AN INTRODUCTION TO SET THEORY

Professor William A. R. Weiss

October 2, 2008

Contents

Preface

These notes for a graduate course in set theory are on their way to becoming a book. They originated as handwritten notes in a course at the University of Toronto given by Prof. William Weiss. Cynthia Church produced the first electronic copy in December 2002. James Talmage Adams produced the copy here in February 2005. Chapters 1 to 9 are close to final form. Chapters 10, 11, and 12 are quite readable, but should not be considered as a final draft. One more chapter will be added.

Chapter 0

Introduction

Set Theory is the true study of infinity. This alone assures the subject of a place prominent in human culture. But even more, Set Theory is the milieu in which mathematics takes place today. As such, it is expected to provide a firm foundation for the rest of mathematics. And it does—up to a point; we will prove theorems shedding light on this issue.

Because the fundamentals of Set Theory are known to all mathematicians, basic problems in the subject seem elementary. Here are three simple statements about sets and functions. They look like they could appear on a homework assignment in an undergraduate course.

1. For any two sets X and Y, either there is a one-to-one function from X into Y or a one-to-one function from Y into X.
2. If there is a one-to-one function from X into Y and also a one-to-one function from Y into X, then there is a one-to-one function from X onto Y.
3. If X is a subset of the real numbers, then either there is a one-to-one function from the set of real numbers into X or there is a one-to-one function from X into the set of rational numbers.

They won't appear on an assignment, however, because they are quite difficult to prove. Statement (2) is true; it is called the Schroder-Bernstein Theorem. The proof, if you haven't seen it before, is quite tricky but nevertheless uses only standard ideas from the nineteenth century. Statement (1) is also true, but its proof needed a new concept from the twentieth century, a new axiom called the Axiom of Choice.

Statement (3) actually was on a homework assignment of sorts. It was the first problem in a tremendously influential list of twenty-three problems posed by David Hilbert to the 1900 meeting of the International Congress of Mathematicians. Statement (3) is a reformulation of the famous Continuum Hypothesis. We don't know if it is true or not, but there is hope that the twenty-first century will bring a solution. We do know, however, that another new axiom will be needed here. All these statements will be discussed later in the book.

Although Elementary Set Theory is well-known and straightforward, the modern subject, Axiomatic Set Theory, is both conceptually more difficult and more interesting. Complex issues arise in Set Theory more than any other area of pure mathematics; in particular, Mathematical Logic is used in a fundamental way. Although the necessary logic is presented in this book, it would be beneficial for the reader to have taken a prior course in logic under the auspices of mathematics, computer science or philosophy. In fact, it would be beneficial for everyone to have had a course in logic, but most people seem to make their way in the world without one.

In order to introduce one of the thorny issues, let's consider the set of all those numbers which can be easily described, say in fewer then twenty English words. This leads to something called Richard's Paradox. The set

$$\{x : x \text{ is a number which can be described}$$
$$\text{in fewer than twenty English words}\}$$

must be finite since there are only finitely many English words. Now, there are infinitely many counting numbers (i.e., the natural numbers) and so there must be some counting number (in fact infinitely many of them) which are not in our set. So there is a smallest counting number which is not in the set. This number can be uniquely described as "the smallest counting number which cannot be described in fewer than twenty English words". Count them—14 words. So the number must be *in* the set. But it can't be in the set. That's a contradiction. What is wrong here?

Our naive intuition about sets is wrong here. Not every collection of numbers with a description is a set. In fact it would be better to stay away from using languages like English to describe sets. Our first task will be to build a new language for describing sets, one in which such contradictions cannot arise.

We also need to clarify exactly what is meant by "set". What is a set? We do not know the complete answer to this question. Many problems are still unsolved simply because we do not know whether or not certain objects constitute a set or not. Most of the proposed new axioms for Set Theory are of this nature. Nevertheless, there is much that we do know about sets and this book is the beginning of the story.

Chapter 1

LOST

We construct a language suitable for describing sets.

The symbols:

variables	$v_0, v_1, v_2, \ldots$
equality symbol	$=$
membership symbol	$\in$
logical connectives	$\wedge, \vee, \neg, \rightarrow, \leftrightarrow$
quantifiers	$\forall, \exists$
parentheses	$($, $)$

The atomic formulas are strings of symbols of the form:

$$(v_i \in v_j) \text{ or } (v_i = v_j)$$

The collection of formulas of set theory is defined as follows:

1. An atomic formula is a formula.
2. If Φ is any formula, then $(\neg\Phi)$ is also a formula.
3. If Φ and Ψ are formulas, then $(\Phi \wedge \Psi)$ is also a formula.

4. If Φ and Ψ are formulas, then $(\Phi \vee \Psi)$ is also a formula.

5. If Φ and Ψ are formulas, then $(\Phi \rightarrow \Psi)$ is also a formula.

6. If Φ and Ψ are formulas, then $(\Phi \leftrightarrow \Psi)$ is also a formula.

7. If Φ is a formula and v_i is a variable, then $(\forall v_i)\Phi$ is also a formula.

8. If Φ is a formula and v_i is a variable, then $(\exists v_i)\Phi$ is also a formula.

Furthermore, any formula is built up this way from atomic formulas and a finite number of applications of the inferences 2 through 8.

Now that we have specified a language of set theory, we could specify a proof system. We will not do this here—see n different logic books for n different proof systems. However, these are essentially all the same—satisfying the completeness theorem (due to K. Gödel) which essentially says that any formula either has a proof or it has an interpretation in which it is false (but not both!). In all these proof systems we have the usual logical equivalences which are common to everyday mathematics. For example:

For any formulas Φ and Ψ:

$$\begin{array}{rll}
(\neg(\neg(\Phi))) & \text{is equivalent to} & \Phi; \\
(\Phi \wedge \Psi) & \text{is equivalent to} & \neg((\neg\Phi) \vee (\neg\Psi)); \\
(\Phi \rightarrow \Psi) & \text{is equivalent to} & ((\neg\Phi) \vee \Psi); \\
(\Phi \leftrightarrow \Psi) & \text{is equivalent to} & ((\Phi \rightarrow \Psi) \wedge (\Psi \rightarrow \Phi)); \\
(\exists v_i)\Phi & \text{is equivalent to} & (\neg(\forall v_i)(\neg\Phi)); \text{ and,} \\
(\Phi \leftrightarrow \Psi) & \text{is equivalent to} & (\Psi \leftrightarrow \Phi).
\end{array}$$

The complete collection of subformulas of a formula Φ is defined as follows:

1. Φ is a subformula of Φ;

2. If $(\neg\Psi)$ is a subformula of Φ, then so is Ψ;

3. If $(\Theta \wedge \Psi)$ is a subformula of Φ, then so are Θ and Ψ;

4. If $(\Theta \vee \Psi)$ is a subformula of Φ, then so are Θ and Ψ;
5. If $(\Theta \rightarrow \Psi)$ is a subformula of Φ, then so are Θ and Ψ;
6. If $(\Theta \leftrightarrow \Psi)$ is a subformula of Φ, then so are Θ and Ψ;
7. If $(\forall v_i)\Psi$ is a subformula of Φ and v_i is a variable, then Ψ is a subformula of Φ; and,
8. If $(\exists v_i)\Psi$ is a subformula of Φ and v_i is a variable, then Ψ is a subformula of Φ.

Note that the subformulas of Φ are those formulas used in the construction of Φ.

To say that a variable v_i occurs bound in a formula Φ means one of the following two conditions holds:

1. For some subformula Ψ of Φ, $(\forall v_i)\Psi$ is a subformula of Φ; or,
2. For some subformula Ψ of Φ, $(\exists v_i)\Psi$ is a subformula of Φ.

The result, Φ^*, of substituting the variable v_j for each bound occurrence of the variable v_i in the formula Φ is defined by constructing a Ψ^* for each subformula Ψ of Φ as follows:

1. If Ψ is atomic, then Ψ^* is Ψ;
2. If Ψ is $(\neg\Theta)$ for some formula Θ, then Ψ^* is $(\neg\Theta^*)$;
3. If Ψ is $(\Gamma \wedge \Theta)$ for some formula Θ, then Ψ^* is $(\Gamma^* \wedge \Theta^*)$;
4. If Ψ is $(\Gamma \vee \Theta)$ for some formula Θ, then Ψ^* is $(\Gamma^* \vee \Theta^*)$;
5. If Ψ is $(\Gamma \rightarrow \Theta)$ for some formula Θ, then Ψ^* is $(\Gamma^* \rightarrow \Theta^*)$;
6. If Ψ is $(\Gamma \leftrightarrow \Theta)$ for some formula Θ, then Ψ^* is $(\Gamma^* \leftrightarrow \Theta^*)$;
7. If Ψ is $(\forall v_k)\Theta$ for some formula Θ then Ψ^* is just $(\forall v_k)\Theta^*$ if $k \neq i$, but if $k = i$ then Ψ^* is $(\forall v_j)\Gamma$ where Γ is the result of substituting v_j for each occurrence of v_i in Θ; and,

8. If Ψ is $(\exists v_k)\Theta$ for some formula Θ then Ψ^* is just $(\exists v_k)\Theta^*$ if $k \neq i$, but if $k = i$ then Ψ^* is $(\exists v_j)\Gamma$ where Γ is the result of substituting v_j for each occurrence of v_i in Θ.

That a variable v_i occurs free in a formula Φ means that at least one of the following is true:

1. Φ is an atomic formula and v_i occurs in Φ;
2. Φ is $(\neg\Psi)$, Ψ is a formula and v_i occurs free in Ψ;
3. $(\Theta \wedge \Psi)$, Θ and Ψ are formulas and v_i occurs free in Θ or occurs free in Ψ;
4. Φ is $(\Theta \vee \Psi)$, Θ and Ψ are formulas and v_i occurs free in Θ or occurs free in Ψ;
5. Φ is $(\Theta \rightarrow \Psi)$, Θ and Ψ are formulas and v_i occurs free in Θ or occurs free in Ψ;
6. Φ is $(\Theta \leftrightarrow \Psi)$, Θ and Ψ are formulas and v_i occurs free in Θ or occurs free in Ψ;
7. Φ is $(\forall v_j)\Psi$ and Ψ is a formula and v_i occurs free in Ψ and $i \neq j$; or,
8. Φ is $(\exists v_j)\Psi$ and Ψ is a formula and v_i occurs free in Ψ and $i \neq j$.

As in the example below, a variable can occur both free and bound in a formula. However, notice that if a variable occurs in a formula at all it must occur either free, or bound, or both (but not at the same occurrence).

We define the important notion of the substitution of a variable v_j for each free occurrence of the variable v_i in the formula Φ. This procedure is as follows.

1. Substitute a new variable v_l for all bound occurrences of v_i in Φ.
2. Substitute another new variable v_k for all bound occurrences of v_j in the result of (1).

3. Directly substitute v_j for each occurrence of v_i in the result of (2).

Example. Let us substitute v_2 for all free occurrences of v_1 in the formula

$$((\forall v_1)((v_1 = v_2) \to (v_1 \in v_0)) \wedge (\exists v_2)(v_2 \in v_1))$$

The steps are as follows.

1. $((\forall v_1)((v_1 = v_2) \to (v_1 \in v_0)) \wedge (\exists v_2)(v_2 \in v_1))$
2. $((\forall v_3)((v_3 = v_2) \to (v_3 \in v_0)) \wedge (\exists v_2)(v_2 \in v_1))$
3. $((\forall v_3)((v_3 = v_2) \to (v_3 \in v_0)) \wedge (\exists v_4)(v_4 \in v_1))$
4. $((\forall v_3)((v_3 = v_2) \to (v_3 \in v_0)) \wedge (\exists v_4)(v_4 \in v_2))$

For the reader who is new to this abstract game of formal logic, step (2) in the substitution proceedure may appear to be unnecessary. It is indeed necessary, but the reason is not obvious until we look again at the example to see what would happen if step (2) were omitted. This step essentially changes $(\exists v_2)(v_2 \in v_1)$ to $(\exists v_4)(v_4 \in v_1)$. We can agree that each of these means the same thing, namely, "v_1 is non-empty". However, when v_2 is directly substituted into each we get something different: $(\exists v_2)(v_2 \in v_2)$ and $(\exists v_4)(v_4 \in v_2)$. The latter says that "$v_2$ is non-empty" and this is, of course what we would hope would be the result of substituting v_2 for v_1 in "v_1 is non-empty". But the former statement, $(\exists v_2)(v_2 \in v_2)$, seems quite different, making the strange assertion that "v_2 is an element of itself", and this is not what we have in mind. What caused this problem? An occurrence of the variable v_2 became bound as a result of being substituted for v_1. We will not allow this to happen. When we substitute v_2 for the free v_1 we must ensure that this freedom is preserved for v_2.

For a formula Φ and variables v_i and v_j, let $\Phi(v_i|v_j)$ denote the formula which results from substituting v_j for each free occurance of v_i. In order to make $\Phi(v_i|v_j)$ well defined, we insist that in steps (1) and (2) of the substitution process, the first new variable available is used. Of course, the use of any other new variable gives an equivalent formula. In the example, if Φ is the formula on the first line, then $\Phi(v_1|v_2)$ is the formula on the fourth line.

As a simple application we can show how to express "there exists a unique element". For any formula Φ of the language of set theory we denote by $(\exists! v_j)\Phi$ the formula

$$((\exists v_j)\Phi \wedge (\forall v_j)(\forall v_l)((\Phi \wedge \Phi(v_j|v_l)) \rightarrow (v_j = v_l)))$$

where v_l is the first available variable which does not occur in Φ. The expression $(\exists! v_j)$ can be considered as an abbreviation in the language of set theory, that is, an expression which is not actually part of the language. However, whenever we have a formula containing this expression, we can quickly convert it to a proper formula of the language of set theory.

A class is just a string of symbols of the form $\{v_i : \Phi\}$ where v_i is a variable and Φ is a formula. Two important and well-known examples are:

$$\{v_0 : (\neg(v_0 = v_0))\}$$

which is called the empty set and is usually denoted by $\emptyset$, and

$$\{v_0 : (v_0 = v_0)\}$$

which is called the universe and is usually denoted by $\mathbb{V}$.

A term is defined to be either a class or a variable. Terms are the names for what the language of set theory talks about. A grammatical analogy is that terms correspond to nouns and pronouns—classes to nouns and variables to pronouns. Continuing the analogy, the predicates, or verbs, are $=$ and $\in$. The atomic formulas are the basic relationships among the predicates and the variables.

We can incorporate classes into the language of set theory by showing how the predicates relate to them. Let Ψ and Θ be formulas of the language of set theory and let v_j, v_k and v_l be variables. We write:

$$\begin{array}{rll}
v_k \in \{v_j : \Psi\} & \text{instead of} & \Psi(v_j|v_k) \\
v_k = \{v_j : \Psi\} & \text{instead of} & (\forall v_l)((v_l \in v_k) \leftrightarrow \Psi(v_j|v_l)) \\
\{v_j : \Psi\} = v_k & \text{instead of} & (\forall v_l)(\Psi(v_j|v_l) \leftrightarrow (v_l \in v_k)) \\
\{v_j : \Psi\} = \{v_k : \Theta\} & \text{instead of} & (\forall v_l)(\Psi(v_j|v_l) \leftrightarrow \Theta(v_k|v_l)) \\
\{v_j : \Psi\} \in v_k & \text{instead of} & (\exists v_l)((v_l \in v_k) \wedge (\forall v_j)((v_j \in v_l) \leftrightarrow \Psi)) \\
\{v_j : \Psi\} \in \{v_k : \Theta\} & \text{instead of} & (\exists v_l)(\Theta(v_k|v_l) \wedge (\forall v_j)((v_j \in v_l) \leftrightarrow \Psi))
\end{array}$$

whenever v_l is neither v_j nor v_k and occurs in neither Ψ nor Θ.

We can now show how to express, as a proper formula of set theory, the substitution of a term t for each free occurrence of the variable v_i in the formula Φ. We denote the resulting formula of set theory by $\Phi(v_i|t)$. The case when t is a variable v_j has already been discussed. Now we turn our attention to the case when t is a class $\{v_j : \Psi\}$ and carry out a proceedure similar to the variable case.

1. Substitute the first available new variable for all bound occurrences of v_i in Φ.
2. In the result of (1), substitute, in turn, the first available new variable for all bound occurrences of each variable which occurs free in Ψ.
3. In the result of (2), directly substitute $\{v_j : \Psi\}$ for v_i into each atomic subformula in turn, using the table above.

For example, the atomic subformula $(v_i \in v_k)$ is replaced by the new subformula

$$(\exists v_l)((v_l \in v_k) \wedge (\forall v_j)((v_j \in v_l) \leftrightarrow \Psi))$$

where v_l is the first available new variable. Likewise, the atomic subformula $(v_i = v_i)$ is replaced by the new subformula

$$(\forall v_l)(\Psi(v_j|v_l) \leftrightarrow \Psi(v_j|v_l))$$

where v_l is the first available new variable (although it is not important to change from v_j to v_l in this particular instance).

Chapter 2

FOUND

The language of set theory is very precise, but it is extremely difficult for us to read mathematical formulas in that language. We need to find a way to make these formulas more intelligible.

In order to avoid the inconsistencies associated with Richard's paradox, we must ensure that the formula Φ in the class $\{v_j : \Phi\}$ is indeed a proper formula of the language of set theory—or, at least, can be converted to a proper formula once the abbreviations are eliminated. It is not so important that we actually write classes using proper formulas, but what *is* important is that whatever formula we write down *can* be converted into a proper formula by eliminating abbreviations and slang.

We can now relax our formalism if we keep the previous paragraph in mind. Let's adopt these conventions.

1. We can use any letters that we like for variables, not just $v_0, v_1, v_2, \ldots$.
2. We can freely omit parentheses and sometimes use brackets] and [instead.
3. We can write out "and" for "$\wedge$", "or" for "$\vee$", "implies" for "$\rightarrow$" and use the "if...then..." format as well as other common English expressions for the logical connectives and quantifiers.

4. We will use the notation $\Phi(x, y, w_1, \ldots, w_k)$ to indicate that all free variables of Φ lie among $x, y, w_1, \ldots, w_k$. When the context is clear we use the notation $\Phi(x, t, w_1, \ldots, w_k)$ for the result of substituting the term t for each free occurrence of the variable y in Φ, i.e., $\Phi(y|t)$.

5. We can write out formulas, including statements of theorems, in any way easily seen to be convertible to a proper formula in the language of set theory.

For any terms r, s, and t, we make the following abbreviations of formulas.

$$\begin{array}{rll} (\forall x \in t)\Phi & \text{for} & (\forall x)(x \in t \rightarrow \Phi) \\ (\exists x \in t)\Phi & \text{for} & (\exists x)(x \in t \wedge \Phi) \\ s \notin t & \text{for} & \neg(s \in t) \\ s \neq t & \text{for} & \neg(s = t) \\ s \subseteq t & \text{for} & (\forall x)(x \in s \rightarrow x \in t) \end{array}$$

Whenever we have a finite number of terms $t_1, t_2, \ldots, t_n$ the notation $\{t_1, t_2, \ldots, t_n\}$ is used as an abbreviation for the class:

$$\{x : x = t_1 \vee x = t_2 \vee \cdots \vee x = t_n\}.$$

Furthermore, $\{t : \Phi\}$ will stand for $\{x : x = t \wedge \Phi\}$, while $\{x \in t : \Phi\}$ will represent $\{x : x \in t \wedge \Phi\}$.

We also abbreviate the following important classes.

$$\begin{array}{rcll} \text{Union} & s \cup t & \text{for} & \{x : x \in s \vee x \in t\} \\ \text{Intersection} & s \cap t & \text{for} & \{x : x \in s \wedge x \in t\} \\ \text{Difference} & s \setminus t & \text{for} & \{x : x \in s \wedge x \notin t\} \\ \text{Symmetric Difference} & s \triangle t & \text{for} & (s \setminus t) \cup (t \setminus s) \\ \text{Ordered Pair} & \langle s, t\rangle & \text{for} & \{\{s\}, \{s, t\}\} \\ \text{Cartesian Product} & s \times t & \text{for} & \{p : \exists x\ \exists y\ (x \in s \wedge y \in t \wedge p = \langle x, y\rangle)\} \\ \text{Domain} & dom(f) & \text{for} & \{x : \exists y\ \langle x, y\rangle \in f\} \\ \text{Range} & rng(f) & \text{for} & \{y : \exists x\ \langle x, y\rangle \in f\} \end{array}$$

$$
\begin{array}{rccl}
\text{Image} & f^{\rightarrow}A & \text{for} & \{y : \exists x \in A\ \langle x,y\rangle \in f\} \\
\text{Inverse Image} & f^{\leftarrow}B & \text{for} & \{x : \exists y \in B\ \langle x,y\rangle \in f\} \\
\text{Restriction} & f|A & \text{for} & \{p : p \in f \wedge \exists x \in A\ \exists y\ p = \langle x,y\rangle\} \\
\text{Inverse} & f^{-1} & \text{for} & \{p : \exists x\ \exists y\ \langle x,y\rangle \in f\ \wedge \langle y,x\rangle = p\}
\end{array}
$$

These latter abbreviations are most often used when f is a <u>function</u>. We write

$$f \text{ is a function}$$

for

$$\forall p \in f\ \exists x\ \exists y\ p = \langle x,y\rangle\ \wedge\ (\forall x)(\exists y\ \langle x,y\rangle \in f \rightarrow \exists! y\ \langle x,y\rangle \in f)$$

and we write

$$
\begin{array}{rcl}
f\colon X \rightarrow Y & \text{for} & f \text{ is a function } \wedge dom(f) = X \wedge rng(f) \subseteq Y \\
f \text{ is one} - \text{to} - \text{one} & \text{for} & \forall y \in rng(f)\ \exists! x\ \langle x,y\rangle \in f \\
f \text{ is onto Y} & \text{for} & Y = rng(f)
\end{array}
$$

We also use the terms injection (for a one-to-one function), surjection (for an onto function), and bijection (for both properties together).

<u>Russell's Paradox</u>

The following is a theorem.

$$\neg\exists z\ z = \{x : x \notin x\}.$$

The proof of this is simple. Just ask whether or not $z \in z$.

The paradox is only for the naive, not for us. $\{x : x \notin x\}$ is a class—just a description in the language of set theory. There is no reason why what it describes should exist. In everyday life we describe many things which don't exist, fictional characters for example. Bertrand Russell did exist and Peter Pan did not, although they each have descriptions in English. Although Peter Pan does not exist, we still find it worthwhile to speak about him. The same is true in mathematics.

Upon reflection, you might say that in fact, nothing is an element of itself so that

$$\{x : x \notin x\} = \{x : x = x\} = \mathbb{V}$$

and so Russell's paradox leads to:

$$\neg\exists z\ z = \mathbb{V}.$$

It seems we have proved that the universe does not exists. A pity!

The mathematical universe fails to have a mathematical existence in the same way that the physical universe fails to have a physical existence. The things that have a physical existence are exactly the things *in* the universe, but the universe itself is not an object in the universe. This does bring up an important point—do any of the usual mathematical objects exist? What about the other things we described as classes? What about $\emptyset$? Can we prove that $\emptyset$ exists?

Actually, we can't; at least not yet. You can't prove very much, if you don't assume something to start. We could prove Russell's Paradox because, amazingly, it only required the basic rules of logic and required nothing mathematical—that is, nothing about the "real meaning" of $\in$. Continuing from Russell's Paradox to "$\neg\exists z\ z = \mathbb{V}$" required us to assume that "$\forall x\ x \notin x$"—not an unreasonable assumption by any means, but a mathematical assumption none the less. The existence of the empty set "$\exists z\ z = \emptyset$" may well be another necessary assumption.

Generally set theorists, and indeed all mathematicians, are quite willing to assume anything which is obviously true. It is, after all, only the things which are not obviously true which need some form of proof. The problem, of course, is that we must somehow know what is "obviously true". Naively, "$\exists z\ z = \mathbb{V}$" would seem to be true, but it is not and if it or any other false statement is assumed, all our proofs become infected with the virus of inconsistency and all of our theorems become suspect.

Historically, considerable thought has been given to the construction of the basic assumptions for set theory. All of mathematics is based on these assumptions; they are the foundation upon which everything else is built. These assumptions are called axioms and this system is called the $\mathfrak{ZFC}$ Axiom System. We will begin to study it in the next chapter.

Chapter 3

The Axioms of Set Theory

We will explore the $\mathfrak{ZFC}$ Axiom System. Each axiom should be "obviously true" in the context of those things that we desire to call sets. Because we cannot give a mathematical proof of a basic assumption, we must rely on intuition to determine truth, even if this feels uncomfortable. Beyond the issue of truth is the question of consistency. Since we are unable to prove that our assumptions are true, can we at least show that together they will not lead to a contradiction? Unfortunately, we cannot even do this—it is ruled out by the famous incompleteness theorems of K. Gödel. Intuition is our only guide. We begin.

