines a group action on $\mathbb{R}^3$.*

Proof. Let $T_1, T_2 \in G_3$ and $x \in \mathbb{R}^3$. Then the function $h : G_3 \times \mathbb{R}^3 \to \mathbb{R}^3$ such that $h(T_1, x) = T_1 \cdot x = T_1(x)$ is a group action of G_3 on $\mathbb{R}^3$ since:

$$\begin{aligned} &T_1 \cdot (T_2 \cdot x) = T_1 \cdot T_2(x) = T_1(T_2(x)) = (T_1 \circ T_2)(x) = (T_1 \circ T_2) \cdot x \\ &I \cdot x = I(x) = x. \end{aligned}$$

□

We will often say that $SO(3)$ is a subgroup of G_3. This is not formally correct since $SO(3)$ is the set of orthogonal matrices M while the subgroup of G_3 we are actually referring to is the set $SO^*(3)$ of functions $T : \mathbb{R}^3 \to \mathbb{R}^3$ of the form Mx with $M \in SO(3)$. However, this distinction is not important since $SO(3)$ and $SO^*(3)$ are isomorphic under $\varphi : SO(3) \to SO^*(3)$ defined by $\varphi(M) = T(x) = Mx$. Note that we could have defined G_3 as the group consisting of elements $M + b$ with $M \in SO(3)$ and $b \in \mathbb{R}^3$ with the group operation $\cdot$ defined by $(M_1 + b_1) \cdot (M_2 + b_2) = M_1(M_2 + b_2) + b_1 = M_1M_2 + (M_1b_2 + b_1)$. We could then have let this G_3 act on $\mathbb{R}^3$ by the group action $\cdot$ defined by $(M_1 + b_1) \cdot x = M_1x + b_1$. With this alternative definition, $SO(3)$ would truly be a subgroup of G_3.

3.1.3 Equidecomposability & Paradoxicality

Let G act on X. If $g \in G$ and $A \subseteq X$, then $g \cdot A$ denotes $\{g \cdot a \in X \mid a \in A\}$. If we also have $H \subseteq G$, then $H \cdot A$ denotes $\{h \cdot a \in X \mid h \in H \text{ and } a \in A\}$.

Definition 3.1.3.1. Let G be a group acting on a non-empty set X and let $A, B \subseteq X$. Then A and B are called *G-equidecomposable* if there exists a bijection $f : A \to B$, an $n \in \mathbb{N}$ and a partition $\{A_i \mid 1 \leq i \leq n\}$ of A such that: For all i, there exists $g_i \in G$ such that for all $a \in A_i$, the equality $f(a) = g_i \cdot a$ holds. We denote that A and B are G-equidecomposable by $A \sim_G B$.

We say that A and B are equidecomposable when it is clear what the acting group G is and we then write $A \sim B$. We say that $(\{A_i \mid 1 \leq i \leq n\}, \{B_i \mid 1 \leq i \leq n\})$ is an *A-B-equidecomposition*. We often just write $(\{A_i\}, \{B_i\})$ and call it an equidecomposition. Moreover, we say that f is an *A-B-equibijection* or simply an equibijection and say that this equibijection defines, or is defined by, $A \sim B$. Note that there is not necessarily a unique equibijection f which is defined by $A \sim B$.

Proposition 3.1.3.2. *Let G be a group acting on a non-empty set X and let $A, B \subseteq X$. Then $A \sim B$ if and only if there exist partitions $\{A_i \mid 1 \leq i \leq n\}$ and $\{B_i \mid 1 \leq i \leq n\}$ of A and B and elements g_i of G such that $g_i \cdot A_i = B_i$.*

Proof. Assume $A \sim B$ and let f be a corresponding equibijection. Define $B_i = f(A_i) = g_i \cdot A_i$ for all $i \leq n$. The B_i are disjoint since f is injective. Since f is surjective, $\bigcup\limits_{i \leq n} B_i = B$ holds, finishing the proof of the right direction.

Conversely, assume there exist partitions $\{A_i \mid 1 \leq i \leq n\}$ and $\{B_i \mid 1 \leq i \leq n\}$ of A and B and elements g_i of G such that $g_i \cdot A_i = B_i$. Now for $a \in A_i$, define $f : A \to B$ by $f(a) = g_i \cdot a$. Then f is surjective since:

$$b \in B \Rightarrow \exists i \leq n \ (b \in B_i) \Rightarrow \exists a \in A_i \ (b = g_i \cdot a = f(a))$$

Moreover, $f \restriction A_i$ is injective since if $a_i, a_i' \in A_i$:

$$f(a_i) = f(a_i') \iff g_i \cdot a_i = g_i \cdot a_i' \iff (g_i^{-1} g_i) \cdot a_i = (g_i^{-1} g_i) \cdot a_i' \iff a_i = a_i'$$

Therefore f is injective since for all $i \neq j$ and all $a_i \in A_i$ and all $a_j \in A_j$, the relations $f(a_i) = g_i \cdot a_i \in g_i \cdot A_i = B_i$ and $f(a_j) = g_j \cdot a_j \in g_j \cdot A_j = B_j$ hold with B_i and B_j being disjoint by assumption. $\square$

Proposition 3.1.3.3. *Let G be a group acting on a non-empty set X. Then the relation of being equidecomposable induces an equivalence relation on $\mathcal{P}(X)$.*

Proof. We prove that $\sim$ satisfies reflexivity, symmetry and transitivity:

- Reflexivity: For every $A \in \mathcal{P}(X)$, $e \cdot A = A$ holds which gives $A \sim A$.
- Symmetry: Assume $A \sim B$ and let $(\{A_i\}, \{B_i\})$ be an equidecomposition. For all i, $B_i = g_i \cdot A_i$ gives $g_i^{-1} \cdot B_i = A_i$, thus proving $B \sim A$.
- Transitivity: Assume $A \sim B$ and $B \sim C$. Let f and h be the respective equibijections. Let $(\{A_i\}, \{B_i\})$ and $(\{B_j\}, \{C_j\})$ be the equidecompositions defined by f and h and denote the respective elements of the group G by g_i and h_j. Assume there are n distinct i and m distinct j, it follows that

there are at most nm distinct products $h_j g_i = g_{i,j}$. Since $h \circ f : A \to C$ is a bijection (as it is the composition of two bijections), it follows that $h \circ f$ is an A-C-equibijection since we can partition A into at most nm sets $A_{i,j} = \{a \in A \mid (h \circ f)(a) = g_{i,j}\}$.

$\square$

By the proof of symmetry, we see that f is an A-B-equibijection if and only if f^{-1} is a B-A-equibijection. Moreover, if $A \sim B$ and $A' \subseteq A$, then there exists $B' \subseteq B$ such that $A' \sim B'$ as we may simple restrict the A-B-equibijection f to A' and redefine its codomain to $f(A')$ in order to obtain an A'-B'-equbijection g.

Proposition 3.1.3.4. [2] *Let G be a group acting on a non-empty set X and let $A, B \subseteq X$. If there exists $A' \subseteq A$ and $B' \subseteq B$ such that $A \sim B'$ and $B \sim A'$, then $A \sim B$.*

Proof. Since $A \sim B'$ holds, there exists an injective $f : A \to B$ defined by $A \sim B'$. Similarly there exists an injective $g : B \to A$ defined by $B \sim A'$.

For $a \in A$, call $b \in B$ a 1-parent of a if $b = g^{-1}(a)$. By recursion, call b a $2n+1$-parent of a if $b^* \in B$ is a $2n-1$-parent of a and $b = g^{-1}(f^{-1}(b^*))$.

For $a \in A$, call $a' \in A$ a 0-parent if $a' = a$. By recursion, call a' a $2n$-parent if $a^* \in A$ is a $2n-2$-parent of a and $a' = f^{-1}g^{-1}(a^*)$.

Define the corresponding concepts for elements of B.

For $a \in A$ and $m \in \mathbb{N}$, define a_m to be the set of m-parents of a and let:

$$A_e = \{a \in A \mid \text{the greatest non-empty } a_m \text{ occurs for an even } m\}$$

Define A_o similarly to be the set of a such that the greatest non-empty a_m occurs when m is odd. Furthermore let A_∞ be the set of a such that a_m is non-empty for all m. Now repeat the process for B by defining B_e, B_o and B_∞ similarly.

Since a given element $a \in A$ can be sent back and forth by the inverse functions g^{-1} and f^{-1} successively at most either an even, an uneven or infinitely many times (and since these options are mutually exclusive), it follows that $\{A_e, A_o, A_\infty\}$ partitions A. Similarly $\{B_e, B_o, B_\infty\}$ partitions B.

Since A_e is the set of elements of A which can be sent back and forth successively by g^{-1} and f^{-1} an even number of times, $f(A_e)$ will be the set of elements of B which can be sent back and forth by f^{-1} and g^{-1} an odd number of times, so $f(A_e) = B_o$. Similarly we have $g(B_e) = A_o$ which implies $B_e = g^{-1}(A_o)$. Again using similar reasoning, we also have $f(A_\infty) = B_\infty$.

Now let:

$$h(x) : A \to B, \ h(x) = \begin{cases} f(x), \ x \in A_e \cup A_\infty \\ g^{-1}(x), \ x \in A_o \end{cases}$$

[2] Corresponds to Proposition G.2 in [Coh2013].

By the above considerations, h is clearly a bijection. Moreover, h is an equibijection: By the remarks preceding this proposition, f defines an equidecomposition of $A_e \cup A_\infty$ and $B_o \cup B_\infty$, while g^{-1} defines an equidecomposition of A_o and B_e. We visually illustrate why this implies that h is an A-B-equibijection:

$$\begin{array}{ccc}
\left.\begin{array}{r} A_e \\ A_\infty \end{array}\right\} & \{A_i\} \xrightarrow{g_i} \{B_i\} & \left\{\begin{array}{l} B_o \\ B_\infty \end{array}\right. \\[2ex]
A_o\} & \{A'_i\} \xrightarrow{g'_i} \{B'_i\} & \{B_e
\end{array}$$

As there are only finitely many A_i and A'_i, it follows that

$$(\{A_1, ..., A_n, A'_1, ...A'_m\}, \{f(A_1), ..., f(A_n), f(A'_1), ..., f(A'_m)\})$$

is an A-B-equipartition. $\square$

Before formally stating the seemingly more general version of BTP, we remember some basic topological ideas about $\mathbb{R}^3$: For $x \in \mathbb{R}^3$ and $r \in \mathbb{R}$, the *open ball* $B(x, r)$ is the set $\{y \in \mathbb{R}^3 \mid |y - x| < r\}$. A set $X \subseteq \mathbb{R}^3$ is *bounded* if there exists $r \in \mathbb{R}$ such that $X \subseteq B(0, r)$. The *interior* of X is the set of points x such that for some $\varepsilon > 0$, the open ball $B(x, \varepsilon)$ is contained in X.

Definition 3.1.3.5 (**Banach-Tarski Paradox in Equidecomposable Form - BTP**). Let A and B be subsets of $\mathbb{R}^3$ that are bounded and have non-empty interiors. Then A and B are G_3-equidecomposable.

We will state the paradox in its usual characterization in Definition 3.1.3.8 below, we will then show that the different forms of BTP are equivalent in Theorem 3.1.3.9.

Definition 3.1.3.6. Let G be a group acting on a non-empty set X and let $A \subseteq X$. Then A is *G-paradoxical* if there exists disjoint subsets A_1 and A_2 such that $\{A_1, A_2\}$ partitions A and $A \sim A_1$ and $A \sim A_2$ hold.

If a set A is G-paradoxical we often say that it is paradoxical and call $\{A_1, A_2\}$ a *paradoxical partition* of A.

Proposition 3.1.3.7. [3] *Let G be a group acting on a non-empty set X. Let $A \subseteq X$ and let A_1 and A_2 be disjoint subsets of A. If $A \sim A_1$ and $A \sim A_2$, then A is paradoxical.*

Proof. We have to find two disjoint subsets of A (not necessarily A_1 and A_2) such that they partition A and are each equidecomposable with A:

Consider the set $A \setminus A_1$, we have $A_2 \subseteq (A \setminus A_1) \subseteq A$. By assumption,

$$A \sim A_2 \subseteq (A \setminus A_1)$$

[3] Corresponds to Corollary G.4 in [Coh2013].

Since equidecomposability is reflexive,

$$A \setminus A_1 \sim A \setminus A_1 \subseteq A$$

We can thus apply Proposition 3.1.3.4 to conclude that $A \sim A \setminus A_1$. By assumption we also have $A \sim A_1$ and since $\{A \setminus A_1, A_1\}$ partitions A, the proof is finished. $\square$

Definition 3.1.3.8 (Banach-Tarski Paradox in Paradoxical Form - BTP). The closed unit ball $B = \{x \in \mathbb{R}^3 \mid |x| \leq 1\}$ is G_3-paradoxical.

This is the form of BTP we will prove. The proof requires some preparation and the final proof of the theorem is delayed until subsection 3.1.5. We now prove that this form of BTP is equivalent to the earlier stated form:

Theorem 3.1.3.9. [4] *The Equidecomposable and Paradoxical forms of BTP are equivalent.*

Proof. For the right direction, split B into two disjoint subsets, for example by cutting it in the y-z-plane so we have: $B_1 = \{x \in B \mid x \leq 0\}$ and $B_2 = \{x \in B \mid x > 0\}$. Then the Equidecomposable form of BTP implies that $B \sim B_1$ and $B \sim B_2$.

For the left direction, first note that if B is G_3-paradoxical, then so are all closed balls of $\mathbb{R}^3$. This is intuitively obvious since the geometry of B does not change depending on what numbers we label our axes with (as long as every axis is scaled by the same magnitude). Neither where we put the origin affect the geometry of B. However, to be rigorous we prove this formally:

Assume that B is G_3-paradoxical and let B' be a ball with radius $r > 0$ having its center at $c \in \mathbb{R}^3$. Thus $B' = rB + c$. Define $\tau : \mathbb{R}^3 \to \mathbb{R}^3$ by $\tau(x) = rx + c$. Let $\{C, D\}$ be a paradoxical partition of B and let f_B be defined by $B \sim C$. Furthermore, let $(\{B_i\}, \{C_i\})$ be an B-C-equidecomposition with n distinct i. Define $\{\tau(B_i) = B'_i\}$ and $\{\tau(C_i) = C'_i\}$. We will prove that for each $i \leq n$, there exists $g'_i \in G_3$ such that $g'_i \cdot B'_i = C'_i$ holds. If $C_i = f_B(B_i) = g_i \cdot B_i = M_i B_i + b_i$, then define:

$$f_{B'}(x) = M_i x + (rb_i + c - M_i c)$$

Denote $rb_i + c - M_i c$ by b'_i and let $M_i + b'_i = g'_i$, then:

$$\begin{aligned} g'_i \cdot B'_i = M_i B'_i + b'_i &= M_i(rB_i + c) + rb_i + c - M_i c = \\ &= r(M_i B_i + b_i) + c = rC_i + c = C'_i \end{aligned}$$

Implying that $(\{B'_i\}, \{C'_i\})$ is an B'-C'-equidecomposition. A similar results holds for B' and $\tau(D) = D'$ and since $\{C', D'\}$ partitions B' (this argument holds since τ is bijective), it follows that B' is paradoxical.

[4] Corresponds to the text at the end of p. 420 and the beginning of p. 421 in [Coh2013].

We have now proved that every closed ball of $\mathbb{R}^3$ is G_3-paradoxical. Now consider any two sets E and F with the properties described in the Equidecomposable form of BPT. We want to prove $E \sim F$. Since E has a non-empty interior, there exists x in E such that some open ball $B(x, r)$ is contained in E. $B(x, r)$ is a ball of $\mathbb{R}^3$ and thus paradoxical, let $\{P_1, P_2\}$ be a paradoxical partition of $B(x, r)$. Now define the sequence $(B_i)_{i=0}^{\infty}$ such that $B_i = B(x + i2re_1, r)$, then $\{B_i\}$ is a sequence of disjoint balls. For any i, B_i is just a translation of $B(x, r) = B_0$ and thus B_i is paradoxical and has a paradoxical partition $\{P_1^i, P_2^i\}$ where P_1^i and P_2^i simply are P_1 and P_1 translated. We use the observation of the preceding sentence to prove that $B_0 \sim \bigcup_{i \leq n} B_i$ is true for all n:

$B_0 \sim B_0$ holds. Assume $B_0 \sim \bigcup_{i \leq n-1} B_i$ holds, then:

$$\bigcup_{i \leq n-1} B_i = \bigcup_{i \leq n-2} B_i \cup B_{n-1} = \bigcup_{i \leq n-2} B_i \cup P_1^{n-1} \cup P_2^{n-1}$$

Since $\sim$ is reflexive and since the following relations hold:

$$P_1^{n-1} \sim B_{n-1}$$
$$P_2^{n-1} \sim P_2^n \sim B_n$$

And all involved sets are disjoint, we obtain:

$$\bigcup_{i \leq n-1} B_i \sim \bigcup_{i \leq n} B_i$$

Thus by induction, $B_0 \sim \bigcup_{i \leq n} B_i$ holds for all $n \in \mathbb{N}$.

Since F is bounded, we can choose n so large that we can move each individual B_i of the union $\bigcup_{i \leq n} B_i$ to a B_i' so that $F \subseteq \bigcup_{i \leq n} B_i'$. We note that $B_i \sim B_i'$ holds for each i (since B_i' is simply a translation of B_i) but that the different B_i' are not necessarily disjoint. They are however easily made disjoint by investigating all intersections of the form $B_i' \cap B_j'$ and removing the intersection from one of the involved sets. Since n is finite, this process of removing intersections will come to an end. We call the newly constructed disjoint sets H_i'. Note that each H_i' is a subset of a unique B_i'.

