AN INTRODUCTION TO SET THEORY

Professor William A. R. Weiss

November 21, 2014

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 a major revision in February 2005. The manuscript has seen many changes since then, often due to generous comments by students, each of whom I here thank. Chapters 1 to 11 are now close to final form. Chapters 12 and 13 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 all 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 elementary undergraduate course.

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

They won't appear on an assignment, however, because they are quite difficult to prove. Statement (1) is true; it is called the Schröder-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 (2) 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 whether 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. Each of 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 more difficult and more interesting. It seems that complicated conceptual issues arise in Set Theory more than any other area of pure mathematics; in particular, Mathematical Logic must be used in a fundamental way. Although all the necessary material from Logic is presented in this book, it would be beneficial for the reader to already have had an introduction to Logic under the auspices of Mathematics, Computer Science or Philosophy. In fact, this would be beneficial for everyone, but most people seem to make their way in the world without it and I do not require it of the reader.

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 the Berry Paradox, attributed to G. G. Berry, an English librarian, by B. Russell in 1908. 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, 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 natural languages like English to describe sets. Our first task will be to build a new language for describing sets, one in which such contradictions do not arise.

We also need to clarify exactly what is meant by "set". What is a set? In truth, we do not know the complete answer to this question. Some problems are still unsolved simply because we do not know whether or not certain things constitute a set. Many 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 formal language suitable for describing sets. Those who have already studied logic will find that most of this chapter is quite familiar. Those who have not may find the notation too pedantic for effective mathematical communication. But worry not, we will soon relax the notation. It is much more important to know that statements in Set Theory *can* be precisely written as formulas of this language than to physically write them out. Because the formal Language Of Set Theory is at first quite perplexing, former students gave it the acronym which is the title of this chapter.

Balance is achieved in the next chapter and for a first perusal of this book, the reader may want to skip to it immediately after reading only the first part of this one — up to but not including the section entitled "Substitution for a Variable".

LOST, the language of set theory will consist of symbols and some ways of stringing them together to make formulas. We will need mathematical symbols as well as purely logical symbols. Logical symbols include the conjunction symbol $\wedge$ read as "and", the disjunction symbol $\vee$ read as "or", the negation symbol $\neg$ read as "not", the implication symbol $\rightarrow$ read as "implies", the universal quantifier $\forall$ read as "for all" and the existential quantifier $\exists$ read as "there exists".

Here is a complete list of all the symbols of the language:

variables	$v_0, v_1, v_2, \ldots$
equality symbol	$=$
membership symbol	$\in$
connectives	$\neg, \wedge, \vee, \rightarrow$
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 is a formula and v_i is a variable, then $(\forall v_i)\Phi$ is also a formula.
7. 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 7. That is, each formula of LOST is either atomic or built up from atomic formulas in a sequence of construction steps.

We have the usual logical equivalences which are common to everyday mathematics and can be easily verified using "common sense logic". In particular, for any formulas Φ and Ψ:

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

Other than the supply of variables, there are only a small number of symbols of LOST. However there are quite a number of useful abbreviations that have been introduced. One such is the symbol $\leftrightarrow$. It is a symbol not formally part of LOST; however, whenever we have a formula containing this expression, we can quickly convert it to a proper formula of the language of set theory by replacing

$$(\Phi \leftrightarrow \Psi) \text{ with } ((\Phi \to \Psi) \wedge (\Psi \to \Phi))$$

whenever Φ and Ψ are formulas. Another example of an abbreviation concerns restricted, sometimes called bounded, quantifiers

$$
\begin{array}{lll}
(\exists v_i \in v_j)\Phi & \text{abbreviates} & (\exists v_i)((v_i \in v_j) \wedge \Phi);\ \text{and,} \\
(\forall v_i \in v_j)\Phi & \text{abbreviates} & (\forall v_i)((v_i \in v_j) \to \Phi).
\end{array}
$$

where Φ any formula of LOST. We will introduce many more abbreviations later.

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

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

and we will see many more in the next chapter. It is important to realise that "class" and "set" are two quite different notions. A class is just a description of, or a name for, something in the languge of set theory; it may or may not exist. A set, on the other hand, is a mathematical object of study. A class is a string of symbols — you can see it. You can't see a set with your eyes; you have to use your imagination.

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 are able to incorporate classes into our language by showing how the predicates relate to them. For example, when Ψ is a formula of LOST, we write $(v_k \in \{v_j : \Psi\})$ to stand for the statement that Ψ holds when v_k is substituted for v_j in Ψ. Continuing in this manner, we write $(v_k = \{v_j : \Psi\})$ to mean that

$$(\forall v_i)((v_i \in v_k) \leftrightarrow (v_i \in \{v_j : \Psi\}))$$

when the variable v_i is distinct from the others. In order to make this sufficiently complete and precise, we must carefully specify the notion of substitution. There are, however, surprising technical complications.

If the reader is comfortable with an intuitive grasp of the concept of substitution, then the remainder of this chapter could be skipped for now. However, the material is a prerequisite for fully understanding Chapter 11.

Substitution for a Variable

Variables can occur in a formula in two ways: as parameter variables whose values affect the truth of the formula or otherwise as "dummy" variables. For example, the truth of the formula

$$(\exists v_0)(\forall v_1)((v_1 = v_0) \rightarrow ((v_1 = v_2) \wedge (v_1 = v_3)))$$

depends entirely upon the values taken by the variables v_2 and v_3; in fact, in this example, whether or not they take on equal values. However, the variables v_0 and v_1 do not have this property and it even seems meaningless to speak of their precise values. The variables v_2 and v_3 are free to take on values but v_0 and v_1 are bound up with a quantifier. Let's make this observation precise.

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. Φ is $(\Theta \wedge \Psi)$, 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 $(\forall v_j)\Psi$ and Ψ is a formula and v_i occurs free in Ψ and $i \neq j$; or,
7. Φ is $(\exists v_j)\Psi$ and Ψ is a formula and v_i occurs free in Ψ and $i \neq j$.

Notice that the determination of whether or not a variable occurs free in a formula is reduced to such a determination in slightly simpler formulas. In fact these simpler formulas appear in the construction of the initial formula. We may thus retrace the construction of our formula back to the original atomic formulas before finally being able to declare whether or not the variable occurred free in the initial formula.

Those formulas which appear somewhere in the construction of a formula Φ are called subformulas of Φ. The complete collection of subformulas of a formula Φ is precisely 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 $(\forall v_i)\Psi$ is a subformula of Φ and v_i is a variable, then Ψ is a subformula of Φ; and,

7. If $(\exists v_i)\Psi$ is a subformula of Φ and v_i is a variable, then Ψ is a subformula of Φ.

To say that a variable v_i occurs bound in a formula Φ means that either 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.

A variable can occur both free and bound in a formula as in this example.

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

However, we usually avoid writing such formulas because it is easier to understand formulas in which this double role does not occur.

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).

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 $(\forall v_k)\Theta$ 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,

7. 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.

If the variable v_j does not originally occur in the formula Φ the result of this substitution is a formula Φ^* which is logically equivalent to Φ. However, notice that substituting v_2 for each bound occurrence of v_1 in the formula $(\exists v_1)(\neg(v_1 = v_2))$ gives rise to something quite different.

We can now formally define the important notion of the <u>substitution of a variable</u> 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 procedure 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: $(\exists v_2)(v_2 \in v_2)$ instead of $(\exists v_4)(v_4 \in v_2)$. 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 $(\exists! v_j)$ is another abbreviation in the language of set theory — whenever we have a formula containing this expression we can quickly convert it to a proper formula of LOST.

We can use substitution to express the Equality Principle which states that for any variable v_i of a formula Φ of LOST, all of whose variables lying among $v_0, \ldots, v_n$ and for any variables v_j and v_k we have:

$$(\forall v_0)\ldots(\forall v_i)\ldots(\forall v_n)(\forall v_j)(\forall v_k)((v_j = v_k) \rightarrow (\Phi(v_i|v_j) \leftrightarrow \Phi(v_i|v_k)))$$

Notice that the Equality Principle is not just one formula, but a scheme of formulas, one for each appropriate Φ, v_i, v_j and v_k. These formulas are basic assumptions of virtually any logical system and we shall assume them here as well.

We are now able to formally 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:

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

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 <u>substitution of a term</u> 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 in Φ 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

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

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).

Relativisation of a Formula to a Term

Let t be a term and Φ any formula of the language of set theory.

$$\begin{array}{lll} (\exists v_i \in t)\Phi & \text{abbreviates} & (\exists v_i)((v_i \in t) \wedge \Phi);\ \text{and,} \\ (\forall v_i \in t)\Phi & \text{abbreviates} & (\forall v_i)((v_i \in t) \to \Phi). \end{array}$$

This allows us to define the relativisation of Φ to t, denoted by Φ^t, as follows:

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

Informally, the relativisation of a formula Φ to a term t states that Φ holds under the interpretation for which everything is required to be in t. A formula and its relativisation may express quite different ideas. For example, the formula:

$$(\exists v_1)(v_1 = v_1)$$

is true by logical assumption. Indeed, if it were not true then there would be very little mathematics!

However, the relativisation to the class $\{v_0 : (\neg(v_0 = v_0))\}$ is:

$$(\exists v_1 \in \{v_0 : (\neg(v_0 = v_0))\})(v_1 = v_1)$$

which is an abbreviation for:

$$(\exists v_1)((v_1 \in \{v_0 : (\neg(v_0 = v_0))\}) \wedge (v_1 = v_1))$$

which when written as a proper formula of LOST is:

$$(\exists v_1)((\neg(v_1 = v_1)) \wedge (v_1 = v_1))$$

which is, of course, absolutely false.

On the other hand, the formula $(\forall v_1)(v_1 = v_1)$ is also true and yet the reader can now easily verify that its relativisation to the same class is the formula

$$(\forall v_1)((\neg(v_1 = v_1)) \to (v_1 = v_1)).$$

which is, in fact, logically equivalent to the original formula $(\forall v_1)(v_1 = v_1)$.