We have the following axioms:

The Axiom of Equality	$\forall x\ \forall y\ [x = y\ \rightarrow\ \forall z\ (x \in z\ \leftrightarrow\ y \in z)]$
The Axiom of Extensionality	$\forall x\ \forall y\ [x = y\ \leftrightarrow\ \forall u\ (u \in x\ \leftrightarrow\ u \in y)]$
The Axiom of Existence	$\exists z\ z = \emptyset$
The Axiom of Pairing	$\forall x\ \forall y\ \exists z\ z = \{x, y\}$

Different authors give slightly different formulations of the $\mathfrak{ZFC}$ axioms. All formulations are equivalent. Some authors omit the Axiom of Equality and Axiom of Existence because they are consequences of the usual logical background to all mathematics. We include them for emphasis. Redundancy is not a bad thing and there is considerable redundancy in this system.

The following theorem gives some results that we would be quite willing to assume outright, were they not to follow from the axioms. The first three parts are immediate consequences of the Axiom of Extensionality.

Theorem 1.

1. $\forall x \; x = x$.
2. $\forall x \; \forall y \; x = y \rightarrow y = x$.
3. $\forall x \; \forall y \; \forall z \; [(x = y \wedge y = z) \rightarrow x = z]$.
4. $\forall x \; \forall y \; \exists z \; z = \langle x, y \rangle$.
5. $\forall u \; \forall v \; \forall x \; \forall y \; [\langle u, v \rangle = \langle x, y \rangle \leftrightarrow (u = x \wedge v = y)]$.

Exercise 1. Prove parts (4) and (5) of Theorem 1

We now assert the existence of unions and intersections. No doubt the reader has experienced a symmetry between these two concepts. Here however, while the Union Axiom is used extensively, the Intersection Axiom is redundant and is omitted in most developments of the subject. We include it here because it has some educational value (see Exercise 4).

The Union Axiom $\qquad \forall x \; [x \neq \emptyset \rightarrow \exists z \; z = \{w : (\exists y \in x)(w \in y)\}]$

The class $\{w : (\exists y \in x)(w \in y)\}$ is abbreviated as $\bigcup x$ and called the "big union".

The Intersection Axiom $\qquad \forall x \; [x \neq \emptyset \rightarrow \exists z \; z = \{w : (\forall y \in x)(w \in y)\}]$

The class $\{w : (\forall y \in x)(w \in y)\}$ is abbreviated as $\bigcap x$ and called the "big intersection".

The Axiom of Foundation $\qquad \forall x \; [x \neq \emptyset \; \rightarrow \; (\exists y \in x)(x \cap y = \emptyset)]$

This axiom, while it may be "obviously true", is not certainly obvious. Let's investigate what it says: suppose there were a non-empty x such that $(\forall \; y \; \in \; x) \; (x \cap y \neq \emptyset)$. For any $z_1 \in x$ we would be able to get $z_2 \in z_1 \cap x$. Since $z_2 \in x$ we would be able to get $z_3 \in z_2 \cap x$. The process continues forever:

$$\cdots \in z_4 \in z_3 \in z_2 \in z_1 \in x$$

We wish to rule out such an infinite regress. We want our sets to be founded: each such sequence should eventually end with the empty set. Hence the name of the axiom, which is also known as the Axiom of Regularity. It is nevertheless best understood by its consequences.

Theorem 2.

1. $\forall x\ \forall y\ \exists z\ z = x \cup y$.
2. $\forall x\ \forall y\ \exists z\ z = x \cap y$.
3. $\forall x\ \forall y\ x \in y \rightarrow y \notin x$.
4. $\forall x\ x \notin x$.

Exercise 2. Prove Theorem 2.

Let $f(x)$ denote the class $\bigcup\{y : \langle x, y\rangle \in f\}$.

Exercise 3. Suppose f is a function and $x \in dom(f)$. Prove that

$$\langle x, y\rangle \in f \text{ iff } y = f(x).$$

Suppose that x is a set and that there is some way of removing each element $u \in x$ and replacing u with some element v. Would the result be a set? Well, of course—provided there are no tricks here. That is, there should be a well defined replacement procedure which ensures that each u is replaced by only one v. This well defined procedure should be described by a formula, Φ, in the language of set theory. We can guarantee that each u is replaced by exactly one v by insisting that $\forall u \in x\ \exists! v\ \Phi(x, u, v)$.

We would like to obtain an axiom, written in the language of set theory stating that for each set x and each such formula Φ we get a set z. However, this is impossible. We cannot express "for each formula" in the language of set theory—in fact this formal language was designed for the express purpose of avoiding such expressions which bring us perilously close to Richard's Paradox.

The answer to this conundrum is to utilise not just one axiom, but infinitely many—one axiom for each formula of the language of set theory. Such a system is called an axiom scheme.

The Replacement Axiom Scheme

For each formula $\Phi(x, u, v, w_1, \ldots, w_n)$ of the language of set theory, we have the axiom:

$$\forall w_1 \ldots \forall w_n \ \forall x \ [\forall u \in x \ \exists! v \ \Phi \rightarrow \exists z \ z = \{v : \exists u \in x \ \Phi\}]$$

Note that we have allowed Φ to have $w_1, \ldots, w_n$ as parameters, that is, free variables which may be used to specify various procedures in various contexts within a mathematical proof. This is illustrated by the following theorem.

Theorem 3. *$\forall x \ \forall y \ \exists z \ z = x \times y$.*

Proof. From Theorem 1 parts (4) and (5), for all $t \in y$ we get

$$\forall u \in x \ \exists! v \ v = \langle u, t \rangle.$$

We now use Replacement with the formula "$\Phi(x, u, v, t)$" as "$v = \langle u, t \rangle$"; t is a parameter. We obtain, for each $t \in y$:

$$\exists q \ q = \{v : \exists u \in x \ v = \langle u, t \rangle\}.$$

By Extensionality, in fact $\forall t \in y \ \exists! q \ q = \{v : \exists u \in x \ v = \langle u, t \rangle\}$.

We again use Replacement, this time with the formula $\Phi(y, t, q, x)$ as "$q = \{v : \exists u \in x \ v = \langle u, t \rangle\}$"; here x is a parameter. We obtain:

$$\exists r \ r = \{q : \exists t \in y \ q = \{v : \exists u \in x \ v = \langle u, t \rangle\}\}$$

By the Union Axiom $\exists z \ z = \bigcup r$ and so we have:

$$\begin{aligned} z &= \{p : \exists q \ [q \in r \wedge p \in q]\} \\ &= \{p : \exists q \ [(\exists t \in y) \ q = \{v : \exists u \in x \ v = \langle u, t \rangle\} \wedge p \in q]\} \\ &= \{p : (\exists t \in y)(\exists q)[q = \{v : \exists u \in x \ v = \langle u, t \rangle\} \wedge p \in q]\} \\ &= \{p : (\exists t \in y) p \in \{v : \exists u \in x \ v = \langle u, t \rangle\}\} \\ &= \{p : (\exists t \in y)(\exists u \in x) \ p = \langle u, t \rangle\} \\ &= x \times y \end{aligned}$$

□

Exercise 4. Show that the Intersection Axiom is indeed redundant.

It is natural to believe that for any set x, the collection of those elements $y \in x$ which satisfy some particular property should also be a set. Again, no tricks—the property should be specified by a formula of the language of set theory. Since this should hold for any formula, we are again led to a scheme.

The Comprehension Scheme

For each formula $\Phi(x, y, w_1, \ldots, w_n)$ of the language of set theory, we have the statement:

$$\forall w_1 \ldots \forall w_n \; \forall x \; \exists z \;\; z = \{y : y \in x \wedge \Phi(x, y, w_1, \ldots, w_n)\}$$

This scheme could be another axiom scheme (and often is treated as such). However, this would be unnecessary, since the Comprehension Scheme follows from what we have already assumed. It is, in fact, a theorem scheme—that is, infinitely many theorems, one for each formula of the language of set theory. Of course we cannot write down infinitely many proofs, so how can we prove this theorem scheme?

We give a uniform method for proving each instance of the scheme. So to be certain that any given instance of the theorem scheme is true, we consider the uniform method applied to that particular instance. We give this general method below.

For each formula $\Phi(x, u, w_1, \ldots, w_n)$ of the language of set theory we have:

Theorem 4. Φ

$\forall w_1 \ldots \forall w_n \; \forall x \; \exists z \;\; z = \{u : u \in x \wedge \Phi\}.$

Proof. Apply Replacement with the formula $\Psi(x, u, v, w_1, \ldots, w_n)$ given by:

$$(\Phi(x, u, w_1, \ldots, w_n) \rightarrow v = \{u\}) \wedge (\neg\Phi(x, u, w_1, \ldots, w_n) \rightarrow v = \emptyset)$$

to obtain:

$$\exists y \;\; y = \{v : (\exists u \in x)[(\Phi \rightarrow v = \{u\}) \wedge (\neg\Phi \rightarrow v = \emptyset)]\}.$$

Note that $\{\{u\} : \Phi(x, u, w_1, \ldots, w_n)\} \subseteq y$ and the only other possible element of y is $\emptyset$. Now let $z = \bigcup y$ to finish the proof.

$\square$

Theorem 4 Φ can be thought of as infinitely many theorems, one for each Φ. The proof of any one of those theorems can be done in a finite number of steps, which invoke only a finite number of theorems or axioms. A proof cannot have infinite length, nor invoke infinitely many axioms or lemmas.

We state the last of the "set behavior" axioms.

The Axiom of Choice

$$\forall X\ [(\forall x \in X\ \forall y \in X\ (x = y \leftrightarrow x \cap y \neq \emptyset)) \rightarrow \exists z\ (\forall x \in X\ \exists! y\ y \in x \cap z)]$$

In human language, the Axiom of Choice says that if you have a collection X of pairwise disjoint non-empty sets, then you get a set z which contains one element from each set in the collection. Although the axiom gives the existence of some "choice set" z, there is no mention of uniqueness—there are quite likely many possible sets z which satisfy the axiom and we are given no formula which would single out any one particular z.

The Axiom of Choice can be viewed as a kind of replacement, in which each set in the collection is replaced by one of its elements. This leads to the following useful reformulation which will be used in Theorem 22.

Theorem 5. *There is a choice function on any set of non-empty sets; i.e.,*

$$\forall X\ [\emptyset \notin X \rightarrow (\exists f)(f \colon X \rightarrow \bigcup X \wedge (\forall x \in X)(f(x) \in x))].$$

Proof. Given such an X, by Replacement there is a set

$$Y = \{\{x\} \times x : x \in X\}$$

which satisfies the hypothesis of the Axiom of Choice. So, $\exists z\ \forall y \in Y\ \exists! p\ p \in y \cap z$. Let $f = z \cap (\bigcup Y)$. Then $f \colon X \rightarrow \bigcup X$ and each $f(x) \in x$.

$\square$

We state the last of the "set creation" axioms.

The Power Set Axiom $\quad \forall x \; \exists z \; z = \{y : y \subseteq x\}$

We denote $\{y : y \subseteq x\}$ by $\mathcal{P}(x)$, called the power set of x. For reasons to be understood later, it is important to know explicitly when the Power Set Axiom is used. This completes the list of the $\mathfrak{ZFC}$ Axiom System with one exception to come later—higher analogues of the Axiom of Existence.

Chapter 4

The Natural Numbers

We now construct the natural numbers. That is, we will represent the natural numbers in our universe of set theory. We will construct a number system which behaves mathematically exactly like the natural numbers, with exactly the same arithmetic and order properties. We will not claim that what we construct are the actual natural numbers—whatever they are made of. But we will take the liberty of calling our constructs "the natural numbers". We begin by taking 0 as the empty set $\emptyset$. We write

$$\begin{array}{rcl} 1 & \text{for} & \{0\} \\ 2 & \text{for} & \{0, 1\} \\ 3 & \text{for} & \{0, 1, 2\} \\ succ(x) & \text{for} & x \cup \{x\} \end{array}$$

We write "n is a natural number" for

$$[n = \emptyset \vee (\exists l \in n)(n = succ(l))] \wedge (\forall m \in n)[m = \emptyset \vee (\exists l \in n)(m = succ(l))]$$

and write:

$$\mathbb{N} \quad \text{for} \quad \{n : n \text{ is a natural number}\}$$

The reader can gain some familiarity with these definitions by checking that $succ(n) \in \mathbb{N}$ for all $n \in \mathbb{N}$.

We now begin to develop the basic properties of the natural numbers by introducing an important concept. We say that a term t is transitive whenever we have $(\forall x \in t)(x \subseteq t)$.

Theorem 6.

1. *Each natural number is transitive.*
2. $\mathbb{N}$ *is transitive; i.e., every element of a natural number is a natural number.*

Proof. Suppose that (1) were false; i.e., some $n \in \mathbb{N}$ is not transitive, so that:

$$\{k : k \in n \text{ and } \neg(k \subseteq n)\} \neq \emptyset.$$

By Comprehension $\exists x\ x = \{k \in n : \neg(k \subseteq n)\}$ and so by Foundation there is $y \in x$ such that $y \cap x = \emptyset$. Note that since $\emptyset \notin x$ and $y \in n$ we have that $y = succ(l)$ for some $l \in n$. But since $l \in y$, $l \notin x$ and so $l \subseteq n$. Hence $y = l \cup \{l\} \subseteq n$, contradicting that $y \in x$.

We also prove (2) indirectly; suppose $n \in \mathbb{N}$ with

$$\{m : m \in n \text{ and } m \notin \mathbb{N}\} \neq \emptyset.$$

By Comprehension $\exists x\ x = \{m \in n : m \notin \mathbb{N}\}$ and so Foundation gives $y \in x$ such that $y \cap x = \emptyset$. Since $y \in n$, we have $y = succ(l)$ for some $l \in n$. Since $l \in y$ and $y \cap x = \emptyset$ we must have $l \in \mathbb{N}$. But then $y = succ(l) \in \mathbb{N}$, contradicting that $y \in x$.

□

Theorem 7. *(Trichotomy of Natural Numbers)*

Let $m, n \in \mathbb{N}$. Exactly one of three situations occurs:

$$m \in n,\ \ n \in m,\ \ m = n.$$

Proof. That at most one occurs follows from Theorem 2. That at least one occurs follows from this lemma.

□

Lemma. Let $m, n \in \mathbb{N}$.

1. If $m \subseteq n$, then either $m = n$ or $m \in n$.
2. If $m \notin n$, then $n \subseteq m$.

Proof. We begin the proof of (1) by letting S denote

$$\{x \in \mathbb{N} : (\exists y \in \mathbb{N})(y \subseteq x \text{ and } y \neq x \text{ and } y \notin x)\}.$$

It will suffice to prove that $S = \emptyset$. We use an indirect proof—pick some $n_1 \in S$. If $n_1 \cap S \neq \emptyset$, Foundation gives us $n_2 \in n_1 \cap S$ with $n_2 \cap (n_1 \cap S) = \emptyset$. By transitivity, $n_2 \subseteq n_1$ so that $n_2 \cap S = \emptyset$. Thus, we always have some $n \in S$ such that $n \cap S = \emptyset$.

For just such an n, choose $m \in \mathbb{N}$ with $m \subseteq n$, $m \neq n$, and $m \notin n$. Using Foundation, choose $l \in n \setminus m$ such that $l \cap (n \setminus m) = \emptyset$. Transitivity gives $l \subseteq n$, so we must have $l \subseteq m$. We have $l \neq m$ since $l \in n$ and $m \notin n$. Therefore we conclude that $m \setminus l \neq \emptyset$.

Using Foundation, pick $k \in m \setminus l$ such that $k \cap (m \setminus l) = \emptyset$. Transitivity of m gives $k \subseteq m$ and so we have $k \subseteq l$. Now, because $l \in n$ we have $l \in \mathbb{N}$ and $l \notin S$ so that either $k = l$ or $k \in l$. However, $k = l$ contradicts $l \notin m$ and $k \in l$ contradicts $k \in m \setminus l$.

We prove the contrapositive of (2). Suppose that n is not a subset of m; using Foundation pick $l \in n \setminus m$ such that $l \cap (n \setminus m) = \emptyset$. By transitivity, $l \subseteq n$ and hence $l \subseteq m$. Now by (1) applied to l and m, we conclude that $l = m$. Hence $m \in n$.

$\square$

These theorems show that "$\in$" behaves on $\mathbb{N}$ just like the usual ordering "$<$" on the natural numbers. In fact, we often use "$<$" for "$\in$" when writing about the natural numbers. We also use the relation symbols $\leq$, $>$, and $\geq$ in their usual sense.

The next theorem scheme justifies ordinary mathematical induction. For brevity let us write $\vec{w}$ for $w_1, \ldots, w_n$.

For each formula $\Phi(v, \vec{w})$ of the language of set theory we have:

Theorem 8. Φ

For all $\vec{w}$, if

$$\forall n \in \mathbb{N}\ [(\forall m \in n\ \Phi(m, \vec{w})) \to \Phi(n, \vec{w})]$$

then

$$\forall n \in \mathbb{N}\ \Phi(n, \vec{w}).$$

Proof. We will assume that the theorem is false and derive a contradiction. We have $\vec{w}$ and a fixed $l \in \mathbb{N}$ such that $\neg\Phi(l, \vec{w})$.

Let t be any transitive subset of $\mathbb{N}$ containing l (e.g., $t = l \cup \{l\}$). By Comprehension, $\exists s\ s = \{n \in t : \neg\Phi(n, \vec{w})\}$. By Foundation, we get $y \in s$ such that $y \cap s = \emptyset$. Transitivity of t guarantees that $(\forall n \in y)\ \Phi(n, \vec{w})$. This, in turn, contradicts that $y \in s$.

$\square$

The statement $\forall m \in n\ \Phi(m, \vec{w})$ in Theorem 8 Φ is usually called the inductive hypothesis.

Exercise 5. Prove or disprove that for each formula $\Phi(v, \vec{w})$ we have

$$\forall \vec{w}\ [(\forall n \in \mathbb{N})((\forall m > n\ \Phi(m, \vec{w})) \to \Phi(n, \vec{w})) \to \forall n \in \mathbb{N}\ \Phi(n, \vec{w})].$$

Recursion on $\mathbb{N}$ is a way of defining new terms (in particular, functions with domain $\mathbb{N}$). Roughly speaking, values of a function F at larger numbers are defined in terms of the values of F at smaller numbers.

We begin with the example of a function F, where we set $F(0) = 3$ and $F(succ(n)) = succ(F(n))$ for each natural number n. We have set out a short recursive procedure which gives a way to calculate $F(n)$ for any $n \in \mathbb{N}$. The reader may carry out this procedure a few steps and recognise this function F as $F(n) = 3 + n$. However, all this is a little vague. What exactly is F?

In particular, is there a formula for calculating F? How do we verify that F behaves like we think it should?

In order to give some answers to these questions, let us analyse the example. There is an implicit formula for the calculation of $y = F(x)$ which is

$$[x = 0 \rightarrow y = 3] \wedge (\forall n \in \mathbb{N})[x = succ(n) \rightarrow y = succ(F(n))]$$

However the formula involves F, the very thing that we are trying to describe. Is this a vicious circle? No — the formula only involves the value of F at a number n less than x, not $F(x)$ itself. In fact, you might say that the formula doesn't really involve F at all; it just involves $F|x$. Let's rewrite the formula as

$$[x = 0 \rightarrow y = 3] \wedge (\forall n)[x = succ(n) \rightarrow y = succ(f(n))]$$

and denote it by $\Phi(x, f, y)$. Our recursive procedure is then described by

$$\Phi(x, F|x, F(x)).$$

In order to describe F we use functions f which approximate F on initial parts of its domain, for example $f = \{\langle 0, 3\rangle\}$, $f = \{\langle 0, 3\rangle, \langle 1, 4\rangle\}$ or

$$f = \{\langle 0, 3\rangle, \langle 1, 4\rangle, \langle 2, 5\rangle\},$$

where each such f satisfies $\Phi(x, f|x, f(x))$ for the appropriate x's. We will obtain F as the amalgamation of all these little f's. F is

$$\{\langle x, y\rangle : (\exists n \in \mathbb{N})(\exists f)[f\colon n \rightarrow \mathbb{V} \wedge f(x) = y \wedge \forall m \in n\ \Phi(m, f|m, f(m))]\}.$$

But in order to justify this we will need to notice that

$$(\forall x \in \mathbb{N})(\forall f)[(f\colon x \rightarrow \mathbb{V}) \rightarrow \exists! y\ \Phi(x, f, y)],$$

which simply states that we have a well defined procedure given by Φ.

Let us now go to the general context in which the above example will be a special case. For any formula $\Phi(x, f, y, \vec{w})$ of the language of set theory, we denote by $REC(\Phi, \mathbb{N}, \vec{w})$ the class

$$\{\langle x, y\rangle : (\exists n \in \mathbb{N})(\exists f)[f\colon n \rightarrow \mathbb{V} \wedge f(x) = y \wedge \forall m \in n\ \Phi(m, f|m, f(m), \vec{w})]\}.$$

We will show, under the appropriate hypothesis, that $REC(\Phi, \mathbb{N}, \vec{w})$ is the unique function on $\mathbb{N}$ which satisfies the procedure given by Φ. This requires a theorem scheme.

For each formula $\Phi(x, f, y, \vec{w})$ of the language of set theory we have:

Theorem 9. Φ

For all $\vec{w}$, suppose that we have

$$(\forall x \in \mathbb{N})(\forall f)[(f\colon x \to \mathbb{V}) \to \exists! y\ \Phi(x, f, y, \vec{w})].$$

Then, letting F denote $REC(\Phi, \mathbb{N}, \vec{w})$, we have:

1. $F\colon \mathbb{N} \to \mathbb{V}$;
2. $\forall m \in \mathbb{N}\ \Phi(m, F|m, F(m), \vec{w})$;

   and, furthermore, for any $n \in \mathbb{N}$ and any function H with $n \in dom(H)$, we have:

3. *If $\Phi(m, H|m, H(m), \vec{w})$ for all $m \in n \cup \{n\}$, then $H(n) = F(n)$.*

Proof. We first prove the following claim.
Claim.

$$(\forall x \in \mathbb{N})(\forall y_1)(\forall y_2)[(\langle x, y_1\rangle \in F \wedge \langle x, y_2\rangle \in F \to y_1 = y_2]$$

Proof of Claim. By definition of F we have, for $i = 1, 2$, functions f_i with domains $n_i \in \mathbb{N}$ such that $f_i(x) = y_i$ and

$$(\forall m \in n_i)\ \Phi(m, f_i|m, f_i(m), \vec{w}).$$

It suffices to prove that

$$(\forall m \in \mathbb{N})(m \in x \cup \{x\} \to f_1(m) = f_2(m)),$$

which we do by induction on $m \in \mathbb{N}$. To this end, we assume that $m \in \mathbb{N}$ and

$$(\forall j \in m)(j \in x \cup \{x\} \to f_1(j) = f_2(j))$$

with intent to show that

$$m \in x \cup \{x\} \to f_1(m) = f_2(m).$$

To do this suppose $m \in x \cup \{x\}$. Since $x \in n_1 \cap n_2$ we have $m \in n_1 \cap n_2$ so that we have both

$$\Phi(m, f_1|m, f_1(m), \vec{w}) \text{ and } \Phi(m, f_2|m, f_2(m), \vec{w}).$$

By transitivity $j \in x \cup \{x\}$ for all $j \in m$ and so by the inductive hypothesis $f_1|m = f_2|m$. Now by the hypothesis of this theorem with $f = f_1|m = f_2|m$ we deduce that $f_1(m) = f_2(m)$. This concludes the proof of the claim.

In order to verify (1), it suffices to show that

$$(\forall x \in \mathbb{N})(\exists y) \; [\langle x, y\rangle \in F]$$

by induction. To this end, we assume that

$$(\forall j \in x)(\exists y) \; [\langle j, y\rangle \in F]$$

with intent to show that $\exists y \; \langle x, y\rangle \in F$. For each $j \in x$ there is $n_j \in \mathbb{N}$ and $f_j \colon n_j \to \mathbb{V}$ such that

$$(\forall m \in n_j) \; \Phi(m, f_j|m, f_j(m), \vec{w}).$$

If $x \in n_j$ for some j, then $\langle x, f_j(x)\rangle \in F$ and we are done; so assume that $n_j \leq x$ for all j. Let $g = \bigcup\{f_j : j \in x\}$. By the claim, the f_j's agree on their common domains, so that g is a function with domain x and

$$(\forall m \in x) \; \Phi(m, g|m, g(m), \vec{w}).$$

By the hypothesis of the theorem applied to g there is a unique y such that $\Phi(x, g, y, \vec{w})$. Define f to be the function $f = g \cup \{\langle x, y\rangle\}$. It is straightforward to verify that f witnesses that $\langle x, y\rangle \in F$.

To prove (2), note that, by (1), for each $x \in \mathbb{N}$ there is $n \in \mathbb{N}$ and $f \colon n \to \mathbb{V}$ such that $F(x) = f(x)$ and, in fact, $F|n = f$. Hence,

$$(\forall m \in n) \; \Phi(m, f|m, f(m), \vec{w}).$$

We prove (3) by induction. Assume that

$$(\forall m \in n)\ H(m) = F(m)$$

with intent to show that $H(n) = F(n)$. We assume $\Phi(n, H|n, H(n), \vec{w})$ and by (2) we have $\Phi(n, F|n, F(n), \vec{w})$. By the hypothesis of the theorem applied to $H|n = F|n$ we get $H(n) = F(n)$.