Since $B_i \sim B_i'$ and $H_i' \subseteq B_i'$, we see that there is a subset H_i of B_i such that H_i is equidecomposable with H_i'. Moreover:

$$\bigcup_{i \leq n} H_i \subseteq \bigcup_{i \leq n} B_i$$

Since $B_0 \sim \bigcup_{i \leq n} B_i$, it follows that there is a subset S of B_0 which is equidecomposable with $\bigcup_{i \leq n} H_i$. Thus by disjointness properties, the following holds:

$$\bigcup_{i \leq n} H_i' \sim \bigcup_{i \leq n} H_i \sim S$$

Moreover, $F \subseteq \bigcup_{i \leq n} H'_i$ holds. Thus there is a $Z \subseteq S \subseteq B_0$ such that $F \sim Z$.

Similarly proving that E is equidecomposable with a subset of an open ball of F lets us apply Proposition 3.1.3.4, giving $E \sim F$. $\square$

We will now prepare for our proof of BTP by discussing some algebraic concepts.

3.1.4 Generators

Definition 3.1.4.1. Let G be a group. If S is a non-empty subset of G, then the smallest subgroup of G which includes S is called the subgroup *generated* by S, we denote it by $\langle S \rangle$. Moreover, if G is a group, $S \subseteq G$ and $\langle S \rangle = G$, then we say that the group G is generated by S.

If G is a group and $S \subseteq G$, define $S^{-1} = \{s^{-1} \in G \mid s \in S\}$. We have $S^{-1} \subseteq \langle S \rangle$ since every element of $\langle S \rangle$ must have its inverse in $\langle S \rangle$. Note that $\langle S \rangle$ is unique since if there would exist two distinct $\langle S \rangle$, then they would be subsets of each other and thus equal.

Definition 3.1.4.2. Let G be a group and $S \subseteq G$. A *word* over S is a sequence $(s_i)_{i=1}^n$ denoted as $s_1...s_n$ such that $n \geq 1$ and $s_i \in S \cup S^{-1}$. We also define the empty sequence to represent e. We define Seq_S to be the set of words over S. We define the function $\Upsilon : Seq_S \to G$ by letting Υ map an element $s_1...s_n \in Seq_S$ to the element g of G which is the result of the group product of the elements of the sequence.

Note that $\Upsilon(Seq_S)$ is a subgroup of G since $e \in \Upsilon(Seq_S)$ and given $g, h \in \Upsilon(Seq_S)$, there exist sequences $s_g, s_h \in Seq_S$ representing s and z and thus the group product sz can also be represented by such a sequence. Furthermore, each element $g \in \Upsilon(Seq_S)$ has its inverse in $\Upsilon(Seq_S)$ since if $s_g = s_0...s_n$ is a representation of g, then there exists $g_s^{-1} = s_n^{-1}...s_0^{-1} \in Seq_S$ and thus $\Upsilon(g_s^{-1}) = g^{-1}$.

Proposition 3.1.4.3. *Let G be a group and let S be a non-empty subset of G, then $\Upsilon(Seq_S) = \langle S \rangle$.*

Proof. Let $s_1...s_n \in Seq_S$. Since all s_i satisfy $s_i \in (S \cup S^{-1}) \subseteq \langle S \rangle$ and $\langle S \rangle$ is closed under the group operation, we get $\Upsilon(s_1...s_n) \in \langle S \rangle$. Since $e \in \langle S \rangle$, we obtain $\Upsilon(Seq_S) \subseteq \langle S \rangle$.

To prove that equality holds, assume $\Upsilon(Seq_S) \subset \langle S \rangle$. Then since $S \subseteq \Upsilon(Seq_S)$ and $\Upsilon(Seq_S)$ is a subgroup of G, we have reached a contradiction against $\langle S \rangle$ being the least subgroup of G containing S. $\square$

Definition 3.1.4.4. Let G be a group generated by a non-empty S such that $S \cap S^{-1} = \emptyset$. Then a *reduced word* over S is a word $s_1...s_n$ such that for no $1 \leq i \leq n-1, s_i^{-1} = s_{i+1}$ holds.

A reduced word is simply a word with no redundant identity elements. The condition $S \cap S^{-1}$ guarantees that such identity elements do not arise from combining elements of S and that $e \notin S$.

Definition 3.1.4.5. Let G be generated by a non-empty S such that $S \cap S^{-1} = \emptyset$. If each $g \in G$ can be represented by a unique reduced word over S, then we say that G is *free* on S or that G is *freely generated* by S. If S has n elements we say that G is *free on n generators*.

For the following proposition, remember that the operation of a group G can be seen as a group action on itself.

Proposition 3.1.4.6. [5] *Let F be a group free on two generators. Then the group set F is F-paradoxical.*

Proof. Let σ and τ generate F. Define F_σ as the set of elements which can be represented by words of F that begin with σ. Define $F_{\sigma^{-1}}, F_\tau$ and $F_{\tau^{-1}}$ similarly. Then $\{\{e\}, F_\sigma, F_{\sigma^{-1}}, F_\tau, F_{\tau^{-1}}\}$ is a partition of F since F is free on $\{\sigma, \tau\}$.

By Proposition 3.1.3.7, to finish the proof we only have to find disjoint subsets A_1 and A_2 of F such that $F \sim A_1$ and $F \sim A_2$: Let $A_1 = F_\sigma \cup F_{\sigma^{-1}}$, then since $e \cdot F_{\sigma^{-1}} = F_{\sigma^{-1}}$ and $\sigma \cdot (\{e\} \cup F_\sigma \cup F_\tau \cup F_{\tau^{-1}}) = F_\sigma$, we get $F \sim A_1$. We prove $F \sim A_2 = F_\tau \cup F_{\tau^{-1}}$ analogously. □

Let G be a group generated by S. Then every element of G can be represented by a word over S and by successively removing redundant identity elements, we obtain a reduced word over S (remember that a word is finite, thus this process works). Thus we only need to prove that the representation of elements of G by reduced words over S is unique in order to prove that G is free on S.

Proposition 3.1.4.7. [6] *SO(3) has a subgroup free on two generators.*

Proof. We want to prove that there exists a subgroup of $SO(3)$ which is generated by some $\{\sigma, \tau\} = X \subseteq SO(3)$ such that every element of this subgroup can be uniquely represented by a reduced word over X. Suppose we were given σ and τ and wanted to know whether $\langle X \rangle$ was free or not, then as noted above it would be sufficient to prove that the representation of an element of $\langle X \rangle$ by a reduced word over X is unique. We could in turn prove the uniqueness of such a representation by proving that if two reduced words v and v' of Seq_X were distinct, then they would represent distinct elements $\Upsilon(v)$ and $\Upsilon(v')$ of $SO(3)$. We now describe a method which we, if we we are given specific σ and τ, could use to prove that distinct reduced words over X represent distinct elements of $SO(3)$. We will thereafter apply this method to specific σ and τ.

---

[5] Corresponds to Proposition G.5 in [Coh2013].
[6] Corresponds to Proposition G.6 in [Coh2013].

Let v, v' be distinct reduced words over X. Note that one of v and v' may be the empty sequence e. We may assume that the leftmost elements of v and v' are distinct: If both words begin with the same $\sigma, \tau, \sigma^{-1}$ or τ^{-1}, then we can remove the first letter of the words since this does not affect whether the elements of $SO(3)$ represented by the words are distinct or not as it corresponds to multiplying the elements represented by the words from the left with the same inverse. Since the words are finitely long, we can repeat this process and rename our words so that we end up with distinct reduced words v, v' such that their first letters are also distinct.

Now suppose that we can find an element $u \in \mathbb{R}^3$ and disjoint subsets S_+, S_-, T_+, T_- of $\mathbb{R}^3$ such that the following condition θ is satisfied:

$$\begin{aligned} u \notin S_+ \cup S_- \cup T_+ \cup T_- \\ w = \Upsilon(\sigma ... w_n) \Rightarrow w \cdot u \in S_+ \\ w = \Upsilon(\sigma^{-1} ... w_n) \Rightarrow w \cdot u \in S_- \\ w = \Upsilon(\tau ... w_n) \Rightarrow w \cdot u \in T_+ \\ w = \Upsilon(\tau^{-1} ... w_n) \Rightarrow w \cdot u \in T_- \end{aligned}$$

If we can find such subsets, then we know that $\Upsilon(v)u$ and $\Upsilon(v')u$ will be distinct elements since they will lie in different disjoint subsets of $\mathbb{R}^3$. Since $\Upsilon(v)$ and $\Upsilon(v')$ are elements of $SO(3)$ and thus matrices representing linear transformations, it would follow that $\Upsilon(v)$ and $\Upsilon(v')$ are distinct.

So we will want to find σ and τ and corresponding subsets S_+, S_-, T_+, T_- of $\mathbb{R}^3$ satisfying θ. However, instead of proving that the elements and subsets satisfy θ, we could prove that they satisfy the following condition which implies θ. We label this condition ϕ:

$$\begin{aligned} u \notin S_+ \cup S_- \cup T_+ \cup T_- \\ \sigma \cdot (S_+ \cup S_- \cup T_+ \cup T_- \cup \{u\}) \subseteq S_+ \\ \sigma^{-1} \cdot (S_+ \cup S_- \cup T_+ \cup T_- \cup \{u\}) \subseteq S_- \\ \tau \cdot (S_+ \cup S_- \cup T_+ \cup T_- \cup \{u\}) \subseteq T_+ \\ \tau^{-1} \cdot (S_+ \cup S_- \cup T_+ \cup T_- \cup \{u\}) \subseteq T_- \end{aligned}$$

Note that we need to know that u is not in the union since the sets S_+, S_-, T_+ and T_- will not be disjoint otherwise: For example, assume $u \in S_+$. Then since $\sigma \cdot S_+ \subseteq S_+$, also $\sigma \cdot u \in S_+$ holds. Moreover, $\sigma^{-1} \cdot (\sigma \cdot u) = (\sigma^{-1}\sigma) \cdot u = u \in S_-$ should also hold. Then $S_+ \cap S_- \neq \emptyset$, contrary to our wish of defining them as disjoint sets.

The following rows prove that ϕ implies θ: $w = \Upsilon(\sigma ... w_n)$ gives $w \cdot u = \Upsilon(\sigma ... w_n) \cdot u = \Upsilon(\sigma ... w_{n-1}) \cdot (w_n \cdot u)$ with $w_n \cdot u \in S_+ \cup S_- \cup T_+ \cup T_-$. Rename $w_n \cdot u$ to u' and we have $w \cdot u = \Upsilon(\sigma ... w_{n-1}) \cdot u'$. Repeating this process n times proves the result for S_+: At the last step we will obtain $\sigma \cdot u'$ with u' in the union as above, thus $\sigma \cdot u' \in S_+$ will hold by our specified conditions. Note that if $w = \sigma$, then the assumption $\sigma \cdot \{u\} \subseteq S_+$ ensures $\sigma \cdot u \in S_+$. The implication is proved similarly for the other subsets.

We now define the different objects. Define σ, τ and u as:

$$\sigma = \begin{pmatrix} 3/5 & 4/5 & 0 \\ -4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \tau = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & -4/5 \\ 0 & 4/5 & 3/5 \end{pmatrix}, \qquad u = (0,1,0)^t$$

The definition of a free group requires $X \cap X^{-1} = \emptyset$. Since $\sigma\tau \neq I$, $\sigma\sigma \neq I$ and $\tau\tau \neq I$, disjointness follows.

We also define the subsets S_+, S_-, T_+ and T_- as:

$$\begin{aligned}
S_+ &= \{\frac{1}{5^k}(x,y,z)^t \in \mathbb{R}^3 \mid k \geq 1,\ x \not\equiv 0 \bmod 5,\ 3y \equiv x \bmod 5,\ z \equiv 0 \bmod 5\} \\
S_- &= \{\frac{1}{5^k}(x,y,z)^t \in \mathbb{R}^3 \mid k \geq 1,\ x \not\equiv 0 \bmod 5,\ -3y \equiv x \bmod 5,\ z \equiv 0 \bmod 5\} \\
T_+ &= \{\frac{1}{5^k}(x,y,z)^t \in \mathbb{R}^3 \mid k \geq 1,\ x \equiv 0 \bmod 5,\ 3y \equiv z \bmod 5,\ z \not\equiv 0 \bmod 5\} \\
T_- &= \{\frac{1}{5^k}(x,y,z)^t \in \mathbb{R}^3 \mid k \geq 1,\ x \equiv 0 \bmod 5,\ -3y \equiv z \bmod 5,\ z \not\equiv 0 \bmod 5\}
\end{aligned}$$

For the rest of this proof we will drop mod 5 just to ease notation. If we prove that the subsets do not contain u, are disjoint and satisfy ϕ, then the proof is finished:

- $u \notin S_+ \cup S_- \cup T_+ \cup T_-$:

$(0,1,0)^t$ is not in the union since for all k, $(0,1,0)^t = \frac{1}{5^k}(5^k * 0, 5^k * 1, 5^k * 0)^t$ and $5^k * 0 = 0 \equiv 0$, disqualifying u from $S_+ \cup S_-$ through its x-coordinate and from $T_+ \cup T_-$ through its z-coordinate.

- The subsets are disjoint:

Assume there is an element in the intersection between any two of the sets. Such an element could be in the different subsets through different representations, i.e. $\frac{1}{5^k}(x,y,z)^t$ to qualify for one set and $\frac{1}{5^{k+j}}(5^j x, 5^j y, 5^j z)^t$ with $j \in \mathbb{Z}$ to qualify for another. However, this cannot be the case because this would imply that all coordinates are zero modulo 5 in one of the representations, contradicting that the element is in the set associated with that representation. We may thus assume that if an element is in two or more of the sets, it is so through a single representation $\frac{1}{5^k}(x,y,z)^t$:

An element in both S_+ and S_- which would satisfy $3y \equiv -3y \Rightarrow 6y \equiv 0 \Rightarrow y \equiv 0 \Rightarrow x \equiv 3y \equiv 0$, contrary to $x \not\equiv 0$. Thus $S_+ \cap S_- = \emptyset$ holds and $T_+ \cap T_- = \emptyset$ is proved analogously. Furthermore, obviously $x \equiv 0$ and $x \not\equiv 0$ is not possible so $S_+ \cap T_+ = S_+ \cap T_- = S_- \cap T_+ = S_- \cap T_- = \emptyset$.

- ϕ is satisfied:

$\sigma \frac{1}{5^k}(x,y,z)^t = \frac{1}{5^{k+1}} \begin{pmatrix} 3x+4y \\ 3y-4x \\ 5z \end{pmatrix}$ and if $\frac{1}{5^k}(x,y,z)^t \in S_+$, then:

$$\begin{cases} 3x+4y \equiv 3x+9y \equiv 3x+3x = 6x \\ 3y-4x \equiv x-4x = -3x \equiv 2x \\ 5z \equiv 0 \end{cases}$$

Since $x \not\equiv 0$ we have $6x \not\equiv 0$ (we may prove this by investigating all cases, $x \equiv i \in \{1,2,3,4\}$) and since $3 * 2x = 6x$, we have that $\sigma \frac{1}{5^k}(x,y,z)^t \in S_+$. Proving the similar result for elements $\frac{1}{5^k}(x,y,z)^t$ in S_-, $\in T_+$ and T_- finishes to part of the proof concerning S_+. We could then do corresponding process for T_+. After that, we could find σ^{-1} and τ^{-1} with the help of Gaussian elimination and we could then do the corresponding processes for S_- and T_-, which would finish the proof. For the sake of completeness, we state σ^{-1} and τ^{-1}:

$$\sigma^{-1} = \begin{pmatrix} 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \tau^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & 4/5 \\ 0 & -4/5 & 3/5 \end{pmatrix}.$$

□

3.1.5 Proving the Paradox

In this subsection we work in ZFC.

Definition 3.1.5.1. Let G act on X. Then $x \in X$ is a non-trivial *fixed element* of X under $g \in G$ if $g \neq e$ and $g \cdot x = x$.

Definition 3.1.5.2. Let G act on X. The *orbit* $\mathscr{O}$ of $x \in X$ under G is $\mathscr{O}(x) = \{g \cdot x \in X \mid g \in G\}$.

Proposition 3.1.5.3 (**ZFC**). [7] *Let G be a group such that the group set G is G-paradoxical. Assume that G acts on a set X in such a way that X has no non-trivial fixed points when being acted on by G, then X is G-paradoxical.*

Proof. Define an equivalence relation $-$ on X by $x - y$ if and only if $x \in \mathscr{O}(y)$. We prove that $-$ really is an equivalence relation:

- Reflexivity: $x = e \cdot x \Rightarrow x \in \mathscr{O}(x) \Rightarrow x - x$.
- Symmetry: $x - y \Rightarrow x \in \mathscr{O}(y) \Rightarrow x = g \cdot y \Rightarrow g^{-1} \cdot x = y \Rightarrow y \in \mathscr{O}(x) \Rightarrow y - x$.
- Transitivity: $x - y$ and $y - z \Rightarrow x \in \mathscr{O}(y)$ and $y \in \mathscr{O}(z) \Rightarrow x = g_y \cdot y$ and $y = g_z \cdot z \Rightarrow x = g_y g_z \cdot z = g \cdot z \in \mathscr{O}(z) \Rightarrow x - z$.

[7] Corresponds to Proposition G.7 in [Coh2013].

For $x \in X$, let $[x] = \{y \in X \mid y - x\}$. Obviously $[x] = \mathscr{O}(x)$ holds since $y - x$ if and only if $y \in \mathscr{O}(x)$. Define $\mathscr{X} = \{[x] \in \mathcal{P}(X) \mid x \in X\}$, all of these sets are non-empty since $-$ is an equivalence relation. Thus by AC, there exists a choice function $f : \mathscr{X} \to \bigcup \mathscr{X} = X$ such that for all $[x] \in \mathscr{X}$, we have $f([x]) \in [x]$.