Suppose that from the formula Θ we were able to logically infer the formula Φ. This means that there could be no counterexample to the fact that Θ implies Φ. That is, Φ would hold under any interpretation in which Θ would hold. So for any term t, Φ^t would hold whenever Θ^t would hold.

But wait! This contradicts the examples with Θ as $(\forall v_1)((v_1 = v_1)$, Φ as $(\exists v_1)(v_1 = v_1)$ and t as $\{v_0 : (\neg(v_0 = v_0))\}$. What causes this problem?

Relativising to the class $\{v_0 : (\neg(v_0 = v_0))\}$ is really a very special case because there is no value of v_0 which satisfies the condition to be in the class — the class is empty. As such, it cannot satisfy the logically true formula $(\exists v_0)(v_0 = v_0)$. However, the class $\{v_0 : (\neg(v_0 = v_0))\}$ is, in fact, the only one with this odd property.

Indeed, if t is any term such that $(\exists v_0)(v_0 \in t)$ then Ψ^t holds whenever Ψ is a basic assumption of pure logic, including, for example, any instance of the Equality Principle. Furthermore, if we are able to infer Φ from Θ by purely logical means, then Φ^t holds whenever Θ^t holds.

Those familiar with Mathematical Logic will notice that the discussion above can be made precise by specifying a formal deduction system and invoking the Soundness Theorem of Mathematical Logic which essentially says that any formula which is the product of a logical deduction has no interpretation in which it is false. Trusting that the conclusion of the previous paragraph is intuitively satisfactory, we will not need to do this.

With this in mind traditional (informal) mathematical proofs will do as well for Set Theory as for other areas of mathematics. But in order to avoid paradoxes it is crucial that all of our mathematical assumptions and all of our theorems are able to be stated in LOST.

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, yet still avoiding inconsistencies associated with Berry's paradox.

In the previous chapter we defined a term to be either a variable $v_0, v_1, v_2, \ldots$ or a class, which is something of the form $\{v_j : \Phi\}$ where Φ is a formula of the language of set theory, LOST. In order to avoid inconsistencies we need to ensure that the formula Φ in the class $\{v_j : \Phi\}$ is indeed a proper formula of LOST — or, at least, can be converted to a proper formula once 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.

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. We can sometimes use "$\Rightarrow$" and "$\Leftrightarrow$" instead of "$\rightarrow$" and "$\leftrightarrow$", respectively.

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 $t(w_1, \ldots, w_k)$ to indicate that all of the variables of the term t lie among $w_1, \ldots, w_k$. As well, we will use the notation $\Phi(x, y, w_1, \ldots, w_k)$ to indicate that all variables of the formula Φ lie among $x, y, w_1, \ldots, w_k$.

5. 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; that is, we don't substitute for those occurences where y is under the scope of a quantifier. In Chapter 1 this was denoted by $\Phi(y|t)$.

6. 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 s and t, we make the following abbreviations of formulas.

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

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:

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

We also have abbreviations for certain special classes; we denote the class $\{x : x = x\}$ by $\mathbb{V}$ and call it the universe. We denote the class $\{x : x \neq x\}$ by $\emptyset$ and call it the empty set. This terminology comes from the fact that $(x = x)$ holds for all x and, as such, $(\neg(x = x))$ holds for no x.

Whenever t is a term and Φ is a formula of LOST, $\{x \in t : \Phi\}$ will abbreviate $\{x : (x \in t) \wedge \Phi\}$. Furthermore, for a class term $t(w_1, \ldots, w_k)$, $\{t : \Phi\}$ will denote $\{x : \exists w_1, \ldots, \exists w_k[(x = t) \wedge \Phi]\}$.

We make the standard abbreviations for these often-used classes.

Union	$s \cup t$	for	$\{x : x \in s \vee x \in t\}$
Intersection	$s \cap t$	for	$\{x : x \in s \wedge x \in t\}$
Difference	$s \setminus t$	for	$\{x : x \in s \wedge x \notin t\}$
Symmetric Difference	$s \triangle t$	for	$(s \setminus t) \cup (t \setminus s)$
Ordered Pair	$\langle s, t \rangle$	for	$\{\{s\}, \{s, t\}\}$
Cartesian Product	$s \times t$	for	$\{\langle x, y \rangle : x \in s \wedge y \in t\}$
Domain	$dom(f)$	for	$\{x : \exists y \ \langle x, y \rangle \in f\}$
Range	$rng(f)$	for	$\{y : \exists x \ \langle x, y \rangle \in f\}$
Image	$f^{\rightarrow} A$	for	$\{y : \exists x \in A \ \langle x, y \rangle \in f\}$
Inverse Image	$f^{\leftarrow} B$	for	$\{x : \exists y \in B \ \langle x, y \rangle \in f\}$
Restriction	$f \mid A$	for	$\{\langle x, y \rangle : \langle x, y \rangle \in f \wedge x \in A\}$
Inverse	f^{-1}	for	$\{\langle y, x \rangle : \langle x, y \rangle \in f\}$
Composition	$g \circ f$	for	$\{\langle x, z \rangle : \exists y \ \langle x, y \rangle \in f \wedge \langle y, z \rangle \in g\}$

These latter abbreviations are most often, but not always, used in the context of functions. We say f is a <u>function</u> provided

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

We write $f \colon X \rightarrow Y$ for

$$f \text{ is a function and } dom(f) = X \text{ and } rng(f) \subseteq Y.$$

The class $\{f : (f \colon X \rightarrow Y)\}$ is denoted by ${}^{X}Y$ or alternatively by Y^X.

We also write:

f is one to one	for	$\forall y \in rng(f) \ \exists! x \ \langle x, y \rangle \in f$
f is onto Y	for	$Y = rng(f)$

and use the common terms <u>injection</u> for a one to one function, <u>surjection</u> for an onto function and <u>bijection</u> for both properties together.

There are several ways other authors denote the image and the inverse image; $f^{\rightarrow}A$ is sometimes written as $f''A$. However, notation like $f(A)$ can become confusing whenever, as frequently happens in Set Theory, the set A is both an element of and a subset of the domain of the function f. Of course, we all prefer notation which is straightforward and unambiguous.

The famous statement known as Russell's Paradox is the following 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, but they both 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 exist. 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 issue — 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 anything 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

$$\forall x\ x \notin x$$

which is not an unreasonable assumption by any means, but a mathematical assumption none-the-less. The existence of the empty set $\emptyset$ 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 obvious that require 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 because all of the rest of mathematics is based on them. They are the foundation upon which everything else is built. These assumptions are called axioms and the system is called the $\mathcal{ZFC}$ Axiom System. We shall begin to study it in the next chapter.

Chapter 3

The Axioms of Set Theory

We will explore the $\mathcal{ZFC}$ 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 Existence	$\exists z\ z = z$
The Axiom of Extensionality	$\forall x\ \forall y\ [\forall u\ (u \in x \ \leftrightarrow\ u \in y) \leftrightarrow\ x = y]$
The Axiom of Pairing	$\forall x\ \forall y\ \exists z\ z = \{x, y\}$

Different authors give slightly different formulations of the $\mathcal{ZFC}$ axioms but these formulations are all equivalent. Here, the Axiom of Existence is only stated for emphasis. It is unnecessary since it is an axiom of pure logic and we already implicitly assume all such logical axioms such as $\forall z\ z = z$ and the Equality Principle from Chapter 1.

As well, the Axiom of Pairing will follow from other axioms to be stated later. So there is considerable redundancy in this system, but redundancy is not always a bad thing.

We now assert the existence of unions and intersections. No doubt the reader has experienced a symmetry between these two concepts. However, while the Union Axiom is used extensively, the Intersection Axiom is omitted in many developments of the subject because it follows from the rest of the $\mathfrak{ZFC}$ axioms. We include it here because it adds educational value; see Theorem 1 and the remarks after Theorem 4.

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

For any term t the class $\{w : (\exists y \in t)(w \in y)\}$ is abbreviated as $\bigcup t$ and called the "big union of t".

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

For any term t the class $\{w : (\forall y \in t)(w \in y)\}$ is abbreviated as $\bigcap t$ and called the "big intersection of t".

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

This axiom, while it may be "obviously true", is not immediately obvious, so let's investigate what it says.

Suppose, for the sake of argument, that 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 $z_2 \in x \cap z_1$. Since $z_2 \in x$ we would be able to get $z_3 \in x \cap z_2$. 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 since it never occurs in other areas of mathematics. We want our sets to be founded: each such sequence should eventually end. Hence the name of the axiom, which is also known as the Axiom of Regularity. In truth, it must be admitted that an important practical reason for the adoption of this axiom is that it allows us to develop a clear and elegant theory, so perhaps the Axiom of Foundation is best understood by its consequences.

Our first theorem states a number of simple results which we would be quite willing to assume outright, were they not to follow readily from the axioms. Notice that the third statement (with $y = x$) leads to the result

$$\forall x \ x \notin x$$

which is exactly what we needed to extend Russell's Paradox in order to obtain

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

Theorem 1.

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

Exercise 1. Prove Theorem 1.

Our next theorem gives the basic facts about ordered pairs. Let $f(x)$ denote the class $\bigcup\{w : \langle x, w\rangle \in f\}$.

Theorem 2.

1. $\forall x \ \forall y \ \exists z \ z = \langle x, y\rangle$.
2. $\forall u \ \forall v \ \forall x \ \forall y \ [\langle u, v\rangle = \langle x, y\rangle \leftrightarrow (u = x \wedge v = y)]$.
3. *For any function f and $x \in dom(f)$ we have: $\langle x, y\rangle \in f \Leftrightarrow y = f(x)$.*

Exercise 2. Prove this theorem.

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 precise purpose of avoiding such expressions which bring us perilously close to Berry'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 $\Phi(x, u, v, w_1, \ldots, w_k)$ of the language of set theory, we have the axiom:

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

Note that we have allowed Φ to have $w_1, \ldots, w_k$ as parameters, that is, free variables which may be used to specify various objects 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. Heuristically, for a given $t \in y$ we first replace each $u \in x$ with $\langle u, t \rangle$, which is like a "horizontal line" of $x \times y$. Next, we replace each $t \in y$ with $x \times \{t\}$ to obtain the complete collection of all "horizontal lines". The union of this collection is $x \times y$.