□

By applying this theorem to our specific example we see that $REC(\Phi, \mathbb{N}, \vec{w})$ does indeed give us a function F. Since F is defined by recursion on $\mathbb{N}$, we use induction on $\mathbb{N}$ to verify the properties of F. For example, it is easy to use induction to check that $F(n) \in \mathbb{N}$ for all $n \in \mathbb{N}$.

We do not often explicitly state the formula Φ in a definition by recursion. The definition of F would be more often given by:

$$F(0) = 3$$
$$F(succ(n)) = succ(F(n))$$

This is just how the example started; nevertheless, this allows us to construct the formula Φ immediately, should we wish. Of course, in this particular example we can use the plus symbol and give the definition by recursion by the following formulas.

$$3 + 0 = 3$$
$$3 + succ(n) = succ(3 + n)$$

Now, let's use definition by recursion in other examples. We can define general addition on $\mathbb{N}$ by the formulas

$$a + 0 = a$$
$$a + succ(b) = succ(a + b)$$

for each $a \in \mathbb{N}$. Here a is a parameter which is allowed by the inclusion of $\vec{w}$ in our analysis. The same trick can be used for multiplicaton:

$$a \cdot 0 = 0$$
$$a \cdot (succ(b)) = a \cdot b + a$$

for each $a \in \mathbb{N}$, using the previously defined notion of addition. In each example there are two cases to specify—the zero case and the successor case. Exponentiation is defined similarly:

$$a^0 = 1$$
$$a^{succ(b)} = a^b \cdot a$$

The reader is invited to construct, in each case, the appropriate formula Φ, with a as a parameter, and to check that the hypothesis of the previous theorem is satisfied.

G. Peano developed the properties of the natural numbers from zero, the successor operation and induction on $\mathbb{N}$. You may like to see for yourself some of what this entails by proving that multiplication is commutative.

A set X is said to be finite provided that there is a natural number n and a bijection $f\colon n \to X$. In this case n is said to be the size of X. Otherwise, X is said to be infinite.

Exercise 6. Use induction to prove the "pigeon-hole principle": for $n \in \mathbb{N}$ there is no injection $f\colon (n+1) \to n$. Conclude that a set X cannot have two different sizes.

Do not believe this next result:

Proposition. All natural numbers are equal.

Proof. It is sufficient to show by induction on $n \in \mathbb{N}$ that if $a \in \mathbb{N}$ and $b \in \mathbb{N}$ and $\max\,(a, b) = n$, then $a = b$. If $n = 0$ then $a = 0 = b$. Assume the inductive hypothesis for n and let $a \in \mathbb{N}$ and $b \in \mathbb{N}$ be such that

$$\max\,(a, b) = n + 1.$$

Then $\max\,(a-1, b-1) = n$ and so $a-1 = b-1$ and consequently $a = b$.

$\square$

Chapter 5

The Ordinal Numbers

The natural number system can be extended to the system of ordinal numbers.

An ordinal is a transitive set of transitive sets. More formally: for any term t, "t is an ordinal" is an abbreviation for

$$(t \text{ is transitive}) \wedge (\forall x \in t)(x \text{ is transitive}).$$

We often use lower case Greek letters to denote ordinals. We denote $\{\alpha : \alpha \text{ is an ordinal}\}$ by $\mathbb{ON}$.

From Theorem 6 we see immediately that $\mathbb{N} \subseteq \mathbb{ON}$.

Theorem 10.

1. $\mathbb{ON}$ *is transitive.*
2. $\neg(\exists z)(z = \mathbb{ON})$.

Proof.

1. Let $\alpha \in \mathbb{ON}$; we must prove that $\alpha \subseteq \mathbb{ON}$. Let $x \in \alpha$; we must prove that

(a) x is transitive; and,

(b) $(\forall y \in x)(y \text{ is transitive})$.

Clearly (a) follows from the definition of ordinal. To prove (b), let $y \in x$; by transitivity of α we have $y \in \alpha$; hence y is transitive.

2. Assume $(\exists z)(z = \mathbb{ON})$. From (1) we have that $\mathbb{ON}$ is a transitive set of transitive sets, i.e., an ordinal. This leads to the contradiction $\mathbb{ON} \in \mathbb{ON}$.

□

Theorem 11. *(Trichotomy of Ordinals)*

$$(\forall \alpha \in \mathbb{ON})(\forall \beta \in \mathbb{ON})(\alpha \in \beta \vee \beta \in \alpha \vee \alpha = \beta).$$

Proof. The reader may check that a proof of this theorem can be obtained by replacing "$\mathbb{N}$" with "$\mathbb{ON}$" in the proof of Theorem 7.

□

Because of this theorem, when α and β are ordinals, we often write $\alpha < \beta$ for $\alpha \in \beta$.

Since $\mathbb{N} \subseteq \mathbb{ON}$, it is natural to wonder whether $\mathbb{N} = \mathbb{ON}$. In fact, we know that "$\mathbb{N} = \mathbb{ON}$" can be neither proved nor disproved from the axioms that we have stated (provided, of course, that those axioms are actually consistent). We find ourselves at a crossroads in Set Theory. We can either add "$\mathbb{N} = \mathbb{ON}$" to our axiom system, or we can add "$\mathbb{N} \neq \mathbb{ON}$".

As we shall see, the axiom "$\mathbb{N} = \mathbb{ON}$" essentially says that there are no infinite sets and the axiom "$\mathbb{N} \neq \mathbb{ON}$" essentially says that there are indeed infinite sets. Of course, we go for the infinite!

The Axiom of Infinity $\quad \mathbb{N} \neq \mathbb{ON}$

As a consequence, there is a set of all natural numbers; in fact, $\mathbb{N} \in \mathbb{ON}$.

Theorem 12. $(\exists z)(z \in \mathbb{ON} \wedge z = \mathbb{N})$.

Proof. Since $\mathbb{N} \subseteq \mathbb{ON}$ and $\mathbb{N} \neq \mathbb{ON}$, pick $\alpha \in \mathbb{ON} \setminus \mathbb{N}$. We claim that for each $n \in \mathbb{N}$ we have $n \in \alpha$; in fact, this follows immediately from the trichotomy of ordinals and the transitivity of $\mathbb{N}$. Thus $\mathbb{N} = \{x \in \alpha : x \in \mathbb{N}\}$ and by Comprehension $\exists z\ z = \{x \in \alpha : x \in \mathbb{N}\}$. The fact that $\mathbb{N} \in \mathbb{ON}$ now follows immediately from Theorem 6.

□

The lower case Greek letter ω is reserved for the set $\mathbb{N}$ considered as an ordinal; i.e., $\omega = \mathbb{N}$. Theorems 6 and 12 now show that the natural numbers are the smallest ordinals, which are immediately succeeded by ω, after which the rest follow. The other ordinals are generated by two processes illustrated by the next lemma.

Lemma.

1. $\forall \alpha \in \mathbb{ON}\ \exists \beta \in \mathbb{ON}\ \beta = succ(\alpha)$.
2. $\forall S\ [S \subseteq \mathbb{ON} \rightarrow \exists \beta \in \mathbb{ON}\ \beta = \bigcup S]$.

Exercise 7. Prove this lemma.

For $S \subseteq \mathbb{ON}$ we write $\sup S$ for the least element of

$$\{\beta \in \mathbb{ON} : (\forall \alpha \in S)(\alpha \leq \beta)\}$$

if such an element exists.

Lemma. $\forall S\ [S \subseteq \mathbb{ON} \rightarrow \bigcup S = \sup S]$

Exercise 8. Prove this lemma.

An ordinal α is called a <u>successor ordinal</u> whenever $\exists \beta \in \mathbb{ON}\ \alpha = succ(\beta)$. If $\alpha = \sup\ \alpha$, then α is called a <u>limit ordinal</u>.

Lemma. Each ordinal is either a successor ordinal or a limit ordinal, but not both.

Exercise 9. Prove this lemma.

We can perform induction on the ordinals via a process called transfinite induction. In order to justify transfinite induction we need a theorem scheme.

For each formula $\Phi(v, \vec{w})$ of the language of set theory we have:

Theorem 13. Φ

For all $\vec{w}$, *if*

$$\forall n \in \mathbb{ON}\ [(\forall m \in n\ \Phi(m, \vec{w})) \to \Phi(n, \vec{w})]$$

then

$$\forall n \in \mathbb{ON}\ \Phi(n, \vec{w}).$$

Proof. The reader may check that a proof of this theorem scheme can be obtained by replacing "$\mathbb{N}$" with "$\mathbb{ON}$" in the proof of Theorem Scheme 8.

□

We can also carry out recursive definitions on $\mathbb{ON}$. This process is called transfinite recursion. For any formula $\Phi(x, f, y, \vec{w})$ of the language of set theory, we denote by $REC(\Phi, \mathbb{ON}, \vec{w})$ the class

$$\{\langle x, y\rangle : (\exists n \in \mathbb{ON})(\exists f)[f : n \to \mathbb{V} \wedge f(x) = y \wedge \forall m \in n\ \Phi(m, f|m, f(m), \vec{w})]\}.$$

Transfinite recursion is justified by the following theorem scheme.

For each formula $\Phi(x, f, y, \vec{w})$ of the language of set theory we have:

Theorem 14. Φ

For all $\vec{w}$, *suppose that we have*

$$(\forall x \in \mathbb{ON})(\forall f)[(f : x \to \mathbb{V}) \to \exists! y\ \Phi(x, f, y, \vec{w})].$$

Then, letting F *denote* $REC(\Phi, \mathbb{ON}, \vec{w})$, *we have:*

1. *$F : \mathbb{ON} \to \mathbb{V}$;*

2. *$\forall x \in \mathbb{ON}\ \Phi(x, F|x, F(x), \vec{w})$;*

   and, furthermore, for any $n \in \mathbb{ON}$ and any function H with $n \in dom(H)$ we have:

3. *If $\Phi(x, H|x, H(x), \vec{w})$ for all $x \in n \cup \{n\}$ then $H(n) = F(n)$.*

Proof. The reader may check that a proof of this theorem scheme can be obtained by replacing "$\mathbb{N}$" with "$\mathbb{ON}$" in the proof of Theorem Scheme 9.

□

When applying transfinite recursion on $\mathbb{ON}$ we often have three separate cases to specify, rather than just two as with recursion on $\mathbb{N}$. This is illustrated by the recursive definitions of the arithmetic operations on $\mathbb{ON}$.

Addition:

$$\alpha + 0 = \alpha;$$
$$\alpha + succ(\beta) = succ(\alpha + \beta);$$
$$\alpha + \delta = \sup\ \{\alpha + \eta : \eta \in \delta\},\ \text{for a limit ordinal } \delta.$$

Multiplication:

$$\alpha \cdot 0 = 0;$$
$$\alpha \cdot succ(\beta) = (\alpha \cdot \beta) + \alpha;$$
$$\alpha \cdot \delta = \sup\ \{\alpha \cdot \eta : \eta \in \delta\},\ \text{for a limit ordinal } \delta.$$

Exponentiation:

$$\alpha^0 = 1;$$
$$\alpha^{succ(\beta)} = (\alpha^\beta) \cdot \alpha;$$
$$\alpha^\delta = \sup\ \{\alpha^\eta : \eta \in \delta\},\ \text{for a limit ordinal } \delta.$$

Note that, in each case, we are extending the operation from $\mathbb{N}$ to all of $\mathbb{ON}$. The following theorem shows that these operations behave somewhat similarly on $\mathbb{N}$ and $\mathbb{ON}$.

Theorem 15. *Let α, β, and δ be ordinals and S be a non-empty set of ordinals. We have,*

1. $0 + \alpha = \alpha$;
2. *If* $\beta < \delta$ *then* $\alpha + \beta < \alpha + \delta$;
3. $\alpha + \sup S = \sup \{\alpha + \eta : \eta \in S\}$;
4. $\alpha + (\beta + \delta) = (\alpha + \beta) + \delta$;
5. *If* $\alpha < \beta$ *then* $\alpha + \delta \leq \beta + \delta$;
6. $0 \cdot \alpha = 0$;
7. $1 \cdot \alpha = \alpha$;
8. *If* $0 < \alpha$ *and* $\beta < \delta$ *then* $\alpha \cdot \beta < \alpha \cdot \delta$;
9. $\alpha \cdot \sup S = \sup \{\alpha \cdot \eta : \eta \in S\}$;
10. $\alpha \cdot (\beta + \delta) = (\alpha \cdot \beta) + (\alpha \cdot \delta)$;
11. $\alpha \cdot (\beta \cdot \delta) = (\alpha \cdot \beta) \cdot \delta$;
12. *If* $\alpha < \beta$ *then* $\alpha \cdot \delta \leq \beta \cdot \delta$;
13. $1^\alpha = 1$;
14. *If* $1 < \alpha$ *and* $\beta < \delta$ *then* $\alpha^\beta < \alpha^\delta$;
15. $\alpha^{\sup S} = \sup \{\alpha^\eta : \eta \in S\}$;
16. $\alpha^{(\beta+\delta)} = \alpha^\beta \cdot \alpha^\delta$;
17. $(\alpha^\beta)^\delta = \alpha^{\beta \cdot \delta}$; *and,*
18. *If* $\alpha < \beta$ *then* $\alpha^\delta \leq \beta^\delta$.

Exercise 10. Build your transfinite induction skills by proving two parts of this theorem.

However, ordinal addition and multiplication are not commutative. This is illustrated by the following examples, which are easy to verify from the basic definitions.

Examples.

1. $1+\omega = 2+\omega$
2. $1+\omega \neq \omega+1$
3. $1\cdot\omega = 2\cdot\omega$
4. $2\cdot\omega \neq \omega\cdot 2$
5. $2^\omega = 4^\omega$
6. $(2\cdot 2)^\omega \neq 2^\omega \cdot 2^\omega$

Lemma. If β is a non-zero ordinal then ω^β is a limit ordinal.

Exercise 11. Prove this lemma.

Lemma. If α is a non-zero ordinal, then there is a largest ordinal β such that $\omega^\beta \leq \alpha$.

Exercise 12. Prove this lemma. Show that the $\beta \leq \alpha$ and that there are cases in which $\beta = \alpha$. Such ordinals β are called epsilon numbers (The smallest such ordinal $\alpha = \omega^\alpha$ is called ϵ_0.)

Lemma. $\forall\alpha \in \mathbb{ON}\ \forall\beta \in \alpha\ \exists!\gamma \in \mathbb{ON}\ \alpha = \beta+\gamma$.

Exercise 13. Prove this lemma.

Commonly, any function f with $dom(f) \subseteq \omega$ is called a sequence. If $dom(f) \subseteq n+1$ for some $n \in \omega$, we say that f is a finite sequence; otherwise f is an infinite sequence. As usual, we denote the sequence f by $\{f_n\}$, where each $f_n = f(n)$.

Theorem 16. *There is no infinite descending sequence of ordinals.*

Proof. Let's use an indirect proof. Suppose $x \subseteq \omega$ is infinite and $f\colon x \to \mathbb{ON}$ such that if $n < m$ then $f(n) > f(m)$. Let $X = \{f(n) : n \in x\}$. By Foundation there is $y \in X$ such that $y \cap X = \emptyset$; i.e., there is $n \in x$ such that $f(n) \cap X = \emptyset$. However, if $m \in x$ and $m > n$ then $f(m) \in f(n)$, which is a contradiction.

□

If $n \in \omega$ and $s\colon (n+1) \to \mathbb{ON}$ is a finite sequence of ordinals, then the sum $\sum_{i=0}^{n} s(i)$ is defined by recursion as follows.

$$\sum_{i=0}^{0} s(i) = s(0); \text{ and,}$$
$$\sum_{i=0}^{m+1} s(i) = \sum_{i=0}^{m} s(i) + s(m+1), \text{ for } m < n.$$

This shows that statements like the following theorem can be written precisely in the language of set theory.

Theorem 17. *(Cantor Normal Form)*

For each non-zero ordinal α there is a unique $n \in \omega$ and finite sequences $m_0, \ldots, m_n$ of positive natural numbers and $\beta_0, \ldots, \beta_n$ of ordinals which satisfy $\beta_0 > \beta_1 > \cdots > \beta_n$ such that

$$\alpha = \omega^{\beta_0} m_0 + \omega^{\beta_1} m_1 + \cdots + \omega^{\beta_n} m_n.$$

Proof. Using the penultimate lemma, let

$$\beta_0 = \max \{\beta : \omega^{\beta} \leq \alpha\}$$

and then let

$$m_0 = \max \{m \in \omega : \omega^{\beta_0} m \leq \alpha\}$$

which must exist since $\omega^{\beta_0} m \leq \alpha$ for all $m \in \omega$ would imply that $\omega^{\beta_0 + 1} \leq \alpha$.

By the previous lemma, there is some $\alpha_0 \in \mathbb{ON}$ such that

$$\alpha = \omega^{\beta_0} m_0 + \alpha_0$$

where the maximality of m_0 ensures that $\alpha_0 < \omega^{\beta_0}$. Now let

$$\beta_1 = \max \{\beta : \omega^{\beta} \leq \alpha_0\}$$

so that $\beta_1 < \beta_0$. Proceed to get

$$m_1 = \max \{m \in \omega : \omega^{\beta_1} m \leq \alpha_0\}$$

and $\alpha_1 < \omega^{\beta_1}$ such that $\alpha_0 = \omega^{\beta_1} m_1 + \alpha_1$. We continue in this manner as long as possible. We must have to stop after a finite number of steps or else $\beta_0 > \beta_1 > \beta_2 > \dots$ would be an infinite decreasing sequence of ordinals. The only way we could stop would be if some $\alpha_n = 0$. This proves the existence of the sum. Uniqueness follows by induction on $\alpha \in \mathbb{ON}$.

□

Exercise 14. Verify the last statement of this proof.

Lemma.

1. If $0 < m < \omega$ and α is a non-zero ordinal, then $m \cdot \omega^{\alpha} = \omega^{\alpha}$.
2. If $k \in \omega$, and $m_0, \dots, m_k < \omega$, and $\alpha_0, \dots, \alpha_k < \beta$, then

$$m_0 \cdot \omega^{\alpha_0} + \cdots + m_k \cdot \omega^{\alpha_k} < \omega^{\beta}.$$

Exercise 15. Prove this lemma and note that it implies that $m \cdot \delta = \delta$ for each positive integer m and each limit ordinal δ.

There is an interesting application of ordinal arithmetic to Number Theory. Pick a number—say $x = 54$. We have $54 = 2^5 + 2^4 + 2^2 + 2$ when it is written as the simplest sum of powers of 2. In fact, we can write out 54 using only the the arithmetic operations and the numbers 1 and 2. This will be the first step in a recursively defined sequence of natural numbers, $\{x_n\}$. It begins with $n = 2$ and is constructed as follows.

$$x_2 = 54 = 2^{(2^2+1)} + 2^{2^2} + 2^2 + 2.$$

Subtract 1.

$$x_2 - 1 = 2^{(2^2+1)} + 2^{2^2} + 2^2 + 1.$$

Change all 2's to 3's, leaving the 1's alone.

$$x_3 = 3^{(3^3+1)} + 3^{3^3} + 3^3 + 1.$$

Subtract 1.

$$x_3 - 1 = 3^{(3^3+1)} + 3^{3^3} + 3^3.$$

Change all 3's to 4's, leaving any 1's or 2's alone.

$$x_4 = 4^{(4^4+1)} + 4^{4^4} + 4^4.$$

Subtract 1.

$$x_4 - 1 = 4^{(4^4+1)} + 4^{4^4} + 3 \cdot 4^3 + 3 \cdot 4^2 + 3 \cdot 4 + 3.$$

Change all 4's to 5's, leaving any 1's, 2's or 3's alone.

$$x_5 = 5^{(5^5+1)} + 5^{5^5} + 3 \cdot 5^3 + 3 \cdot 5^2 + 3 \cdot 5 + 3.$$

Subtract 1 and continue, changing 5's to 6's, subtracting 1, changing 6's to 7's and so on. One may ask the value of the limit

$$\lim_{n\to\infty} x_n.$$

What is your guess? The answer is surprising.

Theorem 18. *(Goodstein)*

For any initial choice of x there is some n such that $x_n = 0$.

Proof. We use an indirect proof; suppose $x \in \mathbb{N}$ and for all $n \geq 2$ we have $x_n \neq 0$. From this sequence, we construct another sequence. For each $n \geq 2$ we let g_n be the result of replacing each occurrence of n in x_n by ω. So, in the example above we would get:

$$\begin{aligned}
g_2 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^\omega + \omega, \\
g_3 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^\omega + 1, \\
g_4 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^\omega, \\
g_5 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + 3 \cdot \omega^3 + 3 \cdot \omega^2 + 3 \cdot \omega + 3, \\
g_6 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + 3 \cdot \omega^3 + 3 \cdot \omega^2 + 3 \cdot \omega + 2,
\end{aligned}$$

etc. The previous lemma can now be used to show that $\{g_n\}$ would be an infinite decreasing sequence of ordinals.

$\square$

It is interesting that, although the statement of the theorem does not mention infinity in any way, we used the Axiom of Infinity in its proof. We do not need the Axiom of Infinity in order to verify the theorem for any one particular value of x—we just need to carry out the arithmetic. The reader can do this for $x = 4$; however, finishing our example $x = 54$ would be tedious. Moreover, the calculations are somewhat different for different values of x. Mathematical logicians have proved that, in fact, there is no uniform method of finitary calculations which will give a proof of the theorem for all x. The Axiom of Infinity is necessary for the proof.

Chapter 6

Relations and Orderings

In the following definitions, R and C are terms.

1. We say R is a <u>relation</u> on C whenever $R \subseteq C \times C$.
2. We say a relation R is <u>irreflexive</u> on C whenever $\forall x \in C\ \langle x, x\rangle \notin R$.
3. We say a relation R is <u>transitive</u> on C whenever

$$\forall x\ \forall y\ \forall z\ [(\langle x, y\rangle \in R \wedge \langle y, z\rangle \in R) \to \langle x, z\rangle \in R.]$$

4. We say a relation R is <u>well founded</u> on C whenever

$$\forall X\ [(X \subseteq C \wedge X \neq \emptyset) \to (\exists x \in X\ \forall y \in X\ \langle y, x\rangle \notin R)].$$

   Such an x is called <u>minimal</u> for X.

5. We say a relation R is <u>total</u> on C whenever

$$\forall x \in C\ \forall y \in C\ [\langle x, y\rangle \in R \vee \langle y, x\rangle \in R \vee x = y].$$

6. We say R is <u>extensional</u> on C whenever

$$\forall x \in C\ \forall y \in C\ [x = y \leftrightarrow \forall z \in C\ (\langle z, x\rangle \in R \leftrightarrow \langle z, y\rangle \in R)].$$

An example of a relation R on an ordinal is given by the membership relation:

$$\langle x, y\rangle \in R \text{ iff } x \in y.$$

R satisfies all the above properties. On the other hand, if $\alpha \notin \omega$, the reverse relation R' given by

$$\langle x, y\rangle \in R' \text{ iff } y \in x$$

is not well founded but nevertheless has all the other properties.

Any well founded relation is irreflexive. Any total relation is extensional. Any relation which is both well founded and total is also transitive.

Exercise 16. Prove that a transitive set α is an ordinal iff the membership relation is total on α.

Suppose δ is an ordinal and $f : X \to \delta$. A relation R on X with the property that $f(x) < f(y)$ whenever $\langle x, y\rangle \in R$ must be a well founded relation. In fact, this turns out to be a characterisation.

Theorem 19. *Let R be a relation on a set X. R is well founded iff there is an ordinal δ and a surjection $f : X \to \delta$ such that $f(x) < f(y)$ whenever $\langle x, y\rangle \in R$.*

Proof. We treat only the forward implication. Using recursion on $\mathbb{ON}$ we define $g : \mathbb{ON} \to \mathcal{P}(X)$ by

$$g(\beta) = \left\{x : x \text{ is a minimal element of } X \setminus \bigcup\{g(\alpha) : \alpha < \beta\}\right\}.$$

From g we obtain $f : X \to \mathbb{ON}$ by

$$f(x) = \begin{cases} \text{the unique } \alpha \in \mathbb{ON} \text{ with } x \in g(\alpha), & \text{if possible;} \\ 0, & \text{otherwise.} \end{cases}$$

By Theorem 10 and the Axiom of Replacement there must be some least $\delta \in \mathbb{ON}$ such that $\delta \notin rng(f)$. This means that $g(\delta) = \emptyset$, and since R is well founded we must have

$$X = \bigcup\{g(\alpha) : \alpha < \delta\}.$$

To finish the proof suppose $\langle x, y\rangle \in R$ and $f(y) = \beta$. We have $y \in g(\beta)$ so that y is a minimal element of

$$X \setminus \bigcup\{g(\alpha) : \alpha < \beta\},$$

and hence we have that

$$x \notin X \setminus \bigcup\{g(\alpha : \alpha < \beta\}.$$

In other words $x \in g(\alpha)$ for some $\alpha < \beta$ and so $f(x) = \alpha < \beta = f(y)$.