Since G is G-paradoxical, there exists a paradoxical partition $\{G_1, G_2\}$ of G. Let I index $f(\mathscr{X})$ so $f(\mathscr{X}) = \{x_i \mid i \in I\}$. Note that $G \cdot \{x_i \mid i \in I\} = \bigcup\limits_{i \in I} G \cdot x_i = \bigcup\limits_{i \in I} [x_i] = X$, the last equality is valid since f chooses one element from each equivalence class. Since also $G_1 \cup G_2 = G$, we obtain $X = (G_1 \cup G_2) \cdot f(\mathscr{X}) = \big(G_1 \cdot f(\mathscr{X})\big) \cup \big(G_2 \cdot f(\mathscr{X})\big)$. We want to prove that these two sets of the last equality form a paradoxical partition of X, thus we first prove that they are disjoint:

Assume they are not. Then there exists x in the intersection such that $x = g_1 \cdot x_i = g_2 \cdot x_j$ with $g_1 \neq g_2$ since $G_1 \cap G_2 = \emptyset$. Apply g_2^{-1} to both sides of the equality to obtain $g_2^{-1} g_1 \cdot x_i = x_j$. Now assume $x_i = x_j$. Then $g_2^{-1} g_1 x_i = x_i$ and since X by assumption has no non-trivial fixed points, we get $g_2^{-1} g_1 = e$ so $g_1 = g_2$ contrary to the fact that G_1 and G_2 are disjoint. Instead assume $x_i \neq x_j$, then $g_2^{-1} g_1 x_i = x_j$ gives $g x_i = x_j$ for some $g \in G$ which gives $x_j \in \mathscr{O}(x_i) = [x_i]$ so $[x_j] = [x_i]$ which contradicts that x_j and x_i are chosen from distinct equivalences classes.

Thus $\big(G_1 \cdot f(\mathscr{X})\big) \cap \big(G_2 \cdot f(\mathscr{X})\big) = \emptyset$ holds. Furthermore, since $\{G_1, G_2\}$ is a paradoxical partition of G, there exist an G-G_1-equidecomposition $(\{A_i\}, \{A_i^1\})$. Thus $X = G \cdot f(\mathscr{X}) = \bigcup\limits_{i \leq n} (A_i \cdot f(\mathscr{X}))$ holds and we can prove that the sets $A_i \cdot f(\mathscr{X})$ are disjoint from each other by similar reasoning as the last paragraph. Thus the sets $A_i \cdot f(\mathscr{X}) = X_i$ defines a partition of X. Similarly, $G_1 \cdot f(\mathscr{X}) = \bigcup\limits_{i \leq n} (A_i^1 \cdot f(\mathscr{X}))$ holds and the sets of the right-hand side defines a partition of the set of the left-hand side. Since $(\{A_i\}, \{A_i^1\})$ is an G-G_1-equidecomposition, there exist $n \in \mathbb{N}$ distinct g_i such that:

$$\bigcup_{i \leq n} \big(A_i^1 \cdot f(\mathscr{X})\big) = \bigcup_{i \leq n} \big(g_i A_i \cdot f(\mathscr{X})\big) = \bigcup_{i \leq n} g_i \cdot (A_i \cdot f(\mathscr{X})) = \bigcup_{i \leq n} g_i X_i$$

Thus $X \sim G_1 \cdot f(\mathscr{X})$ and we prove $X \sim G_2 \cdot f(\mathscr{X})$ analogously to finish the proof. $\square$

Before proceeding with our proof of BTP we remember some fundamental principles about the cardinality of sets. Note that if countable sets are involved in any proposition or definition, we will generally consider them to be countably infinite since the finite cases become trivial. Also note that every subset of a countable set is countable (simply restrict both the domain and the codomain of the concerned bijection to prove this statement).

Definition 3.1.5.4 (CUT). Countable unions of countable sets are countable.

The above statement will be discussed more in Section 4.1. We will see that it cannot be proved in ZF without CC. For now, we view it as a valid proposition as proved in Proposition 2.12 in [Rud76].

The following proposition is provable in ZF by using induction, see Proposition 2.13 in [Rud76]:

Proposition 3.1.5.5. *Let X be a finite family of countable sets and let I be a finite index set of X. Then the Cartesian product of X with respect to I is countable.*

Proposition 3.1.5.6. *If X is uncountable and $S \subseteq X$ is countable, then $X \setminus S$ is uncountable.*

Proof. Assume $X \setminus S$ is countable. Then X is the union of two countable sets, $X = S \bigcup (X \setminus S)$, so by CUT, X is countable contrary to assumption. □

For the following propositions, let S denote the unit sphere of $\mathbb{R}^3$, i.e. $S = \{x \in \mathbb{R}^3 \mid |x| = 1\}$. Also note that the fixed points of a rotation of $\mathbb{R}^3$ is the points on the rotation axis.

Proposition 3.1.5.7 (ZFC). [8] *Let F be a subgroup of $SO(3)$ free on two generators. Then there exists a countable $D \subseteq S$ such that $S \setminus D$ is F-paradoxical and thus $SO(3)$-paradoxical.*

Proof. Let F be generated by $\{\sigma, \tau\} = G$ and define $W = \{\sigma, \tau, \sigma^{-1}, \tau^{-1}\}$. Every $f \in F$ represents a rotation of $\mathbb{R}^3$ about a line through the origin since F is a subset of $SO(3)$. A non-trivial rotation f has exactly two fixed points on S, the two points where the axis of rotation intersects the sphere. Define D to be the set of non-trivial fixed points on S under elements of F. An arbitrary element of F can be represented by a unique word over G of some length n, this word can in turn be uniquely represented by an element of the Cartesian product $\underbrace{W \times ... \times W}_{n \text{ times}}$. Proposition 3.1.5.5 implies that this Cartesian product is countable. By CUT, the union of all such finite Cartesian products is countable and it follows that F is countable since we can define an injection from F to this union by sending an element in F to the associated ordered pair. Moreover, since each element in F (except the identity) corresponds to exactly two elements in D, it follows from CUT that D is countable.

By construction, $S \setminus D$ has no non-trivial fixed points with respect to F and is uncountable by Proposition 3.1.5.6. Also, the set F is F-paradoxical by Proposition 3.1.4.6. Thus we can apply Proposition 3.1.5.3 and conclude that $S \setminus D$ is F-paradoxical and hence $SO(3)$-paradoxical since F is a subgroup of $SO(3)$. □

The quoted parts of the following proof are taken literally from Proposition G.9 in [Coh2013]:

Proposition 3.1.5.8 (ZFC). *The sphere S is $SO(3)$-paradoxical.*

[8] Corresponds to Proposition G.8 in [Coh2013].

Proof. "Let F be a subgroup of $SO(3)$ that is free on two generators, and let D be a countable subset of S such that $S \setminus D$ is F-paradoxical [...]. We begin the proof by constructing an element ρ_0 of $SO(3)$ such that the sets $D, \rho_0(D), \rho_0^2(D), ...$ are disjoint. First we choose as axis for ρ_0 a line L that passes through the origin but through none of the points in D."

Such a line L exists, which is proved by the following: Assume no such line exists. Then for every $x \in S \setminus D$, the line L_x of $\mathbb{R}^3$ that goes through x and the origin passes through one unique point $y \in D$ (uniqueness follows from the geometry of the problem, more specifically from that $D \subseteq S$ and that L only intersects S at two points). Define a function $f : S \setminus D \to D$ by $f(x) = y$, where y is as specified in the previous sentence. Then f is injective since distinct $x_1, x_2 \in S \setminus D$ yield distinct lines L_{x_1} and L_{x_2} and thus distinct values $f(x_1) = y_1$ and $f(x_2) = y_2$. Now restrict the codomain of f to $f(S \setminus D)$ to construct a bijection between $S \setminus D$ and a subset of D, implying that $S \setminus D$ is countable contrary to the fact the it is uncountable.

> "We can describe the nontrivial rotations with axis L in terms of values (i.e., angles) in the interval $(0, 2\pi)$. For each pair of points x, y in $S \setminus D$ there is at most one rotation about L that takes x to y. Thus there are only countably many rotations ρ about L for which $D \cap \rho(D)$ is nonempty. A similar argument show that for each n there are at most countably many rotations ρ for which $D \cap \rho^n(D)$ is nonempty."

Let r_1 denote the set of rotations ρ about L such that $D \cap \rho(D) \neq \emptyset$. Denote the rotation ρ about L which takes some $x \in D$ to some $y \in D$ by (x, y), this notation induces a surjection from a subset A of $D \times D$ to r_1. Since $D \times D$ is countable and thus well-orderable, it follows that there exists an injection from r_1 to a subset of A, r_1 is therefore also countable. More generally, denote the rotation ρ about L such that ρ^n takes some $x \in D$ to some $y \in D$ by (x, y) and denote the set of rotations ρ about L such that $D \cap \rho^n(D) \neq \emptyset$ by r_n, then repeat the above argument to obtain the countability of r_n.

> "Since there are uncountably many rotations about L, we can choose a rotation ρ_0 such that for every n the sets D and $\rho_0^n(D)$ are disjoint."

Denote the set of rotations about L by $\mathscr{R}(L)$. This set is uncountable since it has the same cardinality as $(0, 2\pi)$. By CUT, $\bigcup_{n \in \mathbb{N}^+} r_n$ is countable, so $\mathscr{R}(L) \setminus \bigcup_{n \in \mathbb{N}^+} r_n$ is uncountable by Proposition 3.1.5.6. All elements of this set will satisfy the condition described in the last quoted sentence.

> "It follows that for all k and n the sets $\rho_0^k(D)$ and $\rho_0^{k+n}(D)$ are disjoint, and hence that the sequence $D, \rho_0(D), \rho_0^2(D), ...$ consists of disjoint sets."

To prove the above sentence, we prove that an arbitrary element ρ_0 of $\mathscr{R}(L) \setminus \bigcup_{n\in\mathbb{N}^+} r_n$ satisfies $\rho_0^j(D)\cap\rho_0^k(D) = \emptyset$ for all j and all k, this far we have only proved that ρ_0 satisfies $D\cap\rho_0^n(D) = \emptyset$ for all n. Assume that for some distinct j and k, we have $\rho_0^j(D)\cap\rho_0^k(D) \neq \emptyset$, then there exists (not necessarily distinct) $d_1, d_2 \in D$ such that $\rho_0^j(d_1) = \rho_0^k(d_2)$. Assume w.l.o.g that $j < k$ and apply ρ_0^{-j} to both sides of the equality, then $d_1 = \rho_0^{k-j}(d_2)$ implying $\rho_0 \in r_{k-j}$ contradicting $\rho_0 \notin \bigcup_{n\in\mathbb{N}} r_n$.

> "Let $D^{1,\infty} = \bigcup_{i=1}^{\infty} \rho_0^i(D)$ and let $D^{0,\infty} = \bigcup_{i=0}^{\infty} \rho_0^i(D) = D \cup D^{1,\infty}$. Then $S = (S \setminus D^{0,\infty}) \cup D^{0,\infty}$ and $S \setminus D = (S \setminus D^{0,\infty}) \cup D^{1,\infty}$. Since $D^{1,\infty} = \rho_0 \cdot D^{0,\infty}$, it follows that S and $S \setminus D$ are $SO(3)$-equidecomposable [...]."

Obviously, the equidecomposition $S \sim S \setminus D$ defined above is:

$$e \cdot S \setminus D^{0,\infty} = S \setminus D^{0,\infty}$$

$$\rho_0 \cdot D^{0,\infty} = D^{1,\infty}$$

> "Since S and $S\setminus D$ are equidecomposable, while $S\setminus D$ is paradoxical, it follows from [Proposition 3.1.3.7] that S is paradoxical."

Since $S \setminus D$ is paradoxical, it has a paradoxical partition $\{S_1, S_2\}$. Thus:

$$S \sim (S \setminus D) \sim S_1$$

Implying that $S \sim S_1 \subseteq S$ holds. Similarly $S \sim S_2 \subseteq S$ holds and S_1 and S_2 are disjoint by assumption, letting us apply Proposition 3.1.3.7 to finish the proof. □

Proposition 3.1.5.9 (ZFC). [9] *The ball with its center removed, $B' = \{x \in \mathbb{R}^3 \mid 0 < |x| \leq 1\}$, is $SO(3)$-paradoxical.*

Proof. For $E \subseteq S$, define:

$$c(E) = \{s \in \mathbb{R}^3 \mid s = tx \text{ for some } t \in (0, 1] \text{ and some } x \in E\}$$

Clearly $B' = c(S)$ holds. Let $\{S_1, S_2\}$ be a paradoxical partition of S so we have $c(S) = c(S_1\cup S_2) = c(S_1)\cup c(S_2)$ with $c(S_1)\cap c(S_2) = \emptyset$. Furthermore, there exists an S-S_1-equibijection f. Let $(\{A_i\}, \{A_i^1\})$ be the S-S_1-equidecomposition defined by f so $g_i \cdot A_i = A_i^1$. Note that if $T = (0, 1]$, then $\{TA_i\}$ partitions $c(S)$ and $\{TA_i^1\}$ partitions $c(S_1)$. Thus $(\{TA_i\}, \{TA_i^1\})$ is an $c(S)$-$c(S_1)$-equidecomposition since:

$$g_i \cdot (TA_i) = T(g_i \cdot A_i) = TA_i^1$$

The first equality is valid since $\cdot$ represents matrix multiplication. An analogous argument for $c(S_2)$ finishes the proof. □

[9] Corresponds to Proposition G.10 in [Coh2013].

We now prove that the unit ball $B = \{x \in \mathbb{R}^3 \mid |x| \leq 1\}$ is G_3-paradoxical:

Theorem 3.1.5.10 (ZFC). [10] BTP *holds.*

Proof. Let L be a line in $\mathbb{R}^3$ not passing through the origin but lying close enough to the origin so that for any possible rotation ρ about L, we have $\rho(0) \in B$. Such a line obviously exist, just let L be a line such that the minimum distance from L to the origin is strictly less than $\frac{1}{2}$. Now define ρ to be a rotation about L by $\theta \in (0, 2\pi)$ so that $\frac{2\pi}{\theta}$ is irrational, then $0, \rho(0), \rho^2(0), ...$ is a sequence of distinct elements, we prove this statement: Assume w.l.o.g that there exists $i < j$ such that $\rho^i(0) = \rho^j(0)$. Then $0 = \rho^{j-i}(0)$ so ρ^{j-i} is a rotation about L by $2n\pi$. Thus ρ is a rotation about L by $\theta = \frac{2n\pi}{j-i}$ implying $\frac{2\pi}{\theta} = \frac{j-i}{n} \in \mathbb{Q}$ which contradicts our choice of θ.

Note that since ρ is rotation about a line not passing through the origin, $\rho \in G_3 \setminus SO(3)$ holds.

The proof now succeeds similarly as the last part of Proposition 3.1.5.8: Define $0^{0,\infty} = \bigcup_{i=0}^{\infty} \rho^i(0)$ with $\rho^0(0) = 0$ and $0^{1,\infty} = \bigcup_{i=1}^{\infty} \rho^i(0)$. Then $B = (B \setminus 0^{0,\infty}) \cup 0^{0,\infty}$ and $B' = (B \setminus 0^{0,\infty}) \cup 0^{1,\infty}$, thus B and B' are G_3-equidecomposable:

$$e \cdot B \setminus 0^{0,\infty} = B \setminus 0^{0,\infty}$$

$$\rho \cdot 0^{0,\infty} = 0^{1,\infty}$$

Since B' is $SO(3)$-paradoxical and thus G_3-paradoxical, the proof finishes along the same lines as Proposition 3.1.5.8. ☐

3.2 ZF + AD + DC ⇒ ¬ BTP

BTP is probably considered a highly unintuitive result by most people, even many mathematicians consider it unnatural (in [Her2006], Herrlich quotes famous mathematicians expressing their negative attitudes towards BTP). Borel for example considered BTP to constitute a proof of contradiction against AC (more precisely he considered a statement called Hausdorff's Paradox to constitute a contradiction, see p. 188 in [Moo82], but this this paradox is simply a light version of BTP). However, it is important to remember that even though BTP is counterintuitive, it does not prove any inconsistency of ZFC.

In 2005, Wilson even proved that BTP can be strengthened (see [Wil2005]): He proved that in any equidecomposition of A and B, the A_i can be moved continuously to become the B_i without ever colliding with each other (i.e. the pieces will stay disjoint during their paths).

However, BTP becomes more acceptable when you realize that the sets needed in non-trivial equidecompositions are very strange subsets of $\mathbb{R}^3$, namely non-measurable sets. In this section, we will prove that the existence of non-measurable sets is provable only if a choice principle of similar strength as AC

[10] Corresponds closely to Theorem G.11 in [Coh2013].

is present. We will establish this result as follows: We will first prove that BTP implies the existence of non-measurable sets of $\mathbb{R}^3$. Then we will prove that such sets do not exist in ZF + AD + DC, given that the theory is consistent (AD will be defined later in this section). This in turn implies that the existence of non-measurable sets of $\mathbb{R}^3$ is independent of ZF + DC, thus also yielding BTP independent of ZF + DC.

3.2.1 Measure Theory

We refer the reader to [Coh2013] for a proper development of the fundamental concepts of measure theory, here we only state the results we need for our purposes. In this subsection we work in ZF + CC. Moreover, all results we refer to in this subsection can be proven in this theory.