More precisely, from Theorem 2, 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:

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

$\square$

The statement $\mathbb{V} \times \mathbb{V} \subseteq \mathbb{V}$ is a true formula of LOST, since $\forall x\ x \in \mathbb{V}$. However, it is not true that $\exists z\ z = \mathbb{V} \times \mathbb{V}$. Why not? Because, if so we could use the first two parts of Theorem 2 and an instance of the Replacement Scheme to obtain

$$\exists z\ z = \{v : (\exists u \in \mathbb{V} \times \mathbb{V})[(\exists w)\ u = \langle v, w\rangle]\} = \mathbb{V}.$$

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 of LOST, we are again led to a scheme.

The Comprehension Scheme

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

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

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

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

Theorem 4. Φ

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

Proof. Fix such a formula $\Phi(x, u, w_1, \ldots, w_k)$. Apply Replacement on the set x with the formula $\Psi(x, u, v, w_1, \ldots, w_k)$ given by:

$$(\Phi \to v = \{u\}) \wedge (\neg\Phi \to v = \emptyset)$$

to obtain:

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

Notice that $\{\{u\} : u \in x \wedge \Phi(x, u, w_1, \ldots, w_n)\} \subseteq y$ and that 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 instances of axioms. A proof cannot have infinite length, nor invoke infinitely many axioms or lemmas.

Notice that the Intersection Axiom can be shown to follow from an instance of the Comprehension Scheme:

$$\exists z \ z = \{u : (u \in \bigcup x) \wedge (\forall y \in x)(u \in y)\} = \bigcap x$$

so that one is now tempted to declare the Intersection Axiom to be redundant. However, in the proof of Comprehension, the existance of the empty set $\emptyset$ was used for the (unwritten) verification of the hypothesis of Replacement. And didn't you use the Intersection Axiom in your proof of the existance of the empty set?

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 exactly 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 21.

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

$$\forall X\ [\emptyset \notin X \rightarrow (\exists f)(f\colon X \rightarrow \bigcup X \text{ and } (\forall x \in X)(f(x) \in x))].$$

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 $f(x) \in x$.

$\square$

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

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 $\mathfrak{ZFC}$ axioms with one exception to come later — Infinity.

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 do not claim that what we construct are the "actual" natural numbers — whatever they are. But since what we shall define will have exactly those *mathematical* properties which the "actual" natural numbers have, we will take the liberty of calling our constructs simply, 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 new concept. We say that a term t is <u>transitive</u> whenever we have

$$(\forall x)(x \in t \to x \subseteq t).$$

The Axiom of Foundation ensures that $\emptyset$ is an element of each non-empty transitive set.

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 1. 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 $n \notin m$, then $m \subseteq n$.

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

□

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 each formula $\Phi(n, w_1, \ldots, w_k)$ of the language of set theory we have:

Theorem 8. Φ

For all $w_1, \ldots, w_k$*, if*

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

then

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

Proof. For brevity, we have suppressed explicit mention of the parameters $w_1, \ldots, w_k$ in the formula Φ and from now on we will frequently do this.

We will assume that the theorem is false and derive a contradiction. Take any fixed $w_1, \ldots, w_k$ and a fixed $l \in \mathbb{N}$ such that $\neg\Phi(l)$. Let t be any transitive subset of $\mathbb{N}$ containing l, e.g. $t = l \cup \{l\}$.

By Comprehension, $\exists s\ s = \{m \in t : \neg\Phi(m)\}$. By Foundation, we get $n \in s$ such that $n \cap s = \emptyset$. Transitivity of t guarantees that $(\forall m \in n)\ \Phi(m)$. This, in turn, contradicts that $n \in s$.

□

The statement $\forall m \in n\ \Phi(m)$ in the above Theorem 8 Φ is usually called the inductive hypothesis. When $n = 0$ the inductive hypothesis is trivially true, so verifying

$$(\forall m \in n\ \Phi(m)) \to \Phi(n)$$

when $n = 0$ just amounts to proving $\Phi(0)$.

Our first application of induction will be to show that ordinary counting makes sense. 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 3. Use induction to prove the pigeon-hole principle: for $n \in \mathbb{N}$ there is no injection $f\colon succ(n) \to n$. Conclude that a set X cannot have two different sizes when counted two different ways.

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 $a = b$.

□

Exercise 4. Prove or disprove that for each formula $\Phi(n)$ if

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

then

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

Hint: prove the Proposition!

The Recursion Principle for the Natural Numbers

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

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

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

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

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 the union of

$$\{f : (\exists n \in \mathbb{N})[f \colon n \to \mathbb{V} \wedge \forall m \in n\ \Phi(m, f|m, f(m))]\}.$$

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

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

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 $\Phi(x, f, y, w_1, \ldots, w_k)$ of the language of set theory, we denote by $REC(\Phi, \mathbb{N}, w_1, \ldots, w_k)$ the class

$$\bigcup\{f : (\exists n \in \mathbb{N})[f \colon n \to \mathbb{V} \wedge \forall m \in n\ \Phi(m, f|m, f(m), w_1, \ldots, w_k)]\}$$

We will show, under appropriate hypotheses, that $REC(\Phi, \mathbb{N}, w_1, \ldots, w_k)$ is a function on $\mathbb{N}$ which satisfies the procedure given by Φ. This requires a theorem scheme.

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

Theorem 9. Φ

For all $w_1, \ldots, w_k$ suppose that we have

$$(\forall x \in \mathbb{N})(\forall f)[(f\colon x \to \mathbb{V}) \to \exists! y\ \Phi(x, f, y, w_1, \ldots, w_k)].$$

Then, letting F denote the class $REC(\Phi, \mathbb{N}, w_1, \ldots, w_k)$, we have:

1. $F\colon \mathbb{N} \to \mathbb{V}$;
2. $\forall m \in \mathbb{N}\ \Phi(m, F|m, F(m), w_1, \ldots, w_k)$.

Proof. For the sake of brevity, we will not always explicitly mention the parameters $w_1, \ldots, w_k$ occurring in the formula Φ. 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. Since $\langle x, y_1\rangle \in F$ we have a function f_1 with domain $n_1 \in \mathbb{N}$ such that $f_1(x) = y_1$ and

$$(\forall m \in n_1)\ \Phi(m, f_1|m, f_1(m)).$$

Similarly, we get f_2 and n_2 with $f_2(x) = y_2$ and

$$(\forall m \in n_2)\ \Phi(m, f_2|m, f_2(m)).$$

Let $n_0 = n_1 \cap n_2$. We have $x \in n_0 \in \mathbb{N}$. It suffices to prove that

$$(\forall m \in n_0)(f_1(m) = f_2(m)),$$

which we do by induction on $m \in \mathbb{N}$ using the inductive hypothesis

$$(\forall j \in m)(j \in n_0 \to f_1(j) = f_2(j))$$

with intent to show that

$$m \in n_0 \to f_1(m) = f_2(m).$$

To do this suppose $m \in n_0 = n_1 \cap n_2$ so that we have both

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

By transitivity $m \subseteq n_0$ 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 on $x \in \mathbb{N}$. To this end, we use the inductive hypothesis

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

with intent to show that $\exists y\ \langle x, y\rangle \in F$.

By the inductive hypothesis, for each $j \in x$ there is $n_j \in \mathbb{N}$ with $j \in n_j$ and a function $f_j \colon n_j \to \mathbb{V}$ such that

$$(\forall m \in n_j)\ \Phi(m, f_j|m, f_j(m)).$$

Let $h = \bigcup\{f_j : j \in x\}$. By the claim the f_j's agree on their common domains, so that h is a function with domain including x as a subset. Furthermore, n_j is transitive for each $j \in x$, so that $dom(h)$ is also transitive Hence each $h|j = f_j|j$ so that

$$(\forall j \in x)\ \Phi(j, h|j, h(j)).$$

By the hypothesis of the theorem applied to $g = h|x$ there is a unique y such that $\Phi(x, g, y)$. Define f to be the function $f = h \cup \{\langle x, y\rangle\}$. 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

$$(\forall m \in n)\ \Phi(m, f|m, f(m)).$$

In fact, $F|n = f$ so that (2) follows immediately. $\square$

Return now to the example. By applying this theorem, we see that $REC(\Phi, \mathbb{N}, w_1, \ldots, w_k)$ 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:

$$\begin{aligned} F(0) &= 3 \\ F(succ(n)) &= succ(F(n)) \end{aligned}$$

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 in the usual way to denote the defined function and give the recursion by the following formulas.

$$\begin{aligned} 3 + 0 &= 3 \\ 3 + succ(n) &= succ(3 + n) \end{aligned}$$

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

$$\begin{aligned} a + 0 &= a \\ a + succ(b) &= succ(a + b) \end{aligned}$$

for each $a \in \mathbb{N}$. Here a is a parameter which is allowed by the inclusion of $w_1, \ldots, w_k$ in our analysis. The same trick can be used for multiplicaton:

$$\begin{aligned} a \cdot 0 &= 0 \\ a \cdot (succ(b)) &= a \cdot b + a \end{aligned}$$

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:

$$\begin{aligned} a^0 &= 1 \\ a^{succ(b)} &= a^b \cdot a \end{aligned}$$

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. For example, for addition the formula $\Phi(x, f, y, a)$ is

$$(x = 0 \to y = a) \wedge (\forall b \in x)(x = succ(b) \to y = succ(f(b))$$

and it is easy to see that

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

The properties of the natural numbers may now be verified by induction on $\mathbb{N}$ in a straightforward manner.