$\square$

A relation R on a set A is said to be isomorphic to a relation S on a set B provided that there is a bijection $f : A \to B$, called an isomorphism, such that for all x and y in A we have

$$\langle x, y\rangle \in R \text{ iff } \langle f(x), f(y)\rangle \in S.$$

Exercise 17. Prove that any isomorphism between transitive sets is the identity. Of course, the relation on the transitive set is the membership relation, which is extensional and well founded.

This unexpected result leads to the important Mostowski Collapsing Theorem.

Theorem 20. *Let R be a well founded extensional relation on a set X. There is a unique transitive set M and a unique isomorphism $h : X \to M$.*

Proof. Obtain $f : X \to \delta$ directly from the previous theorem. By recursion on the ordinals we define for each $\beta < \delta$ a function $h_\beta : f^{\leftarrow}\{\beta\} \to \mathbb{V}$ such that

$$h_\beta(y) = \{h_\alpha(x) : \alpha < \beta \text{ and } \langle x, y\rangle \in R\}.$$

Let $h = \bigcup\{h_\beta : \beta < \delta\}$. Clearly h is a function with domain X.

Note that if $y \in f^{\leftarrow}\{\beta\}$ and $\langle x, y\rangle \in R$ then $x \in f^{\leftarrow}\{\alpha\}$ for some $\alpha < \beta$, so that in fact

$$h_\beta(y) = \{h_\alpha(x) : \alpha < \delta \text{ and } \langle x, y\rangle \in R\},$$

and hence we have

$$h(y) = \{h(x) : \langle x, y \rangle \in R\}.$$

Since R is extensional, h is an injection. Letting $M = rng(h)$, it is now straightforward to use the previous exercise to complete the proof.

□

Exercise 18. Verify the last two sentences in this proof.

We say that a relation R is a partial ordering or partial order whenever it is both irrefexive and transitive; if in addition it is total, then it is called a linear ordering or linear order; furthermore, if in addition it is well founded, then it is called a well ordering or well order. For those orderings we usually write $<$ instead of R and we write $x < y$ for $\langle x, y \rangle \in R$.

Whenever

$$\exists z \; z = \langle X, < \rangle \text{ and } < \text{ is a partial ordering on } X,$$

we say that $\langle X, < \rangle$ is a parially ordered set. We similarly have the concepts of linearly ordered set and well ordered set.

The study of partially ordered sets continues to be a major theme in contemporary Set Theory and the construction of elaborate partial orders is of great technical importance. In contrast, well orders have been thoroughly analysed and we shall now classify all well ordered sets.

Theorem 21. *Each well ordered set is isomorphic to a unique ordinal.*

Proof. Since well orders are extensional and well founded we can use the Mostowski Collapsing Theorem. By Exercise 16 the resulting transitive set is an ordinal.

□

The unique ordinal given by this theorem is called the order type of the well ordered set. We denote the order type of $\langle X, < \rangle$ by type$\langle X, < \rangle$.

We now come to the Well Ordering Principle, which is the fundamental theorem of Set Theory due to E. Zermelo. In order to prove it we use the Axiom of Choice and, for the first time, the Power Set Axiom.

Theorem 22. $(\forall X)(\exists <)$ $[\langle X, <\rangle$ is a well ordered set$]$.

Proof. We begin by using Theorem 5 to obtain a choice function

$$f : \mathcal{P}(X) \setminus \{\emptyset\} \to X$$

such that for each nonempty $A \subseteq X$ we have $f(A) \in A$.

By recursion on $\mathbb{ON}$ we define $g : \mathbb{ON} \to X \cup \{X\}$ as:

$$g(\beta) = \begin{cases} f(X \setminus \{g(\alpha) : \alpha < \beta\}), & \text{if } X \setminus \{g(\alpha) : \alpha < \beta\} \neq \emptyset; \\ X, & \text{otherwise.} \end{cases} \tag{6.1}$$

Now replace each $x \in X \cap ran(g)$ by the unique ordinal β such that $g(\beta) = x$. The Axiom of Replacement gives the resulting set $S \subseteq \mathbb{ON}$, where

$$S = \{\beta \in \mathbb{ON} : g(\beta) \in X\}.$$

By Theorem 10 there is a $\delta \in \mathbb{ON} \setminus S$. Choosing any such, we must have $g(\delta) \notin X$; that is, $g(\delta) = X$ and so $X \subseteq \{g(\alpha) : \alpha < \delta\}$. It is now straightforward to verify that

$$\{\langle x, y\rangle \in X \times X : x = g(\alpha) \text{ and } y = g(\beta) \text{ for some } \alpha < \beta < \delta\}$$

is a well ordering of X, which completes the proof.

$\square$

This Well Ordering Principle is used frequently in modern Set Theory. In fact, most uses of the Axiom of Choice are via the Well Ordering Principle. The Power Set Axiom, first used in this proof, will now also be used frequently without special mention.

Chapter 7

Cardinality

In this chapter, we investigate a concept which aims to translate our intuitive notion of size into formal language. By Zermelo's Well Ordering Principle (Theorem 22) every set can be well ordered. By Theorem 21, every well ordered set is isomorphic to an ordinal. Therefore, for any set x there is some ordinal $\kappa \in \mathbb{ON}$ and a bijection $f \colon x \to \kappa$.

We define the cardinality of x, $|x|$, to be the least $\kappa \in \mathbb{ON}$ such that there is some bijection $f : x \to \kappa$. Every set has a cardinality.

Those ordinals which are $|x|$ for some x are called cardinals.

Exercise 19. Prove that each $n \in \omega$ is a cardinal and that ω is a cardinal. Show that $\omega + 1$ is not a cardinal and that, in fact, each other cardinal is a limit ordinal.

Theorem 23. *The following are equivalent.*

1. *κ is a cardinal.*

2. $(\forall \alpha < \kappa)(\neg\exists$ bijection $f \colon \kappa \to \alpha)$*; i.e.,* $|\kappa| = \kappa$.

3. $(\forall \alpha < \kappa)(\neg\exists$ injection $f \colon \kappa \to \alpha)$.

Proof. We prove the negations of each are equivalent:

$\neg(1)$ $(\forall x)[(\exists \text{ bijection } g\colon x \to \kappa) \to (\exists \alpha < \kappa)(\exists \text{ bijection } h\colon x \to \alpha)]$
$\neg(2)$ $\exists \alpha < \kappa \; \exists \text{ bijection } f\colon \kappa \to \alpha$
$\neg(3)$ $\exists \alpha < \kappa \; \exists \text{ injection } f\colon \kappa \to \alpha$

$\neg(2) \Rightarrow \neg(1)$ Just take $h = f \circ g$.

$\neg(1) \Rightarrow \neg(3)$ Just consider $x = \kappa$.

$\neg(3) \Rightarrow \neg(2)$ Suppose $\alpha < \kappa$ and $f\colon \kappa \to \alpha$ is an injection. By Theorem 21, there is an isomorphism $g\colon \beta \to f^{\to}\kappa$ for some $\beta \in \mathbb{ON}$. Since g is order preserving, we must have $\gamma \leq g(\gamma)$ for each $\gamma \in \beta$ and hence $\beta \leq \alpha$. Now $g^{-1} \circ f\colon \kappa \to \beta$ is the desired bijection.

□

The following exercises are applications of the theorem. The first two statements relate to the questions raised in the introduction.

Exercise 20. Prove the following:

1. $|x| = |y|$ iff $\exists$ bijection $f\colon x \to y$.
2. $|x| \leq |y|$ iff $\exists$ injection $f\colon x \to y$.
3. $|x| \geq |y|$ iff $\exists$ surjection $f\colon x \to y$. Assume here that $y \neq \emptyset$.

Theorem 24. *(G. Cantor)*

$\forall x \; |x| < |\mathcal{P}(x)|.$

Proof. First note that if $|x| \geq |\mathcal{P}(x)|$, then there would be a surjection

$$g\colon x \to \mathcal{P}(x).$$

But this cannot happen, since $\{a \in x : a \notin g(a)\} \notin g^{\to}(x)$.

□

For any ordinal α, we denote by α^+ the least cardinal greater than α. This is well defined by Theorem 24.

Exercise 21. Prove the following:

1. The supremum of a set of cardinals is a cardinal.

2. $\neg\exists z\ z = \{\kappa : \kappa \text{ is a cardinal}\}$.

Theorem 25. *For any infinite cardinal κ, $|\kappa \times \kappa| = \kappa$.*

Proof. Let κ be an infinite cardinal. The formulas $|\kappa| = |\kappa \times \{0\}|$ and $|\kappa \times \{0\}| \leq |\kappa \times \kappa|$ imply that $\kappa \leq |\kappa \times \kappa|$. We now show that $|\kappa \times \kappa| \leq \kappa$. We use induction and assume that $|\lambda \times \lambda| = \lambda$ for each infinite cardinal $\lambda < \kappa$.

We define an ordering on $\kappa \times \kappa$ by:

$$\langle \alpha_0, \beta_0 \rangle < \langle \alpha_1, \beta_1 \rangle \text{ iff } \begin{cases} \max\{\alpha_0, \beta_0\} < \max\{\alpha_1, \beta_1\}; \\ \max\{\alpha_0, \beta_0\} = \max\{\alpha_1, \beta_1\} \wedge \alpha_0 < \alpha_1; \text{ or,} \\ \max\{\alpha_0, \beta_0\} = \max\{\alpha_1, \beta_1\} \wedge \alpha_0 = \alpha_1 \wedge \beta_0 < \beta_1. \end{cases}$$

It is easy to check that $<$ well orders $\kappa \times \kappa$.

Let $\theta = \mathit{type}\ \langle \kappa \times \kappa, < \rangle$. It suffices to show that $\theta \leq \kappa$. And for this it suffices to show that for all $\langle \alpha, \beta \rangle \in \kappa \times \kappa$,

$$\mathit{type}\ \langle <^{\leftarrow} \{\langle \alpha, \beta \rangle\}, < \rangle < \kappa.$$

To this end, pick $\langle \alpha, \beta \rangle \in \kappa \times \kappa$. Let $\delta \in \kappa$ be such that $<^{\leftarrow} \{\langle \alpha, \beta \rangle\} \subseteq \delta \times \delta$. This is possible since κ is a limit ordinal. It now suffices to prove that

$$\mathit{type}\ \langle <^{\leftarrow} \{\langle \delta, \delta \rangle\}, < \rangle < \kappa.$$

By Theorem 23 it suffices to prove that $|<^{\leftarrow} \{\langle \delta, \delta \rangle\}| < \kappa$ and so it suffices to prove that $|\delta \times \delta| < \kappa$.

Since $|\delta \times \delta| = ||\delta| \times |\delta||$, it suffices to prove that $|\lambda \times \lambda| < \kappa$ for all cardinals $\lambda < \kappa$. If λ is infinite, this is true by inductive hypothesis. If λ is finite, then $|\lambda \times \lambda| < \omega \leq \kappa$.

$\square$

Corollary. If X is infinite, then $|X \times Y| = \max\{|X|, |Y|\}$.

Theorem 26. *For any X we have $|\bigcup X| \leq \max\{|X|, \sup\{|a| : a \in X\}\}$, provided that at least one element of $X \cup \{X\}$ is infinite.*

Proof. Let $\kappa = \sup\{|a| : a \in X\}$. Using Exercise 20, for each $a \in X$ there is a surjection $f_a \colon \kappa \to a$. Define a surjection $f \colon X \times \kappa \to \bigcup X$ by

$$f(\langle a, \alpha \rangle) = f_a(\alpha).$$

Using Exercise 20 and the previous corollary, the result follows.

$\square$

Define ${}^{B}A$ as $\{f : f \colon B \to A\}$ and $[A]^{\kappa}$ as $\{x : x \subseteq A \wedge |x| = \kappa\}$.

Lemma. If κ is an infinite cardinal, then $|{}^{\kappa}\kappa| = |[\kappa]^{\kappa}| = |\mathcal{P}(\kappa)|$.

Proof. We have,

$${}^{\kappa}2 \subseteq {}^{\kappa}\kappa \subseteq [\kappa \times \kappa]^{\kappa} \subseteq \mathcal{P}(\kappa \times \kappa).$$

Using characteristic functions it is easily proved that $|{}^{\kappa}2| = |\mathcal{P}(\kappa)|$. Since $|\kappa \times \kappa| = \kappa$ we have the result.

$\square$

Lemma. If $\kappa \leq \lambda$ and λ is an infinite cardinal, then $|{}^{\kappa}\lambda| = |[\lambda]^{\kappa}|$.

Proof. For each $x \in [\lambda]^{\kappa}$ there is a bijection $f_x \colon \kappa \to x$. Since $f_x \in {}^{\kappa}\lambda$, we have an injection from $[\lambda]^{\kappa}$ into ${}^{\kappa}\lambda$, so $|[\lambda]^{\kappa}| \leq |{}^{\kappa}\lambda|$.

Now ${}^{\kappa}\lambda \subseteq [\kappa \times \lambda]^{\kappa}$. Thus, $|{}^{\kappa}\lambda| \leq |[\lambda]^{\kappa}|$.

$\square$

A subset S of a limit ordinal α is said to be <u>cofinal</u> whenever $\sup S = \alpha$. A function $f : \delta \to \alpha$ is said to be <u>cofinal</u> whenever $rng(f)$ is cofinal in α. The <u>cofinality</u>, $cf(\alpha)$, of a limit ordinal α is the least $\delta \in \mathbb{ON}$ such that there is a cofinal $f : \delta \to \alpha$.

A cardinal κ is said to be regular whenever $cf(\kappa) = \kappa$. Otherwise, κ is said to be singular. These notions are of interest when κ is an infinite cardinal. In particular, ω is regular.

Lemma.

1. For each limit ordinal κ, $cf(\kappa)$ is a regular cardinal.
2. For each limit ordinal κ, κ^+ is a regular cardinal.
3. Each infinite singular cardinal contains a cofinal subset of regular cardinals.

Exercise 22. Prove this lemma.

Theorem 27. *(König's Theorem)*

For each infinite cardinal κ, $|{}^{cf(\kappa)}\kappa| > \kappa$.

Proof. We show that there is no surjection $g\colon \kappa \to {}^{\delta}\kappa$, where $\delta = cf(\kappa)$. Let $f\colon \delta \to \kappa$ witness that $cf(\kappa) = \delta$. Define $h\colon \delta \to \kappa$ such that each $h(\alpha) \notin \{g(\beta)(\alpha) : \beta < f(\alpha)\}$. Then $h \notin g^{\rightarrow}(\kappa)$, since otherwise $h = g(\beta)$ for some $\beta < \kappa$; pick $\alpha \in \delta$ such that $f(\alpha) > \beta$.

□

Corollary. For each infinite cardinal κ, $cf(|\mathcal{P}(\kappa)|) > \kappa$.

Proof. Let $\lambda = |\mathcal{P}(\kappa)|$. Suppose $cf(\lambda) \leq \kappa$. Then

$$\lambda = |\mathcal{P}(\kappa)| = |{}^{\kappa}2| = |{}^{(\kappa\times\kappa)}2| = |{}^{\kappa}({}^{\kappa}2)| = |{}^{\kappa}\lambda| \geq |{}^{cf(\lambda)}\lambda| > \lambda.$$

□

Cantor's Theorem guarantees that for each ordinal α there is a set, $\mathcal{P}(\alpha)$, which has cardinality greater than α. However, it does not imply, for example, that $\omega^+ = |\mathcal{P}(\omega)|$. This statement is called the Continuum Hypothesis, and is equivalent to the third question in the introduction.

The aleph function $\aleph\colon \mathbb{ON} \to \mathbb{ON}$ is defined as follows:

$$\begin{aligned}\aleph(0) &= \omega\\ \aleph(\alpha) &= \sup\ \{\aleph(\beta)^+ : \beta \in \alpha\}.\end{aligned}$$

We write $\aleph_\alpha$ for $\aleph(\alpha)$. We also sometimes write ω_α for $\aleph(\alpha)$.

The beth function $\beth\colon \mathbb{ON} \to \mathbb{ON}$ is defined as follows:

$$\begin{aligned}\beth(0) &= \omega\\ \beth(\alpha) &= \sup\ \{|\mathcal{P}(\beth(\beta))| : \beta \in \alpha\}.\end{aligned}$$

We write $\beth_\beta$ for $\beth(\beta)$.

It is apparent that $\aleph_1 \leq \beth_1$. The continuum hypothesis is the statement $\aleph_1 = \beth_1$; this is abbreviated as CH. The generalised continuum hypothesis, GCH, is the statement $\forall\alpha \in \mathbb{ON}\ \aleph_\alpha = \beth_\alpha$.

Exercise 23. Prove that

$$\forall\kappa\ [\kappa \text{ is an infinite cardinal } \to (\exists\alpha \in \mathbb{ON}(\kappa = \aleph_\alpha)].$$

A cardinal κ is said to be inaccessible whenever both κ is regular and

$$\forall\lambda < \kappa\ |\mathcal{P}(\lambda)| < \kappa.$$

An inaccessible cardinal is sometimes said to be strongly inaccessible, and the term weakly inaccesible is given to a regular cardinal κ such that

$$\forall\lambda < \kappa\ \lambda^+ < \kappa.$$

Under the GCH these two notions are equivalent.

Axiom of Inaccessibles $\quad \exists\kappa\ \kappa > \omega$ and κ is an inaccessible cardinal

This axiom is a stronger version of the Axiom of Infinity, but the mathematical community is not quite ready to replace the Axiom of Infinity with it just yet. In fact, the Axiom of Inacessibles is not included in the basic $\mathfrak{ZFC}$ axiom system and is therefore always explicitly stated whenever it is used.

Exercise 24. Are the following two statements true? What if κ is assumed to be a regular cardinal?

1. κ is weakly inaccessible iff $\kappa = \aleph_\kappa$.
2. κ is strongly inaccesssible iff $\kappa = \beth_\kappa$.

Chapter 8

There Is Nothing Real About The Real Numbers

We now formulate three familiar number systems in the language of set theory. From the natural numbers we shall construct the integers; from the integers we shall construct the decimal numbers and the real numbers.

For each $n \in \mathbb{N}$ let $-n = \{\{m\} : m \in n\}$. The integers, denoted by $\mathbb{Z}$, are defined as

$$\mathbb{Z} = \mathbb{N} \cup \{-n : n \in \mathbb{N}\}.$$

For each $n \in \mathbb{N} \setminus \{\emptyset\}$ we denote $\{n\}$ by $-n$.

We can extend the ordering $<$ on $\mathbb{N}$ to $\mathbb{Z}$ by letting $x < y$ iff one of the following holds:

1. $x \in \mathbb{N} \wedge y \in \mathbb{N} \wedge x < y$;
2. $x \notin \mathbb{N} \wedge y \in \mathbb{N}$; or,
3. $x \notin \mathbb{N} \wedge y \notin \mathbb{N} \wedge \bigcup y < \bigcup x$.

To form the reals, first let

$$F = \{f : f \in {}^{\omega}\mathbb{Z}\}.$$

We pose a few restrictions on such functions as follows. Let us write:

$$\begin{array}{rll} A(f) & \text{for} & (\forall n > 0)(-9 \leq f(n) \leq 9); \\ B(f) & \text{for} & (\forall n \in \omega)(f(n) \geq 0) \vee (\forall n \in \omega)(f(n) \leq 0); \\ C(f) & \text{for} & (\forall m \in \omega)(\exists n \in \omega \setminus m)(f(n) \notin \{9, -9\}); \text{ and,} \\ D(f) & \text{for} & (\exists m \in \omega)(\forall n \in \omega \setminus m)(f(n) = 0). \end{array}$$

Finally, let

$$\mathbb{R} = \{f : f \in F \text{ and } A(f) \text{ and } B(f) \wedge C(f)\}$$

to obtain the <u>real numbers</u> and let

$$\mathbb{D} = \{f : f \in \mathbb{R} \text{ and } D(f)\}$$

to obtain the <u>decimal numbers</u>.

We now order $\mathbb{R}$ as follows: let $f < g$ iff

$$(\exists n \in \omega)\ [f(n) < g(n) \wedge (\forall m \in n)(f(m) = g(m))].$$

This ordering clearly extends our ordering on $\mathbb{N}$ and $\mathbb{Z}$, and restricts to $\mathbb{D}$.

In light of these definitions, the operations of addition, multiplication and exponentiation defined in Chapter 4 can be formally extended from $\mathbb{N}$ to $\mathbb{Z}$, $\mathbb{D}$ and $\mathbb{R}$ in a natural—if cumbersome—fashion.

Lemma.

1. $|\mathbb{D}| = \aleph_0$.
2. $\mathbb{R}$ is uncountable.
3. $<$ is a linear order on $\mathbb{D}$.
4. $(\forall p \in \mathbb{R})(\forall q \in \mathbb{R})\ [p < q \rightarrow \exists d \in \mathbb{D}\ p < d < q]$.
   I.e., $\mathbb{D}$ is a countable <u>dense</u> subset of $\mathbb{R}$.
5. $<$ is complete; i.e., bounded subsets have suprema and infima.

Exercise 25. Prove this lemma.

Theorem 28. *Any two countable dense linear orders without endpoints are isomorphic.*

Proof. This method of proof, the back-and-forth argument, is due to G.Cantor. The idea is to define an isomorphism recursively in ω steps, such that at each step we have an order-preserving finite function; at even steps $f(x_i)$ is defined and at odd steps $f^{-1}(y_j)$ is defined.

Precisely, if $X = \{x_i : i \in \omega\}$ and $Y = \{y_j : j \in \omega\}$ are two countable dense linear orders we define $f \colon X \to Y$ by the formulas

$$f_0 = \{\langle x_0, y_0\rangle\}$$
$$f_{n+1} = f_n \cup \{\langle x_i, y_j\rangle\}$$

where

1. if n is even, $i = \min\ \{k \in \omega : x_k \notin dom(f_n)\}$ and j is chosen so that $f_n \cup \{\langle x_i, y_j\rangle\}$ is order-preserving; and,
2. if n is odd, $j = \min\ \{k \in \omega : y_k \notin rng(f_n)\}$ and i is chosen so that $f_n \cup \{\langle x_i, y_j\rangle\}$ is order-preserving.

We then check that for each $n \in \omega$, there is indeed a choice of j in (1) and i in (2) and that $f = \bigcup\{f_n : n \in \omega\}$ is an isomorphism.

□

Any complete dense linear order without endpoints and with a countable dense subset is isomorphic to $\langle \mathbb{R}, <\rangle$. Can "with a countable dense subset" be replaced by "in which every collection of disjoint intervals is countable"?. The affirmation of this is called the Suslin Hypothesis. It is the second most important problem in Set Theory. It too requires new axioms for its solution.

We have begun with $\mathbb{N}$, extended to $\mathbb{Z}$, then extended again to $\mathbb{R}$. We now extend once more to $^*\mathbb{R}$. This is the set of hyperreals. In order to do this, we introduce the important notion of an ultrafilter.

A collection of subsets $\mathcal{U} \subseteq \mathcal{P}(S)$ is a <u>filter</u> provided that it satisfies the first three of the following conditions:

1. $S \in \mathcal{U}$ and $\emptyset \notin \mathcal{U}$.

2. If $A \in \mathcal{U}$ and $B \in \mathcal{U}$, then $A \cap B \in \mathcal{U}$.
3. If $A \in \mathcal{U}$ and $A \subseteq B$, then $B \in \mathcal{U}$.
4. $A \in \mathcal{U}$ or $B \in \mathcal{U}$ whenever $A \cup B = S$.
5. $\forall x \in S$ $\{x\} \notin \mathcal{U}$.

If a filter $\mathcal{U}$ obeys condition (4) it is called an ultrafilter, and if $\mathcal{U}$ satisfies conditions (1) through (5) it is said to be a free or non-principal ultrafilter.

An ultrafilter is a maximal filter under inclusion. Every filter can be extended to an ultrafilter. (We can recursively define it.) There is a free ultrafilter over ω.

Theorem 29. *(Ramsey)*

If $P\colon [\omega]^2 \to \{0,1\}$, *then there is* $H \in [\omega]^\omega$ *such that* $|P''[H]^2| = 1$.

Proof. Let $\mathcal{U}$ be a free ultrafilter over ω. Either:

1. $\{\alpha \in \omega : \{\beta \in \omega : P(\{\alpha, \beta\}) = 0\} \in \mathcal{U}\} \in \mathcal{U}$; or,
2. $\{\alpha \in \omega : \{\beta \in \omega : P(\{\alpha, \beta\}) = 1\} \in \mathcal{U}\} \in \mathcal{U}$.

As such, the proof breaks into two similar cases. We address case (1).

Let $S = \{\alpha \in \omega : \{\beta \in \omega : P(\{\alpha, \beta\}) = 0\} \in \mathcal{U}\}$. Pick $\alpha_0 \in S$ and let

$$S_0 = \{\beta \in \omega : P(\{\alpha_0, \beta\}) = 0\}.$$

Pick $\alpha_1 \in S \cap S_0$ and let

$$S_1 = \{\beta \in \omega : P(\{\alpha_1, \beta\}) = 0\}.$$

In general, recursively choose $\{\alpha_n : n < \omega\}$ such that for each n

$$\alpha_{n+1} \in S \cap S_0 \cap \cdots \cap S_n,$$

where

$$S_n = \{\beta \in \omega : P(\{\alpha_n, \beta\}) = 0\}.$$

Then $H = \{\alpha_n : n < \omega\}$ exhibits the desired property.