Definition 3.2.1.1. We define the *Lebesgue outer measure* $\lambda^* : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$ in the following way:

Let $I_k^{R_i}$ be a subinterval of $\mathbb{R}$ with included or excluded endpoints $a_k^{R_i}$ and $b_k^{R_i}$,

$$\text{Let } R_i = I_1^{R_i} \times ... \times I_n^{R_i} \subseteq \mathbb{R}^n,$$

$$\text{define } \mathrm{vol}(R_i) = (b_1^{R_i} - a_1^{R_i})...(b_n^{R_i} - a_n^{R_i}).$$

$$\text{For all } X \subseteq \mathbb{R}^3, \text{ define } \mathscr{L}_X = \{(R_i)_{i=1}^{\infty} \mid X \subseteq \bigcup_{i=1}^{\infty} R_i\}$$

$$\text{and } \lambda^*(X) = \inf\{\sum_{i=1}^{\infty} \mathrm{vol}(R_i) \mid (R_i)_{i=1}^{\infty} \in \mathscr{L}_X\}.$$

We define an λ^* for each $\mathbb{R}^n$ such that $n \in \mathbb{N}$, we will use the same notation for all these functions even though they are distinct. $\lambda^*(X)$ is intuitively the greatest lower bound for the volume of a union of n-dimensional cubes containing X. The Lebesgue measure of a n-dimensional cube R_i clearly equals $\mathrm{vol}(R_i)$, i.e. what we intuitively consider to be the volume of such a set. By Proposition 1.3.4 in [Coh2013], the Lebesgue outer measure is:

(1) *Zero at the empty set*: $\lambda^*(\emptyset) = 0$ holds.
(2) *Monotone*: For every $A, B \in \mathbb{R}^n$ such that $A \subseteq B$, $\lambda^*(A) \leq \lambda^*(B)$ holds.
(3) *Countably subadditive*: For every sequence $(A_i)_{i=1}^{\infty}$ of sets belonging to $\mathbb{R}^n$, the inequality $\lambda^*(\bigcup\limits_{i=1}^{\infty} A_i) \leq \sum\limits_{i=1}^{\infty} \lambda^*(A_i)$ holds.

Proposition 3.2.1.2. *Lebesgue outer measure λ^* in $\mathbb{R}^3$ is invariant under rigid motions, i.e. for any element $g \in G_3$ and $X \subseteq \mathbb{R}^3$, the equality $\lambda^*(X) = \lambda^*(g \cdot X)$ holds.*

Proof. The volume of a single cube does not change when being translated or rotated. For a translation we may simply prove this formally by adjusting the endpoints of the underlying intervals $I_k^{R_i}$. To prove it formally in the case of a rotation we need to use the characterization of the determinant as the scale factor of the linear transformation it represents. Moreover, a subset X of $\mathbb{R}^3$ will be covered by a union of cubes if and only if the corresponding translated or rotated X is covered by the corresponding translated or rotated cubes. Since the infimum is taken over the sum of volumes of individual cubes, the proposition follows. $\square$

We now define an important restriction of λ^*:

Definition 3.2.1.3. The family of *Lebesgue measurable subsets* of $\mathbb{R}^n$ is the family $\mathscr{A}$ of sets $A \subseteq \mathbb{R}$ satisfying:

$$\lambda^*(B) = \lambda^*(B \cap A) + \lambda^*(B \cap A^c)$$

For all $B \subseteq \mathbb{R}$. We denote the *Lebesgue measure* $\lambda^* \restriction \mathscr{A}$ by λ.

We will call a set *measurable* if it is Lebesgue measurable. Note that there are several equivalent ways of defining measurability, for example by using inner and outer measures. However, we will not discuss these alternative definitions.

Every n-dimensional cube is measurable by Proposition 1.3.8 in [Coh2013]. The family $\mathscr{A}$ constitutes a *σ-algebra* on $\mathbb{R}^n$, i.e. $\mathbb{R}^n$ belongs to $\mathscr{A}$ and $\mathscr{A}$ is closed under complements, countable unions and countable intersections (see Theorem 1.3.6 (a) in [Coh2013]). Moreover, the Lebesgue measure is *countably additive*, i.e. for each sequence $(A_i)_{i=1}^{\infty}$ of disjoint sets belonging to $\mathscr{A}$, $\lambda(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} A_i$ holds (again, see Theorem 1.3.6 (a) in [Coh2013]).

Theorem 3.2.1.4. *If* BTP *holds, then every bounded subset of* $\mathbb{R}^3$ *with non-empty interior includes at least one bounded non-measurable subset of* $\mathbb{R}^3$*.*

Proof. Let A be any bounded subset of $\mathbb{R}^3$ with non-empty interior and let A' be a subset of A with non-empty interior such that $\lambda^*(A) > \lambda^*(A')$ (such a set A' obviously exists since some open ball $B(x, r)$ is contained in A and letting $B(x, \frac{r}{2}) = A'$ yields a valid A'). Since A' is also necessarily bounded, BTP in Equidecomposable Form implies that there exists an equidecomposition of A and A' into $\{A_i \mid i \leq n\}$ and $\{A'_i \mid i \leq n\}$ with associated g_i. Assume that all A_i and A'_i are measurable, then A and A' are also measurable since σ-algebras are closed under countable unions. Thus our initial assumption can be written as $\lambda(A) > \lambda(A')$. Furthermore, by the countable additivity of the Lebesgue measure and the invariance of the Lebesgue outer measure under rigid motions, the following equalities hold:

$$\lambda(A) = \lambda(\bigcup_{i \leq n} A_i) = \sum_{i \leq n} \lambda(A_i) = \sum_{i \leq n} \lambda^*(g_i \cdot A_i) = \sum_{i \leq n} \lambda(A'_i) = \lambda(\bigcup_{i \leq n} A'_i) = \lambda(A')$$

Which obviously contradicts $\lambda(A) > \lambda(A')$, thus some A_i or A'_i is non-measurable. $\square$

Since ZFC implies BTP, there exist non-measurable sets of $\mathbb{R}^3$ in ZFC. Moreover, clearly similar reasoning as in the proof above holds for any two equidecomposable sets $A \sim B$ such that $\lambda^*(A) \neq \lambda^*(B)$: By assuming that the corresponding equidecompositions $\{A_i\}$ and $\{B_i\}$ only consists of measurable subsets, we will be able to derive a contradiction using the countable additivity of λ and the G_3-invariance of λ^*.

Note that not only λ^* but also λ is G_3-invariant in the sense that for all measurable $A \subseteq \mathbb{R}^3$ and $g \in G_3$, the equality $\lambda^*(A) = \lambda^*(g \cdot A)$ holds and A is measurable if and only if $g \cdot A$ is measurable. The latter assertion intuitively follows from the fact that the characteristics of a set does not change depending on where the origin is relative to the set and is formally proved by letting $g = T(x) = \rho_0 x + c$ with $\rho_0 \in SO(3)$ and $c \in \mathbb{R}^3$ and noting that A satisfies:

$$\lambda^*(\rho_0^{-1}(B - c)) = \lambda^*(\rho_0^{-1}(B - c) \cap A) + \lambda^*(\rho_0^{-1}(B - c) \cap A^c)$$

For all $\rho_0^{-1}(B - c)$ (i.e. for all subsets of $\mathbb{R}^3$) if and only if $\rho_0 A + c$ satisfies:

$$\lambda^*(B) = \lambda^*(B \cap (\rho_0 A + c)) + \lambda^*(B \cap (\rho_0 A + c)^c)$$

For all B. Note that this argument relies on the G_3-invariance of λ^* and the fact that $T(x)$ is a bijective function. This argument is a generalized form of the content of Proposition 1.4.4 in [Coh2013] (which says that the Lebesgue measure of $\mathbb{R}$ is translation invariant).

We will need the following definition and proposition for our work in the next subsections:

Definition 3.2.1.5. If $X \subseteq \mathbb{R}^n$ and $\lambda^*(X) = 0$, then we say that X is *null*.

Proposition 3.2.1.6. [11] *Every null set is measurable.*

Proof. Let A be a null set and let B be some subset of $\mathbb{R}^n$. Note that $B = (B \cap A) \cup (B \cap A^c)$. Thus by the subadditivity of λ^*, the inequality $\lambda^*(B) \leq \lambda^*(B \bigcap A) + \lambda^*(B \bigcap A^c)$ holds. Thus if we prove:

$$\lambda^*(B) \geq \lambda^*(B \cap A) + \lambda^*(B \cap A^c) \qquad (\varphi)$$

Then the proof is finished. Since $\lambda^*(A) = 0$ and λ^* is monotone, $\lambda^*(B \cap A) = 0$ holds. Again referring to the monotonicity of λ^*, the inequality $\lambda^*(B) \geq \lambda^*(B \cap A^c)$ holds, thus φ holds. □

3.2.2 Game Theory

Definition 3.2.2.1. The *Baire space* $\mathcal{N}$ is the metric space $(\mathbb{N}^{\mathbb{N}}, d)$ where $\mathbb{N}^{\mathbb{N}}$ is the set of functions $f : \mathbb{N} \to \mathbb{N}$ and the metric d is defined as $d(f, g) = \frac{1}{2^{n+1}}$ for the least n such that $f(n) \neq g(n)$.

Definition 3.2.2.2 (Two Player $\mathcal{N}$-games). [12] Let $A \subseteq \mathcal{N}$. The *game* G_A is played by letting player I and player II successively choose elements of $\mathbb{N}$. If

[11] Corresponds to Proposition 1.3.5 in [Coh2013] and his remarks made before the statement of the proposition.

[12] This definition is written drawing inspiration from the corresponding definitions in [Jec73], [Kle77] and [Her2006].

we denote the $n+1$-th choice of player I by a_n and similarly denote the $n+1$-th choice of player II by b_n, then playing the game yields an infinite sequence $(a_0, b_0, a_1, b_1, ...) \in \mathcal{N}$ which we call the *outcome* of the game and denote by $(a_n, b_n)_{n=0}^{\infty}$. If $(a_n, b_n)_{n=0}^{\infty} \in A$, then player I *wins* the game and otherwise player II wins the game.

A *strategy* for a player is a complete description of how the player will play the game, i.e. a description of what choice the player will make at any stage of the game given the other player's earlier choices in the game. Thus a strategy for player I is a function $\sigma : \bigcup\{\mathbb{N}^{\{i\in\mathbb{N}|i\leq n\}} \mid n \in \mathbb{N}\} \to \mathbb{N}$ describing that given that $(b_0, ..., b_{n-1})$ has been chosen by player II, the $n+1$-th choice of player I under σ is $\sigma((b_0, ..., b_{n-1})) = a_n$. A strategy for player II is defined similarly. A *winning strategy* of a player is a strategy such that the player wins the game no matter what strategy the other player plays. The game G_A is *determined* if one of the players has a winning strategy.

Definition 3.2.2.3 (**Axiom of Determinateness - AD**)**.** For every $A \subseteq \mathcal{N}$, the game G_A is determined.

We note that a restricted version of CC holds in ZF + AD (see Lemma 12.15 in [Jec73]), namely CC($\mathbb{R}$), the statement that every countable family of non-empty set of reals (real numbers) has a choice function. We can thus view AD as a choice principle.

We should now state a relative consistency theorem saying that ZF and ZF + AD are relatively consistent. However, it is still unknown whether ZF and ZF + AD are relatively consistent, set theorists believe this is the case but there is no proof of it yet (p. 150 in [Her2006]). Therefore, we state the following assumption:

Assumption 3.2.2.4. ZF *and* ZF + AD *are relatively consistent.*

On a more positive note, the following theorem is due to Kechris (see [Kec84]) and we state it without proof:

Theorem 3.2.2.5. ZF + AD *and* ZF + AD + DC *are relatively consistent.*

For the rest of this chapter, we will work in ZF + AD + DC. By our Proposition 2.1.2.14, we see that CC holds in ZF + AD + DC.

3.2.3 Measurability

In this subsection we will prove that all sets of reals are measurable in ZF + AD + DC. All definitions and proofs of this subsection generalize to every $\mathbb{R}^n$, we will discuss this generalization more after having proved Lemma 3.2.3.6.

Definition 3.2.3.1. Let X be a set of reals. A measurable set A such that $X \subseteq A$ and $\lambda^*(X) = \lambda(A)$ is called a *measurable cover of X*.

Proposition 3.2.3.2 (ZF + CC). [13] *Every set X of reals such that $\lambda^*(X) < \infty$ has a measurable cover.*

Proof. Let X be as specified. For every $n \in \mathbb{N}^+$, let $\mathscr{R}_n$ be the set of sequences $(R_i)_{i=1}^{\infty}$ such that $X \subseteq \bigcup\limits_{i=1}^{\infty} R_i$ and $\sum\limits_{i=1}^{\infty} \mathrm{vol}(R_i) < \lambda^*(X) + \frac{1}{n}$. If $\mathscr{R}_n$ is empty for some n, then $\lambda^*(X) + \frac{1}{n}$ is a lower bound for the outer measure of X, implying $\lambda^*(X) + \frac{1}{n} \leq \lambda^*(X)$ which clearly does not hold, thus every $\mathscr{R}_n$ is non-empty. Thus by CC, the set $\{\mathscr{R}_n \mid n \in \mathbb{N}^+\}$ has a choice function, denote the choice from each $\mathscr{R}_n$ by $(R_i^n)_{i=1}^{\infty}$. Now let $A = \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{i=1}^{\infty} R_i^n$. Each $\bigcup\limits_{i=1}^{\infty} R_i^n$ is a countable union of intervals and is thus measurable, thus A is measurable since it is an intersection of countably many measurable sets. Moreover, $X \subseteq A$ holds since X is contained in every $\bigcup\limits_{i=1}^{\infty} R_i^n$. Clearly $A \subseteq \bigcup\limits_{i=1}^{\infty} R_i^n$ also holds for each n. Thus the following holds for each n:

$$\lambda^*(X) \leq \lambda(A) \leq \lambda(\bigcup_{i=1}^{\infty} R_i^n) \leq \sum_{i=1}^{\infty} \lambda(R_i) = \sum_{i=1}^{\infty} \mathrm{vol}(R_i) < \lambda^*(X) + \frac{1}{n}$$

Implying $\lambda^*(X) = \lambda(A)$. $\square$

Proposition 3.2.3.3 (ZF + CC). [14] *Let X be a set of reals such that $\lambda^*(X) < \infty$. Then the measurable cover A of X has the property that if $Z \subseteq A \setminus X$ and Z is measurable, then Z is null.*

Proof. Let Z be a measurable subset of $A \setminus X$. As noted in Proposition 1.2.2 in [Coh2013], $\lambda(A \setminus Z) = \lambda(A) - \lambda(Z)$ holds (partition A into $\{Z, A \setminus Z\}$, apply the countably additivity of λ and rearrange the terms). Since $X \subseteq A \setminus Z$, the inequality $\lambda^*(X) \leq \lambda(A) - \lambda(Z)$ holds. By assumption, $\lambda^*(X) = \lambda(A)$, thus $\lambda(Z) \leq 0$ holds, implying $\lambda(Z) = 0$. $\square$

Definition 3.2.3.4. A set of reals A is *analytic* if there exists a continuous function $f : \mathcal{N} \to \mathbb{R}$ such that $A = f(\mathcal{N})$.

We omit the proof of the following theorem (see Theorem 11.18 in [Jec2006]) as it requires a bit too much development of measure theory. According to Jech, the theorem generalizes to $\mathbb{R}^n$ and even more general measure spaces (Polish spaces).

Proposition 3.2.3.5 (ZF + CC). *Every analytic set is measurable.*

The following lemma constitutes the major part of our proof of the measurability of every set of reals:

Lemma 3.2.3.6 (ZF + AD + DC). [15] *If $S \subseteq [0, 1]$ is a set of reals such that every measurable $Z \subseteq S$ is null, then S is null.*

[13] Corresponds to Theorem 11.9 in [WhZy77].
[14] Corresponds to Lemma 11.13 in [Jec2006].
[15] Corresponds to Lemma 33.4 in [Jec2006].

Before proving the lemma, we first state our main theorem and discuss why the lemma implies the theorem:

To prove that every set of reals is measurable it suffices to prove that every subset of $[0,1]$ is measurable, this is seen as follows: Let X be any set. If $X \cap [z, z+1]$ is measurable for any $z \in \mathbb{Z}$, then X is measurable since $X = \bigcup_{z \in \mathbb{Z}} (X \cap [z, z+1])$ is a countable union of supposedly measurable sets. Thus, if X is non-measurable then there exists some non-measurable $X \cap [z, z+1]$. Since the Lebesgue measure of $\mathbb{R}$ is translation invariant by our discussion after Theorem 3.2.1.4, we can simply translate this non-measurable $X \cap [z, z+1]$ so it is contained in $[0,1]$.

Theorem 3.2.3.7 (ZF + AD + DC). [16] *Every set X of reals is measurable.*

Proof. By the previous remarks, assume $X \subseteq [0,1]$. By Proposition 3.2.3.3, there exists a measurable A such that $X \subseteq A \subseteq [0,1]$ and if $Z \subseteq A \setminus X$ holds with Z being measurable, then Z is null. By Lemma 3.2.3.6, this implies that $A \backslash X$ is null and thus measurable. Lastly, note that $X = A \backslash (A \backslash X) = A \cap (A \backslash X)^c$ is measurable since the measurable sets form a σ-algebra. □

Before proving Lemma 3.2.3.6, we prove one more proposition:

Proposition 3.2.3.8 (ZF + CC). *The set $\mathscr{A}$ of finite unions of intervals with rational endpoints contained in $[0,1]$ is countable.*

Proof. Since $\mathbb{Q}$ is countable, $\mathbb{Q} \times \mathbb{Q}$ is also countable by Proposition 3.1.5.5 and an element (q_1, q_2) such that $q_1 \leq q_2$ can be seen as representing the four possible intervals with corresponding endpoints (the open, closed and the two half-open, half-closed intervals). By CUT, the set A of intervals contained in $[0,1]$ having rational endpoints is thus countable. The set $A_n = \underbrace{A \times ... \times A}_{n \text{ times}}$ is therefore countable and an element $a_1 \times ... \times a_n$ of the set can be viewed as representing $a_1 \cup ... \cup a_n$. This notation defines a surjection from A_n to the set A_n^* of unions of intervals of $[0,1]$ having rational endpoints consisting of at most n disjoint intervals. Since A_n is countable, this implies the existence of an injection from A_n^* to A_n, in turn implying that A_n^* is countable. By CUT, the union $\mathscr{A}$ of all A_n^* is thus countable. □

Proof of Lemma 3.2.3.6. [17] By our above considerations: Let $S \subseteq [0,1]$ be a set of reals such that every measurable subset Z of S is null and let $\varepsilon > 0$. We will prove that $\lambda^*(S) < \varepsilon$, it follows that $\lambda^*(S) = 0$.