Another example of recursion on $\mathbb{N}$ gives the famous Fibonacci numbers:

$$\begin{aligned} F(0) &= 1 \\ F(1) &= 1 \\ F(n+2) &= F(n) + F(n+1) \end{aligned}$$

and the relevant formula $\Phi(x, f, y)$ is:

$$(x = 0 \vee x = 1 \to y = 1) \wedge (\forall n \in x)[x = succ(succ(n)) \to y = f(n) + f(succ(n))].$$

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)*

Let $\alpha, \beta \in \mathbb{ON}$. Exactly one of three situations occurs:

$$\alpha \in \beta,\ \ \beta \in \alpha,\ \ \alpha = \beta.$$

Proof. The reader may check that a proof of this theorem can be obtained simply 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 $\alpha \in \beta$ as we do with natural numbers. Note that with this notation, trichotomy implies that $\alpha \leq \beta$ if and only if $\alpha \subseteq \beta$.

Since $\mathbb{N} \subseteq \mathbb{ON}$, it is natural to wonder whether $\mathbb{N} = \mathbb{ON}$. As we shall see, the formula "$\mathbb{N} = \mathbb{ON}$" essentially says that there are no infinite sets whereas the formula "$\mathbb{N} \neq \mathbb{ON}$" says that there are indeed infinite sets. Since it seems that we cannot prove either of these statements, 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 "$\mathbb{N} \neq \mathbb{ON}$".

Of course, we go for the infinite!

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

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. $\forall S\ [S \subseteq \mathbb{ON} \to \exists \beta \in \mathbb{ON}\ \beta = \bigcup S]$.

Exercise 5. 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. $\forall S\ [S \subseteq \mathbb{ON} \to \bigcup S = \sup S]$

Exercise 6. 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 7. Prove this lemma.

We are able to carry out induction on the ordinals in a way similar to induction on the natural numbers via a process called transfinite induction. In order to justify transfinite induction we need a theorem scheme.

For each formula $\Phi(n, w_1, \ldots, w_k)$ of the language of set theory we have:

Theorem 13. Φ

For all $w_1, \ldots, w_k$, if

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

then

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

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 8.

□

The Recursion Principle for the Ordinal Numbers

We can also carry out recursive definitions on $\mathbb{ON}$. This process is called transfinite recursion. For any formula $\Phi(x, f, y, w_1, \ldots, w_k)$ of the language of set theory, we denote by $REC(\Phi, \mathbb{ON}, w_1, \ldots, w_k)$ the class

$$\bigcup\{f : (\exists n \in \mathbb{ON})(\exists f)[f : n \to \mathbb{V} \wedge \forall m \in n\ \Phi(m, f|m, f(m))]\}.$$

Transfinite recursion is justified by the following theorem scheme.

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

Theorem 14. Φ

For all $w_1, \ldots, w_k$ suppose that we have

$$(\forall x \in \mathbb{ON})(\forall f)[(f \colon x \to \mathbb{V}) \to \exists! y\ \Phi(x, f, y, w_1, \ldots, w_k)].$$

Then, letting F denote the class $REC(\Phi, \mathbb{ON}, w_1, \ldots, w_k)$, we have:

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

2. $\forall m \in \mathbb{ON}\ \Phi(m, F|m, F(m), w_1, \ldots, w_k)$.

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. $\square$

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

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

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

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 8. Build your transfinite induction skills by proving two parts of this theorem. Be prepared to use this theorem repeatedly as a lemma for future exercises.

The extension of arithmetic operations from $\mathbb{N}$ to $\mathbb{ON}$ gives something new and different. 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. $(2 \cdot 2)^\omega \neq 2^\omega \cdot 2^\omega$

We do have a form of subtraction, as in this subtraction lemma.

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

Exercise 9. Prove this lemma and show that every ordinal can be written uniquely as $\delta + n$ where δ is a limit ordinal and $n \in \omega$.

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

Exercise 10. 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 11. 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.)

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 $f_n = f(n)$.

Lemma. There is no infinite decreasing sequence of ordinals.

Proof. Let's use an indirect proof. Suppose $x \subseteq \omega$ is infinite and $f: x \to \mathbb{ON}$ such that if $n < m$ then $f(n) > f(m)$. Let $X = \{f(n) : n \in x\}$. By the Axiom of 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 16. *(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

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

Proof. Using the penultimate lemma, let

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

and then let

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

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 subtraction lemma, there is some $\alpha_0 \in \mathbb{ON}$ such that

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

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

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

and $\alpha_1 < \omega^{\beta_1}$ such that $\alpha_0 = \omega^{\beta_1} \cdot m_1 + \alpha_1$. 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, contradicting the previous lemma. The only way we could stop would be if some $\alpha_n = 0$. This proves the existence of the sum. □

Exercise 12. Prove that whenever $k \in \omega$, and $m_0, \dots, m_k < \omega$, and $\alpha_0, \dots, \alpha_k < \beta$, we have

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

and use this to finish the proof of Theorem 16 by showing that the sum is unique.

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.

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

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

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

Subtract 1 and continue, changing 5's to 6's, subtracting 1, changing 6's to 7's and so on, obtaining:

$$\begin{aligned}
x_6 &= 6^{(6^6+1)} + 6^{6^6} + 6^3 \cdot 3 + 6^2 \cdot 3 + 6 \cdot 3 + 2, \\
x_7 &= 7^{(7^7+1)} + 7^{7^7} + 7^3 \cdot 3 + 7^2 \cdot 3 + 7 \cdot 3 + 1, \\
x_8 &= 8^{(8^8+1)} + 8^{8^8} + 8^3 \cdot 3 + 8^2 \cdot 3 + 8 \cdot 3, \\
x_9 &= 9^{(9^9+1)} + 9^{9^9} + 9^3 \cdot 3 + 9^2 \cdot 3 + 9 \cdot 2 + 7,
\end{aligned}$$

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 17. *(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:

$$\begin{aligned}
g_2 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^\omega + \omega, \\
g_3 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^\omega + 1, \\
g_4 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^\omega, \\
g_5 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^3 \cdot 3 + \omega^2 \cdot 3 + \omega \cdot 3 + 3, \\
g_6 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^3 \cdot 3 + \omega^2 \cdot 3 + \omega \cdot 3 + 2, \\
g_7 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^3 \cdot 3 + \omega^2 \cdot 3 + \omega \cdot 3 + 1, \\
g_8 &= \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^3 \cdot 3 + \omega^2 \cdot 3 + \omega \cdot 3,
\end{aligned}$$

$$g_9 = \omega^{(\omega^\omega+1)} + \omega^{(\omega^\omega)} + \omega^3 \cdot 3 + \omega^2 \cdot 3 + \omega \cdot 2 + 7,$$

and so on. Now use the inequality of Exercise 12 to show that $\{g_n\}$ would be an infinite decreasing sequence of ordinals contradicting the previous lemma.

□

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 $x = 4$; doing it for $x = 5$ would be tedious. Finishing our example $x = 54$ would be humanly impossible.

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

We say R is a <u>relation</u> whenever

$$\forall p \in R \; \exists x \; \exists y \; p = \langle x, y \rangle$$

and we say that R is a relation on X provided $R \subseteq X \times X$.

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

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

Another example is the inclusion relation:

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

Let us set out some terminology for properties of relations. In the following definitions, R and X are terms.

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

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

4. We say a relation R is symmetric on X whenever
$$\forall x \in X\ \forall y \in X\ (\langle x, y\rangle \in R \to \langle y, x\rangle \in R).$$

5. We say a relation R is well founded on X whenever
$$\forall Y\ [(Y \subseteq X \wedge Y \neq \emptyset) \to (\exists y \in Y\ \forall x \in Y\ \langle x, y\rangle \notin R)].$$
Such an y is called minimal for Y.

6. We say a relation R is total on X whenever
$$\forall x \in X\ \forall y \in X\ [\langle x, y\rangle \in R \vee \langle y, x\rangle \in R \vee x = y].$$

7. We say R is extensional on X whenever
$$\forall x \in X\ \forall y \in X\ [x = y \leftrightarrow \forall z \in X\ (\langle z, x\rangle \in R \leftrightarrow \langle z, y\rangle \in R)].$$

Notice the use of the word *transitive* again; this time for *transitive relation* rather than *transitive set*. The terminology is unfortunate — the membership relation is not necessarily transitive on a transitive set; for example:
$$\{0, 1, \{1\}\}$$
However the Axiom of Extensionality ensures that the membership relation is extensional on any transitive set and the Axiom of Foundation says that the membership relation is well founded on any set whatsoever.

Exercise 13. Prove that any relation which is both well founded and total is also irreflexive, extensional and transitive.

There is an interesting and useful characterisation of well founded relations, giving an ordinal number "rank" to each element of a well founded set.

Theorem 18. *For a relation R on a set X the following are equivalent:*

1. *R is well founded on X.*

2. *There a function $f : X \to \mathbb{ON}$ with $f(x) < f(y)$ whenever $\langle x, y\rangle \in R$.*

3. *There is no function $f : \omega \to X$ such that $\forall n \in \mathbb{N}\ \langle f(n+1), f(n)\rangle \in R$*

Proof. We first treat the implication from (1) to (2). 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 $f(x) = \alpha < \beta = f(y)$.

The implication from (2) to (3) is quick. If (3) were false and there were such an $f' : \omega \ \to X$ then, using f from (2), we see that $f \circ f'$ would give an infinite decreasing sequence of ordinals.

We now verify the implication from (3) to (1) by proving the contrapositive. Suppose that X is not well founded; this means that there is some non-empty $Y \subseteq X$ with no minimal element. Fix some $y_0 \in Y$. We will construct $f : \omega \to X$ by recursion beginning with $f(0) = y_0$.