□

Theorem 30. *(Sierpinski)*

There is a function $P\colon [\omega_1]^2 \to \{0,1\}$ *such that there is no* $H \in [\omega_1]^{\omega_1}$ *with* $|P''[H]^2| = 1$.

Proof. Let $f\colon \omega_1 \to \mathbb{R}$ be an injection. Define P as follows: for $\alpha < \beta$, let

$$P(\{\alpha,\beta\}) = \begin{cases} 0, & \text{if } f(\alpha) < f(\beta); \\ 1, & \text{if } f(\alpha) > f(\beta). \end{cases} \tag{8.1}$$

The following exercise finishes the proof.

□

Exercise 26. There is no subset of $\mathbb{R}$ with order type ω_1.

Let $\mathcal{U}$ be a free ultrafilter over ω. Form an equivalence relation $\sim$ on ${}^{\omega}\mathbb{R}$ by the rule:

$$f \sim g \text{ whenever } \{n \in \omega : f(n) = g(n)\} \in \mathcal{U}.$$

The equivalence class of f is denoted by

$$[f] = \{g \in {}^{\omega}\mathbb{R} : g \sim f\}.$$

The set of equivalence classes of $\sim$ is called the ultrapower of $\mathbb{R}$ with respect to $\mathcal{U}$. The elements of ${}^{\omega}\mathbb{R}$ are often called the hyperreal numbers and denoted ${}^*\mathbb{R}$.

There is a natural embedding of $\mathbb{R}$ into ${}^*\mathbb{R}$ given by

$$x \mapsto [f_x]$$

where $f_x\colon \omega \to \mathbb{R}$ is the constant function; i.e., $f_x(n) = x$ for all $n \in \omega$; we identify $\mathbb{R}$ with its image under the natural embedding.

We can define an ordering $^* <$ on $^*\mathbb{R}$ by the rule:

$$a <^* b \text{ whenever } \exists f \in a \; \exists g \in b \; \{n \in \omega : f(n) < g(n)\} \in \mathcal{U}.$$

Lemma. $^* <$ is a linear ordering on $^*\mathbb{R}$ which extends the usual ordering of $\mathbb{R}$.

Exercise 27. Prove this lemma.

We usually omit the asterisk, writing $<$ for $^* <$.

Exercise 28. There is a subset of $^*\mathbb{R}$ of order type ω_1.

For each function $F\colon \mathbb{R}^n \to \mathbb{R}$ there is a natural extension

$$^*F\colon (^*\mathbb{R})^n \to {}^*\mathbb{R}$$

given by

$$^*F(\vec{a}) = [F \circ \vec{s}]$$

where $\vec{a} = \langle a_0, \ldots, a_{n-1} \rangle \in (^*\mathbb{R})^n$ and $\vec{s}\colon \omega \to \mathbb{R}^n$ such that for each $j \in \omega$

$$\vec{s}(j) = \langle s_0(j), \ldots, s_{n-1}(j) \rangle$$

and $s_i \in a_i$ for each i.

Theorem 31. *(The Leibniz Transfer Principle)*

Suppose $F\colon \mathbb{R}^n \to \mathbb{R}$ *and* $G\colon \mathbb{R}^n \to \mathbb{R}$.

1. $\forall \vec{a} \in \mathbb{R}^n \; F(\vec{a}) = G(\vec{a})$ iff $\forall \vec{a} \in (^*\mathbb{R})^n \;\; ^*F(\vec{a}) = {}^*G(\vec{a})$.
2. $\forall \vec{a} \in \mathbb{R}^n \; F(\vec{a}) < G(\vec{a})$ iff $\forall \vec{a} \in (^*\mathbb{R})^n \;\; ^*F(\vec{a}) < {}^*G(\vec{a})$.

Exercise 29. Prove the above theorem. This includes verifying that $^*F$ is a function.

As a consequence of this theorem, we can extend $+$ and $\times$ to $^*\mathbb{R}$. For example, $a + b = c$ means that

$$\exists f \in a \; \exists g \in b \; \exists h \in c \; \{n \in \omega : f(n) + g(n) = h(n)\} \in \mathcal{U}.$$

Indeed, the natural embedding embeds $\mathbb{R}$ as an ordered subfield of ${}^*\mathbb{R}$.

In order to do elementary calculus we consider the infinitesimal elements of ${}^*\mathbb{R}$. Note that $\mathbb{R} \neq {}^*\mathbb{R}$; consider

$$\alpha = [\langle 1, 1/2, 1/3, \dots, 1/n, \dots \rangle].$$

Then $\alpha > 0$ but $\alpha < r$ for each positive $r \in \mathbb{R}$; a member of ${}^*\mathbb{R}$ with this property is called a positive infinitesimal. There are negative infinitesimals; 0 is an infinitesimal. Since ${}^*\mathbb{R}$ is a field $1/\alpha$ exists and $1/\alpha > r$ for any real number r; it is an example of a positive infinite number.

A hyperreal number a is said to be finite whenever $|a| < r$ for some real r. Two hyperreal numbers a and b are said to be infinitely close whenever $a - b$ is infinitesimal. We write $a \approx b$.

Lemma. Each finite hyperreal number is infinitely close to a unique real number.

Proof. Let a be finite. Let $s = \sup \{r \in \mathbb{R} : r < a\}$. Then $a \approx s$. If we also have another real t such that $a \approx t$, then we have $s \approx t$ and so $s = t$.

□

If a is finite, the standard part of a, $st(a)$, is defined to be the unique real number which is infinitely close to a.

It is easy to check that for finite a and b,

$$st(a + b) = st(a) + st(b); \text{ and,}$$
$$st(a \times b) = st(a) \times st(b).$$

For a function $F \colon \mathbb{R} \to \mathbb{R}$ we define the derivative, $F'(x)$, of $F(x)$ to be

$$st\left(\frac{F(x + \triangle x) - F(x)}{\triangle x}\right)$$

provided this exists and is the same for each non-zero infinitesimal $\triangle x$.

The fact that $F'(x)$ exists implies that for each infinitesimal $\triangle x$ there is an infinitesimal ϵ such that $F(x+\triangle x) = F(x) + F'(x)\triangle x + \epsilon\triangle x$. That is, for any $\triangle x \approx 0$ there is some $\epsilon \approx 0$ such that

$$\triangle y = F'(x)\triangle x + \epsilon\triangle x$$

where $\triangle y = F(x+\triangle x) - F(x)$.

Theorem 32. *(The Chain Rule)*

Suppose $y = F(x)$ and $x = G(t)$ are differentiable functions. Then $y = F(G(t))$ is differentiable and has derivative $F'(G(t)) \cdot G'(t)$.

Proof. Let $\triangle t$ be any non-zero infinitesimal. Let $\triangle x = G(t+\triangle t) - G(t)$. Since $G'(x)$ exists, $\triangle x$ is infinitesimal. Let $\triangle y = F(x+\triangle x) - F(x)$. We wish to calculate $st(\frac{\triangle y}{\triangle x})$, which will be the derivative of $y = F(G(t))$. We consider two cases.

Case 0: $\triangle x = 0$

$\triangle y = 0$, $st(\frac{\triangle y}{\triangle t}) = 0$, and $G'(t) = st(\frac{\triangle x}{\triangle t}) = 0$.
So $st(\frac{\triangle y}{\triangle t}) = F'(G(t)) \cdot G'(t)$.

Case 1: $\triangle x \neq 0$

$\frac{\triangle y}{\triangle t} = \frac{\triangle y}{\triangle x} \cdot \frac{\triangle x}{\triangle t}$. So $st(\frac{\triangle y}{\triangle t}) = st(\frac{\triangle y}{\triangle x}) \cdot st(\frac{\triangle x}{\triangle t})$, and again,
$st(\frac{\triangle y}{\triangle t}) = F'(x) \cdot G'(t) = F'(G(t)) \cdot G'(t)$.

□

Chapter 9

The Universe

In this chapter we shall discuss two methods of measuring the complexity of a set, as well as their corresponding gradations of the universe. For this discussion it will be helpful to develop both a new induction and a new recursion procedure, this time on the whole universe. Each of these will depend upon the fact that every set is contained in a transitive set, which we now prove.

By recursion on $\mathbb{N}$, we define:

$$\bigcup{}^0 X = X; \text{ and,}$$
$$\bigcup{}^{n+1} X = \bigcup(\bigcup{}^n X).$$

We define the transitive closure of X as,

$$trcl(X) = \bigcup\{\bigcup{}^n X : n \in \omega\}.$$

Theorem 33.

1. $\forall X$ $trcl(X)$ *is the smallest transitive set containing* X.
2. $\forall x$ $trcl(x) = x \cup \bigcup\{trcl(y) : y \in x\}$.

Proof. Note that Y is transitive iff $\bigcup Y \subseteq Y$, so $trcl(X)$ is transitive.

1. If Y is transitive and $X \subseteq Y$ then for each $n \in \mathbb{N}$, $\bigcup^n X \subseteq \bigcup^n Y \subseteq Y$. So $trcl(X) \subseteq Y$.

2. To prove that $trcl(x) \supseteq x \cup \bigcup\{trcl(y) : y \in x\}$, notice first that since $\forall y \in x\ y \subseteq \bigcup x$, we have that $\forall n \in \omega\ \bigcup^n y \subseteq \bigcup^{n+1} x$. Hence $trcl(y) \subseteq trcl(x)$. For the opposite containment, we prove by induction that $\forall n \in \omega\ \bigcup^n x \subseteq x \cup \bigcup\{trcl(y) : y \in x\}$.

□

As with $\mathbb{N}$ and $\mathbb{ON}$, we can perform induction on the universe, called $\in$ −induction, as illustrated by the following theorem scheme.

For each formula $\Phi(v, \vec{w})$ of the language of set theory we have:

Theorem 34. Φ

For all $\vec{w}$, if

$$\forall n\ [(\forall m \in n\ \Phi(m, \vec{w})) \rightarrow \Phi(n, \vec{w})]$$

then

$$\forall n\ \Phi(n, \vec{w}).$$

Proof. We shall assume that the theorem is false and derive a contradiction. We have $\vec{w}$ and a fixed l such that $\neg\Phi(l, \vec{w})$.

Let t be any transitive set containing l. Thanks to the previous theorem, we can let $t = trcl(l \cup \{l\})$. The proof now proceeds verbatim as the proofs of Theorems 8 and 13.

□

We can also carry out recursive definitions on $\mathbb{V}$; this is called $\in$ −recursion. For any formula $\Phi(x, f, y, \vec{w})$ of the language of set theory, we denote by $REC(\Phi, \mathbb{V})$ the class

$$\{\langle x, y\rangle : (\exists n)(\exists f)\ [f : n \rightarrow \mathbb{V} \wedge f(x) = y \wedge \forall m \in n\ \Phi(m, f|m, f(m), \vec{w})]\}.$$

$\in$ −recursion is justified by the next theorem scheme.

For each formula $\Phi(x, f, y, \vec{w})$ of the language of set theory we have:

Theorem 35. Φ

For all $\vec{w}$, suppose that we have

$$(\forall x)(\forall f)\ [(f : x \to \mathbb{V}) \to \exists! y\ \Phi(x, f, y, \vec{w})].$$

Then, letting F denote $REC(\Phi, \mathbb{V})$, we have:

1. $F : \mathbb{V} \to \mathbb{V}$;
2. $\forall x\ \Phi(x, F|x, F(x), \vec{w})$; *and furthermore for any n and any function H with $n \in dom(H)$ we have,*
3. $\Phi(x, H|x, H(x), \vec{w})$ *for all $x \in n \cup \{n\}$ then $H(n) = F(n)$.*

Proof. The proof is similar to that of Theorems 9 and 14.

$\square$

Our first new measure of the size of a set is given by the rank function. This associates, to each set x, an ordinal $rank(x)$ by the following rule:

$$rank(x) = \sup\ \{rank(y) + 1 : y \in x\}$$

Observe that $\forall \alpha \in \mathbb{ON}\ rank(\alpha) = \alpha$.

By recursion on $\mathbb{ON}$ we define the cumulative hierarchy, an ordinal-gradation on $\mathbb{V}$, as follows.

$$\begin{aligned} \mathbf{R}(0) &= \emptyset; \\ \mathbf{R}(\alpha + 1) &= \mathcal{P}(\mathbf{R}(\alpha));\ \text{and,} \\ \mathbf{R}(\delta) &= \bigcup \{\mathbf{R}(\alpha) : \alpha < \delta\} \text{ if } \delta \text{ is a limit ordinal.} \end{aligned}$$

Sometimes we write $\mathbf{R}_\alpha$ or $\mathbb{V}_\alpha$ for $\mathbf{R}(\alpha)$. The next theorem connects and lists some useful properties of the cumulative hierarchy and the rank function.

Theorem 36.

1. *$\forall \alpha \in \mathbb{ON}$ $\boldsymbol{R}(\alpha)$ is transitive.*
2. *$\forall \alpha \in \mathbb{ON}\ \forall \beta \in \mathbb{ON}\ (\beta < \alpha \to \boldsymbol{R}(\beta) \subseteq \boldsymbol{R}(\alpha)$.*
3. *$\forall x\ \forall \alpha \in \mathbb{ON}\ (x \in \boldsymbol{R}(\alpha) \leftrightarrow \exists \beta \in \alpha\ x \subseteq \boldsymbol{R}(\beta))$.*
4. *$\forall x\ \forall \alpha \in \mathbb{ON}\ (x \in \boldsymbol{R}(\alpha + 1) \setminus \boldsymbol{R}(\alpha) \leftrightarrow rank(x) = \alpha)$.*
5. *$\forall x\ \exists \alpha \in \mathbb{ON}\ x \in \boldsymbol{R}(\alpha)$; i.e., $\mathbb{V} = \bigcup\{\boldsymbol{R}(\alpha) : \alpha \in \mathbb{ON}\}$.*

Proof.

1. This is an easy induction on $\alpha \in \mathbb{ON}$.
2. Apply induction on α, using (1).
3. **(→)** Note that if γ is least ordinal such that $x \in \mathbf{R}(\gamma)$, then γ is a successor ordinal, so choose β such that $\gamma = \beta + 1$;

   (←) This uses (2).
4. First show by induction on α that $rank(x) < \alpha$ implies $x \in \mathbf{R}(\alpha)$.
5. This follows from $\forall x\ \exists \alpha \in \mathbb{ON}\ rank(x) = \alpha$.

□

We have discussed cardinality as a way of measuring the size of a set. However, if our set is not transitive, the cardinality function does not tell the whole story since it cannot distinguish the elements of the set. For example, although $\mathbb{N} \in \{\mathbb{N}\}$, $|\mathbb{N}| = \aleph_0$ while $|\{\mathbb{N}\}| = 1$. Intuitively, we think of $\{\mathbb{N}\}$ as no smaller than $\mathbb{N}$. As such, we define the hereditary cardinality, $hcard(x)$, of a set x, as the cardinality of its transitive closure:

$$hcard(x) = |trcl(x)|$$

The corresponding cardinal-gradation is defined as follows.

For each cardinal κ,

$$\mathbf{H}(\kappa) = \{x : |trcl(x) < \kappa\}.$$

The members of $\mathbf{H}(\omega)$ are called the hereditarily finite sets and the members of $\mathbf{H}(\omega_1)$ are called the hereditarily countable sets. The next theorem lists some important properties of $\mathbf{H}$.

Theorem 37.

1. *For any infinite cardinal* κ*,* $\boldsymbol{H}(\kappa)$ *is transitive.*

2. *For any infinite cardinal* κ*,* $\forall x\ hcard(x) < \kappa \rightarrow rank(x) < \kappa$.

3. *For any infinite cardinal* κ*,* $\boldsymbol{H}(\kappa) \subseteq \boldsymbol{R}(\kappa)$.

4. *For any infinite cardinal* κ*,* $\exists z\ z = \boldsymbol{H}(\kappa)$.

5. $\forall x\ \exists \kappa\ (\kappa$ is a cardinal and $x \in \boldsymbol{H}(\kappa))$*; i.e,*
   $\mathbb{V} = \bigcup\{\boldsymbol{H}(\kappa) : \kappa$ is a cardinal$\}$.

Proof.

1. Apply part (2) of Theorem 33.

2. **Case 1:** κ is regular.
   The proof is by $\in$-induction on $\mathbb{V}$, noting that if $rank(y) < \kappa$ for each $y \in x$ and $|x| < \kappa$, then $rank(x) < \kappa$.

   Case 2: κ is singular.
   There is a regular cardinal λ such that $hcard(x) < \lambda < \kappa$; now use Case 1.

3. This follows from (2) and Theorem 36.

4. Apply (3) and Comprehension.

5. This follows since $\forall x\ \exists \kappa\ hcard(x) = \kappa$.

□

Theorem 38. *If* κ *is an inaccessible cardinal, then* $\boldsymbol{H}(\kappa) = \boldsymbol{R}(\kappa)$.

Proof. From Theorem 37 we have $\mathbf{H}(\kappa) \subseteq \mathbf{R}(\kappa)$.

For the reverse containment, let $x \in \mathbf{R}(\kappa)$. By Theorem 36, $\exists \alpha < \kappa$ such that $x \in \mathbf{R}(\alpha + 1) \setminus \mathbf{R}(\alpha)$. So $x \subseteq \mathbf{R}(\alpha)$. Since $\mathbf{R}(\alpha)$ is transitive, $trcl(x) \subseteq \mathbf{R}(\alpha)$.

It suffices to prove by induction that $\forall \alpha < \kappa \; |\mathbf{R}(\alpha)| < \kappa$. For successor $\alpha = \beta + 1$, we note that $|\mathcal{P}(\lambda)| < \kappa$, where $\lambda = |\mathbf{R}(\beta)|$; for limit α we apply Theorem 36, observing that κ is regular.

$\square$

Chapter 10

Reflection

There is a generalisation of the Equality Axiom, called the Equality Principle, which states that for any formula Φ we have $x = y$ implying that Φ holds at x iff Φ holds at y. The proof requires a new technique, called induction on complexity of the formula.

We make this precise. For each formula Φ of set theory all of whose free variables lie among $v_0, \ldots, v_n$ we write $\Phi(v_0, \ldots, v_n)$ and for each i and j with $0 \leq i \leq n$, we denote by $\Phi(v_0, \ldots, v_i/v_j, \ldots, v_n)$ the result of substituting v_j for each free occurance of v_i.

For each formula $\Phi(v_0, \ldots, v_n)$ and each i and j with $0 \leq i \leq n$ we have:

Theorem 39. Φ, i, j

$$\forall v_0 \ldots \forall v_i \ldots \forall v_n \forall v_j [v_i = v_j \rightarrow (\Phi(v_0, \ldots, v_n) \leftrightarrow (\Phi(v_0, \ldots, v_i/v_j, \ldots, v_n))]$$

This is a scheme of theorems, one for each appropriate Φ, i, j.

Proof. The proof will come in two steps.

1. We first prove this for atomic formulas Φ.

Case 1 Φ is $v_i \in v_k$ where $i \neq k$
This is true by Axiom of Equality.

Case 2 Φ is $v_k \in v_i$ where $i \neq k$
This is true by Axiom of Extensionality.

Case 3 Φ is $v_i \in v_i$
This is true since by Theorem 2 both $v_i \in v_j$ and $v_j \in v_i$ are false.

Case 4 Φ is $v_k \in v_k$ where $i \neq k$
This is true since both $\Phi(v_0, \ldots, v_n)$ and $\Phi(v_0, \ldots, v_i/v_j, \ldots, v_n)$ are the same.

Case 5 Φ is $v_i = v_k$ where $i \neq k$
This is true by Theorem 1.

Case 6 Φ is $v_k = v_i$ where $i \neq k$
This is similar to case 5.

Case 7 Φ is $v_i = v_i$
This is true since by Theorem 1 both $v_i = v_i$ and $v_j = v_j$ are true.

Case 8 Φ is $v_k = v_k$ where $i \neq k$
This is similar to case 4.

2. We now show that for any subformula Ω of Φ, if the theorem is true for all proper subformulas of Ω then it is true for Ω. Here Θ and Ψ are subformulas of Ω and Ω is expressed in parentheses for each case according to how Ω is built.

Case 1 $(\neg\Theta)$
This is true since if

$$\Theta(v_0, \ldots, v_n) \leftrightarrow \Theta(v_0, \ldots, v_i/v_j, \ldots, v_n)$$

then

$$\neg\Theta(v_0, \ldots, v_n) \leftrightarrow \neg\Theta(v_0, \ldots, v_i/v_j, \ldots, v_n)$$

Case 2 $(\Theta \wedge \Psi)$
From the hypothesis that

$$\Theta(v_0, \ldots, v_n) \leftrightarrow \Theta(v_0, \ldots, v_i/v_j, \ldots, v_n)$$

and

$$\Psi(v_0, \ldots, v_n) \leftrightarrow \Psi(v_0, \ldots, v_i/v_j, \ldots, v_n)$$

we obain

$$\Theta(v_0, \ldots, v_n) \wedge \Psi(v_0, \ldots, v_n)$$

iff

$$\Theta(v_0, \ldots, v_i/v_j, \ldots, v_n) \wedge \Psi(v_0, \ldots, v_i/v_j, \ldots, v_n).$$

Cases 3 through 5 result from Cases 1 and 2.

Case 3 $(\Theta \vee \Psi)$

Case 4 $(\Theta \rightarrow \Psi)$

Case 5 $(\Theta \leftrightarrow \Psi)$

Case 6 $(\forall v_k\ \Theta)$ and $i \neq k$
We have

$$\forall v_0 \ldots \forall v_n\ \forall v_j\ [v_i = v_j \rightarrow \Theta(v_0, \ldots, v_n) \\ \leftrightarrow \Theta(v_0, \ldots, v_i/v_j, \ldots, v_n)].$$

If v_k is not free in Θ, then $\Theta \leftrightarrow \forall v_k\ \Theta$ and we are done. If v_k is free in Θ, then since $0 \leq k \leq n$ we have

$$\forall v_0 \ldots \forall v_n\ \forall v_j\ \forall v_k\ [v_i = v_j \rightarrow \Theta(v_0, \ldots, v_n) \\ \leftrightarrow \Theta(v_0, \ldots, v_i/v_j, \ldots, v_n)]$$

and so

$$\forall v_0 \ldots \forall v_n\ \forall v_j\ [v_i = v_j \rightarrow (\forall v_k)(\Theta(v_0, \ldots, v_n) \\ \leftrightarrow \Theta(v_0, \ldots, v_i/v_j, \ldots, v_n))]$$

and so

$$\forall v_0 \ldots \forall v_n\ \forall v_j\ [v_i = v_j \rightarrow (\forall v_k)\Theta(v_0, \ldots, v_n) \\ \leftrightarrow (\forall v_k)\Theta(v_0, \ldots, v_i/v_j, \ldots, v_n))].$$

Case 7 $(\exists v_k)\ \Theta)$ and $i \neq k$
This follows from Cases 1 and 6.

Case 8 $(\forall v_i\ \Theta)$
This is true since v_i is not free in Φ, hence $\Phi(v_0, \ldots, v_i/v_j, \ldots, v_n)$ is $\Phi(v_0, \ldots, v_n)$.

Case 9 $(\exists v_i)\ \Theta)$
This is similar to case 8.

We conclude that the theorem scheme holds for all appropriate Φ, i, j.

$\square$

Let M be a term and Φ any formula of the language of set theory. We define the relativisation of Φ to M, denoted by Φ^M, as follows:

1. If Φ is atomic then Φ^M is Φ;
2. If Φ is $\neg\Psi$ then Φ^M is $\neg\Psi^M$;
3. If Φ is $(\Psi_1 \wedge \Psi_2)$ then Φ^M is $(\Psi_1^M \wedge \Psi_2^M)$;
4. If Φ is $(\Psi_1 \vee \Psi_2)$ then Φ^M is $(\Psi_1^M \vee \Psi_2^M)$;
5. If Φ is $(\Psi_1 \to \Psi_2)$ then Φ^M is $(\Psi_1^M \to \Psi_2^M)$;
6. If Φ is $(\Psi_1 \leftrightarrow \Psi_2)$; then Φ^M is $(\Psi_1^M \leftrightarrow \Psi_2^M)$;
7. If Φ is $(\forall v_i)\Psi$ then Φ^M is $(\forall v_i \in M)\Psi^M$; and,
8. If Φ is $(\exists v_i)\Psi$ then Φ^M is $(\exists v_i \in M)\Psi^M$.

We write $M \models \Phi$ for Φ^M and moreover whenever Φ has no free variables we say that M is a model of Φ.

We denote by $\mathfrak{ZFC}$ the collection of axioms which include: Equality, Extensionality, Existence, Pairing, Foundation, Union, Intersection, the Replacement Scheme, Power Set, Choice and Infinity.

For each axiom Φ of $\mathfrak{ZFC}$, except for the Axiom of Infinity, we have:

Lemma. Φ

$\mathbf{R}(\omega) \models \Phi$.

Lemma. If M is transitive, then M models Equality, Extensionality, Existence and Foundation.