Define the following game G: First, remember that $\mathscr{A}$ is the set of finite unions of intervals with rational endpoints contained in $[0,1]$. At each move of

[16] Corresponds to Theorem 33.3 (i) in [Jec2006] and the comments made after his statement of Lemma 33.4.

[17] Corresponds to 33.5 in [Jec2006], the proof is originally due to Leo Harrington according to Jech.

player I, she chooses between the numbers 0 and 1 while player II at his $n+1$-th choice chooses an element H_n from the set:

$$K_n = \{H \in \mathscr{A} \mid \lambda(H) < \frac{\varepsilon}{2^{2(n+1)}}\} \tag{ϕ}$$

Thus the players will generate an outcome $(a_n, H_n)_{n=0}^{\infty}$ of the game.

Moreover, we let the sequence $(a_n)_{n=0}^{\infty}$ generated by player I describe a real number a by letting:

$$a = \sum_{n=0}^{\infty} \frac{a_n}{2^{n+1}} \tag{θ}$$

We define $\pi : \mathcal{N} \to \mathbb{R}$ by letting $\pi((a_n)_{n=0}^{\infty})$ be the corresponding $a \in S$ as described by θ. We define the game G as player I winning if the outcome of the game satisfies the following conditions:

(1) $a_n = 0$ or 1, for all n.
(2) $a \in S$.
(3) $a \notin \bigcup_{n \in \mathbb{N}} H_n$.

We note that our game might not seem to be on the form stated in our definition of a game, we will therefore prove that it actually is on the specified form: By Proposition 3.2.3.8, $\mathscr{A}$ is countable and thus each K_n is at most countable. Moreover, for every n there exists a rational $q \in [0, \frac{\varepsilon}{2^{2(n+1)}})$, for example $q = 1/\left\lceil \frac{2^{2(n+1)}}{\varepsilon} + 1 \right\rceil$. This implies $[0, q] \in K_n$, furthermore an interval with rational endpoints has infinitely many subintervals with rational endpoints and each such interval will have small enough measure. Thus each K_n is countably infinite. By CC, we can choose an enumeration e_n of each K_n, meaning that the game can be coded as player II choosing any natural number $b_n = e_n^{-1}(H_n)$ at each stage of the game just as our definition prescribes.

Furthermore, note that player I is allowed to choose any natural number at each round of the game, however she loses if she chooses any number which is not 0 or 1. By these considerations, the conditions (1)-(3) can be summarized by saying that player I wins the game if the outcome of the game is in the following $X \subseteq \mathcal{N}$:

$$X = \{(a_n, b_n)_{n=0}^{\infty} \in \mathcal{N} \mid \big(\forall n \in \mathbb{N}\ (a_n \in \{0,1\})\big) \land \pi((a_n)_{n=0}^{\infty}) \in S \setminus \bigcup_{n \in \mathbb{N}} e_n(b_n)\}.$$

Thus our game G is on the specified form and can be denoted by G_X. By AD, G_X is determined. We will first prove that player I cannot have a winning strategy:

Assume σ is a winning strategy for player I. If player I uses the strategy σ and player II plays the sequence $(b_n)_{n=0}^{\infty}$, then let $(a_n)_{n=0}^{\infty}$ be the sequence which player I plays. Define the function $f_{\mathcal{N}} : \mathcal{N} \to \mathcal{N}$ by $f_{\mathcal{N}}((b_n)_{n=0}^{\infty}) = (a_n)_{n=0}^{\infty}$. Note that $\pi((a_n)_{n=0}^{\infty})$ must be in S since σ is assumed to be a winning strategy of player I. Now let $f : \mathcal{N} \to \mathbb{R}$ be defined as $\pi \circ f_{\mathcal{N}}$.

Let $Z = f(\mathcal{N}) \subseteq S$. Note that Z contains all possible $a \in S$ which player I can play under σ. Our goal is to define a sequence $(H_n)_{n=0}^{\infty}$ such that $Z \subseteq \bigcup_{n \in \mathbb{N}} H_n$. Then σ cannot be a winning strategy of player I since she loses if she uses the strategy σ and player II plays $(H_n)_{n=0}^{\infty}$.

We begin by proving that f is uniformly continuous and therefore continuous. Thus we will prove that for every $\mu > 0$, there exists $\delta > 0$ such that if b and b' are sequences of player II such that $d(b, b') < \delta$, then $|f(b) - f(b')| < \mu$.

Let $\mu > 0$ be arbitrary. Obviously there is a least $j \in \mathbb{N}$ such that $\frac{1}{2^{j+1}} < \mu$ holds. Define $\delta = \frac{1}{2^j}$ and let b and b' be sequences such that $d(b, b') < \frac{1}{2^j}$. Note that $b_i = b'_i$ holds for all $0 \leq i < j$ by definition of the metric in $\mathcal{N}$. Let $f_{\mathcal{N}}(b) = (a_n)_{n=0}^{\infty}$ and $f_{\mathcal{N}}(b') = (a'_n)_{n=0}^{\infty}$. Since player I uses the strategy σ and the first j choices of player II are the same in the sequences b and b', the equality $a_i = a'_i$ holds for all $0 \leq i \leq j$. Moreover, since σ is a winning strategy of player I, all a_i and a'_i will assume one of the values 0 or 1. Thus, when applying π to the sequences $(a_n)_{n=0}^{\infty}$ and $(a'_n)_{n=0}^{\infty}$, the largest possible distance between the resulting a and a' is obtained by letting $a_i = 0$ and $a'_i = 1$ (or vice versa) for all $i > j$. Thus:

$$|a - a'| \leq \sum_{n=j+1}^{\infty} \frac{1}{2^{n+1}} = \frac{1}{2^{j+1}} \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} = \frac{1}{2^{j+1}} < \mu$$

This finishes the proof of f being uniformly continuous.

Thus Proposition 3.2.3.5 implies that Z is measurable. Since $Z \subseteq S$ holds, Z is null by assumption. Thus we can cover Z with $\bigcup_{n \in \mathbb{N}} H_n$ so each $H_n \in K_n$, we prove this as follows: Let $\frac{\varepsilon}{2^{2(n+1)}} = \varepsilon_n$ and let $(R_i)_{i=0}^{\infty}$ be a sequence of disjoint intervals contained in $[0, 1]$ such that the corresponding union covers Z and the measure ϱ of the union is strictly less than $\frac{\varepsilon_0}{2}$. Note that we may assume that all R_i are disjoint, otherwise we simply make them disjoint by recursion as described later in this thesis in Proposition 4.1.1.1. Let W well-order $\mathbb{Q}$ (remember that $\mathbb{Q}$ is countable). Enlarge each R_i having endpoints v_i and w_i to R'_i having rational endpoints by choosing the W-least v'_i in $(v_i - \frac{\varepsilon_0}{2^{n+3}}, v_i)$ and the W-least w'_i in $(w_i, w_i + \frac{\varepsilon_0}{2^{n+3}})$ as new inclusive endpoints. Note that we have to be careful so these extended intervals do not expand beyond 0 or 1, in that case simply cut them off at 0 or 1. Then $\lambda(R'_i) < \lambda(R_i) + \frac{\varepsilon_0}{2^{n+2}}$, implying:

$$\lambda(\bigcup_{i \in \mathbb{N}} R'_i) \leq \sum_{i \in \mathbb{N}} \lambda(R'_i) < \sum_{i \in \mathbb{N}} \lambda(R_i) + \frac{\varepsilon_0}{2} = \lambda(\bigcup_{i \in \mathbb{N}} R_i) + \frac{\varepsilon_0}{2} = \varrho + \frac{\varepsilon_0}{2} < \varepsilon_0$$

Specifically this implies that all $\lambda(R'_i)$ are strictly less than ε_0. Now make all of the R'_i disjoint, this is possible as earlier noted. All of the endpoints of the intervals will still be rational after the intervals have been made disjoint, we keep the same notation R'_i for the new disjoint intervals. Now rearrange the sequence $(R'_i)_{i=0}^{\infty}$ to a sequence $(T_i)_{i=0}^{\infty}$ of descending measure, i.e. a sequence such that $T_i \leq T_{i+1}$ holds for all i. If two intervals have the same measure, let the one closest to the origin precede the other one (and if they have equal left endpoints, let the right endpoint decide which one comes first). Note that this rearrangement is possible since if the set $\{\lambda(R'_i) \mid i \in \mathbb{N}\}$ would not have a maximal element, then there would exist an infinite sequence of R'_i ascending in measure. Thus there would exist $\nu > 0$ such that for some R'_i of the sequence, $\lambda(R'_i)$ would equal ν and since there would exist infinitely many disjoint R'_i having measure strictly larger than ν, $\lambda(\bigcup_{i\in\mathbb{N}} R'_i)$ would be infinite contradictory to the fact that it is bounded by ε_0. The same reasoning will prove that $\{\lambda(R'_i) \mid i \in \mathbb{N}\}$ minus its maximal element has a maximal element, and so on. We may thus define $(T_i)_{i=0}^{\infty}$ by recursion. Note that:

$$\lim_{n\to\infty} \lambda(\bigcup_{i\geq n} T_i) = \lim_{n\to\infty} \sum_{i\geq n} \lambda(T_i) = \lim_{n\to\infty} \big(\sum_{i\in\mathbb{N}} \lambda(T_i) \; - \sum_{i\leq n-1} \lambda(T_i)\big) = 0$$

Thus for every $\varsigma > 0$, there exists n large enough so $\lambda(\bigcup_{i\geq n} T_i) < \varsigma$. Thus for some n, the measure of $\bigcup_{i\geq n} T_i$ will be less than ε_1. We let $\bigcup_{i\leq n-1} T_i = H_0$, note that $H_0 \in K_0$ since $\lambda(\bigcup_{i\leq n-1} T_i) \leq \lambda(\bigcup_{i\in\mathbb{N}} T_i) < \varepsilon_0$ (since $\bigcup_{i\in\mathbb{N}} T_i = \bigcup_{i\in\mathbb{N}} R'_i$, note that the union of the original R'_i and the union of the disjoint R'_i are equal and thus have equal measure). We define H_1 similarly by removing finitely many T_i from $\bigcup_{i\geq n} T_i$ in successive order until the measure becomes less than ε_2. Formalizing this procedure through recursion defines each H_n and finishes the argument. Note that we have assumed the initial sequence $(R_i)_{i=0}^{\infty}$ to be infinite, however the argument is obviously valid if this sequence would have been finite.

Thus player I cannot have a winning strategy.

Since the game is determined and player I does not have a winning strategy, player II must have a winning strategy τ. Note that player I can construct any $a \in S \subseteq [0,1]$ by choosing 0's and 1's since any real number has a binary representation (see [I-DMS] for a short note about the binary representation of $[0,1]$). Let $\mathscr{H}_n = \{\tau((a_i)_{i=0}^n) \mid (a_i)_{i=0}^n$ is sequence of 0's and 1's$\}$, i.e. $\mathscr{H}_n$ is the set of different choices of player II at the $n+1$-th stage of the game when using the strategy τ and when player I in the previous stages has played 0's and 1's. Then every $a \in S$ must be contained in some H belonging to some $\mathscr{H}_n$. Thus $S \subseteq \bigcup_{n\in\mathbb{N}} \bigcup \mathscr{H}_n$ holds, implying $\lambda^*(S) \leq \lambda^*(\bigcup_{n\in\mathbb{N}} \bigcup \mathscr{H}_n)$. We will prove that the righthand side of the equality is bounded by ε, finishing the proof:

For every $n \in \mathbb{N}$, the number of possible sequences of 0's and 1's of length $n+1$ is 2^{n+1}. Thus $|\mathscr{H}_n| \leq 2^{n+1}$. Moreover, each $H \in \mathscr{H}_n$ satisfies the measure inequality described by ϕ, thus:

$$\lambda(\bigcup \mathscr{H}_n) \leq \sum_{H \in \mathscr{H}_n} \lambda(H) < \frac{2^{n+1}\varepsilon}{2^{2(n+1)}} = \frac{\varepsilon}{2^{n+1}}$$
$$\Longrightarrow$$
$$\lambda(\bigcup_{n\in\mathbb{N}} \bigcup \mathscr{H}_n) \leq \sum_{n\in\mathbb{N}} \lambda(\bigcup \mathscr{H}_n) < \sum_{n\in\mathbb{N}} \frac{\varepsilon}{2^{n+1}} = \varepsilon$$

□

Remark. All definitions and proofs of this subsection generalize to any $\mathbb{R}^n$: The proof of Proposition 3.2.3.2 is valid for $\mathbb{R}^n$. In the proof of 3.2.3.3, the only observation needed for the generalization is to see that $\mathbb{R}^n$ can be partitioned into countably many n-dimensional cubes of unit size with integer coefficients. Such a partition is countable since $\underbrace{\mathbb{Z} \times ... \times \mathbb{Z}}_{n \text{ times}} = \mathbb{Z}_n$ is countable and a cube of unit size in $\mathbb{R}^n$ is defined by 2^n distinct points. Thus $\underbrace{\mathbb{Z}_n \times ... \times \mathbb{Z}_n}_{2^n \text{ times}}$ is a superset of the specified partition and this set is countable. Similar remarks apply to Proposition 3.2.3.8.

Moreover, the game G_X in the proof of Lemma 3.2.3.6 can be generalized to $\mathbb{R}^n$: Letting player I generate the sequence $(a_i)_{i=0}^{\infty}$ of 0's and 1's, we define n new sequences $(a_i^k)_{i=0}^{\infty}$ (i.e. $1 \leq k \leq n$) by:

$$a_i^k = a_{in+k}$$

Moreover, we define:

$$\alpha_k = \sum_{i=0}^{\infty} \frac{a_i^k}{2^{i+1}}$$

And define the $a \in \underbrace{[0,1] \times ... \times [0,1]}_{n \text{ times}} \subseteq \mathbb{R}^n$ generated by the sequence $(a_i)_{i=0}^{\infty}$ of player I as:

$$a = (\alpha_1, ..., \alpha_n)$$

This ends our illustration of our description of how to generalize our result. We we now state the generalized result as a theorem:

Theorem 3.2.3.9 (ZF + AD + DC). *Any subset of any Euclidean space* $\mathbb{R}^n$ *is measurable.*

The above theorem together with the contrapositive of Theorem 3.2.1.4 yield the following corollary:

Corollary 3.2.3.10 (ZF + AD + DC). BTP *fails.*

This directly implies:

Corollary 3.2.3.11. ZFC + AD *is inconsistent.*

Theorem 3.2.3.9 and Theorem 3.2.1.4 also imply:

Corollary 3.2.3.12. *Both the existence of non-measurable sets of* $\mathbb{R}^3$ *and* BTP *are independent of* ZF + DC.

Proof. The existence of non-measurable sets of $\mathbb{R}^3$ in ZF + DC yields ZF + AD + DC inconsistent. Their non-existence yields ZFC inconsistent. Thus the existence of non-measurable sets of $\mathbb{R}^3$ is independent of ZF + DC.

If BTP holds in ZF + DC, then non-measurable sets of $\mathbb{R}^3$ exist contrary to the fact that their existence is independent of ZF + DC. If BTP fails, then ZFC is inconsistent. Thus BTP is independent of ZF + DC. □

Chapter 4

Mathematics without Choice

In this chapter we will give some examples of when choice principles are used in mathematics and when the usage is necessary. We will provide some examples to illustrate that ZF + CC is sufficient to develop much of mathematics but we will also give examples of statements for which AC is necessary, thus illustrating what mathematics without AC is like.

4.1 Set Theory

4.1.1 Elementary Set Theory

As noted in chapter 1 where we discussed the history of AC, the use of a choice principle is sometimes hidden in the sense that it is not obvious that such a principle is being used. This will be illustrated in the following propositions. We let CC($|\mathbb{N}|$) denote CC restricted to countable families of countable sets.

Proposition 4.1.1.1. [1] CC $\Rightarrow$ CUT $\Rightarrow$ CC($|\mathbb{N}|$).

Proof. Let $\{X_n \mid n \in \mathbb{N}\}$ be a countable family of countable sets. Construct a corresponding countable family $\{Y_n \mid n \in \mathbb{N}\}$ of disjoint sets by defining:

$$\begin{cases} Y_0 = X_0 \\ Y_n = X_n \setminus \bigcup\limits_{n<k} \big(X_n \bigcap Y_k\big), \text{ for } n \geq 1 \end{cases}$$

Now set up an array where the n-th column consists of the elements of Y_n, i.e. choose one enumeration of each Y_n, let the n-th column list the elements of Y_n according to the corresponding chosen enumeration and apply Cantor's

[1]The first implication corresponds to Proposition 2.12 [Rud76] together with an obeservation from Proposition 3.5 in [Her2006].

diagonal argument to prove that the elements in the constructed array (that is, the elements of $\bigcup_{n\in\mathbb{N}} X_n$) are countable, proving the first implication. As Herrlich notes on p. 22 in [Her2006], to choose one enumeration of each Y_n we use CC: Our only assumption is that each Y_n can be listed by enumeration so we use CC to choose one specific enumeration of each Y_n.