We have:

$$(\forall y \in Y)(\exists! z)\ z = \{x \in Y : \langle x, y\rangle \in R\}$$

and so by Replacement:

$$\exists Z\ Z = \{z : (\exists y \in Y) z = \{x \in Y : \langle x, y\rangle \in R\}$$

and since Y has no minimal element each set in Z is non-empty. We now invoke Theorem 5 to obtain

$$g : Z \to \bigcup Z \text{ such that for all } z \in Z \text{ we have } g(z) \in z.$$

That is, for each $y \in Y$:

$$g(\{x \in Y : \langle x, y\rangle \in R\}) \in \{x \in Y : \langle x, y\rangle \in R\}$$

and so for each $y \in Y$:

$$\langle g(\{x \in Y : \langle x, y\rangle \in R\}), y\rangle \in R.$$

We can now use Theorem 9 to obtain $f : \omega \to Y$ defined by $f(0) = y_0$ and

$$f(n+1) = g(\{x \in Y : \langle x, f(n)\rangle \in R\}).$$

□

The ordinal-valued function arising in the second part of the theorem is call a <u>rank function</u> for the relation R. It will allow us to prove statements about R by induction on rank, as we shall see below.

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

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

One of the key facts in Set Theory is the Mostowski Collapsing Theorem.

Theorem 19. *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. In order to show the existence of the isomorphism, we first obtain a rank function $f : X \to \mathbb{ON}$ directly from the characterisation of well foundedness, Theorem 18. Let δ be an ordinal such that $rng(f) \subseteq \delta$. By recursion on the ordinals we define for each $\beta \in \delta$ 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

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

Letting $M = rng(h)$, it is now straightforward to see that h is a surjection onto a transitive set and that $h(x) \in h(y)$ iff $\langle x, y\rangle \in R$.

Induction on rank, that is, on $\beta \in \delta$ will show that h is an injection. For $x \in X$, call $f(x)$ the rank of x; the inductive hypothesis is:

$$(\forall y_1)(\forall y_2)((f(y_1) \leq \beta \wedge f(y_2) \leq \beta \wedge h(y_1) = h(y_2)) \;\to\; y_1 = y_2)$$

Assuming this is true for all $\alpha < \beta$, let y_1 and y_2 be in X such that $f(y_1) \leq \beta$ and $f(y_2) \leq \beta$. If $h(y_1) = h(y_2)$, then

$$\{h(x) : \langle x, y_1\rangle \in R\} = \{h(w) : \langle w, y_2\rangle \in R\}.$$

Let $x \in X$ with $\langle x, y_1\rangle \in R$. There is $w \in X$ with $h(x) = h(w)$ and $\langle w, y_2\rangle \in R$. The ranks of x and w must be strictly less than β so by inductive hypothesis we must have $x = w$ and therfore $\langle x, y_2\rangle \in R$.

Similarly, for each $w \in X$ with $\langle w, y_2\rangle \in R$ we see that $\langle w, y_1\rangle \in R$. Since R is extensional we must have $y_1 = y_2$ and so h is an injection and hence an isomorphism.

Uniqueness follows rather immediately from the following surprising result. $\square$

Exercise 14. Prove that any isomorphism between transitive sets is the identity and show how this can be used to finish the proof of the Mostowski Collapsing Theorem.. Of course, the relation on the transitive set is the membership relation, which is extensional and well founded.

We say that a relation R is a partial ordering or partial order whenever it is both irreflexive 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

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

we say that $\langle P, <\rangle$ is a partially ordered set. Whenever the context is clear, we just write P instead of $\langle P, <\rangle$. The concepts of linearly ordered set and well ordered set are defined similarly.

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 20. *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, Theorem 19. The membership relation is transitive on the resulting transitive set. It follows directly that each element of this transitive set is transitive.

□

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 $\text{type}(\langle X, <\rangle)$. In case X is a set of ordinals with the usual ordering we just write $\text{type}(X)$. For example $\text{type}(\{n \in \omega : n \text{ is even}\}) = \omega$.

Corollary. Whenever $\alpha \in \mathbb{ON}$ and $A \subseteq \alpha$ we have $\text{type}(A) \leq \alpha$.

Proof. Let type$(A) = \delta$ and $f\colon \delta \to A$ be the isomorphism given by Theorem 20. Let $\gamma \in \delta$. Since f is order preserving, if $f(\gamma) < \gamma$ we would be able to recursively construct a strictly decreasing sequence of ordinals $\gamma,\ f(\gamma),\ f(f(\gamma))\ \ldots$ etc. Therefore we must have that $\gamma \leq f(\gamma) \in A \subseteq \alpha$ and so $\delta \subset \alpha$. $\square$

A partially ordered set T with a smallest element is said to be a tree provided that for each $t \in T$ the predecessors of t, $<^{\leftarrow} \{t\}$, form a well ordered set; the order type of $<^{\leftarrow} \{t\}$ is called the height of t in the tree T.

Exercise 15. Prove these two results about trees.

1. For any set X and $\beta \in \mathbb{ON}$ the partially ordered set $\langle \bigcup \{{}^{\alpha}X : \alpha \in \beta\}, \subset \rangle$ is a tree.

2. For any tree T there is a set X and an ordinal α such that T is isomorphic to a subset S of the partial order above. Theorem 20 may be helpful.

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

Proof. We begin by immediately using the Power Set Axiom and invoking 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.

□

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. We illustrate this by proving the famous Hausdorff Maximal Principle.

A subset Y of a partially ordered set X is said to be a <u>chain</u> provided that the ordering is total when restricted to just Y. On the other hand, a subset Y of a partially ordered set X is said to be an <u>antichain</u> provided that the ordering is empty when restricted to Y so that no two elements of Y are related. Furthermore, we say that a subset Y of X is <u>centred</u> provided that for each finite $Y' \subseteq Y$ there some $x \in X$ such that $x \leq y$ for all $y \in Y'$. Of course, every chain is centred.

Theorem 22. *Let X be a partially ordered set.*

1. *Let $Y \subseteq X$ be a chain. There is a maximal chain C with $Y \subseteq C \subseteq X$; that is, no larger C' including C is a chain.*
2. *Let $Y \subseteq X$ be an antichain. There is a maximal antichain A with $Y \subseteq A \subseteq X$; that is, no larger A' including A is an antichain.*
3. *Let $Y \subseteq X$ be a centred subset. There is a maximal centred G with $Y \subseteq G \subseteq X$; that is, no larger G' including G is centred.*

Proof. We address the first part of the theorem. Obtain a well ordering of X from Theorem 21. By Theorem 20 there is an ordinal κ and an isomorphism $f : \kappa \to X$. Denoting each $f(\alpha)$ by x_α, we are able to enumerate X in order type κ as $\{x_\alpha : \alpha \in \kappa\}$.

We define $F : (\kappa + 1) \to \mathcal{P}(X)$ by transfinite recursion.

$$F(0) = Y$$

$$
\begin{aligned}
F(\lambda) &= \bigcup\{F(\alpha) : \alpha \in \lambda\} && \text{if } \lambda \text{ is a limit ordinal}\\
F(\alpha+1) &= F(\alpha) \ \cup \ \{x_\alpha\} && \text{if this is a chain}\\
F(\alpha+1) &= F(\alpha) && \text{otherwise}
\end{aligned}
$$

Transfinite induction on β may now be used to verify that:

1. $(\forall\alpha \leq \kappa)(\forall\beta \leq \kappa)\ (\alpha \leq \beta \rightarrow F(\alpha) \subseteq F(\beta) \subseteq X)$.
2. $(\forall\beta \leq \kappa)\ F(\beta)$ is a chain.

Now, if $x \in X$ then $x = x_\alpha$ for some $\alpha \in \kappa$ and furthermore if $x_\alpha \notin F(\kappa)$ it must be that $x_\alpha \notin F(\alpha+1)$. By the definition of F, this means that $F(\alpha) \cup \{x_\alpha\}$ is not a chain and hence neither is $F(\kappa) \cup \{x\}$. So taking $C = F(\kappa)$ gives a proof of the first part of the theorem except for verifying (1) and (2).

The proof of the second part of the theorem is much like the first. $\square$

Exercise 16. Give, for the third part of the theorem, the verifications of (1) and (2).

Chapter 7

Cardinality

In this chapter, we extend the concept of the size of a finite set to include infinite sets as well. By Zermelo's Well Ordering Principle, Theorem 21, every set can be well ordered. By Theorem 20, 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 κ such that $\kappa = |x|$ for some x are called cardinals. Of course, κ is a cardinal iff $\kappa = |\kappa|$.

Exercise 17. Show that the pigeonhole principle implies that each $n \in \omega$ is a cardinal and so the concept of cardinality extends the concept of size. Show also that ω is a cardinal but $\omega + 1$ is not a cardinal and that, in fact, every other cardinal is a limit ordinal.

Theorem 23. *For any x and y we have:*

1. $|x| = |y|$ *iff* $\exists$ bijection $f\colon x \to y$,

2. $|x| \leq |y|$ *iff* $\exists$ injection $f\colon x \to y$, *and*

3. $|x| \geq |y|$ *iff* $\exists$ surjection $f\colon x \to y$. *(Here $y \neq \emptyset$).*

Exercise 18. Look over the corollary to Theorem 20 and prove Theorem 23. Indicate how it answers two of the questions mentioned in the introduction to this book.

A set with cardinality at most ω is said to be countable and otherwise it is said to be uncountable. A ground-breaking theorem of G. Cantor shows that there exist uncountable sets and so there are different infinitudes. Here we must call upon the Power Set Axiom and from now on we will use it without special mention.

Theorem 24. *For all x $|x| < |\mathcal{P}(x)|$.*

Proof. First note that if $|x| \geq |\mathcal{P}(x)|$, then there would, by Theorem 23, be a surjection

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

But this cannot happen, since $\{a \in x : a \notin g(a)\} \notin g^{\to}(x)$. Notice the similarity to the argument for Russell's Paradox which was patterned after this proof.

$\square$

For any ordinal α, we denote by α^+ the least cardinal greater than α; by Cantor's Theorem, Theorem 24, this is guaranteed to exist. We often denote the first uncountable cardinal ω^+ by ω_1, the second uncountable cardinal ω_1^+ by ω_2 and so on.

Cantor's Theorem implies $|\mathcal{P}(\omega)| \geq \omega_1$. The Continuum Hypothesis is the statement $|\mathcal{P}(\omega)| = \omega_1$. This hypothesis, abbreviated as CH, was proposed by G. Cantor in the nineteenth century and has been continually arising in applications of Set Theory ever since. Unfortunately, we do not know whether it is true or false.