For each axiom Φ of $\mathcal{ZFC}$, except for those in the Replacement Scheme, we have:

Theorem 40. *For each uncountable* κ, $\boldsymbol{R}(\kappa) \models \Phi$.

Exercise 30. Prove the above theorem scheme. Use the fact that $\mathbb{V} \models \Phi$.

For each axiom Φ of $\mathcal{ZFC}$, except for Power Set, we have:

Theorem 41. *For each uncountable regular cardinal* κ, $\boldsymbol{H}(\kappa) \models \Phi$.

Exercise 31. Prove the above theorem scheme.

If κ is an inaccessible cardinal, then $\mathbf{R}(\kappa) = \mathbf{H}(\kappa) \models \Phi$, for each axiom Φ of $\mathcal{ZFC}$.

Lemma. If κ is the least inaccessible cardinal, then

$$H(\kappa) \models \neg\exists \text{ an inaccessible cardinal.}$$

From this lemma we can infer that there is no proof, from $\mathcal{ZFC}$, that there is an inaccessible cardinal. Suppose Θ is the conjuction of all the (finitely many) axioms used in such a proof. Then from Θ we can derive $(\exists\lambda)(\lambda$ is an inaccessible). Let κ be the least inaccessible cardinal. Then $H(\kappa) \models \Theta$ so

$$H(\kappa) \models (\exists\lambda)(\lambda \text{ is an inaccessible}).$$

This assumes that our proof system is sound; i.e., if from Θ_1 we can derive Θ_2 and $M \models \Theta_1$, then $M \models \Theta_2$. We conclude that $\mathcal{ZFC}$ plus "$\neg\exists$ an inaccessible" is consistent.

We can also infer that $\mathcal{ZFC}$ minus "Infinity" plus "$\neg(\exists z)(z = \mathbb{N})$" is consistent. Suppose not. Suppose Θ is the conjuction of the finitely many axioms of $\mathcal{ZFC}$ − Inf needed to prove $(\exists z)(z = N)$. Then $\mathbf{H}(\omega) \models \Theta$, so $\mathbf{H}(\omega) \models (\exists z)(z = \mathbb{N})$, which is a contradiction.

In fact, given any collection of formulas without free variables and a class M such that for each Φ in the collection $M \models \Phi$, we can then conclude that the collection is consistent.

For example, $\mathfrak{ZFC}$ minus "Power Set" plus "$(\forall x)(x$ is countable)" is consistent:

Lemma. $\mathbf{H}(\omega_1) \models (\forall x)(x \text{ is countable})$.

If $\Phi(v_0, \ldots, v_k)$ is a formula of the language of set theory and $M = \{x : \chi_M(x, \vec{v})\}$ and $C = \{x : \chi_C(x, \vec{v})\}$ are classes, then we say Φ is <u>absolute between M and C</u> whenever

$$(\forall v_0 \in M \cap C) \ldots (\forall v_k \in M \cap C)\ [\Phi^M \leftrightarrow \Phi^C].$$

This concept is most often used when $M \subseteq C$. When $C = \mathbb{V}$, we say that Φ is <u>absolute</u> for M.

For classes $M = \{x : \chi_M(x, \vec{v})\}$ and $C = \{x : \chi_C(x, \vec{v})\}$ with $M \subseteq C$ and a list $\Phi_0, \ldots, \Phi_m$ of formulas of set theory such that for each $i \leq m$ every subformula of Φ_i is contained in the list, we have:

Lemma. $\chi_M, \chi_C, \Phi_1, \ldots, \Phi_m$

The following are equivalent:

1. Each of $\Phi_1, \ldots, \Phi_m$ are absolute between M and C.

2. Whenever Φ_i is $\exists x\ \Phi_j(x, \vec{v})$ for $i, j \leq m$ we have

$$(\forall v_1 \in M \ldots \forall v_k \in M)(\exists x \in C\ \Phi_j^C(x, \vec{v}) \rightarrow \exists x \in M\ \Phi_j^C(x, \vec{v})).$$

The latter statement is called the <u>Tarski-Vaught Condition</u>.

Proof. $((1) \Rightarrow (2))$

Let $v_0 \in M, \ldots, v_k \in M$ and suppose $\exists x \in C\ \Phi_j^C(x, v_0, \ldots, v_k)$. Then $\Phi_i^C(v_0, \ldots, v_k)$ holds. By absoluteness $\Phi_i^M(v_0, \ldots, v_k)$ holds; i.e., $\exists x \in M\ \Phi_j^M(x, v_0, \ldots, v_k)$, so by absoluteness of $\Phi_j(x, v_0, \ldots, v_k)$ for $x \in M$ we have $\exists x \in M\ \Phi_j^C(x, v_0, \ldots, v_k)$.

(2) $\Rightarrow$ (1)

This is proved by induction on complexity of Φ_i, noting that each subformula appears in the list. There is no problem with the atomic formula step since atomic formulas are always absolute. Similarly, the negation and connective steps are easy. Now if Φ_i is $\exists x\ \Phi_j$ and each of $v_1, \ldots, v_k$ is in M we have

$$\begin{aligned}
\Phi_i^M(v_1, \ldots, v_k) &\Leftrightarrow \exists x \in M\ \Phi_j^M(x, v_1, \ldots, v_k) \\
&\Leftrightarrow \exists x \in M\ \Phi_j^C(x, v_1, \ldots, v_k) \\
&\Leftrightarrow \exists x \in C\ \Phi_j^C(x, v_1, \ldots, v_k) \\
&\Leftrightarrow \Phi_i^C(v_1, \ldots, v_k)
\end{aligned}$$

where the second implication is due to the inductive hypothesis and the third implication is by part (2).

□

The next theorem scheme is called the Levy Reflection Principle.

For each formula Φ of the language of set theory, we have:

Theorem 42. Φ

$\forall \alpha \in \mathbb{ON}\ \exists \beta \in \mathbb{ON}\ [\beta > \alpha$ and Φ is absolute for $R(\beta)]$.

If, in addition, Φ has no free variables then Φ implies $R(\beta) \models \Phi$. This is interpreted as the truth of Φ being reflected to $R(\beta)$.

Proof. Form a collection $\Phi_1, \ldots, \Phi_m$ of all the subformulas of Φ. We will use the Tarski-Vaught Condition for $\Phi_1, \ldots, \Phi_m$ to get absoluteness between $R(\beta)$ and $\mathbb{V}$; but first we must find β.

For each $i = 1, \ldots, m$ such that Φ_i is $\exists x\ \Psi$ for some formula Ψ we define f_i such that $f_i \colon \mathbb{V} \to \mathbb{ON}$ by setting

$$f_i(\langle y_1, \ldots, y_{l_i} \rangle) = \min \{\gamma : \gamma \in \mathbb{ON} \wedge \exists x \in R(\gamma)\ \Psi(x, y_1, \ldots, y_{l_i})\}$$

if such an ordinal exists, and $f_i(\langle y_1, \ldots, y_{l_i} \rangle) = \alpha$ otherwise.

Now define $h \colon \omega \to \mathbb{ON}$ by recursion by the formulas

$$\begin{aligned} h(0) &= \alpha \\ h(n+1) &= \sup \{f_i(\langle y_1, \ldots, y_{l_i} \rangle) : 1 \leq i < m \text{ and each } y_j \in R(h(n))\} \end{aligned}$$

and then let $\beta = \sup \{h(n) : n \in \omega\}$. This β works.

$\square$

The analogous theorem scheme can be proven for the $H(\kappa)$ hierarchy as well.

We can now argue that $\mathfrak{ZFC}$ cannot be finitely axiomatised. That is, there is no one formula without free variables which implies all axioms of $\mathfrak{ZFC}$ and is, in turn, implied by $\mathfrak{ZFC}$. Suppose such a Φ exists. By the Levy Reflection Principle, choose the least $\beta \in \mathbb{ON}$ such that $\Phi^{R(\beta)}$. We have $R(\beta) \models \Phi$. Hence $R(\beta) \models \exists \alpha \in \mathbb{ON}\ \Phi^{R(\alpha)}$, since this instance of the theorem follows from $\mathfrak{ZFC}$. Thus,

$$(\exists \alpha \in \mathbb{ON}\ \Phi^{R(\alpha)})^{R(\beta)}.$$

That is,

$$\exists \alpha \in (\mathbb{ON} \cap R(\beta))\ \Phi^{R(\alpha) \cap R(\beta)}.$$

But $\mathbb{ON} \cap R(\beta) = \beta$, the ordinals of *rank* $< \beta$. So, $\alpha < \beta$ and we have $\exists \alpha < \beta\ \Phi^{R(\alpha)}$, contradicting the minimality of β.

For any formula Φ of the language of set theory with no free variables and classes $M = \{v : \chi_M(v)\}$, $C = \{v : \chi_C(v)\}$, and $F = \{v : \chi_F(v)\}$ we have:

Lemma. $\Phi, \chi_M, \chi_C, \chi_F$

If $F \colon M \to C$ is an isomorphism then Φ is absolute between M and C. That is, $M \models \Phi$ iff $C \models \Phi$.

Proof. It is easy to show by induction on the complexity of Φ that

$$\Psi(x_1, \ldots, x_k) \Leftrightarrow \Psi(F(x_1), \ldots, F(x_k))$$

for each subformula Ψ of Φ.

$\square$

A <u>bounded formula</u> (also called a <u>$\triangle_0$ formula</u>) is one which is built up as usual with respect to atomic formulas and connectives, but where each $\exists x\ \Phi$ clause is replaced by $\exists x \in y\ \Phi$, and the $\forall x\ \Phi$ clause is replaced by $\forall x \in y\ \Phi$.

$\triangle_0$ formulas are absolute for transitive models. That is, for each $\triangle_0$ formula $\Phi(v_0, \ldots, v_k, \vec{w})$ and for each class $M = \{x : \chi_M(x, \vec{w})\}$ we have:

Theorem 43. Φ, χ_M

$\forall \vec{w}$ *if* M *is transitive, then*

$$(\forall v_0 \in M) \ldots (\forall v_k \in M)\ [\Phi^M(v_0, \ldots, v_k, \vec{w}) \leftrightarrow \Phi(v_0, \ldots, v_k, \vec{w})]$$

Proof. We use induction on the complexity of Φ. We only show the $\exists$ step. Suppose $[(\forall v_0 \in M) \ldots (\forall v_k \in M)\ \Phi^M(v_0, \ldots, v_k, \vec{w})] \leftrightarrow \Phi(v_0, \ldots, v_k, \vec{w})$. We wish to consider $(\exists v_i \in v_j)\ \Phi(v_0, \ldots, v_k, \vec{w})$.

Fix any $v_0, \ldots, v_k \in M$, but not v_i; however, v_j is fixed in M. Now

$$(\exists v_i \in v_j\ \Phi(v_0, \ldots, v_k, \vec{w}))^M \rightarrow (\exists v_i \in v_j\ \Phi(v_0, \ldots, v_k, \vec{w}))$$

since $\Phi^M(\vec{v}, \vec{w}) \rightarrow \Phi(\vec{v}, \vec{w})$. Also,

$$\begin{aligned}
&(\exists v_i \in v_j)\ \Phi(\vec{v}, \vec{w}) \\
\rightarrow &(\exists v_i \in v_j)\ \Phi^M(\vec{v}, \vec{w}) \text{ since } \Phi(\vec{v}, \vec{w}) \\
\rightarrow &\Phi^M(\vec{v}, \vec{w}) \\
\rightarrow &\exists v_i \in (v_j \cap M)\ \Phi^M(\vec{v}, \vec{w}) \text{ since } v_j \in M \text{ implies } v_j \subseteq M \\
\rightarrow &((\exists v_i \in v_j)\ \Phi(\vec{v}, \vec{w}))^M
\end{aligned}$$

$\square$

A formula $\Phi(\vec{w})$ is said to be $\triangle_1^T$, where T is a subcollection of $\mathfrak{ZFC}$, whenever there are $\triangle_0$ formulas $\Phi_1(\vec{v}, \vec{w})$ and $\Phi_2(\vec{v}, \vec{w})$ such that using only the axioms from T we can prove that both:

$$(\forall \vec{w})(\Phi_1(\vec{w}) \leftrightarrow \exists v_0 \ldots \exists v_k \ \Phi_1(\vec{v}, \vec{w}))$$
$$(\forall \vec{w})(\Phi_2(\vec{w}) \leftrightarrow \forall v_0 \ldots \forall v_k \ \Phi_2(\vec{v}, \vec{w})).$$

We can now use the above theorem to show that if $\Phi(\vec{w})$ is a $\triangle_1^T$ formula and $M \models \Theta$ for each Θ in T, then $\Phi(\vec{w})$ is absolute for M, whenever M is transitive. We do so as follows, letting Ψ play the role of Ψ_1 above:

$$\begin{aligned}
\Phi(\vec{w}) &\Leftrightarrow \exists v_0 \ldots \exists v_k \ \Psi(\vec{v}, \vec{w}) \\
&\Leftrightarrow (\exists v_0 \in M) \ldots (\exists v_k \in M) \ [\Psi(\vec{v}, \vec{w})] \\
&\Leftrightarrow \exists v_0 \ldots \exists v_k \ \Psi(\vec{v}, \vec{w})^M \\
&\Leftrightarrow \Phi(\vec{w})^M
\end{aligned}$$

The first and third implications accrue from $\Phi(\vec{w})$ being $\triangle_1^T$; the second from the fact that Ψ is $\triangle_0$ and M is transitive and models the axioms of T.

This is often used with T as $\mathfrak{ZFC}$ without Power Set and $M = \mathbf{H}(\kappa)$.

Chapter 11

Elementary Submodels

In this chapter we shall first introduce a collection of set operations proposed by Kurt Gödel which are used to build sets. We shall then discuss the new concept of elementary submodel.

We now define the ordered $\underline{n\text{-tuple}}$ with the following infinitely many formulas, thereby extending the notion of ordered pair.

$$\begin{aligned}
\langle x \rangle &= x \\
\langle x, y \rangle &= \{\{x\}, \{x, y\}\} \\
\langle x, y, z \rangle &= \langle \langle x, y \rangle, z \rangle \\
\langle x_1, \ldots, x_n \rangle &= \langle \langle x_1, \ldots, x_{n-1} \rangle, x_n \rangle
\end{aligned}$$

The following operations shuffle the components of such tuples in a set S.

$$\begin{aligned}
F_0(S) &= \{\langle \langle u, v \rangle, w \rangle : \langle u, \langle v, w \rangle \rangle \in S\} \\
F_1(S) &= \{\langle u, \langle v, w \rangle \rangle : \langle \langle u, v \rangle, w \rangle \in S\} \\
F_2(S) &= \{\langle v, u \rangle : \langle u, v \rangle \in S\} \\
F_3(S) &= \{\langle v, u, w \rangle : \langle u, v, w \rangle \in S\} \\
F_4(S) &= \{\langle t, v, u, w \rangle : \langle t, u, v, w \rangle \in S\}
\end{aligned}$$

Lemma. (Shuffle Lemma Scheme)

For any $m \in \mathbb{N}$ and any permutation $\sigma\colon m \to m$, there is a composition of the operations F_0, F_1, F_2, F_3, F_4 such that for any S,

$$F_\sigma(S) = \{\langle x_{\sigma(0)}, \ldots, x_{\sigma(m-1)}\rangle : \langle x_0, \ldots, x_{m-1}\rangle \in S\}.$$

Proof. Since binary exchanges generate the symmetric group, noting that the identity permutation is given by $F_2 \circ F_2$, it suffices to consider only σ such that for some $l \leq m$,

$$\sigma(i) = \begin{cases} i+1, & \text{if } i = l; \\ i-1, & \text{if } i = l+1; \\ i, & \text{otherwise.} \end{cases}$$

For convenience, let F_i^n denote the $n-$fold composition of F_i. There are several cases.

Case 1: $m = 2$
$F_\sigma = F_2$

Case 2: $m = 3,\ l = 1$
$F_\sigma = F_3$

Case 3: $m = 3,\ l = 2$
$F_\sigma = F_0 \circ F_2 \circ F_3 \circ F_2 \circ F_1$

Case 4: $m \geq 4,\ l = 1$
$F_\sigma = F_0^{m-3} \circ F_3 \circ F_1^{m-3}$

Case 5: $m \geq 4,\ 1 \leq l \leq m-1$
$F_\sigma = F_0^{m-l-2} \circ F_4 \circ F_1^{m-l-2}$

Case 6: $m \geq 4,\ l = m-1$
$F\sigma = F_0 \circ F_2 \circ F_3 \circ F_2 \circ F_1$

$\square$

The Gödel Operations are as follows:

$$
\begin{aligned}
G_1(X,Y) &= \{X,Y\}\\
G_2(X,Y) &= X \setminus Y\\
G_3(X,Y) &= \{\langle u,v\rangle : u \in X \wedge v \in Y\};\ \text{i.e., } X \times Y\\
G_4(X,Y) &= \{\langle u,v\rangle : u \in X \wedge v \in Y \wedge u = v\}\\
G_5(X,Y) &= \{\langle u,v\rangle : u \in X \wedge v \in Y \wedge u \in v\}\\
G_6(X,Y) &= \{\langle u,v\rangle : u \in X \wedge v \in Y \wedge v \in u\}\\
G_7(X,R) &= \{u : \exists x \in X\ \langle u,x\rangle \in R\}\\
G_8(X,R) &= \{u : \forall x \in X\ \langle u,x\rangle \in R\}\\
G_9(X,R) &= F_0(R)\\
G_{10}(X,R) &= F_1(R)\\
G_{11}(X,R) &= F_2(R)\\
G_{12}(X,R) &= F_3(R)\\
G_{13}(X,R) &= F_4(R)
\end{aligned}
$$

G_9 through G_{13} are defined as binary operations for conformity.

We now construct a function which, when coupled with recursion on $\mathbb{ON}$, will enumerate all possible compositions of Gödel operations. $\mathbf{G}\colon \mathbb{N}\times\mathbb{V} \to \mathbb{V}$ is given by the following rule. For each $k \in \omega$ and each $\vec{s}\colon k \to \mathbb{V}$, we define $\mathbf{G}|_{\mathbb{N}\times\vec{s}}$ by recursion on $\mathbb{N}$ as follows:

$$
\mathbf{G}(n,\vec{s}) = \begin{cases} \vec{s}(i), & \text{if } n = 17^i, 0 \le i \le k;\\ \mathbf{G}_m(\mathbf{G}(i,\vec{s}),\mathbf{G}(j,\vec{s})), & \text{if } n = m\cdot 19^{i+1}23^{j+1}, 1 \le m \le 13;\\ \emptyset, & \text{otherwise.}\end{cases}
$$

—

Now $\mathbf{G}$ enumerates all compositions of Gödel operations.

We wish to prove that sets defined by $\triangle_0$ formulas can be obtained through a composition of Gödel operations. First we need a lemma.

Lemma. Each of the following sets is equal to a composition of Gödel operations on $X, \vec{w}$.

1. $\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \wedge w_i = w_j\}$
2. $\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \wedge w_i \in w_j\}$
3. $\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \wedge w_i = x_j\}$
4. $\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \wedge w_i \in x_j\}$
5. $\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \wedge x_i \in w_j\}$
6. $\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \wedge x_i = x_j\}$
7. $\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \wedge x_i \in x_j\}$

Proof. First note that

$$\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X\} = G_3(G_3 \ldots (G_3(X, X), \ldots X), X)$$

where G_3 is composed $(m-1)-$fold; i.e., $(\ldots((X \times X) \times X) \times \cdots \times X)$. Now call this set $P_m(X)$, the m^{th} power of X. Let us now examine each of the seven cases individually.

1. This is either $P_m(X)$ or $\emptyset$ depending upon whether or not $w_i = w_j$.
2. This is either $P_m(X)$ or $\emptyset$ depending upon whether or not $w_i \in w_j$.
3. If $m = 1$, then this set is $\langle w_i\rangle = w_i$. For $m > 1$, we may, thanks to the Shuffle Lemma, without loss of generality assume that $j = m$. The set is equal to
$$G_3(P_{m-1}(X), w_i)$$
if $w_i \in X$ and $\emptyset$ otherwise.
4. Again, without loss of generality, we assume $j = m$. This set is equal to
$$G_3(P_{m-1}(X), G_7(X, G_6(X, \{w_i\}))).$$
5. Again, assume $j = m$. This set is given by
$$G_3(P_{m-1}(X), G_7(X, G_5(X, \{w_i\}))).$$

6. This time, assume $i = m - 1$ and $j = m$. This set is
$$G_3(P_{m-2}(X), G_4(X, X)).$$

7. We assume again that $i = m - 1$ and $j = m$. This set becomes
$$G_3(P_{m-2}(X), G_5(X, X)).$$

□

For each formula $\Phi(x, \vec{w})$ of the language of set theory we have:

Theorem 44. Φ
$$(\forall \vec{w})(\forall X)(\exists n \in \omega)\ [\{x \in X : \Phi^X(x, \vec{w})\} = \boldsymbol{G}(n, \vec{s})]$$
where $\vec{s}(i) = w_i$ *for* $i < k$ *and* $s(k) = X$*.*

Proof. We prove by induction on the complexity of Φ that for all $m \in \omega$ and Φ

$$(\forall \vec{w})(\exists n \in \omega)[\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \text{ and } \Phi^X(x_1, \ldots, x_m, \vec{w})\} = \mathbf{G}(n, \vec{s})]$$

The proof of the theorem for any given Φ will assume the corresponding result for a finite number of simpler formulas, the proper subformulas of Φ.

We begin by looking at atomic formulas Φ. This step is covered by the previous lemma.

Now we proceed by induction on complexity. Suppose that
$$\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \text{ and } \Phi_1^X(x_1, \ldots, x_m, \vec{w})\} = G(n_1, \vec{s})$$
and
$$\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \wedge \cdots \wedge x_m \in X \text{ and } \Phi_2^X(x_1, \ldots, x_m, \vec{w})\} = G(n_2, \vec{s}).$$

Then

$$\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \ldots x_m \in X \text{ and } \neg\Phi_1^X(x_1, \ldots, x_m, \vec{w})\}$$
$$= P_m(X) \setminus G(n_1, \vec{s}) = G_2(P_m(X), G(n_1, \vec{s}))$$

and

$$\{\langle x_1, \ldots, x_m\rangle : x_1 \in X \ldots x_m \in X \text{ and } \Phi_1^X \wedge \Phi_2^X(x_1, \ldots, x_m, \vec{w})\}$$
$$= G(n_1, \vec{s}) \cap G(n_2, \vec{s}).$$

The other connectives can be formed from $\neg$ and $\wedge$; as such, $\forall x_l$ is $\neg\exists x_l$, so it only remains to do the $\Phi = \exists x_l\ \Phi_1$ step. Thanks again to the last lemma, we may assume that $l = m$. Then Φ is $\exists x_m\ \Phi_1$ and

$$\{\langle x_1 \ldots x_{m-1}\rangle : x_1 \in X \wedge \cdots \wedge x_{m-1} \in X \text{ and } \Phi^X(x_1, \ldots, x_{m-1}, \vec{w})\}$$
$$= G_7(G(n_1, \vec{s}))$$

$\square$

Let's write $\mathbf{G}(n, X, \vec{y})$ for $\mathbf{G}(n, \vec{s})$, where $\vec{s}(0) = X$ and $\vec{s}(k+1) = \vec{y}(k)$ for all $k \in dom(\vec{y}) = dom(\vec{s}) - 1$.

M is said to be an <u>elementary submodel</u> of N whenever

1. $M \subseteq N$; and,
2. $\forall k \in \omega\ \forall \vec{y} \in {}^kM\ \forall n \in \omega\ \mathbf{G}(n, N, \vec{y}) \cap N \neq \emptyset \Leftrightarrow \mathbf{G}(n, N, \vec{y}) \cap M \neq \emptyset$.

We write $M \prec N$.

Justification of the terminology comes from the following theorem scheme. For each formula Φ of the language of set theory we have:

Theorem 45. Φ

Suppose $M \prec N$. Then Φ is absolute between M and N.

Proof. We will use the Tarski-Vaught criterion. Let $\Phi_0, \ldots, \Phi_m$ enumerate Φ and each of its subformulas. Suppose Φ_i is $\exists x\ \Phi_j(x, y_0, \ldots, y_k)$ with $y_0, \ldots, y_k \in M$ in a sequence $\vec{y}$.

Let $n \in \omega$ so that by Theorem 44 we can find $n \in \omega$ such that

$$\mathbf{G}(n, N, \vec{y}) = \{x \in N : \Phi_j^N(x, y_0, \ldots, y_k)\}.$$

Then

$$\begin{aligned}
\exists x \in N\ \Phi_j^N(x, y_0, \ldots, y_k) &\leftrightarrow \mathbf{G}(n, N, \vec{y}) \cap N \neq \emptyset \\
&\leftrightarrow \mathbf{G}(n, N, \vec{y}) \cap M \neq \emptyset \\
&\leftrightarrow \exists x \in M\ \Phi_j^N(x, y_0, \ldots, y_k).
\end{aligned}$$

□

The following is sometimes called the Lowenheim-Skolem-Tarski theorem.