For the second implication, assume $\{X_n \mid n \in \mathbb{N}\}$ is a countable family of countable non-empty sets. Then $X = \bigcup_{n\in\mathbb{N}} X_n$ is countable by assumption and we can use an enumeration of X to well-order X as follows: Let $x, y \in X$ and let g be an enumeration of X. Define the well-order W of X as xWy if $g^{-1}(x) \leq g^{-1}(y)$. Now define $f : \mathbb{N} \to X$ by $f(n) = z$ for the W-least z such that $z \in X_n$, then $f \in \prod_{n\in\mathbb{N}} X_n$. □

Let CC(fin) and CUT(fin) be the principles CC and CUT restricted to countable families X of finite sets X_n. Then CC(fin) is clearly a weakening of CC($|\mathbb{N}|$). Moreover, note that CC(fin) if and only if CUT(fin) holds by similar reasoning as in the proof above: The left direction is proved analogously as the implication from CUT to CC($|\mathbb{N}|$). Note that in the proof of the right direction, we have $|Y_n| = m_n \in \mathbb{N}$ which implies that Y_n can be enumerated in $m_n! \in \mathbb{N}$ ways. Thus by using CC(fin), we can choose one enumeration of each Y_n and then end the proof by successively putting the elements x_k^n of all Y_n in a row to construct a bijection with $\mathbb{N}$:

$$x_1^1 \quad \ldots \quad x_{m_1}^1 \quad x_1^2 \quad \ldots \quad x_{m_2}^2 \quad \ldots$$

By p. 23 in [Her2006], ZF cannot even prove CUT(2), that the union of countable families of two-element sets are countable. Since CUT(2) clearly is a weakening of CUT(fin), all of the principles discussed above are unprovable in ZF.

We continue by giving a precise definition of what we mean by infinite:

Definition 4.1.1.2. A set X is *finite* if there exists $n \in \mathbb{N}$ such that there exists a bijection $f : X \to \{m \in \mathbb{N} \mid m < n\}$. A set which is not finite is *infinite*.

Definition 4.1.1.3 (Fin). For any infinite set X, there exists an injection $f : \mathbb{N} \to X$.

Proposition 4.1.1.4 (ZF + CC). Fin *holds.*

Proof. Define the function f recursively: Choose one element from X, let this element be $f(0)$. Having chosen $n-1$ elements, choose the n-th element $f(n)$ from the set $X \setminus \{f(i) \mid i \leq n-1\}$. This set is non-empty since assuming it is empty implies that X is finite. This f is obviously injective since if $m \neq n$, then w.l.o.g assume $m < n$ so $f(n)$ is chosen from the set $X \setminus \{f(i) \mid i \leq n-1\}$ and thus $f(n) \neq f(m)$. □

Even though it is not obvious, the above proof makes use of CC. This is most easily understood by seeing that the chosen $f(n)$ are completely arbitrary in the sense that the proof provides no description of how they are chosen. Note that even without CC, it is possible to use recursion to define an injective function $f_n : \{m \in \mathbb{N} \mid m \leq n\} \to X$ for every $n \in \mathbb{N}$. However, $(\bigcup_{n\in\mathbb{N}} f_n) : \mathbb{N} \to X$ is not necessarily injective or even a function. The interested reader is referred to [I-Cai] for a longer discussion about the subtleties of recursion and CC.

Moreover, as proved in [Her2006] in Theorem 2.14, CC(fin) is a necessary condition for Fin and by our discussion above, Fin is thus unprovable in ZF.

We continue with a fundamental proposition about functions:

Proposition 4.1.1.5. [2] *The following statements are equivalent:*

(1) AC.
(2) *For every family F (indexed by I) of non-empty pairwise disjoint sets, there exists a set S such that $|S \bigcap F_i| = 1$ for each $i \in I$.*
(3) *Every set X is projective, i.e. for each function $f : X \to Y$ and each surjection $g : Z \to Y$ there exists a function $k : X \to Z$ with $g \circ k = f$.*
(4) *Every surjection $g : A \to B$ is a retraction, i.e. there exists a function $k : B \to A$ with $g \circ k = \mathrm{id}_B$.*

In (4), such a function k is called a *right inverse* of g.

Proof. (1) $\Rightarrow$ (2): Apply AC to F. Thus there exists a function f which chooses one element from each F_i. Letting $S = f(F)$ gives the desired set since the F_i are disjoint by assumption.

(2) $\Rightarrow$ (3): Let f and g be as described in (3). Consider the sets $f^{-1}(y) = \{x \in X \mid f(x) = y\}$ and $g^{-1}(y) = \{z \in Z \mid g(z) = y\}$. Define the family $F = \{g^{-1}(y) \mid y \in f(X)\}$. That g is a surjection ensures the non-emptiness of each set $g^{-1}(y)$ of F and these sets are disjoint since if they're not, then g is not a function. Now apply (2), thus there exists a set S such that $|S \bigcap g^{-1}(y)| = 1$ for all $y \in f(X)$. For a given $y \in f(X)$ and all $x \in f^{-1}(y)$, let $k(x) = s$ for the unique $s \in S$ such that $s \in g^{-1}(y)$. Then $g \circ k\ (x) = g(k(x)) = g(s) = y = f(x)$.

(3) $\Rightarrow$ (4): Let $g : A \to B$ be an arbitrary surjection and in statement (3), let $Z = A$ and $X = Y = B$ and also $f = \mathrm{id}_B$. Then (3) implies that there exists $k : B \to A$ such that $g \circ k = \mathrm{id}_B$.

(4) $\Rightarrow$ (1): Let X be a family of non-empty sets. Define $\mathscr{X} = \{\langle x, Y\rangle \in \bigcup X \times X \mid x \in Y\}$ and $g : \mathscr{X} \to X$ by $g(\langle x, Y\rangle) = Y$. g is surjective since every $Y \in X$ is non-empty. Thus by (4), g has a right inverse $k : X \to \mathscr{X}$. Now define $\pi : \mathscr{X} \to \bigcup X$ by $\pi(\langle x, Y\rangle) = x$. Then $\pi \circ k\ (Y) \in Y$ so $\pi \circ k = f$ is a choice function for X. □

We also see that (4) could be proved from AC: Define $X = \{g^{-1}(b) \mid b \in B\}$, by AC this set X has a choice function f. Define k by $k(b) = f(g^{-1}(b))$.

[2] The formulations of these statements are taken from Exercise E.4 of Section 2.1 in [Her2006].

Moreover, note that every right inverse is injective: Assume k is not injective, then for some distinct $b_1, b_2 \in B$, the equality $k(b_1) = k(b_2)$ holds so $b_1 = g(k(b_1)) = g(k(b_2)) = b_2$, contrary to assumption.

Remembering the different definitions of cardinality from chapter 1, we define the following principle:

Definition 4.1.1.6 (Partition Principle - PP). If $X \preceq^* Y$, then $X \preceq Y$.

We thus see that AC implies PP and that the different definitions of cardinality we have given are equivalent in ZFC. We state this result as a corollary:

Corollary 4.1.1.7 (ZFC). *If $X \neq \emptyset$, then $X \preceq Y$ if and only if $X \preceq^* Y$.*

it is still an open question whether PP implies AC (see [BaMo90]).

4.1.2 Intuitionistic ZF & AC

We finish this section with a connection between AC and intuitionism. IZF denotes ZF expressed in intuitionistic logic, thus IZF is ZF without tertium and RAA. Without tertium, many trivial equivalences of ZF become unprovable in IZF:

Definition 4.1.2.1. Let X be a set, then:

- X is *inhabited* if there exists $x \in X$.
- X is *empty* if for all sets x, the relation $x \notin X$ holds.
- X is *non-empty* if X is not empty.

Classically, a set X is inhabited if and only if it is non-empty. Intuitionistically, only the right direction holds. We thus distinguish between two versions of AC, namely *every family of inhabited sets has a choice function* ($\mathrm{AC_i}$) and *every family of non-empty sets has a choice function* ($\mathrm{AC_{ne}}$). By our preceding remarks, $\mathrm{AC_{ne}}$ implies $\mathrm{AC_i}$. We now prove that $\mathrm{AC_{ne}}$ is incompatible with IZF:

Theorem 4.1.2.2 (IZF). [3] *$\mathrm{AC_{ne}} \Rightarrow \varphi \vee \neg\varphi$.*

Proof. Let φ be any well-formed formula of ZFC. Let x be a variable not occurring in φ and define the set:

$$X = \{x \in \{0, 1\} \mid (x = 0 \wedge \varphi) \vee (x = 1 \wedge \neg\varphi)\}$$

Using tertium, this means:

$$X = \begin{cases} \{0\}, & \text{if } \varphi \text{ is true.} \\ \{1\}, & \text{if } \varphi \text{ is false.} \end{cases}$$

[3] Thanks to Håkon Robbestad Gylterud and Henrik Forssell for the help with this proof.

However, without using tertium, it is unclear which elements belong to X. We label the condition $(x = 0 \wedge \varphi) \vee (x = 1 \wedge \neg\varphi)$ by $P(x)$.

Assume X is empty. If φ holds, then 0 belongs to X, contradicting X being empty. By implication introduction, we conclude $\neg\varphi$ and discharge φ. However, since $\neg\varphi$ holds, 1 belongs to X, again contradicting X being empty. By another use of implication introduction, we conclude that X is non-empty.

Thus by AC, the family $\{X\}$ has a choice function $f : \{X\} \to \bigcup\{X\}$. Moreover, $f(X) = 0 \vee f(X) = 1$ holds since $X \subseteq \{0, 1\}$. If $f(X) = 0$, then 0 belongs to X. Since x belongs to X if and only if $P(x)$ holds, it follows that φ holds if $f(X) = 0$: The righthand condition of $P(0)$ derives false and by false elimination, we may obtain any formula. When applying or-elimination to $P(0)$, we can thus obtain the lefthand condition of $P(0)$. And-elimination yields φ.

Similarly, $f(X) = 1$ implies $\neg\varphi$. Thus, when applying or-elimination to $f(X) = 0 \vee f(X) = 1$, we can apply or-introduction in the subtree for $f(X) = 0$ to obtain $\varphi \vee \neg\varphi$ and similarly for the case $f(X) = 1$. This finishes the proof of tertium. □

Even the weaker $\mathrm{AC_i}$ derives tertium:

Theorem 4.1.2.3 (IZF). [4] $\mathrm{AC_i} \Rightarrow \varphi \vee \neg\varphi$.

Proof. Let φ be as in the proof above and define the sets:

$$\begin{cases} U = \{x \in \{0, 1\} \mid (x = 0) \vee \varphi\} \\ V = \{x \in \{0, 1\} \mid (x = 1) \vee \varphi\} \end{cases}$$

U is inhabited since $0 \in U$ and V is inhabited since $1 \in V$. Thus we may apply $\mathrm{AC_i}$ to $\{U, V\}$, then there exists f such that:

$$f(U) \in U \wedge f(V) \in V$$

Thus $f(U)$ and $f(V)$, being elements of U and V, satisfy:

$$\Big(f(U) = 0 \vee \varphi\Big) \wedge \Big(f(V) = 1 \vee \varphi\Big)$$

Even in IZF, conjunction distributes over disjunction:

$$\Big(f(U) = 0 \wedge f(V) = 1\Big) \vee \Big(f(U) = 0 \wedge \varphi\Big) \vee \Big(f(V) = 1 \wedge \varphi\Big) \vee \varphi \tag{4.1}$$

Assume the leftmost condition holds, then clearly $f(U) \neq f(V)$ holds: Assuming $f(U) = f(V)$ directly leads to a contradiction, thus by implication introduction we conclude $\neg(f(U) = f(V))$, i.e. $f(U) \neq f(V)$.

More generally, if ψ is an arbitrary formula which derives false in IZF, then it is intuitionistically valid to conclude $\neg\psi$ as this corresponds to implication introduction. The prohibited reasoning is to use a derivation from $\neg\psi$ to false in order to conclude ψ (which we in classical logic are allowed to do, using RAA).

[4] Corresponds to the proof at [I-Wiki2].

Furthermore, if any of the other conditions in (4.1) hold, then φ holds. Thus (4.1) implies:

$$f(U) \neq f(V) \vee \varphi \tag{4.2}$$

Assume φ holds, then both U and V equal $\{0,1\}$, thus the implication $\varphi \Rightarrow U = V$ holds. Given $U = V$, of course $f(U) = f(V)$ holds. Thus we have proved the implication $\varphi \Rightarrow f(U) = f(V)$.

Now note that it is intuitionistically valid to turn an implication into its contrapositive form: Assume ψ, $\psi \Rightarrow \sigma$ and $\neg\sigma$. Then we can clearly derive false, thus we conclude $\neg\psi$. One more use of implication introduction yields $\neg\sigma \Rightarrow \neg\psi$, thus we have proved $(\psi \Rightarrow \sigma) \Rightarrow (\neg\sigma \Rightarrow \neg\psi)$. However, note that $(\neg\sigma \Rightarrow \neg\psi) \Rightarrow (\psi \Rightarrow \sigma)$ relies on cancelation of double negation and thus the equivalence $(\psi \Rightarrow \sigma) \iff (\neg\sigma \Rightarrow \neg\psi)$ only holds classically.

We turn our specific implication into its contrapositive: We have proved $\varphi \Rightarrow f(U) = f(V)$, turning it into contrapositive form, we obtain $f(U) \neq f(V) \Rightarrow \neg\varphi$. Thus (4.2) implies tertium, $\varphi \vee \neg\varphi$. □

Note that $\mathrm{AC_i} \Rightarrow \varphi \vee \neg\varphi$ is sufficient for proving $\mathrm{AC_{ne}} \Rightarrow \varphi \vee \neg\varphi$.

4.2 Topology

In the first subsection, we define fundamental topological concepts and propositions. In the second subsection, we prove that all separable metric spaces satisfy the Lindelöf property and in the third subsections we present a proof of the equivalence between AC and Tychonoff's Theorem.

4.2.1 Topological Spaces

Several of the definitions and statements of propositions in the rest of this section are taken from Appendix D of [Coh2013], however no proper proofs are stated there.

As we work in classical logic, $X = (X^c)^c$ holds: Let U be a universe. By definition, $X^c = \{x \in U \mid x \notin X\}$. We show that X and X^c are subsets of each other, assume $x \in U$:

$$x \in X \underset{imp.intro}{\Rightarrow} x \notin X^c \underset{def}{\Rightarrow} x \in (X^c)^c$$
$$x \in (X^c)^c \underset{def}{\Rightarrow} x \notin X^c \underset{RAA}{\Rightarrow} x \in X$$

Definition 4.2.1.1. Let X be a set. A *topology* $\mathscr{O}$ on X is a family of subsets of X such that X is in $\mathscr{O}$, the empty set is in $\mathscr{O}$ and $\mathscr{O}$ is closed under arbitrary unions and finite intersections.

Definition 4.2.1.2. A *topological space* is an ordered pair $(X, \mathscr{O})$ of a set X and a topology $\mathscr{O}$ on X.

A *topological subspace* of $(X, \mathscr{O})$ is an ordered pair of a subset Y of X and the topology $\mathscr{O} \restriction Y = \{O \cap Y \mid O \in \mathscr{O}\}$. The topological space $(X, \mathscr{O})$ is *compact* if every open cover of X has a finite subcover. Moreover, $Y \subseteq X$ is said to be compact if $(Y, \mathscr{O} \restriction Y)$ is compact.

As we see, the definitions used for metric spaces generalize to topological spaces. We state some definitions and refer the reader to [Mun2000] for a more extensive list:

Definition 4.2.1.3. Let $(X, \mathscr{O})$ be a topological space, $S \subseteq X$ and $x \in X$, then:

- S is *open* if $S \in \mathscr{O}$.
- S is *closed* if $S^c \in \mathscr{O}$.
- An *open neighborhood* of x is an open set $O \in \mathscr{O}$ such that $x \in O$.
- x is a *limit point* of S if each open neighborhood O of x contains at least one point s of S such that $x \neq s$.
- S is *dense* in X if every $x \in X$ either belongs to S or is a limit point of S.
- X is *separable* if it contains a countable dense subset.

Using $X = (X^c)^c$, clearly S is open if and only if S^c is closed.

Definition 4.2.1.4. Let $\mathscr{O}_1$ and $\mathscr{O}_2$ be topologies on a set X. If $\mathscr{O}_1 \subseteq \mathscr{O}_2$, then $\mathscr{O}_1$ is said to be *weaker* than $\mathscr{O}_2$ on X.

We will call a set a collection when we want to emphasize that its elements are families of subsets.

Proposition 4.2.1.5. *Let X be a set and let $\mathscr{A}$ be a family of subsets of X. Then $\mathscr{A}$ is included in some topology on X and the intersection of all topologies which $\mathscr{A}$ is included in, is a topology. Moreover, this intersection is the weakest topology on X which includes $\mathscr{A}$.*

Proof. $\mathcal{P}(X)$ is a topology, thus $\mathscr{A}$ is included in some topology on X.

Let $\mathscr{T}$ be the collection of topologies on X which includes $\mathscr{A}$. Both X and $\emptyset$ belong to every $O \in \mathscr{T}$ and thus to $\bigcap \mathscr{T}$. Let $\mathscr{S}$ be a family of sets of $\bigcap \mathscr{T}$, then for $O \in \mathscr{T}$, every $S \in \mathscr{S}$ is in O and since O is closed under arbitrary unions, $\bigcup \mathscr{S}$ is in O. Since O was arbitrary, it follows that $\bigcup \mathscr{S} \in \bigcap \mathscr{T}$ holds. The closure of $\bigcap \mathscr{T}$ under finite intersections is proved similarly.

Let $\mathscr{O}$ be a topology on X which includes $\mathscr{A}$, then $\mathscr{O} \in \mathscr{T}$ and thus $\bigcap \mathscr{T} \subseteq \mathscr{O}$. □

Definition 4.2.1.6. Let X be a set and let $\mathscr{A}$ be a family of subsets of X. Then the weakest topology $\mathscr{O}$ on X which includes $\mathscr{A}$ is called the topology *generated* by $\mathscr{A}$.

A family $\mathscr{A}$ of subsets of X is called a *synthetic subbasis* on X. A family $\mathscr{B}$ of subsets of X is called a *synthetic basis* if for every A and B belonging to $\mathscr{B}$, there exists $\mathscr{C} \subseteq \mathscr{B}$ such that $A \cap B = \cup\mathscr{C}$.