We do know that CH does not follow from the other $\mathfrak{ZFC}$ axioms (P. Cohen, 1963). Nevertheless, we also know that adding CH to $\mathfrak{ZFC}$ will not give rise to any new inconsistencies (Gödel, 1938). So $\mathfrak{ZFC}$ alone will not determine the truth of CH and we find ourselves in a dilemma. Our axiom system $\mathfrak{ZFC}$ is too weak — we need to add some new axiom in order to strengthen it. This new axiom should be powerful enough to decide CH and yet this new axiom should be obviously true. How is it possible that we could have overlooked such a thing?

Exercise 19. Prove these two statements.

1. The supremum of a set of cardinals is a cardinal.
2. $\neg\exists z\ z = \{\kappa : \kappa \text{ is a cardinal}\}$.

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

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

We write $\aleph_\alpha$ for $\aleph(\alpha)$. For small α we sometimes write ω_α for $\aleph(\alpha)$.

Exercise 20. Prove that

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

and use this to conclude that each singular cardinal contains a cofinal subset of regular cardinals.

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

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

We write $\beth_\beta$ for $\beth(\beta)$. It is apparent from Cantor's Theorem 24 that $\aleph_\alpha \leq \beth_\alpha$ for all $\alpha \in \mathbb{ON}$.

The Continuum Hypothesis, CH, can be written as $\aleph_1 = \beth_1$. It can be strengthened to the Generalised Continuum Hypothesis, abbreviated as GCH, which is the statement:

$$\forall\alpha \in \mathbb{ON}\ \aleph_\alpha = \beth_\alpha.$$

Using Exercise 20 it is easy to see that the GCH is equivalent to requiring that $|\mathcal{P}(\kappa)| = \kappa^+$ for all infinite cardinals κ.

Interestingly, there are cardinals at which the GCH holds provided it holds at all smaller cardinals. That is:

$$(\forall \beta \leq \alpha)(\beth_\beta = \aleph_\beta) \;\Rightarrow\; \beth_{\alpha+1} = \aleph_{\alpha+1}$$

is true for all ordinals α such that $\aleph_\alpha$ is a singular cardinal of uncountable cofinality (J. Silver). However the implication does not hold when $\aleph_\alpha$ is a singular cardinal of countable cofinality (M. Magidor) nor does it hold when $\aleph_\alpha$ is a regular cardinal (P. Cohen).

Ideally, we would like to know if, for a limit ordinal δ, we can determine $\beth_{\delta+1}$ in terms of $\beth_\delta$. A substantial amount of work has produced striking results. As examples (S. Shelah):

$$\beth_\omega = \aleph_\omega \;\Rightarrow\; \beth_{\omega+1} < \aleph_\alpha \text{ for some } \alpha < |\mathcal{P}(\omega)|^+$$

and (F. Galvin and A. Hajnal):

$$\beth_{\omega_1} = \aleph_{\omega_1} \;\Rightarrow\; \beth_{\omega_1+1} < \aleph_\alpha \text{ for some } \alpha < |\mathcal{P}(\omega_1)|^+$$

and most remarkable of all (S. Shelah):

$$\beth_\omega = \Rightarrow\; \beth_{\omega+1} < \aleph_{\omega_4}.$$

But the best bounds are still a mystery.

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

Proof. Let κ be an infinite cardinal. Since $\kappa = |\kappa \times \{0\}|$ and

$$\kappa \times \{0\} \subseteq \kappa \times \kappa$$

we have that $\kappa \leq |\kappa \times \kappa|$. In order to show that $|\kappa \times \kappa| \leq \kappa$, we use induction and assume that $|\rho \times \rho| = \rho$ for each infinite cardinal $\rho < \kappa$. We can look at this as a transfinite induction, proving that for all $\beta \in \mathbb{ON}$:

$$\forall \alpha \in \beta \; |\aleph_\alpha \times \aleph_\alpha| \leq \aleph_\beta.$$

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 θ be the order type of $\langle \kappa \times \kappa, < \rangle$. It suffices to show that $\theta \leq \kappa$, which we will prove by contradiction. Suppose $\theta > \kappa$; there must be some $\langle \alpha, \beta \rangle \in \kappa \times \kappa$ such that:

$$<^{\leftarrow} (\{\langle \alpha, \beta \rangle\}) = \{\langle \gamma, \delta \rangle \colon \langle \gamma, \delta \rangle < \langle \alpha, \beta \rangle\}$$

has order type κ and hence cardinality κ.

By Exercise 17 we know that κ is a limit ordinal and so there is some $\lambda \in \kappa$ such that $\{\alpha, \beta\} \subset \lambda$ and hence

$$<^{\leftarrow} (\{\langle \alpha, \beta \rangle\}) \subseteq \lambda \times \lambda.$$

In order to get a contradiction and complete the proof, it suffices to show that $|\lambda \times \lambda| < \kappa$ for every $\lambda \in \kappa$. Since

$$|\lambda \times \lambda| = ||\lambda| \times |\lambda||$$

this follows by induction if λ is infinite and trivially if λ is finite. $\square$

Corollary. Let X be an infinite set.

1. For any non-empty Y $|X \times Y| = \max\{|X|, |Y|\}$.
2. $|\bigcup X| \leq \max\{|X|, \sup\{|x| : x \in X\}\}$.

Proof. The first part follows immediately from Theorem 23 and Theorem 25.

For the second part let $\kappa = \sup\{|x| : x \in X\}$. Use Theorem 23 and the Axiom of Choice to obtain, for each $x \in X$, a surjection $f_x \colon \kappa \to x$. Define a surjection $f \colon X \times \kappa \to \bigcup X$ by

$$f(\langle x, \beta \rangle) = f_x(\beta).$$

The result now follows from the first part and Theorem 23. $\square$

The second part of the corollary is a generalization of the famous dictum "the countable union of countable sets is countable". The inequality can be strict, as when $X = \mathcal{P}(\omega)$.

Exercise 21.

1. Prove an infinite version of the pigeonhole principle: if $f: \omega_1 \to \omega$ then there is some $n \in \omega$ such that $f^{\leftarrow}(\{n\})$ is uncountable.
2. Use this and induction on $\mathbb{N}$ to prove the famous Delta System Lemma which is stated below.

Lemma. Whenever $\mathcal{A}$ is an uncountable collection of finite sets, there is an uncountable $\mathcal{D} \subseteq \mathcal{A}$ and an R such that $A \cap B = R$ for all distinct A and B in $\mathcal{D}$.

Exercise 22.

1. Prove that if $g\colon \omega_1 \to \omega_1$ then for some $0 < \alpha \in \omega_1$ we get $g^{\rightarrow}(\alpha) \subseteq \alpha$.
2. Use this to prove the famous Pressing Down Lemma stated below.

Lemma. If $f: \omega_1 \to \omega_1$ and $f(\alpha) < \alpha$ whenever $\alpha \neq 0$ then for some $\beta \in \omega_1$ we have $|f^{\leftarrow}\{\beta\}| = \omega_1$.

Along with the Delta System Lemma and the Pressing Down Lemma we have the Free Set Lemma below.

Lemma. Whenever $h : \omega_1 \to \{x : x \text{ is finite}\}$ there is an uncountable $S \subseteq \omega_1$ such that for all distinct α and β in S we have $\alpha \notin h(\beta)$.

Proof. Let's use the Pressing Down Lemma. Define $f : \omega_1 \to \omega_1$ by letting $f(\alpha) = 0$ if $h(\alpha) \cap \alpha = \emptyset$ and otherwise let $f(\alpha)$ be the maximum element of $h(\alpha) \cap \alpha$.

By the Pressing Down Lemma there is some uncountable $S \subseteq \omega_1$ such that $\alpha \notin h(\beta)$ whenever $\alpha < \beta$ and $\{\alpha, \beta\} \subseteq S$. Therefore each $\alpha \in S$ is in at most countably many members of $\{h(\beta) : \beta \in S\}$.

Now recursively define $t_\alpha \in S$ for $\alpha \in \omega_1$ by:

$$t_\alpha \text{ is the least element of } S \setminus \bigcup \{h(t_\gamma) : \gamma < \alpha\}$$

and check that $\{t_\alpha : \alpha < \omega_1\}$ is the set we seek. $\square$

Let's use the bracket notation. For a set X and a cardinal κ, $[X]^\kappa$ will denote $\{a \in \mathcal{P}(X) : |a| = \kappa\}$ and $[X]^{<\kappa}$ will denote $\{a \in \mathcal{P}(X) : |a| < \kappa\}$.

Theorem 26. *Let κ and λ be cardinals.*

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

2. *If $\kappa \geq \lambda \geq 2$ and κ is infinite, then $|{}^\kappa\lambda| = |\mathcal{P}(\kappa)|$.*

Proof. To prove the first part notice that ${}^\kappa\lambda \subseteq [\kappa \times \lambda]^\kappa$ and so $|{}^\kappa\lambda| \leq |[\lambda]^\kappa|$. Furthermore, using the Axiom of Choice, for each $x \in [\lambda]^\kappa$ we can pick a bijection $f_x \colon \kappa \to x$. Since each $f_x \in {}^\kappa\lambda$, we have an injection from $[\lambda]^\kappa$ into ${}^\kappa\lambda$, so $|[\lambda]^\kappa| \leq |{}^\kappa\lambda|$.

To prove the second part, we use characteristic functions of subsets of κ and the result of the first part with κ equal to λ to get

$$|\mathcal{P}(\kappa)| = |{}^\kappa 2| \leq |{}^\kappa\lambda| \leq |{}^\kappa\kappa| = |[\kappa]^\kappa| \leq |\mathcal{P}(\kappa)|$$

which finishes the proof. $\square$

Exercise 23. Verify the following facts:

1. for any infinite cardinal λ we have $|[\lambda]^{<\omega}| = \lambda$,

2. for any infinite cardinal κ we have $|[\kappa]^\kappa| = |\mathcal{P}(\kappa)|$,

3. $|[\mathcal{P}(\omega)]^\omega| = |\mathcal{P}(\omega)|$.