Theorem 46. *Suppose $X \subseteq N$. Then there is an M such that*

1. *$M \prec N$;*
2. *$X \subseteq M$; and,*
3. *$|M| \leq max\{\omega, |X|\}$.*

Proof. Define $F\colon \omega \times \bigcup\{{}^kN : k \in \omega\} \to N$ by choice:

$$F(n, \vec{s}) = \begin{cases} \text{some element of } \mathbf{G}(n, N, \vec{s}) & \text{if } \mathbf{G}(n, N, \vec{s}) \neq \emptyset \\ \text{any element of } \mathbb{N} & \text{otherwise} \end{cases}$$

Now define $\{X_n\}_{n\in\omega}$ by recursion on $\mathbb{N}$ as follows:

$$\begin{aligned}
X_0 &= X \\
X_{m+1} &= X_m \cup F''(\omega \times \bigcup\{{}^k(X_m) : k \in \omega\})
\end{aligned}$$

Let $M = \bigcup_{m\in\omega} X_m$. As such, (2) and (3) are clearly satisfied. To check (1), let $\vec{y} \in {}^k(X_m)$. If $\mathbf{G}(n, N, \vec{y}) \cap N \neq \emptyset$, then $F(n, \vec{y}) \in X_{m+1} \subseteq M$ and $\mathbf{G}(n, N, \vec{y}) \cap M \neq \emptyset$.

□

The use of elementary submodels of the $\mathbf{H}(\theta)$ can be illustrated.

Theorem 47. *(Pressing Down Lemma)*

Let $f : \omega_1 \setminus \{0\} \to \omega_1$ be regressive; i.e., $f(\alpha) < \alpha$ for all α.

Then $\exists \beta \in \omega_1$ such that $f^{\leftarrow}\{\beta\}$ is uncountable.

Theorem 48. *(Delta System Lemma)*

Let $\mathcal{A}$ be an uncountable collection of finite sets.

Then $\exists \mathcal{D} \subseteq \mathcal{A}\ \exists R$ such that

1. *$\mathcal{D}$ is uncountable, and*
2. *$\forall D_1, D_2 \in \mathcal{D}\ D_1 \cap D_2 = R$.*

We need some lemmas. Assume $M \prec \mathbf{H}(\theta)$ where θ is an uncountable regular cardinal. For each $\triangle_0$ formula $\Phi(v_0, \ldots, v_k)$ we have:

Lemma. Φ

$$(\forall y_0 \in M) \ldots (\forall y_k \in M)\ [M \models \Phi(y_0, \ldots, y_k) \Leftrightarrow \Phi(y_0, \ldots, y_k)].$$

Proof.

$$\begin{aligned} M \models \Phi(y_0, \ldots, y_k) &\Leftrightarrow \mathbf{H}(\theta) \models \Phi(y_0, \ldots, y_k) \text{ by elementarity,} \\ &\Leftrightarrow \Phi(y_0, \ldots, y_k) \text{ since } H(\theta) \text{ is transitive.} \end{aligned}$$

□

Remark. The same is true for $\triangle_1^T$ formulas where T is $\mathfrak{ZFC}$ without Power Set.

For any formula $\Phi(v_0, \ldots, v_k)$ of LOST, we have:

Lemma. Φ

$$\begin{gathered}\forall y_0 \in M\ \forall y_2 \in M \ldots \forall y_k \in M\ \forall x \in \mathbf{H}(\theta) \\ [\mathbf{H}(\theta) \models\ z = \{x : \Phi(x, y_0, \ldots, y_k)\} \rightarrow z \in M].\end{gathered}$$

Proof. Let $y_0, \ldots, y_k \in M$ and $z \in \mathbf{H}(\theta)$ be given such that

$$\mathbf{H}(\theta) \models z = \{x : \Phi(x, y_0, \ldots, y_k)\}.$$

Then,

$$\begin{aligned}&\mathbf{H}(\theta) \models\ \exists u\ u = \{x : \Phi(x, y_0, \ldots, y_k)\} \\ \Rightarrow &M \models\ \exists u\ u = \{x : \Phi(x, y_0, \ldots, y_k)\} \\ \Rightarrow &\exists p \in M\ [M \models p = \{x : \Phi(x, y_0, \ldots, y_k)\}] \\ \Rightarrow &\mathbf{H}(\theta) \models p = \{x : \Phi(x, y_0, \ldots, y_k)\} \\ \Rightarrow &\mathbf{H}(\theta) \models p = z.\end{aligned}$$

$\mathbf{H}(\theta)$ is transitive; therefore, $p = z$ and hence $z \in M$.

□

Corollaries. 1. If $M \prec \mathbf{H}(\theta)$, then

(a) $\emptyset \in M$;

(b) $\omega \in M$; and,

(c) $\omega \subseteq M$.

2. If also $\theta > \omega_1$, then $\omega_1 \in M$.

Proof. $\emptyset$ and ω are direct. For $\omega \subseteq M$ show that $y \in M \Rightarrow y \cup \{y\} \in M$.

□

Lemma. Suppose $M \prec \mathbf{H}(\theta)$ where θ is regular and uncountable. Suppose p is countable and $p \in M$. Then $p \subseteq M$.

Proof. Let $q \in p$; we must show that $q \in M$. Let $f_0 \colon \omega \to p$ be a surjection. Since $\{\omega, p\} \subset \mathbf{H}(\theta)$ we must have $f_0 \in \mathbf{H}(\theta)$. Since the formula "$f \colon \omega \to p$ and p is surjective" is a $\triangle_0$ formula and $\{f_0, \omega, p\} \subset \mathbf{H}(\theta)$, we have $\mathbf{H}(\theta) \models (f_0 \colon \omega \to p$ and p is surjective$)$. So

$$\mathbf{H}(\theta) \models (\exists f)(f \colon \omega \to \text{ and } p \text{ is surjective}).$$

Since $\{\omega, p\} \subset M$ we have,

$$M \models (\exists f)(f \colon \omega \to p \text{ and } p \text{ is surjective}).$$

That is, $(\exists f_p \in M)(f_p \colon \omega \to p$ and p is surjective$)$.

Pick $n \in \omega$ such that $f_p(n) = q$, and again use the first lemma as follows. Since $\{p, f_p, n\} \subset M$ and $(\exists! x)(x \in p$ and $f_p(n) = x)$ is a $\triangle_0$ formula

$$M \models (\exists! x)(x \in p \text{ and } f_p(n) = x).$$

That is, $(\exists! x)(x \in p \cap M$ and $f_p(n) = x)$. Since x is unique, $x = q$ and thus $q \in M$.

$\square$

Corollary. $\omega_1 \cap M \in \omega_1$.

Proof. It is enough to show that $\omega_1 \cap M$ is a countable initial segment of ω_1. If $\alpha \in \omega_1 \cap M$, then by the above lemma, $\alpha \subseteq M$.

$\square$

<u>Proof of Pressing Down Lemma</u>

Let $M \prec \mathbf{H}(\omega_2)$ such that M is countable and $f \in M$. Let $\delta = \omega_1 \cap M$ and let $\beta = f(\delta) < \delta$. Then,

$$(\forall \alpha < \delta)(\exists x \in \omega_1)\ [x > \alpha \ \wedge \ f(x) = \beta].$$

So $\forall\alpha < \delta$ $\mathbf{H}(\omega_2) \models (\exists x \in \omega_1)(x > \alpha \ \wedge \ f(x) = \beta)$, since everything relevant is in $\mathbf{H}(\omega_2)$. Hence,

$$\forall\alpha < \delta \ M \models (\exists x \in \omega_1)(x > \alpha \ \wedge \ f(\alpha) = \beta)$$

since $\{\alpha, \beta, \omega_1, f\} \subset M$. Now, since $\delta = \omega_1 \cap M$ we have,

$$M \models (\forall\alpha \in \omega_1)(\exists x \in \omega_1) \ [x > \alpha \ \wedge \ f(\alpha) = \beta].$$

So $\mathbf{H}(\omega_2) \models (\forall\alpha \in \omega_1)(\exists x \in \omega_1) \ [x > \alpha \ \wedge \ f(\alpha) = \beta]$. Thus we have

$$\mathbf{H}(\omega_2) \models f^{\leftarrow}\{\beta\} \text{ is uncountable.}$$

Again, since everything relevant is in $\mathbf{H}(\omega_2)$ we conclude that $f^{\leftarrow}\{\beta\}$ is uncountable.

□

Proof of the Delta System Lemma

Let $\mathcal{A}$ be as given. We may, without loss of generosity, let

$$\mathcal{A} = \{a(\alpha) : \alpha < \omega_1\}$$

where $a \colon \omega_1 \to \mathbb{V}$. We may also assume that $a \colon \omega_1 \to \mathcal{P}(\omega_1)$.

Let M be countable with $\{\mathcal{A}, \dashv\} \subseteq \mathcal{M}$ and $M \prec \mathbf{H}(\omega_2)$. Let $\delta = \omega_1 \cap M$. Let $R = a(\delta) \cap \delta$. Since $R \subseteq M$, we know $R \in M$ by the second lemma. So,

$$\begin{aligned}
&\forall\alpha < \delta \ \exists\beta > \alpha \ a(\beta) \cap \beta = R \\
\Rightarrow &\mathbf{H}(\omega_2) \models (\forall\alpha < \delta)(\exists\beta > \alpha) \ [a(\beta) \cap \beta = R] \\
\Rightarrow &(\forall\alpha < \delta) \ [\mathbf{H}(\omega_2) \models (\exists\beta > \alpha)(a(\beta) \cap \beta = R)] \\
\Rightarrow &(\forall\alpha < \delta) \ [M \models (\exists\beta > \alpha)(a(\beta) \cap \beta = R)] \\
\Rightarrow &M \models (\forall\alpha < \omega_1)(\exists\beta > \alpha) \ [a(\beta) \cap \beta = R] \\
\Rightarrow &(\forall\alpha < \omega_1)(\exists\beta > \alpha) \ [a(\beta) \cap \beta = R].
\end{aligned}$$

Now recursively define $D \colon \omega_1 \to \mathcal{A}$ as follows:

$$\begin{aligned}
D(\alpha) &= a(0); \\
D(\gamma) &= a(\beta)
\end{aligned}$$

where β is the least ordinal such that

$$\beta > \sup\,\{D(\gamma) : \gamma < \alpha\} \text{ and } a(\beta) \cap \beta = R.$$

Now if $\gamma_1 < \gamma_2 < \omega_1$, then $D(\gamma_1) \subseteq \gamma_2$. So,

$$R \subseteq D(\gamma_1) \cap D(\gamma_2) \subseteq \gamma_2 \cap D(\gamma_2) = R.$$

Thus we let $\mathcal{D} = \{D(\alpha) : \alpha < \omega_1\}$.

□

Theorem 49. *(Elementary Chain Theorem)*

Suppose that δ is a limit ordinal and $\{M_\alpha : \alpha < \delta\}$ is a set of elementary submodels of $\boldsymbol{H}(\theta)$ such that

$$\forall\alpha\ \forall\alpha'\ (\alpha < \alpha' < \delta \to M_\alpha \subset M_{\alpha'}).$$

Let

$$M_\delta = \bigcup\{M_\alpha : \alpha < \delta\}.$$

Then $M_\delta \prec \boldsymbol{H}(\theta)$.

Proof. Let $k \in \omega$, let $\vec{y} \in {}^kM_\delta$, and let $n \in \omega$. We need to show that

$$\mathbf{G}(n, \mathbf{H}(\theta), \vec{y}) \cap \mathbf{H}(\theta) \neq \emptyset \Rightarrow \mathbf{G}(n, \mathbf{H}(\theta), \vec{y}) \cap M_\delta \neq \emptyset.$$

But this is easy since $\vec{y} \in {}^kM_\alpha$ for some $\alpha < \delta$.

□

Chapter 12

Constructibility

The Gödel closure of a set X is denoted by

$$cl(X) = \{X \cap \mathbf{G}(n, \vec{y}) : n \in \omega \text{ and } \exists k \in \omega \ \vec{y} \in \ {}^{k}(X)\}.$$

The constructible sets are obtained by first defining a function

$$\mathbb{L} \colon \mathbb{ON} \to \mathbb{V}$$

by recursion as follows:

$$\begin{aligned} \mathbf{L}(0) &= \emptyset \\ \mathbf{L}(\alpha + 1) &= cl(\mathbf{L}(\alpha) \cup \{\mathbf{L}(\alpha)\}) \\ \mathbf{L}(\delta) &= \bigcup \{\mathbf{L}(\alpha) : \alpha < \delta\} \text{ for a limit ordinal } \delta \end{aligned}$$

We denote by $\mathbb{L}$ the class $\bigcup\{\mathbf{L}(\alpha) : \alpha \in \mathbb{ON}\}$. Sets in $\mathbb{L}$ are said to be constructible.

Lemma. For each ordinal α, $\mathbf{L}(\alpha) \subseteq \mathbf{R}(\alpha)$.

Proof. This is proved by induction. $\mathbf{L}(0) = \emptyset = \mathbf{R}(0)$ and for each $\alpha \in \mathbb{ON}$ we have, by definition,

$$\begin{aligned} \mathbf{L}(\alpha + 1) &\subseteq \mathcal{P}(\mathbf{L}(\alpha)) \\ &\subseteq \mathbf{R}(\alpha + 1) \end{aligned}$$

□

Lemma.

1. $\forall X\ X \subseteq cl(X)$.
2. If X is transitive, then $cl(X)$ is transitive.
3. For each ordinal α, $\mathbf{L}(\alpha)$ is transitive.

Proof.

1. For any $w \in X$, $w = \mathbf{G}(1, \vec{s})$, where $\vec{s}(0) = w$.
2. Now, if $z \in cl(X)$ then $z \subseteq X$ so $z \subseteq cl(X)$.
3. This follows from (1) by induction on $\mathbb{ON}$.

□

Lemma.

1. For all ordinals $\alpha < \beta$, $\mathbf{L}(\alpha) \in \mathbf{L}(\beta)$.
2. For all ordinals $\alpha < \beta$, $\mathbf{L}(\alpha) \subseteq \mathbf{L}(\beta)$.

Proof.

1. For each α, $\mathbf{L}(\alpha) \in \mathbf{L}(\alpha + 1)$ by Part (1) of the previous lemma. We then apply induction on β.
2. This follows from (1) by transitivity of $\mathbf{L}(\beta)$.

□

Lemma.

1. For each ordinal β, $\beta \notin \mathbf{L}(\beta)$.
2. For each ordinal β, $\beta \in \mathbf{L}(\beta + 1)$.

Proof.

1. This is proved by induction on β. The case $\beta = 0$ is easy. If $\beta = \alpha + 1$ then $\beta \in \mathbf{L}(\alpha + 1)$ would imply that $\beta \subseteq \mathbf{L}(\alpha)$ and hence
$$\alpha \in \beta \subseteq \mathbf{L}(\alpha)$$
contradicting the inductive hypothesis. If β is a limit ordinal and $\beta \in \mathbf{L}(\beta)$ then $\beta \in \mathbf{L}(\alpha)$ for some $\alpha \in \beta$ and hence $\alpha \in \beta \in \mathbf{L}(\alpha)$, again a contradiction.

2. We employ induction on β. The $\beta = 0$ case is given by $0 \in \{0\}$. We do the sucessor and limit cases uniformly. Assume that
$$\forall \alpha \in \beta \ \alpha \in \mathbf{L}(\alpha + 1).$$
Claim 1. $\beta = \mathbf{L}(\beta) \cap \mathbb{ON}$.

Proof of Claim 1. If $\alpha \in \beta$, then $\alpha \in \mathbf{L}(\alpha + 1) \subseteq \mathbf{L}(\beta)$. If $\alpha \in \mathbf{L}(\beta)$, then $\alpha \in \beta$ because otherwise $\alpha = \beta$ or $\beta \in \alpha$, which contradicts $\beta \notin \mathbf{L}(\beta)$ from (1).

Claim 2. $\forall x \ x \in \mathbb{ON}$ iff
$$\begin{aligned}&[(\forall u \in x \ \forall v \in u \ v \in x) \wedge (\forall u \in x \ \forall v \in x \ (u \in v \ \vee \ v \in u \ \vee \ u = v))\\ &\qquad \wedge (\forall u \in x \ \forall v \in x \ \forall w \in x \ (u \in v \ \wedge \ v \in w \rightarrow u \in w))].\end{aligned}$$
Proof of Claim 2. The statement says that x is an ordinal iff x is a transitive set and the ordering $\in$ on x is transitive and satisfies trichotomy. This is true since $\in$ is automatically well founded.

The importance of this claim is that this latter formula, call it $\Phi(x)$, is $\triangle_0$ and hence absolute for transitive sets.

We have:
$$\begin{aligned}\beta = \mathbf{L}(\beta) \cap \mathbb{ON} &= \{x \in \mathbf{L}(\beta) : x \text{ is an ordinal}\}\\ &= \{x \in \mathbf{L}(\beta) : \Phi(x)\}\\ &= \{x \in \mathbf{L}(\beta) : \Phi^{\mathbf{L}(\beta+1)}(\alpha)\}\\ &\in cl(\mathbf{L}(\beta)) \text{ using Theorem 44}\\ &= \mathbf{L}(\beta + 1).\end{aligned}$$

□

Lemma.

1. For each ordinal β, $\beta = \mathbf{L}(\beta) \cap \mathbb{ON}$.
2. $\mathbb{ON} \subseteq \mathbb{L}$.

Proof. This is easy from the previous lemmas.

□

Lemma.

1. If W is a finite subset of X then $W \in cl(X)$.
2. If W is a finite subset of $\mathbf{L}(\beta)$ then $W \in \mathbf{L}(\beta + 1)$.

Proof.

1. We apply Theorem 44 to the formula "$x = w_0 \vee \cdots \vee x = w_n$", where $W = \{w_1, w_2, \ldots, w_n\}$.
2. This follows immediately from (1).

□

Lemma.

1. If X is infinite then $|cl(X)| = |X|$.
2. $\alpha \geq \omega$ then $|\mathbf{L}(\alpha)| = |\alpha|$.

Proof.

1. By Theorem 44 we can construct an injection $cl(X) \to \omega \times \bigcup\{{}^kX : k \in \omega\}$. Hence, $|X| \leq |cl(X)| \leq \max\,(\aleph_0, |\{{}^kX : k \in \omega\}|) = |X|$.

2. We proceed by induction, beginning with the case $\alpha = \omega$. We first note that from the previous lemma, we have $\mathbf{L}(n) = R(n)$ for each $n \in \omega$. Therefore,

$$\begin{aligned} |\mathbf{L}(\omega)| &= |\bigcup\{\mathbf{L}(n) : n \in \omega\}| \\ &= \max\left(\aleph_0, \sup\{|\mathbf{L}(n)| : n \in \omega\}\right) \\ &= \max\left(\aleph_0, \sup\{|R(n)| : n \in \omega\}\right) \\ &= \aleph_0. \end{aligned}$$

For the successor case,

$$\begin{aligned} |\mathbf{L}(\beta+1)| &= |\mathbf{L}(\beta)| \text{ by (1)} \\ &= |\beta| \text{ by inductive hypothesis} \\ &= |\beta+1| \text{ since } \beta \text{ is infinite .} \end{aligned}$$

And if δ is a limit ordinal then

$$\begin{aligned} |\mathbf{L}(\delta)| &= |\bigcup\{\mathbf{L}(\beta) : \beta \in \delta\}| \\ &= \max\left(|\delta|, \sup\{|\mathbf{L}(\beta)| : \beta \in \delta\}\right) \\ &= \max\left(|\delta|, \sup\{|(\beta)| : \beta \in \delta\}\right) \text{ by inductive hypothesis} \\ &= |\delta|. \end{aligned}$$

□

Lemma. $(\forall x)\ [x \subseteq \mathbb{L} \to (\exists y \in \mathbb{L})(x \subseteq y)]$.

Proof. $x \subseteq \mathbb{L}$ means that $\forall u \in x\ \exists \alpha \in \mathbb{ON}\ x \in \mathbf{L}(\alpha)$. By the Axiom of Replacement,

$$\exists z\ z = \{\alpha : (\exists u \in x)(\alpha \text{ is the least ordinal such that } u \in \mathbf{L}(\alpha))\}.$$

Let $\beta = \sup\ z$; then $\beta \in \mathbb{ON}$ and for each $u \in x$, there is $\alpha \leq \beta$ such that $u \in \mathbf{L}(\alpha) \subseteq \mathbf{L}(\beta)$. Since $\mathbf{L}(\beta) \in \mathbf{L}(\beta+1) \subseteq \mathbb{L}$, we can take $y = \mathbf{L}(\beta)$.

□

Remark. The above lemma is usually quoted as "$\mathbb{L}$ is almost universal".

Lemma. $\mathbb{L} \models \mathbb{V} = \mathbb{L}$.

Proof. This is *not* the trivial statement

$$\forall x \in \mathbb{L}\ x \in \mathbb{L}$$

but rather

$$\forall x \in \mathbb{L}\ (x \in \mathbb{L})^{\mathbb{L}}$$

which is equivalent to $(\forall x \in \mathbb{L})(\exists \alpha \in \mathbb{ON}\ x \in \mathbf{L}(\alpha))^{\mathbb{L}}$; which is, in turn, since $\mathbb{ON} \subseteq \mathbb{L}$, equivalent to $(\forall x \in \mathbb{L})(\exists \alpha \in \mathbb{ON})(x \in \mathbf{L}(\alpha))^{\mathbb{L}}$.

This latter statement is true since "$x \in \mathbf{L}(\alpha)$" is a $\triangle_0$ formula when written out in full in LOST, and since $\mathbb{L}$ is transitive.

□

For each Axiom Φ of $\mathfrak{ZFC}$ we have:

Theorem 50. Φ

$\mathbb{L} \models \Phi$.

Proof. Transitivity of $\mathbb{L}$ automatically gives Equality, Extensionality, Existence and Foundation. We get Infinity since $\omega \in \mathbb{L}$ and "$z = \mathbb{N}$" is a $\triangle_0$ formula.

For Comprehension, let Φ be any formula of LOST; we wish to prove

$$\forall y \in \mathbb{L}\ \forall w_0 \in \mathbb{L} \ldots \forall w_n \in \mathbb{L}\ \exists z \in \mathbb{L}\ z = \{x \in y : \Phi^{\mathbb{L}}(x, y, w_0, \ldots, w_n)\}$$

since $\mathbb{L}$ is transitive.

Fix $y, w_0, \ldots, w_n$ and $\alpha \in \mathbb{ON}$ such that $\{y, \vec{w}\} \subseteq \mathbf{L}(\alpha)$. By the Levy Reflection Principle, there is some $\beta > \alpha$ such that Φ is absolute between $\mathbb{L}$ and $\mathbf{L}(\beta)$.

By Theorem 44, there is an $n \in \omega$ such that

$$\mathbf{G}(n, \mathbf{L}(\beta), y, \vec{w}) = \{x \in \mathbf{L}(\beta) : \Phi^{\mathbf{L}(\beta)}(x, y, \vec{w})\}.$$

and so by definition, $\{x \in \mathbf{L}(\beta) : \Phi^{\mathbf{L}(\beta)}(x, y, \vec{w})\} \in \mathbf{L}(\beta + 1)$. Now by absoluteness, $\{x \in \mathbf{L}(\beta) : \Phi^{\mathbf{L}(\beta)}(x)\} = \{x \in \mathbf{L}(\beta)\Phi^{\mathbb{L}}(x)\}$. So we have

$$\{x \in \mathbf{L}(\beta) : \Phi^{\mathbb{L}}(x, y, \vec{w})\} \in \mathbf{L}(\beta + 1).$$

Moreover, since $y \in \mathbf{L}(\beta + 1)$,

$$\begin{aligned} \{x \in y : \Phi^{\mathbb{L}}(x, y, \vec{w})\} &= y \cap \{x \in \mathbf{L}(\beta + 1) : \Phi^{\mathbb{L}}(x, y, \vec{w})\} \\ &\in \mathbf{L}(\beta + 2) \end{aligned}$$

and since $\mathbf{L}(\beta + 2) \subseteq \mathbb{L}$ we are done.

For the Power Set Axiom, we must prove that $(\forall x\ \exists z\ z = \{y : y \subseteq x\})^{\mathbb{L}}$. That is, $\forall x \in \mathbb{L}\ \exists z \in \mathbb{L}\ z = \{y : y \in \mathbb{L}$ and $y \subseteq x\}$. Fix $x \in \mathbb{L}$; by the Power Set Axiom and the Axiom of Comprehension we get

$$\exists z'\ z' = \{y \in \mathcal{P}(x) : y \in \mathbb{L} \wedge y \subseteq x\} = \{y : y \in \mathbb{L} \wedge y \subseteq x\}.$$

By the previous lemma $\mathbb{L}$ is almost universal and $z' \subseteq \mathbb{L}$ so

$$\exists z'' \in \mathbb{L}\ z' \subseteq z''.$$

So $z' = z' \cap z'' = \{y \in z'' : y \in \mathbb{L} \cap y \subseteq x\}$. By the fact that the Axiom of Comprehension holds relativised to $\mathbb{L}$ we get

$$(\exists z\ z = \{y \in z'' : y \subseteq x\})^{\mathbb{L}};$$

i.e.,

$$\begin{aligned} \exists z \in \mathbb{L}\ z &= \{y \in z'' : y \in \mathbb{L} \wedge y \subseteq x\} \\ &= \{y : y \in \mathbb{L} \wedge y \subseteq x\} \end{aligned}$$

The Union Axiom and the Replacement Scheme are treated similarly. To prove (the Axiom of Choice) $^{\mathbb{L}}$, we will show that the Axiom of Choice follows from the other axioms of $\mathfrak{ZFC}$ with the additional assumption that $\mathbb{V} = \mathbb{L}$.