For a synthetic subbasis $\mathscr{A}$ on X, we see that the family $\mathscr{B} = \{\bigcap \mathscr{A}' \mid \mathscr{A}' \subseteq \mathscr{A}$ and $\mathscr{A}'$ is finite$\}$ is a synthetic basis on X: If A and B are finite intersections of sets belonging to $\mathscr{A}$, then so is $A \cap B$. By definition of $\mathscr{B}$, the set $A \cap B = C$ belongs to $\mathscr{B}$ so $A \cap B = \cup\{C\}$.

Moreover, given a synthetic subbasis $\mathscr{A}$ and the synthetic basis $\mathscr{B}$ consisting of finite intersections of $\mathscr{A}$, the set $\mathscr{O} = \{\bigcup \mathscr{B}' \mid \mathscr{B}' \subseteq \mathscr{B}\}$ can be proven to be precisely the topology generated by $\mathscr{A}$. We refer the reader to [I-PW1] and [I-PW2] for proofs of these claims.

The following observation will be of use to prove that Tychonoff's Theorem holds in ZFC: Let $(X, \mathscr{O})$ be a topological space such that $\mathscr{O}$ is generated by some $\mathscr{A}$. From the above considerations it follows directly that for an open set O containing a point x, there exists finitely many sets A_i belonging to $\mathscr{A}$ such that $x \in \bigcap\limits_{i \in I} A_i \subseteq O$.

Proposition 4.2.1.7 (De Morgan's Laws). *Let X be any family of sets and I be an index set of X. Then the following equalities hold:*

$$\Big(\bigcap_{i \in I} X_i\Big)^c = \bigcup_{i \in I} X_i^c \qquad\qquad \Big(\bigcup_{i \in I} X_i\Big)^c = \bigcap_{i \in I} X_i^c$$

Proof. Here X_i^c denotes $X_i^c \setminus U$ where U is some universe. Let $y \in U$. The left equality is proved by:

$$y \in \Big(\bigcap_{i \in I} X_i\Big)^c \iff \exists i \in I \ (y \notin X_i) \iff y \in \bigcup_{i \in I} X_i^c$$

And the right by:

$$y \in \Big(\bigcup_{i \in I} X_i\Big)^c \iff \nexists i \in I \ (y \in X_i) \iff \forall i \in I \ (y \in X_i^c) \iff y \in \bigcap_{i \in I} X_i^c$$

□

For any topological space $(X, \mathscr{O})$, X is closed since $X^c = \emptyset \in \mathscr{O}$. Thus every subset S of X is contained in a closed subset of X. Taking the intersection of all closed subsets C of X such that $S \subseteq C$ yields the least closed subset S^- which contains S (note that S^- is closed by the De Morgan's Laws).

Definition 4.2.1.8. Let $(X, \mathscr{O})$ be a topological space and $S \subseteq X$, then the least closed subset containing S is called the *closure* of S and is denoted by S^-.

Definition 4.2.1.9. Let $(X, \mathscr{O}_x)$ and $(Y, \mathscr{O}_y)$ be topological spaces. A function $f : X \to Y$ is *continuous* if $f^{-1}(U)$ is an open subset of X whenever U is an open subset of Y.

Thus f is continuous if and only if $f^{-1}(U)$ is closed whenever U is closed:

$$x \in f^{-1}(U)^c \iff x \notin f^{-1}(U) \iff$$
$$f(x) \notin U \iff f(x) \in U^c \iff x \in f^{-1}(U^c)$$

Implying $f^{-1}(U)^c = f^{-1}(U^c)$. Thus $f^{-1}(U^c)$ is closed whenever U^c is closed if and only if $f^{-1}(U)$ is open whenever U is open. Since $U \mapsto U^c$ is a surjective map, the result follows.

Definition 4.2.1.10. Let X be a set and let $\{Y_i \mid i \in I\}$ be a family of sets with associated topologies $\mathscr{O}_i$. Moreover, let $\{f_i \mid i \in I\}$ be a family of functions such that for each i, $f_i : X \to Y_i$. Then the topology $\mathscr{O}$ *generated by* $\{f_i \mid i \in I\}$ is the weakest topology on X that makes each f_i continuous on X.

Since f_i by definition is continuous if and only if $f_i^{-1}(U)$ is open in X whenever U is open in Y_i, it follows that the topology generated by the functions f_i is the topology generated by the subsets $f_i^{-1}(U)$ such that $i \in I$ and U is open in Y_i.

Definition 4.2.1.11 (Finite Intersection Property - FIP). Let X be a set. A family $\mathscr{S}$ of subsets of X satisfies the *finite intersection property* if every finite subfamily of $\mathscr{S}$ has a non-empty intersection.

Proposition 4.2.1.12. *Let $(X, \mathscr{O})$ be a topological space. Then X is compact if and only if each family $\mathscr{S}$ of closed subsets of X satisfying* FIP *has a non-empty intersection.*

Proof. Let X be compact and let $\mathscr{S}$ be a family of closed subsets of X satisfying FIP. Assume $\big(\bigcap \mathscr{S}\big)^c = X$. Then $X = \big(\bigcap_{S \in \mathscr{S}} S\big)^c = \bigcup_{S \in \mathscr{S}} S^c$ and since S is closed, S^c is open so $G = \{S^c \mid S \in \mathscr{S}\}$ is an open cover of X. Since X is compact, there exists a finite subcover G^* of G indexed by I:

$$X = \bigcup G^* = \bigcup_{i \in I} S_i^c = \big(\bigcap_{i \in I} S_i\big)^c$$

But then $\bigcap_{i \in I} S_i = \emptyset$, contradicting FIP of $\mathscr{S}$. Thus we conclude $\big(\bigcap \mathscr{S}\big)^c \neq X$, i.e. $\bigcap \mathscr{S} \neq \emptyset$.

Coversely, assume each family $\mathscr{S}$ of closed subsets of X satisfies FIP but X is not compact. Let G be an open cover of X without a finite subcover and let G^* be an arbitrary finite subset of G indexed by I. Then the following statements hold:

$$\emptyset \neq \big(\bigcup_{i \in I} g_i\big)^c = \bigcap_{i \in I} g_i^c \tag{1}$$
$$\emptyset = \big(\bigcup_{g \in G} g\big)^c = \bigcap_{g \in G} g^c \tag{2}$$

By (1), $\{g^c \mid g \in G\}$ satisfies FIP but by (2), the intersection of $\{g^c \mid g \in G\}$ is empty. Moreover, each g^c is closed since each g is open, thus we have reached a contradiction to our assumption that each family of closed subsets satisfying FIP has a non-empty intersection. $\square$

Definition 4.2.1.13. Let $\{X_i \mid i \in I\}$ be a family of sets. For each $j \in I$, the j-th *canonical projection* π_j of $\prod_{i\in I} X_i$ is $\pi_j : \prod_{i\in I} X_i \to X_j$ such that $\pi_j(f) = f(j)$.

Expressed more intuitively, the j-th canonical projection gives the j-th coordinate.

Definition 4.2.1.14. Let $\{(X_i, \mathscr{O}_i) \mid i \in I\}$ be family of topological spaces. Then the *product topology* on $\prod_{i\in I} X_i$ is the topology generated by the canonical projections $\{\pi_i \mid i \in I\}$.

FIP is invariant under canonical projection to compact spaces. This is intuitively clear since given a family of subsets of $\prod_{i\in I} X_i$ satisfying FIP, each individual coordinate-position must satisfy FIP. We formalize the argument for the sake of rigour:

Proposition 4.2.1.15. *Let $\{(X_i, \mathscr{O}_i) \mid i \in I\}$ be a family of topological spaces and let $\mathscr{S}$ be a family of subsets such that $\mathscr{S}$ satisfies* FIP *in $\prod_{i\in I} X_i$. Then for every $i \in I$, $\{\pi_i(S) \mid S \in \mathscr{S}\}$ satisfies* FIP *in X_i. Moreover, $\bigcap_{S\in\mathscr{S}} \pi_i(S)^-$ is non-empty in each compact $(X_i, \mathscr{O}_i)$.*

Proof. Let $\mathscr{Z}$ be a finite subfamily of $\mathscr{S}$, then $\bigcap \mathscr{Z}$ is non-empty so there exists $f \in \bigcap \mathscr{Z}$. Thus $f \in Z$ holds for every Z of $\mathscr{Z}$, implying that $\pi_i(f) \in \pi_i(Z)$ holds for every Z, yielding $\pi_i(f) \in \bigcap\{\pi_i(Z) \mid Z \in \mathscr{Z}\}$.

Moreover, since $\pi_i(S) \subseteq \pi_i(S)^-$ holds, it follows that the set $\{\pi_i(S)^- \mid S \in \mathscr{S}\}$ satisfies FIP. Since X_i is compact, we can apply Proposition 4.2.1.12 to conclude that $\bigcap_{S\in\mathscr{S}} \pi_i(S)^-$ is non-empty. $\square$

We say that two sets *intersects* each other if their intersection is non-empty.

Proposition 4.2.1.16. [5] *Let $(X, \mathscr{O})$ be a topological space, $x \in X$ and $S \subseteq X$. Then $x \in S^-$ holds if and only if every open neighborhood of x intersects S.*

Proof. Turning the statements into their contrapositive forms, we will instead prove $x \notin S^-$ if and only if there exists an open neighborhood of x which does not intersect S:

Let $x \notin S^-$. Then $x \in (S^-)^c \in \mathscr{O}$, the last set membership relation following from that S^- is closed. Obviously $(S^-)^c \cap S$ is empty, proving the right direction.

Instead, let S' be an open neighborhood containing x but not intersecting S. Then $x \notin (S')^c \supseteq S$. Since $(S')^c$ is a closed subset containing S, it follows that $S^- \subseteq (S')^c$ holds and thus $x \notin S^-$, finishing the proof. $\square$

[5] Corresponds to Theorem 17.5 on p. 96 in [Mun2000] and the remark further down on the same page.

4.2.2 All Separable Metric Spaces are Lindelöf

Definition 4.2.2.1. A metric space (X, d) is a *Lindelöf space* if every open cover G of X has a countable subcover H.

Proposition 4.2.2.2 (ZF + CC). *All separable metric spaces are Lindelöf spaces.*

We note that according to Herrlich, all separable metric spaces are Lindelöf spaces if and only if CC($\mathbb{R}$) holds (to prove this equivalence is part of Exercise E.3 in Section 4.6 of [Her2006]). However, we only prove a weaker version of the right direction of the equivalence.

Proof. Let (X, d) be a separable metric space, let E be a countable dense subset of X and let G be an open cover of X. For each $e \in E$, define:

$$R_e = \{r \in \mathbb{R} \mid \exists g \in G\ (B_r(e) \subseteq g)\},$$
$$r_e = \sup(R_e),$$
$$\mathscr{G}_e = \{(g_n)_{n=0}^{\infty} \mid B_{r_e}(e) \subseteq \bigcup_{n\in\mathbb{N}} g_n\}.$$

Each $\mathscr{G}_e$ is non-empty: Since G is an open cover, every $e \in E$ is contained in an open $g \in G$, implying $r_e > 0$. Let $0 < \varepsilon < r_e$. Then for each $n \in \mathbb{N}$ there exists $g \in G$ containing $B_{r_e - \frac{\varepsilon}{n}}(e)$, otherwise r_e would not be the least upper bound of R_e. Let $\mathscr{B}_n$ be the set of $g \in G$ containing $B_{r_e - \frac{\varepsilon}{n}}(e)$. By CC, the family $\{\mathscr{B}_n \mid n \in \mathbb{N}\}$ has a choice function, denote the choice from $\mathscr{B}_n$ by g_n. Moreover, $(g_n)_{n=0}^{\infty}$ is in $\mathscr{G}_e$ since if $x \in B_{r_e}(e)$, then $d(e, x) < r_e$ and for some least n, the inequality $d(e, x) < (r_e - \frac{\varepsilon}{n})$ holds so $x \in g_n$.

Since E is countable, there exists an enumeration f of E. Letting $(g_n^m)_{n=0}^{\infty}$ be the sequence corresponding to $f(m) \in E$, we see that the family $H = \{g_n^m \mid n, m \in \mathbb{N}\}$ is countable by CUT. We prove that H is a cover of X, this establishes that H is a countable subcover of G for X since $H \subseteq G$ clearly holds:

Assume $x \in X$. If $x \in E$, then $x = f(m)$ for some m and thus $x \in B_{r_x}(x) \subseteq \bigcup_{n\in\mathbb{N}} g_n^m$. Instead assume x is a limit point of E, then consider an arbitrary $h \in G$ such that $x \in h$. Since h is open, there exists $r > 0$ such that $B_r(x) \subseteq h$. Thus also $B_{r/2}(x) \subseteq h$ and since x is a limit point of E, there exists $e \in B_{r/2}(x)$. Now consider $B_{r/2}(e)$: Since $d(x, e) < r/2$ we have $x \in B_{r/2}(e)$. Moreover, each element z of $B_{r/2}(e)$ satisfy $d(e, z) < r/2$ so for each such z we also have $d(x, z) \leq d(x, e) + d(e, z) < r/2 + r/2 = r$ and thus $B_{r/2}(e) \subseteq B_r(x) \subseteq h$. Thus $r/2 \in R_e$ and therefore $r/2 \leq r_e$ so $B_{r/2}(e) \subseteq B_{r_e}(e)$. Since $e = f(m)$ for some m and $B_{r_e}(e) \subseteq \bigcup_{n\in\mathbb{N}} g_n^m$, we obtain $x \in H$. □

4.2.3 Tychonoff's Theorem

As illustrated by several quotes on p. 85 in [Her2006], Tychonoff's Theorem is considered one of the most important theorems of topology.

Definition 4.2.3.1 (Tychonoff's Theorem - TT). Let $\{(X_i, \mathscr{O}_i) \mid i \in I\}$ be family of topological spaces. If each $(X_i, \mathscr{O}_i)$ is compact, then $\prod_{i \in I} X_i$ is also compact.

Theorem 4.2.3.2. [6] AC $\iff$ TT.

Proof. For the right direction, we will prove that every family $\mathscr{S}$ of closed subsets of $\prod_{i \in I} X_i$ that satisfies FIP has a non-empty intersection $\bigcap \mathscr{S}$, thus compactness is implied by Proposition 4.2.1.12.

Let $\mathscr{S}$ be any such family and consider the collection $\mathscr{S}^* = \{\mathscr{A} \mid \mathscr{A}$ satisfies FIP and $\mathscr{S} \subseteq \mathscr{A}\}$ partially ordered by the subset relation. Note that $\mathscr{S}^*$ may contain families in which all subsets are not closed in $\prod_{i \in I} X_i$. Let $\mathscr{C}^*$ be a chain in $\mathscr{S}^*$, then $\bigcup \mathscr{C}^* = \mathscr{C}$ is an upper bound for $\mathscr{C}^*$ in $\mathscr{S}^*$: Clearly $\mathscr{A} \subseteq \mathscr{C}$ holds for every $\mathscr{A} \in \mathscr{C}^*$, thus we only have to establish $\mathscr{C} \in \mathscr{S}^*$. The relation $\mathscr{S} \subseteq \mathscr{C}$ holds as a result of the previous sentence and given a finite subfamily $\mathscr{F}$ of $\mathscr{C}$, since $\mathscr{C}^*$ is a chain there exists an element $\mathscr{A} \in \mathscr{C}^*$ such that $\mathscr{F} \subseteq \mathscr{A}$. Since $\mathscr{A}$ satisfies FIP, the result follows. Thus every chain in $\mathscr{S}^*$ has an upper bound in $\mathscr{S}^*$ and by Zorn's Lemma, $\mathscr{S}^*$ has a maximal element $\mathscr{M}$.

We now manipulate $\mathscr{M}$: For every $i \in I$, applying the second part of Proposition 4.2.1.15 to $\{\pi_i(M)^- \mid M \in \mathscr{M}\}$ gives that $\bigcap_{M \in \mathscr{M}} \pi_i(M)^- = Y_i$ is non-empty. Thus by AC, there exists a function f in the Cartesian product $\prod_{i \in I} Y_i$.

We will prove that this f is in the closure of every set in $\mathscr{M}$. Since $\mathscr{S} \subseteq \mathscr{M}$ holds with all elements S of $\mathscr{S}$ being closed, $f \in \bigcap \mathscr{S}$ then follows. Thus we let U be any open neighborhood of f and prove that U intersects every set in $\mathscr{M}$, then Proposition 4.2.1.16 implies that f is in the closure of every set in $\mathscr{M}$.

Since the product topology on $\prod_{i \in I} X_i$ is generated by $T = \{\pi_i^{-1}(V) \mid i \in I$ and V is open in $X_i\}$, the remarks after Definition 4.2.1.10 implies that there exists a finite subset T^* of T such that $f \in \bigcap T^* \subseteq U$. Let $\pi_i^{-1}(V)$ be an arbitrary element of T^*. Then the relation $\pi_i(f) = f(i) \in V$ holds so V is an open neighborhood of $f(i)$. By construction of f, remember that also $f(i) \in \bigcap_{M \in \mathscr{M}} \pi_i(M)^-$ holds, thus Proposition 4.2.1.16 implies that V intersects every set in $\{\pi_i(M) \mid M \in \mathscr{M}\}$. Moreover, $V \cap \pi_i(M) \neq \emptyset$ implies $\pi_i^{-1}(V) \cap M \neq \emptyset$, so $\pi_i^{-1}(V)$ intersects every set M in $\mathscr{M}$. Since $\pi_i^{-1}(V)$ was an arbitrary element of T^*, every element of T^* intersects every set M in $\mathscr{M}$.