A subset S of an ordinal α is said to be cofinal whenever

$$\forall \beta \in \alpha \; \exists \sigma \in S \; \beta \leq \sigma.$$

The cofinality of an ordinal α, $cf(\alpha)$, is the smallest cardinality of a cofinal subset of α.

An infinite cardinal κ is said to be a regular cardinal whenever $cf(\kappa) = \kappa$. Other infinite cardinals are said to be singular cardinals. It should be obvious that ω is a regular cardinal. From the second part of the corollary to Theorem 25 it is easy to check that ω_1 is a regular cardinal. In fact, for each infinite ordinal κ, κ^+ is a regular cardinal.

Exercise 24. Prove this last statement and also prove that $cf(\delta)$ is a regular cardinal for any infinite limit ordinal δ.

This next theorem of Gy. König will lead to a strengthening of Cantor's Theorem.

Theorem 27. *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 $h \notin g^{\to}(\kappa)$, since otherwise $h = g(\beta)$ for some $\beta < \kappa$; pick $\alpha \in \delta$ such that $f(\alpha) > \beta$. $\square$

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

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

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

$\square$

A regular uncountable cardinal κ is said to be <u>inaccessible</u> whenever

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

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

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

Under the GCH these two notions are equivalent.

Axiom of Inaccessibles $\quad \exists \kappa\ \kappa$ is an inaccessible cardinal

Since ω would be inaccessible if only it were uncountable, the Axiom of Inaccessibles is a stronger version of the Axiom of Infinity. However, the mathematical community is not quite ready to replace the Axiom of Infinity with it just yet. The Axiom of Inacessibles is not included in the basic $\mathfrak{ZFC}$ axiom system and so it is always explicitly stated whenever it is used.

This exercise will help you to get a feeling for the size of these cardinals.

Exercise 25. 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

What's So 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 and from the integers we shall construct the decimal numbers and the real numbers.

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

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

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

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

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

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

Finally, let

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

to obtain the real numbers and let

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

to obtain the decimal numbers. The number $f(n)$ is the n^{th} decimal place of the real number f. As usual we will identify the integers with a subset of the decimal numbers, identifying each $z \in \mathbb{Z}$ with the $f \in \mathbb{D}$ such that:

$$f(0) = z \text{ and } f(n) = 0 \text{ for all other } n \in \omega.$$

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 is clearly a linear ordering which 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 then to $\mathbb{R}$ in a natural — if cumbersome — fashion. The real numbers are complete; each bounded subset has a supremum and an infimum which can be constructed recursively, decimal place by decimal place.

Theorem 28.

1. *$\mathbb{D}$ is a countable dense subset of $\mathbb{R}$.*
2. *$\mathbb{R}$ is uncountable; in fact $|\mathbb{R}| = |\mathcal{P}(\omega)|$.*
3. *There is no subset of $\mathbb{R}$ with order type ω_1.*

Exercise 26. Prove this theorem; *dense* is defined below.

That $\mathbb{X}$ is a dense subset of a linear order $\mathbb{Y}$ means that

$$(\forall p \in \mathbb{Y})(\forall q \in \mathbb{Y})\ [p < q \to \exists d \in \mathbb{X}\ p < d < q].$$

To simply say that a linear order is dense is to say that it is a dense subset of itself, as in the following theorem.

Theorem 29. *Any two non-empty 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

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

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.

□

This theorem leads to the fact that any non-empty complete dense linear order without endpoints and with a countable dense subset is isomorphic to $\langle \mathbb{R}, < \rangle$. It is natural to wonder if "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, one of the most important issues in Set Theory. The next theorem, due to D. Kurepa, connects it to the study of trees.

Theorem 30. *The following are equivalent.*

1. *Suslin's Hypothesis fails (there is a Suslin line).*
2. *There is an uncountable tree with no uncountable chains or antichains (there is a Suslin tree).*

Exercise 27. Prove this theorem.

First hint: construct the Suslin tree from intervals of the Suslin line with an ordering of reverse inclusion. Second hint: construct the Suslin line from maximal chains of the Suslin tree. Third hint: don't attempt to prove either statement directly. The Suslin Hypothesis, like the Continuum Hypothesis, will require new axioms for its resolution.

A collection A of subsets of a set X is said to be almost disjoint provided that $a \cap b$ is finite for all distinct a and b in A. Although any pairwise disjoint family of subsets of ω is countable, there is an uncountable almost disjoint subcollection of $\mathcal{P}(\omega)$. In fact there is one of cardinality $|\mathcal{P}(\omega)| = |\mathbb{R}|$. Since $|\mathbb{D}| = \omega$, it suffices to find the almost disjoint collection as a subset of $\mathcal{P}(\mathbb{D})$. For each $f \in \mathbb{R}$, let a_f be

$$\{d \in \mathbb{D} : (\exists n \in \omega)[(\forall m < n)\ d(m) = f(m) \wedge (\forall m > n)\ d(m) = 0]$$

which is the set of decimal approximations to f. Clearly, if $f \in \mathbb{R}$ and $g \in \mathbb{R}$ with $f \neq g$ then $a_f \cap a_g$ must be finite.

The famous American philosopher Yogi Berra once said, "It's very difficult to predict what is going to happen — especially in the future".

We will use Set Theory to show that predicting the future most of the time is theoretically possible, but that it is very difficult. This example of "applied mathematics" is due to C. S. Hardin and A. D. Taylor.

Let T denote the real interval $[0, \infty)$. We think of $t \in T$ as a moment in time. We are trying to predict the values of a fixed but unknown function $f\colon T \to \mathbb{R}$. At any time t we know the "history" of f up until time t, that is, we know $f|[0,t)$ and we wish to predict $f|[t,\infty)$.

To simplify matters we make all our predictions as functions with full domain; so at time $t \in T$ we will make a prediction $g\colon T \to \mathbb{R}$ from the set of functions $H(t)$ which historically agree with f up to t.

$$H(t) = \{g : (g\colon T \to \mathbb{R}) \text{ and } g|[0,t) = f|[0,t)\}.$$

The correct prediction would be f itself, but how can there be a way to correctly guess f with no other information provided?

Surprisingly, there is a good strategy for correctly predicting the short term future — that is, for predicting, at each time $t \geq 0$ a function g such that $g|[t, t+\epsilon) = f|[t, t+\epsilon)$ for some $\epsilon > 0$. We construct such a strategy below.

From Theorem 21 we get a well ordering $\prec$ of ${}^T\mathbb{R}$ which we now fix. From this well ordering we make a strategy σ for predicting f:

$$\sigma\colon T \to {}^T\mathbb{R}$$

where for each $t \in T$:

$$\sigma(t) \text{ is the least element of } H(t) \text{ according to } \prec .$$

Let $W = \{t \in T : \sigma(s) \neq \sigma(t) \text{ for all } s > t\}$. We claim that the strategy σ correctly predicts the short term future of the function f at each time t not in W.

Claim. For all $t \in T \setminus W$ there is $\epsilon > 0$ such that

$$\sigma(t)(x) = f(x) \text{ for all } x \in [t, t+\epsilon).$$

Proof of Claim. Let $t \in T \setminus W$ and fix $s > t$ with $\sigma(s) = \sigma(t)$. For all x with $t \leq x < s$ we have:

$$\begin{aligned}
\sigma(t)(x) &= \sigma(s)(x) \text{ and} \\
\sigma(s)(x) &= f(x) \text{ since } \sigma(s) \in H(s).
\end{aligned}$$

This proves the claim for $\epsilon = s - t$.

In order to show that σ is a good prediction strategy we need to show that the set W is a small subset of T.

Claim. W is a well founded subset of $\mathbb{R}$.

Proof of Claim. By Theorem 18 it suffices to show that W contains no infinite decreasing sequence. Suppose then, for the sake of argument, that there were such a sequence $\{t_n : n \in \omega\}$. Let $\sigma(t_n) = g_n$ for each $n \in \omega$; since each $t_{n+1} \in W$ we have that $g_{n+1} \neq g_n$.

Moreover, since $t_{n+1} < t_n$ we have

$$H(t_n) \subseteq H(t_{n+1}) \text{ and so } \{g_n, g_{n+1}\} \subseteq H(t_{n+1}).$$

Since g_{n+1} is the least member of $H(t_{n+1})$ according to $\prec$, we must have that $g_{n+1} \prec g_n$. Since $\prec$ is a well ordering, we get our desired contradiction.

Since it contains no infinite decreasing sequences, W cannot be dense in any real interval; it is topologically small. It is also small in other ways. W is a well ordered set and so by Theorem 20 it is isomorphic to an ordinal. It cannot be isomorphic to an ordinal greater than or equal to ω_1 because that would contradict Theorem 28. Therefore W is countable and hence of Lebesgue measure zero.

This latter fact means that there is some positive real number ϵ such that the strategy σ correctly predicts the function f over the interval $[t, t + \epsilon)$ for at least 99% of the time t in $[0, 1]$. Unfortunately, the proof gives no idea how large (or how very tiny) this ϵ might be.

Furthermore, implementing the prediction strategy σ seems to be impractical because we don't have a good working knowledge of the well ordering $\prec$ on ${}^T\mathbb{R}$. Perhaps we should end this example with another quote attributed to Yogi Berra. When asked to distinguish between *theory* and *practice* he said: "In theory, theory and practice are the same; in practice — they ain't."

Let's now turn to ordinary Euclidean space and prove a geometric theorem of P. Komjáth and V. Totik in order to demonstrate a method of proof using transfinite induction and transfinite recursion simultaneously. From $\mathbb{R}$ and $n \in \mathbb{N}$ we construct $\mathbb{R}^n$ as ${}^n\mathbb{R}$, the set of sequences of reals of length n and as usual, we identify $\mathbb{R}^1$ with $\mathbb{R}$ when no confusion can possibly arise.