It suffices to prove that for each $\alpha \in \mathbb{ON}$ there is a $\beta \in \mathbb{ON}$ and a surjection $f_\alpha \colon b_\alpha \to \mathbf{L}(\alpha)$.

To do this we define f_α recursively. Of course $f_0 = \beta_0 = \emptyset = \mathbf{L}(0)$. If α is a limit ordinal, then we let

$$\beta_\alpha = \sum\{\beta_\epsilon : \epsilon < \alpha\}$$

and $f_\alpha(\sigma) = f_\delta(\tau)$ where $\sigma = \sum\{\beta_\epsilon : \epsilon < \delta\} + \tau$ and $\tau < \beta_\delta$.

If $\alpha = \gamma + 1$ is a successor ordinal, use f_γ to generate a well ordering of $\mathbf{L}(\gamma)$ and use this well ordering to generate a lexicographic well ordering of $\bigcup\{^k(\mathbf{L}(\gamma)) : k \in \omega\}$ and use this to obtain an ordinal $\bar{\beta_\alpha}$ and a surjection

$$\bar{f_\gamma} \colon \bar{\beta_\gamma} \to \bigcup\{^k(\mathbf{L}(\gamma)) : k \in \omega\}.$$

Now let $\beta_\alpha = \beta_{\gamma+1} = \bar{\beta_\gamma} \times \omega$ and let

$$f_\alpha \colon \beta_\alpha \to \mathbf{L}(\alpha) = \{\mathbf{G}(n, \mathbf{L}(\alpha), \vec{y}) : n \in \omega \text{ and } \exists k \in \omega\ \vec{y} \in^k \mathbf{L}(\gamma)\}$$

be defined by $f_\alpha(\sigma) = \mathbf{G}(n, \mathbf{L}(\gamma), \bar{f_\gamma}(\tau))$ where $\sigma = \bar{\beta_\gamma} \times n + \tau$, $\tau < \bar{\beta_\gamma}$.

This completes the proof of Theorem 50 Φ and motivates calling "$\mathbb{V} = \mathbb{L}$" the <u>Axiom of Constructibility</u>.

□

Remark. $\mathbb{V} = \mathbb{L}$ is consistent with $\mathfrak{ZFC}$ in the sense that no finite subcollection of $\mathfrak{ZFC}$ can possibly prove $\mathbb{V} \neq \mathbb{L}$; To see this, suppose

$$\{\Psi_0, \dots, \Psi_n\} \vdash \mathbb{V} \neq \mathbb{L}.$$

Then

$$\Psi_0^{\mathbb{L}}, \dots, \Psi_n^{\mathbb{L}}\} \vdash (\mathbb{V} \neq \mathbb{L})^{\mathbb{L}}.$$

by Theorem 50. This contradicts the preceding lemma.

Remark. Assuming $\mathbb{V} = \mathbb{L}$ we actually can find a formula $\Psi(x, y)$ which gives a well ordering of the universe.

We denote by $\Phi_{\mathbb{L}}$ the conjunction of a finite number of axioms of $\mathfrak{ZFC}$ conjoined with "$\mathbb{V} = \mathbb{L}$" such that $\Phi_{\mathbb{L}}$ implies all our lemmas and theorems about ordinals and ensures that $x \in \mathbf{L}(\alpha)$ is equivalent to some $\triangle_0$ formula (but I think we have already defined it to be $\triangle_0$). In particular, $x \in \mathbb{ON}$ will be equivalent to a $\triangle_0$ formula.

Furthermore, we explicitly want $\Phi_{\mathbb{L}}$ to imply that $\forall\alpha \in \mathbb{ON}\ \exists z\ z = \mathbf{L}(\alpha)$ and that there is no largest ordinal.

We shall use the abbreviation $o(M) = \mathbb{ON} \cap M$.

Lemma. $\forall M(M$ is transitive and $\Phi_{\mathbb{L}}^M \to M = \mathbf{L}(o(M)))$.

Proof. Let M be transitive such that $M \models \Phi_{\mathbb{L}}$. Note that $o(M) \in \mathbb{ON}$. We have $M \models \forall\alpha \in \mathbb{ON}\ \exists z\ z = \mathbf{L}(\alpha)$. So,

$$\begin{aligned}
&\forall\alpha \in o(M)\ M \models \exists z\ z = \mathbf{L}(\alpha)\\
\Rightarrow&\forall\alpha \in o(M)\ \exists z \in M\ M \models z = \mathbf{L}(\alpha)\\
\Rightarrow&\forall\alpha \in o(M)\ \exists z \in M\ z = \mathbf{L}(\alpha)\\
\Rightarrow&\forall\alpha \in o(M)\ \mathbf{L}(\alpha) \in M\\
\Rightarrow&\forall\alpha \in o(M)\ \mathbf{L}(\alpha) \subseteq M.
\end{aligned}$$

Since $M \models \Phi_{\mathbb{L}}$, $o(M)$ is a limit ordinal and hence

$$\mathbf{L}(o(M)) = \bigcup\{\mathbf{L}(\alpha) : \alpha \in o(M)\} \subseteq M.$$

Now let $a \in M$. Since $M \models \mathbb{V} = \mathbb{L}$ we have

$$\begin{aligned}
&M \models \forall x\ \exists y \in \mathbb{ON}\ x \in \mathbf{L}(y)\\
\Rightarrow&M \models \exists y \in \mathbb{ON}\ a \in \mathbf{L}(y)\\
\Rightarrow&\exists\alpha \in o(M)\ M \models a \in \mathbf{L}(\alpha)\\
\Rightarrow&\exists\alpha \in o(M)\ a \in \mathbf{L}(\alpha)\\
\Rightarrow&a \in \mathbf{L}(o(M)).
\end{aligned}$$

$\square$

Lemma. χ_C

If $\mathbb{ON} \subseteq C$, C is transitive, and $\Phi_{\mathbb{L}}^C$, then $C = \mathbb{L}$.

Proof. The proof is similar to that of the previous lemma.

□

Theorem 51. *(K. Gödel)*

If $\mathbb{V} = \mathbb{L}$ *then* GCH *holds.*

Proof. We first prove the following:
Claim. $\forall \alpha \in \mathbb{ON}\ \mathcal{P}(\mathbf{L}(\alpha)) \subseteq \mathbf{L}(\alpha^+)$.
Proof of Claim. This is easy for finite α, since $\mathbf{L}(n) = R(n)$ for each $n \in \omega$.

Let's prove the claim for infinite $\alpha \in \mathbb{ON}$. Let $X \in \mathcal{P}(\mathbf{L}(\alpha))$; we will show that $X \in \mathbf{L}(\alpha^+)$.

Let $A = \mathbf{L}(\alpha) \cup \{X\}$. A is transitive and $|A| = |\alpha|$.

By the Levy Reflection Principle, there is a $\beta \in \mathbb{ON}$ such that both $A \subseteq \mathbf{L}(\beta)$ and $\mathbf{L}(\beta) \models \Phi_{\mathbb{L}}$, where $\Phi_{\mathbb{L}}$ is the formula introduced earlier.

Now use the Lowenheim-Skolem-Tarski Theorem to obtain an elementary submodel $K \prec \mathbf{L}(\beta)$ such that $A \subseteq K$ and $|K| = |A| = |\alpha|$ so by elementarily we have $K \models \Phi_{\mathbb{L}}$.

Now use the Mostowski Collapsing Theorem to get a transitive M such that $K \cong M$. Since A is transitive, the isomorphism is the indentity on A and hence $A \subseteq M$. We also get $M \models \Phi_{\mathbb{L}}$ and $|M| = |\alpha|$.

Now we use the penultimate lemma to infer that $M = \mathbf{L}(o(M))$. Since $|M| = |\alpha|$ we have $|o(M)| = |\alpha|$ so that $o(M) < |\alpha^+|$.

Hence $A \subseteq M = \mathbf{L}(o(M)) \subseteq \mathbf{L}(\alpha^+)$, so that $X \in \mathbf{L}(\alpha^+)$.

We now see that the GCH follows from the claim. For each cardinal κ we have $\kappa \subseteq \mathbf{L}(\kappa)$ so that $|\mathcal{P}(\kappa)| \leq |\mathcal{P}(\mathbf{L}(\kappa))| \leq |\mathbf{L}(\kappa^+)$.

Since $|\mathbf{L}(\kappa^+)| = \kappa^+$ we have $|P(\kappa)| = \kappa^+$.

□

We now turn our attention to whether $\mathbb{V} = \mathbb{L}$ is true.

Let μ be a cardinal and let $\mathcal{U}$ be an ultrafilter over μ. Recalling that ${}^{\mu}\mathbb{V} = \{f : f\colon \mu \to \mathbb{V}\}$, let $\sim_{\mathcal{U}}$ be a binary relation on ${}^{\mu}\mathbb{V}$ defined by

$$f \sim_{\mathcal{U}} g \text{ iff } \{\alpha \in \mu : f(\alpha) = g(\alpha)\} \in \mathcal{U}.$$

It is easy to check that $\sim_{\mathcal{U}}$ is an equivalence relation.

For each $f \in {}^{\mu}\mathbb{V}$ let $\rho(f)$ be the least element of

$$\{\alpha \in \mathbb{ON} : rank(g) = \alpha \wedge f \sim_{\mathcal{U}} g\}.$$

Let $[f] = \{g \in \mathbf{R}(\rho(f) + 1) : g \sim_{\mathcal{U}} f\}$ and let $ULT_{\mathcal{U}}\ \mathbb{V} = \{[f] : f \in {}^{\mu}\mathbb{V}\}$.

Define a relation $\in_{\mathcal{U}}$ on $ULT_{\mathcal{U}}\ \mathbb{V}$ by

$$[f] \in_{\mathcal{U}} [g] \text{ iff } \{\alpha \in \mu : f(\alpha) \in g(\alpha)\} \in \mathcal{U}.$$

It is easy to check that $\in_{\mathcal{U}}$ is well defined.

For each cardinal κ, we use the abbreviation

$$[X]^{<\kappa} = \{Y \subseteq X : |Y| < \kappa\}.$$

Given an uncountable cardinal κ, an ultrafilter $\mathcal{U}$ is said to be κ–complete if $\forall X \in [\mathcal{U}]^{<\kappa}\ \bigcap X \in \mathcal{U}$.

An uncountable cardinal κ is said to be measurable whenever there exists a κ–complete free ultrafilter over κ.

Lemma. If $\mathcal{U}$ is a countably complete ultrafilter (in particular if $\mathcal{U}$ is a μ–complete ultrafilter) then $\in_{\mathcal{U}}$ is set-like, extensional and well founded.

Proof. To see that $\in_{\mathcal{U}}$ is set-like, just note that

$$\{[g] : [g] \in_{\mathcal{U}} [f]\} \subseteq \mathbf{R}(\rho(f) + 2).$$

For extentionality, suppose $[f] \neq [g]$; i.e., $\{\alpha \in \mu : f(\alpha) = g(\alpha)\} \notin \mathcal{U}$. Then either $\{\alpha \in \mu : \neg f(\alpha) \subseteq g(\alpha)\} \in \mathcal{U}$ or $\{\alpha \in \mu : \neg g(\alpha) \subseteq f(\alpha)\} \in \mathcal{U}$. This leads to two similar cases; we address the first.

Pick any $h \in {}^{\mu}\mathbb{V}$ such that $h(\alpha) \in f(\alpha) \setminus g(\alpha)$ whenever $\neg f(\alpha) \subseteq g(\alpha)$. Then $[h] \in_{\mathcal{U}} [f]$ and $[h] \in_{\mathcal{U}} [g]$.

To see that $\in_{\mathcal{U}}$ is well founded, suppose $\exists\{f_n\}_{n\in\omega}$ such that

$$\forall n \in \omega\ [f_{n+1}] \in_{\mathcal{U}} [f_n].$$

Let

$$A = \bigcap\{\{\alpha \in \mu : f_{n+1}(\alpha) \in f_n(\alpha)\} : n \in \omega\} \in \mathcal{U}.$$

$A \in \mathcal{U}$ by the countable completeness of $\mathcal{U}$, so that $A \neq \emptyset$. Pick any $\beta \in A$. Then $F_{n+1}(\beta) \in f_n(\beta)$ for each $n \in \omega$, which is a contradiction.

$\square$

We now create a Mostowski collapse of $ULT_{\mathcal{U}}\ \mathbb{V}$

$$h_{\mathcal{U}} \colon ULT_{\mathcal{U}}\ \mathbb{V} \to M_{\mathcal{U}}$$

given by the recursion

$$h_{\mathcal{U}}([f]) = \{h_{\mathcal{U}}([g]) : [g] \in_{\mathcal{U}} [f]\}$$

As per the Mostowski Theorem, h is an isomorphism and $M_{\mathcal{U}}$ is transitive.

The natural embedding $i_{\mathcal{U}} \colon \mathbb{V} \to ULT_{\mathcal{U}}\ \mathbb{V}$ is given by $i_{\mathcal{U}}(x) = [f_x]$ where $f_x \colon \mu \to \mathbb{V}$ such that $f_x(\alpha) = x$ for all $\alpha \in \mu$.

This natural embedding $i_{\mathcal{U}}$ combines with the unique isomorphism $h_{\mathcal{U}}$ to give

$$j_{\mathcal{U}} \colon \mathbb{V} \to M_{\mathcal{U}}$$

given by $j_{\mathcal{U}}(x) = h_{\mathcal{U}}(i_{\mathcal{U}}(x))$.

$j_{\mathcal{U}}$ is called the elementary embedding generated by $\mathcal{U}$, since for all formulas $\Phi(v_0, \ldots, v_n)$ of LOST we have:

Lemma. $\forall v_0 \ldots \forall v_n\ \Phi(v_0, \ldots, v_n) \leftrightarrow \Phi^{M_{\mathcal{U}}}(j_{\mathcal{U}}(v_0), \ldots, j_{\mathcal{U}}(v_n))$.

Proof. This follows from two claims, each proved by induction on the complexity of Φ.

Claim 1. $\forall v_0 \ldots \forall v_n \; \Phi(v_0, \ldots, v_n) \leftrightarrow \bar{\Phi}(i_{\mathcal{U}}(v_0), \ldots, i_{\mathcal{U}}(v_n))$.
Claim 2. $\forall v_0 \ldots \forall v_n \; \bar{\Phi}(i_{\mathcal{U}}(v_0), \ldots, i_{\mathcal{U}}(v_n)) \leftrightarrow \Phi^{M_{\mathcal{U}}}(j_{\mathcal{U}}(v_0), \ldots, j_{\mathcal{U}}(v_n))$, where $\bar{\Phi}$ is Φ with $\in$ replaced by $\in_{\mathcal{U}}$ and all quantifiers restricted to $ULT_{\mathcal{U}}\ \mathbb{V}$.

We leave the proofs to the reader.

$\square$

Theorem 52. *Every measurable cardinal is inaccessible.*

Proof. We first prove that κ is regular. If $cf(\kappa) = \lambda < \kappa$, then κ is the union of λ sets each smaller than κ. This contradicts the existence of a $\kappa-$complete free ultrafilter over κ.

We now prove that if $\lambda < \kappa$, then $|\mathcal{P}(\lambda)| < \kappa$. Suppose not; then there is $X \in [\mathcal{P}(\lambda)]^\kappa$ and a $\kappa-$complete free ultrafilter $\mathcal{U}$ over X. Now, for each $\alpha \in \lambda$ let $A_\alpha = \{x \in X : \alpha \in x\}$ and $B_\alpha = \{x \in X : \alpha \notin x\}$. Let $I = \{\alpha \in \lambda : A_\alpha \in \mathcal{U}\}$ and $J = \{\alpha \in \lambda : B_\alpha \in \mathcal{U}\}$. Since $\mathcal{U}$ is an ultrafilter, $I \cup J = \lambda$. Since $\mathcal{U}$ is $\kappa-$complete and $\lambda < \kappa$ we have

$$\bigcap \{A_\alpha : \alpha \in I\} \cap \bigcap \{B_\alpha : \alpha \in J\} \in \mathcal{U}.$$

But this intersection is equal to $X \cup \{I\}$, which is either empty or a singleton, contradicting that $\mathcal{U}$ is a free filter.

$\square$

Lemma. Let $\mathcal{U}$ be a $\mu-$complete ultrafilter over an measurable cardinal μ. Let $M = M_{\mathcal{U}}$, $h = h_{\mathcal{U}}$, $i = i_{\mathcal{U}}$ and $j = j_{\mathcal{U}}$ as above. Then for each $\beta \in \mathbb{ON}$ we have $j(\beta) \in \mathbb{ON}$ and $j(\beta) \geq \beta$. Furthermore, if $\beta < \mu$ then $j(\beta) = \beta$ and $j(\mu) > \mu$.

Proof. For each $\beta \in \mathbb{ON}$ we get, by the elementary embedding property of j, that $M \models j(\beta) \in \mathbb{ON}$; since M is transitive, $j(\beta) \in \mathbb{ON}$.

Let β be the least ordinal such that $j(\beta) \in \beta$. Then $M \models j(j(\beta)) \in j(\beta)$ by elementarity, and $j(j(\beta)) \in j(\beta)$ by transitivity of M. This contradicts the minimality of β.

Now let's prove that $j(\beta) = \beta$ for all $\beta < \mu$ by induction on β. Suppose that $j(\gamma) = \gamma$ for all $\gamma < \beta < \mu$. We have

$$\begin{aligned}
j(\beta) &= h(i(\beta)) \\
&= \{h([g]) : [g] \in_{\mathcal{U}} i(\beta)\} \\
&= \{h([g]) : [g] \in_{\mathcal{U}} [f_\beta]\} \text{ where } f_\beta(\alpha) = \beta \text{ for all } \alpha \in \mu \\
&= \{h([g]) : \{\alpha \in \mu : g(\alpha) \in f_\beta(\alpha)\} \in \mathcal{U}\} \\
&= \{h([g]) : \{\alpha \in \mu : g(\alpha) \in \beta\} \in \mathcal{U}\} \\
&= \{h([g]) : \exists \gamma \in \beta \ \{\alpha \in \mu : g(\alpha) = \gamma\} \in \mathcal{U}\} \text{ by } \mu - \text{completeness of } \mathcal{U} \\
&= \{h([g]) : \exists \gamma \in \beta \ [g] = [f_\gamma]\} \text{ where } f_\gamma(\alpha) = \gamma \text{ for all } \alpha \in \mu \\
&= \{h([f_\gamma]) : \gamma \in \beta\} \\
&= \{h(i(\gamma)) : \gamma \in \beta\} \\
&= \{j(\gamma) : \gamma \in \beta\} \\
&= \{\gamma : \gamma \in \beta\} \text{ by inductive hypothesis}
\end{aligned}$$

Hence $j(\beta) = \beta$.

We now show that $j(\mu) > \mu$. Let $g\colon \mu \to \mathbb{ON}$ such that $g(\alpha) = \alpha$ for each α. We will show that $\beta \in h([g])$ for each $\beta \in \mu$ and that $h([g]) \in j(\mu)$.

Let $\beta \in \mu$.

$$\begin{aligned}
\{\alpha \in \mu : f_\beta(\alpha) \in g(\alpha)\} &= \{\alpha \in \mu : \beta \in \alpha\} \\
&= \mu \setminus (\beta + 1) \\
&\in \mathcal{U}
\end{aligned}$$

Hence $[f_\beta] \in_{\mathcal{U}} [g]$ and so $h([f_\beta]) \in h([g])$. But since $\beta \in \mu$,

$$\begin{aligned}
\beta &= j(\beta) \\
&= h(i(\beta)) \\
&= h([f(\beta)])
\end{aligned}$$

Hence $\beta \in h([g])$.

Now, $\{\alpha \in \mu : g(\alpha) \in f_\mu(\alpha)\} = \{\alpha \in \mu : \alpha \in \mu\} = \mu \in \mathcal{U}$. Hence $[g] \in_{\mathcal{U}} [f_\mu]$ and so $h[g] \in h([f_\mu]) = h(i(\mu)) = j(\mu)$.

$\square$

Theorem 53. *(D. Scott)*

If $\mathbb{V} = \mathbb{L}$ *then there are no measurable cardinals.*

Proof. Assume that $\mathbb{V} = \mathbb{L}$ and that μ is the least measurable cardinal; we derive a contradiction. Let $\mathcal{U}$ be a $\mu-$complete ultrafilter over μ and consider $j = j_{\mathcal{U}}$ and $M = M_{\mathcal{U}}$ as above.

Since $\mathbb{V} = \mathbb{L}$ we have $\Phi_{\mathbb{L}}$ and by elementarity of j we have $\Phi_{\mathbb{L}}^{M}$. Note that $\Phi_{\mathbb{L}}$ is a <u>sentence</u>; i.e., it has no free variables.

Since M is transitive, $\mathbb{ON} \subseteq M$ by the previous lemma. So, by an earlier lemma $M = \mathbb{L}$. So we have

$$\mathbb{L} = \mathbb{V} \models (\mu \text{ is the least measurable cardinal})$$

and

$$\mathbb{L} = M \models (j(\mu) \text{ is the least measurable cardinal}).$$

Thus $\mathbb{L} \models j(\mu) = \mu$; i.e., $j(\mu) = \mu$, contradicting the previous theorem.

□

Remark. We have demonstrated the existence of an elementary embedding $j\colon \mathbb{V} \to M$. K. Kunen has shown that there is no elementary $j\colon \mathbb{V} \to \mathbb{V}$.

Large cardinal axioms are often formulated as embedding axioms. For example, κ is said to be <u>supercompact</u> whenever

$$\forall\lambda\ \exists j\ [j\colon \mathbb{V} \to M \text{ and } j(\kappa) > \lambda \text{ and } j|_{\mathbf{R}(\lambda)} = id|_{\mathbf{R}(\lambda)} \text{ and } {}^{\lambda}M \subseteq M].$$

Chapter 13

Appendices

.1 The Axioms of ZFC

Zermelo-Frankel (with Choice) Set Theory, abbreviated to $\mathfrak{ZFC}$, is constituted by the following axioms.

1. **Axiom of Equality**

$$\forall x \; \forall y \; [x = y \rightarrow \forall z \; (x \in z \leftrightarrow y \in z)]$$

2. **Axiom of Extensionality**

$$\forall x \; \forall y \; [x = y \leftrightarrow \forall u \; (u \in x \leftrightarrow u \in y)]$$

3. **Axiom of Existence**

$$\exists z \; z = \emptyset$$

4. **Axiom of Pairing**

$$\forall x \; \forall y \; \exists z \; z = \{x, y\}$$

5. **Union Axiom**

$$\forall x \; [x \neq \emptyset \rightarrow \exists z \; z = \{w : (\exists y \in x)(w \in y)]$$

6. **Intersection Axiom**

$$\forall x\ [x \neq \emptyset \rightarrow \exists z\ z = \{w : (\forall y \in x)(w \in y)]$$

7. **Axiom of Foundation**

$$\forall x\ [x \neq \emptyset \rightarrow (\exists y \in x)(x \cap y = \emptyset)]$$

8. **Replacement Axiom Scheme**
   For each formula $\Phi(x, u, v, w_1, \dots, w_n)$ of the language of set theory,

$$\forall w_1 \dots \forall w_n\ \forall x\ [\forall u \in x\ \exists! v\ \Phi \rightarrow \exists z\ z = \{v : \exists u \in x\ \Phi\}]$$

9. **Axiom of Choice**

$$\forall X\ [(\forall x \in X\ \forall y \in X\ (x = y \leftrightarrow x \cap y \neq \emptyset)) \rightarrow \exists z\ (\forall x \in X\ \exists! y\ y \in x \cap z)]$$

10. **Power Set Axiom**

$$\forall x\ \exists z\ z = \{y : y \subseteq x\}$$

11. **Axiom of Infinity**

$$\mathbb{N} \neq \mathbb{ON}$$

.2 Tentative Axioms

Here is a summary of potential axioms which we have discussed but which lie outside of $\mathcal{ZFC}$.

1. **Axiom of Inaccessibles**

$$\exists \kappa\ \kappa > \omega \text{ and } \kappa \text{ is an inaccessible cardinal}$$

2. **Continuum Hypothesis**

$$|\mathcal{P}(\omega)| = \omega_1$$

3. **Generalised Continuum Hypothesis**

$$\forall\kappa\ [\kappa \text{ is a cardinal} \rightarrow |\mathcal{P}(\kappa)| = \kappa^+$$

4. **Suslin Hypothesis**

   Suppose that R is a complete dense linear order without endpoints in which every collection of disjoint intervals is countable.

   Then $R \cong \mathbb{R}$.

5. **Axiom of Constructibility**

$$\mathbb{V} = \mathbb{L}$$