Since $\mathscr{M}$ is maximal with respect to FIP, it is closed under finite intersections, this is proved as follows: Let $\{M_i \mid i \leq n\}$ with $n \in \mathbb{N}$ be an arbitrary

[6] The right direction corresponds to Theorem 5.D in [Loo53] and is originally due to the collective Bourbaki while the left direction corresponds to Theorem 4.68 in [Her2006].

subfamily of $\mathscr{M}$. Let $\{M'_i \mid i \leq m\}$ with $m \in \mathbb{N}$ be another subfamily of $\mathscr{M}$. Then:

$$M_1 \cap ... \cap M_n \cap M'_1 \cap ... \cap M'_n$$

is non-empty since $\mathscr{M}$ satisfies FIP. Letting $N = \bigcap_{i \leq n} M_i$, we see that $N \in \mathscr{M}$ holds since $\{M'_i \mid i \leq m\}$ was arbitrary and $\mathscr{M}$ is maximal in $\mathscr{S}^*$.

As every $\pi_i^{-1}(V) \in T^*$ intersects every set in $\mathscr{M}$, letting $\{M_i \mid i \leq n\}$ be an arbitrary subcollection of $\mathscr{M}$ as above, $\bigcap_{i \leq n} M_i$ belongs to $\mathscr{M}$ and is thus intersected by every $\pi_i^{-1}(V) \in T^*$. Thus FIP would still be satisfied if every $\pi_i^{-1}(V) \in T^*$ would belong to $\mathscr{M}$, and since $\mathscr{M}$ is maximal, this relation holds. Since $T^* \subseteq \mathscr{M}$ and T^* is finite with $\mathscr{M}$ satisfying FIP, $\bigcap T^* \in \mathscr{M}$ holds. Since $\bigcap T^* \subseteq U$ holds, U intersects every set in $\mathscr{M}$.

For the left direction, let $\{X_i \mid i \in I\}$ be a family of non-empty sets and let ∞ be a set which is not in $\bigcup_{i \in I} X_i$ (if no such set exists, then X is the class of all sets and thus not a set by Russell's paradox). Define the sets $Y_i = X_i \cup \{\infty\}$ with associated topologies $\tau_i = \{\emptyset, Y_i, \{\infty\}\}$. Considering all different combinations for possible unions and intersections of sets in a given τ_i proves that each τ_i is in fact a topology. Let $Z_i = (Y_i, \tau_i)$. Each Z_i is obviously compact since $G = \{Y_i\}$ is the only possible open cover of Z_i, thus TT implies that $\prod_{i \in I} Y_i$ is compact.

For each i, define $A_i = \pi_i^{-1}(X_i) = \{f \in \prod_{i \in I} Y_i \mid f(i) \in X_i\}$. X_i is closed in Y_i since $X_i^c = \{\infty\} \in \tau_i$, moreover each π_i is continuous by definition of the product topology so by the remarks after Definition 4.2.1.9, each A_i is closed.

We now prove that $\mathscr{A} = \{A_i \mid i \in I\}$ satisfies FIP: Let $\mathscr{A}_n = \{A_k \mid k \leq n\}$ be a finite subfamily of $\mathscr{A}$. Since each corresponding X_k is non-empty, we can choose one $x_k \in X_k$ for each $k \leq n$ and define:

$$f(i) = \begin{cases} x_k, & \text{if there exists } k \text{ such that } X_i = X_k \\ \infty, & \text{if there does not exist } k \text{ such that } X_i = X_k \end{cases}$$

Since $f(k) = x_k \in X_k$ holds for all k, $f \in \bigcap \mathscr{A}_n$ holds. Thus $\mathscr{A}$ satisfies FIP and since $\prod_{i \in I} Y_i$ is compact and each A_i is closed, Proposition 4.2.1.12 implies that $\bigcap \mathscr{A}$ is non-empty. Finally, note that $\bigcap \mathscr{A} = \bigcap_{i \in I} A_i = \{f \in \prod_{i \in I} Y_i \mid \forall i \ (f(i) \in X_i)\} = \prod_{i \in I} X_i$, finishing the proof. $\square$

Chapter 5

Outroduction

Last Examples

The results of this thesis illustrate the following well-known fact: AC has some counterintuitive consequences but the presence of a choice principle is necessary to prove a wide variety of fundamental mathematical statements. For many of these statements, AC is the necessary choice principle. The reader is referred to [Her2006] for a more extensive account of this fact.

To further illustrate this point, we provide a few more examples of theorems which are equivalent to different choice principles:

Definition 5.0.3.3 (**Hypothesis of Cardinal Trichotomy - HCT**). For any two cardinals $\mathfrak{M}$ and $\mathfrak{N}$, either $\mathfrak{M} \preceq \mathfrak{N}$ or $\mathfrak{N} \prec \mathfrak{M}$ holds.

Cardinals are representatives of cardinality. In short, in ZFC the construction of the cardinals goes as follows: We define the *ordinals* to be the class of $\in$-well-ordered transitive sets (a set X is *transitive* if $x \in X$ implies $x \subseteq X$). Every well-ordered set $(X, \leq)$ can be proven to be order-isomorphic to a unique ordinal α. Moreover, the class *Ord* of ordinals is well-ordered by $\in$ and thus the set of ordinals being bijective with α has a least element β, we define β to be the *cardinal number* of X, denoted by $|X| = \beta$. This definition extends our previous notion of cardinality, i.e. two sets X and Y are bijective if and only if $|X| = |Y|$ in our new sense. This procedure of assigning cardinality is called *von Neumann's Cardinal Assignment.*

Note that in the formal construction of the natural numbers, we define $0 = \emptyset$ and the *successor* of n by $n + 1 = n \cup \{n\}$. From this definition, it follows that every natural number n as well as the set of natural numbers $\mathbb{N}$ are ordinals. Seeing that $\mathbb{N}$ cannot be written as $n + 1$ for any natural number (and proving that the natural numbers are the only finite ordinals), we see that $\mathbb{N}$ cannot be obtained as a successor of any ordinal. Generalizing these notions, we define the ordinal α to be a *successor ordinal* if $\alpha = \beta \cup \{\beta\}$ for some ordinal β. If α is neither zero nor a successor ordinal, then we call it α a *limit ordinal.*

In the absence of AC, there exists non-well-orderable sets and it becomes troublesome to define the cardinality of these sets. The cardinal number of a non-well-orderable set can be rigorously defined using *Scott's Trick*: We begin by defining the *von Neumann universe* $\mathcal{V} = \bigcup_{\alpha \in Ord} \mathcal{V}_\alpha$ with:

$$\begin{cases} \mathcal{V}_0 = \emptyset \\ \mathcal{V}_{\alpha+1} = \mathcal{P}(\mathcal{V}_\alpha) \\ \mathcal{V}_\beta = \bigcup_{\alpha<\beta} \mathcal{V}_\alpha, \text{ if } \beta \text{ is a limit ordinal.} \end{cases}$$

Intuitively, we are building a universe of sets by successively applying the power set operation and in the limit cases taking the union of all preceding cases. We define the *rank* of a set X to be the least ordinal α such that $X \in V_\alpha$. This enables us to define the cardinality of a non-well-orderable set X as the set of $Y \in \mathcal{V}$ of minimal rank for which there exists a bijection between X and Y.

As the cardinal numbers are supposed to represent the possible sizes a set can have, it is intuitively clear that we expect all cardinal numbers to be comparable. However, this is not the case since the following theorem holds (see Theorem 4.20 in [Her2006]):

Theorem 5.0.3.4. AC $\iff$ HCT.

The situation is not resolved by replacing $\preceq$ with $\preceq^*$ in HCT, the theorem still holds.

The following statement is similarly to TT important in topology:

Definition 5.0.3.5 (Baire Category Theorem for $\mathbb{R}$). If $(D_n)_{n=0}^\infty$ is a sequence of dense open subsets of $\mathbb{R}$, then $D = \bigcap_{n\in\mathbb{N}} D_n$ is dense in $\mathbb{R}$.

We can generalize the above statement by defining some topological concepts: Given a metric space (X, d), the set $\mathscr{O}$ of open subsets of X is called the topology *induced* by d. A topological space $(X, \mathscr{O})$ is said to be *metrizable* if there exists a metric d on X such that d induces $\mathscr{O}$. Remember that a metric space (X, d) is Cauchy-complete if every Cauchy sequence in X converges, i.e. every sequence $(x_n)_{n=0}^\infty$ in X such that the distance between its elements are diminishing, is also convergent towards some $x \in X$. A topological space $(X, \mathscr{O})$ is said to be *completely metrizable* if there exists a Cauchy-complete metric d which induces $\mathscr{O}$. Moreover, we say that $(X, \mathscr{O})$ is *Baire* if every countable intersection of dense open subsets of X is dense.

Definition 5.0.3.6 (Baire Category Theorem - BCT). Every completely metrizable topological space is Baire.

There are other alternative formulations of BCT. However, with BCT stated as above, the following theorem holds (see [I-TA]):

Theorem 5.0.3.7. DC $\iff$ BCT.

Finally, we note that the statement Fin (*for every infinite set X, there exists an injection $f : \mathbb{N} \to X$*) is related to equivalence between different notions of finiteness: We say that a set X is *Dedekind-infinite*, or simply D-infinite, if there exists a proper subset Y of X such that there exists a bijection between Y and X. If X is not D-infinite then X is *D-finite*. Our hope is of course that finite and D-finite are equivalent concepts (which coincides with infinite and D-infinite being equivalent concepts), however this is not the case in ZF due to the following theorem (see Proposition 4.2, 4.10 and 4.13 in [Her2006]):

Theorem 5.0.3.8. CC $\Rightarrow$ *X is finite if and only if X is D-finite* $\iff$ Fin.

Remembering our discussion in section 4.1 about Fin being unprovable in ZF, we see that the equivalence between these concepts of finiteness also is unprovable in ZF.

Reflections

The most remarkable result of this thesis is seeing that BTP holds in ZFC but is independent of ZF + DC. That is, even if we accept the choice principle DC which is strictly stronger than CC, we can neither prove nor disprove BTP. However, when generalizing CC to families of higher cardinality, BTP becomes provable. We note that AC is not necessary for BTP, there exist models of ZF where BTP holds and AC fails (see Diagram A.3 in [Her2006]).

The result described above together with the other results presented in this thesis build a case for being hesitant towards AC but accepting DC (moreover, DC restricted to relations on $\mathbb{N} \times \mathbb{N}$ is also intuitionistically acceptable according to [I-S], furthering our acceptance of the principle). Ultimately, ZF is our attempt to define a foundation of classical mathematics and the question is if we should consider ZF sufficient for this purpose or if we should add more axioms. This question is surrounded by philosophical considerations: What is mathematics supposed to describe? Are we allowed to change our axioms if the implications of them are not to our liking, is such behavior justified by our view of what mathematics is describing? Do we demand our axioms to be self-evident, what does self-evident even mean?

To try to avoid problems at this philosophical level, we may say that mathematics merely is the investigation of necessary relations between statements given specified axioms and rules of inference. It is satisfactory if these investigations yield results having real world applications but we are not claiming mathematics is neither absolute truth nor a description of nature, it is simply the human perception of necessity.

However, even with such a view of mathematics, the question about which axioms we should investigate the consequences of still has to be answered. It seems feasible that the answer will be *axioms having consequences consistent with our perception of reality and extending our understanding of reality to cases we lack intuitive understanding of*, it would moreover be pleasant if there was a unique such set of axioms. Our idea is to find these axioms through a trial and error procedure by finding a priori reasonable axioms such as AC and after having investigated their consequences decide if they are worthy of being axioms or not.

The real difficulty with the above procedure of finding a suitable extension of ZF lies at the crossroads: We cannot have both AC and AD, as ZFC + AD is inconsistent, so if we were to define an absolute foundation of classical mathematics, which way should we choose? Should we even add AC or AD to ZF at all? Given the weakness of ZF, we definitely should.

From our intuition about the real number line, we believe it to be desirable that every set of reals is measurable. However, measurability also has paradoxical consequences: For example, as stated and proved at [I-Ka] and [I-DS], if all sets of reals are measurable in ZF + DC then there is a partition of $\mathbb{R}$ into strictly more parts than elements.

Our results regarding BTP and its independence of ZF + DC (and thus also of ZF + CC) tells us that something happens when we generalize ideas from the countable to the uncountable. This observation can be used to argue for as well as against accepting AC: For example, our belief that every surjection have a right inverse is based on our intuition derived from the countable case, how do we know that the generalization of this statement to higher cardinalities is as trivial as we expect it to be? Similarly, how can we be expected to have any intuition about the general behavior of the real number line, such as the non-existence of non-measurable sets, when it is difficult for us to even understand what the continuum actually is?

The current state of mathematics is somewhat simplified that one accepts the variant of ZF which best fits one's needs: Those who work in classical analysis and other traditional areas of mathematics generally accept AC and probably do not question the axiom actively. Those who work in set theory investigating the existence and properties of large cardinals may be more prone to accepting AD as the axiom affects questions in this area (see [Jec2006] for a presentation of the connection between AD and large cardinals). If one is searching for actual (i.e. computable) solutions to real world problems, then ironically the way forward is probably the intuitionistic view of mathematics which historically was developed for philosophical rather than practical reasons.

References

Articles

[BaMo90] B. Banaschewski & G. H. Moore, *The Dual Cantor-Bernstein Theorem and the Partition Principle*, Notre Dame Journal of Formal Logic, v. 31 (3), 1990.

[Göd38] K. Gödel, *The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis*, Proceedings of the National Academy of Sciences of the United States of America 24 (12), 1938.

[Kec84] A. S. Kechris, *The Axiom of Determinancy Implies Dependent Choices in L(R)*, The Journal of Symbolic Logic, v. 70 (1), 1984, Association for Symbolic Logic.

[Wil2005] T. M. Wilson, *A continuous movement version of the Banach-Tarski paradox: A solution to de Groot's Problem*, Journal of Symbolic Logic, v. 70 (3), 2005, DOI:10.2178/jsl/1122038921, Association for Symbolic Logic.

Books

[BeBl2006] J. A. Beachy & W. D. Blair, *Abstract Algebra*, 3rd ed. 2006, ISBN: 1577664434, Waveland Press, Inc.

[Car2013] J. Carlström, *Logic*, Revised English edition 2013, Course material, Department of Mathematics, Stockholm University.

[Coh2013] D. L. Cohn, *Measure Theory*, 2nd ed. 2013, ISBN: 978-1-4614-6956-8 Springer New York Heidelberg Dordrecht London.

[Gol96] D. Goldrei, *Classic Set Theory - For Guided Independent Study*, 1st ed. 1996, ISBN: 0-412-60610-0, Chapman & Hall.

[Her2006] H. Herrlich, *Axiom of Choice*, Electronic edition 2006, ISSN: 1617-9692, Springer-Verlag Berlin Heidelberg New York.

[HrJe99] K. Hrbacek & T. Jech, *Introduction to Set Theory*, Third Edition, Revised and Expanded, 1999, ISBN: 0-8247-7915-0, Taylor & Francis Group.

[Jec73] T. Jech, *The Axiom of Choice*, 1973, ISBN: 0486466248, North-Holland Publishing Company.

[Jec2006] T. Jech, *Set Theory*, The Third Millennium Edition revised and expanded, corrected 4th printing 2006, ISBN: 978-3-540-37097-0, Springer-Verlag Berlin Heidelberg New York.

[Kle77] E. M. Kleinberg, *Infinitary Combinatorics and the Axiom of Determinateness*, 1st ed. 1987, ISBN: 3-540-44085-2, Springer Berlin Heidelberg.

[Loo53] L. H. Loomis, *An Introduction to Abstract Harmonic Analysis*, 1953, Library of Congress Catalogue Card No. 53-5459, D. Van Nostrand Company, Inc.

[Man2010] Y. I. Manin, *A Course in Mathematical Logic for Mathematicians*, 2nd ed. 2010, ISBN: 978-1-4419-0614-4, Springer New York Dordrecht Heidelberg London.

[Moo82] G. H. Moore, *Zermelo's Axiom of Choice - Its Origins, Development and Influence*, 1st ed. 1982, ISBN: 978-1-4613-9480-8, Springer-Verlag Berlin Heidelberg New York.

[Mun2000] J. Munkres, *Topology*, 2nd ed. 2000, ISBN: 0131816292, Pearson.

[PaPa2007] B. Palais & R. Palais, *Euler's fixed point theorem: The axis of a rotation*, Journal of Fixed Point Theory and Applications, v. 2, 2007, DOI: 10.1007/s11784-007-0042-5, Birkhäuser Verlag.

[RuRu85] H. Rubin & J. E. Rubin, *Equivalents of the Axiom of Choice, II*, Studies in logic and the foundations of mathematics, v. 116, 1985, ISBN: 0-444-87706-8, Elsevier Science Publishers B.V.

[Rud76] W. Rudin, *Principles of Mathematical Analysis*, 3rd int. ed. 1976, ISBN: 9780070856134, McGraw Hill Higher Education.

[WhZy77] R. Wheeden & A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, 1st ed. 1977, ISBN: 0824764994, Chapman & Hall.

[Whi88] E. T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed. 1988, ISBN: 9780511608797, Cambridge Mathematical Library.

Internet

[I-Cai] `http://math.stackexchange.com/questions/537461/axiom-of-countable-choice/537495#537495`

[I-DMS] http://www.dms.uaf.edu/~maxwell/math215/BinaryExpansion.pdf

[I-DS] http://mathoverflow.net/questions/22927/why-worry-about-the-axiom-of-choice/22935#22935

[I-Ka] http://boolesrings.org/asafk/2014/anti-anti-banach-tarski-arguments

[I-PW1] https://proofwiki.org/wiki/Union_from_Synthetic_Basis_is_Topology

[I-PW2] https://proofwiki.org/wiki/Equivalence_of_Definitions_of_Generated_Topology

[I-S] http://plato.stanford.edu/entries/intuitionism/

[I-TA] http://www.math.tau.ac.il/~tsirel/Courses/MeasCateg/lect4.pdf

[I-Wiki1] http://en.wikipedia.org/wiki/Well-ordering_theorem#Statement_and_sketch_of_proof

[I-Wiki2] http://en.wikipedia.org/wiki/Diaconescu's_theorem