Theorem 31.

$\mathbb{R}^3$ is the disjoint union of straight lines, no two of which are parallel.

Proof. Induction is a method of proof, whereas recursion is a method of definition. Until now, we would define something by recursion (for example, ordinal addition) and then use induction to prove some properties about it (for example, associativity). For this proof, however, we will use induction concurrently with recursion, inductively verifying properties of our new objects as we recursively define them.

Let $\lambda = |\mathbb{R}|$; since $\mathbb{R}^3 \subseteq [3 \times \mathbb{R}]^3$ we also have $|\mathbb{R}^3| = \lambda$ by Exercise 23. Using a bijection between λ and $\mathbb{R}^3$ we can enumerate $\mathbb{R}^3$ as $\{p_\alpha : \alpha < \lambda\}$; formally each $p_\alpha = P(\alpha)$ where $P \colon \lambda \to \mathbb{R}^3$ is a bijection.

We will now recursively choose straight lines $l_\beta \subseteq \mathbb{R}^3$ for $\beta < \lambda$ such that for each β we have:

1. $\{p_\alpha : \alpha < \beta\} \subseteq \bigcup\{l_\alpha : \alpha < \beta\}$,
2. l_β is disjoint from l_α for all $\alpha < \beta$ and
3. l_β is not parallel to any l_α for any $\alpha < \beta$.

It is clear that if we can do this then $\{l_\alpha : \alpha < \lambda\}$ will satisfy the theorem.

To begin, we let l_0 be any straight line through p_0. Suppose now that we are at stage β and that for all $\alpha < \beta$ we have properly chosen l_α.

Let p be defined as the first point in the enumeration of $\mathbb{R}^3$ which is not covered by the lines already defined. In particular if $p_\beta \notin \bigcup\{l_\alpha : \alpha < \beta\}$ then $p = p_\beta$. For each $\alpha < \beta$ let S_α be the plane containing $l_\alpha \cup \{p\}$.

Claim. $\mathbb{R}^3 \neq \bigcup\{S_\alpha : \alpha < \beta\}$.

Proof of Claim. Since $|\beta| < \lambda$ there is at least one horizontal plane, S, not in the collection $\{S_\alpha : \alpha < \beta\}$. The non-horizontal members of the collection intersect S in at most $|\beta|$ straight lines. This resulting collection of straight lines in S does not include every line in S and any line in S not in the collection is not covered by them.

Now choose $q \in \mathbb{R}^3 \setminus \bigcup\{S_\alpha : \alpha < \beta\}$ and let l_β be the straight line containing $\{p, q\}$. Property (1) holds by definition.

If l_β would intersect some l_α then they would both lie in the plane S_α; since $q \notin S_\alpha$ we must have Property (2). Similarly, Property (3) holds as well, since if l_β would be parallel to l_α they would both have to lie in the plane S_α. $\square$

We often use this "concurrent recursion with induction" method, rather than the "first recursion and then induction" method of proof. Technically, what is being done here? In fact we are simply using induction to verify the hypothesis of the theorem on recusion (that is, the existence of a unique set with desired properties) and then taking the object that the theorem gives us. More precisely, we verify

$$(\forall x \in \mathbb{ON})(\forall f)[(f\colon x \to \mathbb{V}) \to \exists! y\ \Phi(x, f, y, w_1, \ldots, w_k)].$$

for those f which satisfy $(\forall j < x)\ \Phi(j, f, y, w_1, \ldots, w_k)$. The reader may verify that in the proof of Theorem Scheme 14 this was all that was required.

Chapter 9

Ultrafilters Are Useful

It is intuitively clear that some infinite subsets of $\mathbb{N}$ may be considered "bigger" than others, for example $\{n \in \mathbb{N} : n \geq 9\}$ may well be considered "bigger" than $\{n \in \mathbb{N} : n \text{ is even}\}$. However, comparing other infinite sets may be more problematic. It turns out to be useful to have a device which can measure all subsets of $\mathbb{N}$, even to divide them into just two types: big and little, and even if some of the comparisons are done quite arbitrarily.

In order to accomplish this, we introduce the important notion of an ultrafilter, which picks out the big subsets of a set S.

A collection of subsets $\mathcal{U} \subseteq \mathcal{P}(S)$ is a filter 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 $A \subseteq B \subseteq S$, then $B \in \mathcal{U}$.

4. If $A \subseteq S$, then either $A \in \mathcal{U}$ or $S \setminus A \in \mathcal{U}$.

5. For all $x \in S$, $S \setminus \{x\} \in \mathcal{U}$.

If a filter $\mathcal{U}$ obeys condition (4) it is called an ultrafilter, If $\mathcal{U}$ also satisfies condition (5) it is said to be a free or non-principal ultrafilter; otherwise it is a fixed or principal ultrafilter. The fixed ulrafilters over S are of the form $\{A \subseteq S : x \in A\}$ for some $x \in S$.

One classic example of a filter is the Fréchet filter $\mathcal{F}$ which is:

$$\{A \subseteq \omega : \omega \setminus A \text{ is finite}\}$$

This can be generalised to any infinite S. Clearly, an ultrafilter $\mathcal{U}$ over ω is free iff the Fréchet filter $\mathcal{F} \subseteq \mathcal{U}$. Fixed ultrafilters are trivial, but examples of free ultrafilters are not so easy to come by.

Exercise 28. Consider the partial ordering of inclusion on $\mathcal{P}(\omega)\setminus\{\emptyset\}$. Apply Theorem 22 to the Fréchet filter to prove the existence of a free ultrafilter over ω. Generalise this from ω to any infinite set S..

Theorem 32. *(F. Ramsey)*

If $P\colon [\omega]^2 \to \{1, 2\}$, then there is an infinite $H \subseteq \omega$ such that P is constant on $[H]^2$.

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

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

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

Let $S = \{\alpha \in \omega : \{\beta \in \omega : P(\{\alpha, \beta\}) = 1\} \in \mathcal{U}\}$. Pick $\alpha_0 \in S$ and let

$$S_0 = \{\beta \in \omega : P(\{\alpha_0, \beta\}) = 1\}.$$

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

$$S_1 = \{\beta \in \omega : P(\{\alpha_1, \beta\}) = 1\}.$$

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

$$S_n = \{\beta \in \omega : P(\{\alpha_n, \beta\}) = 1\}.$$

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

$\square$

Ramsey used this in order to prove a finitary version.

Theorem 33. *For each $n \in \mathbb{N}$ there is $r \in \mathbb{N}$ such whenever $P\colon [r]^2 \to \{1, 2\}$ there is an $H \subseteq r$ of size n such that P is constant on $[H]^2$.*

Proof. We use an indirect proof. Fix $n \in \mathbb{N}$ such that for each $r \in \mathbb{N}$ there is a partition function $P_r\colon [r]^2 \to \{1, 2\}$ such that P_r is non-constant on $[H]^2$ for any $H \subseteq r$ of size n.

Let F be $\{f : (\exists r \in \mathbb{N}) f\colon [r]^2 \to \{1, 2\}\}$. F becomes a tree with the ordering of proper function extention (equivalently: proper set inclusion). For any $r \in \mathbb{N}$ the elements of T with height r are functions with domain $[r]^2$ and range contained in $\{1, 2\}$. There are only finitely many (in fact exactly $2^{r(r+1)/2}$) possibilities for such functions. Therefore there are only finitely many elements of the tree with height r.

Let T be the subtree of F given by:

$$\{f \in F : f \text{ is not constant on } [H]^2 \text{ for any } H \subseteq r \text{ of size } n\}$$

This tree is infinite because each P_r is in T. However, it has no infinite chains because the union of an infinitely long chain in T would be a partition function contradicting the previous theorem. So the following exercise completes the proof. $\square$

Exercise 29. Prove that if T is an infinite tree with only finitely many elements of each finite height, then T must contain an infinite chain. Hint: let $\mathcal{U}$ be a free ultrafilter over T and consider $C = \{s \in T : \{t \in T : s < t\} \in \mathcal{U}\}$.

Theoren 33 gives us an $r \in \mathbb{N}$ for each $n \in \mathbb{N}$, but does not tell us whether r is large or small compared to n. The smallest such $r = R(n)$ is called the n^{th} Ramsey number and must be determined by different means.

It is elementary to show that $R(3) = 6$, somewhat more difficult to see that $R(4) = 18$ and quite challenging to show that $43 \leq R(5) \leq 49$ The exact value of $R(5)$ or the higher Ramsey numbers is unknown and determination of the values of even the first few Ramsey numbers is thought to be one of the most difficult problems in mathematics.

Ramsey's Theorem is *not* true if ω is replaced by ω_1 as this example of W. Sierpinski shows.

Theorem 34. *There is a function $P\colon [\omega_1]^2 \to \{1, 2\}$ such that P is non-constant on $[H]^2$ for any uncountable $H \subseteq \omega_1$.*

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

$$P(\{\alpha, \beta\}) = \begin{cases} 1, & \text{if } f(\alpha) < f(\beta); \\ 2, & \text{if } f(\alpha) > f(\beta). \end{cases} \tag{9.1}$$

Appealing to Theorem 28 now finishes the proof. □

S. Todocevic has extended this by constructing a function $P\colon [\omega_1]^2 \to \omega_1$ such that P takes all values on $[H]^2$ for any uncountable $H \subseteq \omega_1$. That is, for any $\gamma \in \omega_1$ there are α and β in H such that $P(\{\alpha, \beta\}) = \gamma$.

As we have seen, the intersection of finitely many members of an ultrafilter is also a member of the ultrafilter. Given the rough analogy between ultrafilters and two-valued measures, it is natural to ask for a free ultrafilter over an uncountable set S such that the intersection of countably many members of the ultrafilter is also a member of the ultrafilter. Such an ultrafilter will be called <u>countably complete</u>.

More generally, given an uncountable cardinal κ, an ultrafilter $\mathcal{U}$ is said to be <u>$\kappa-$complete</u> if $\forall X \in [\mathcal{U}]^{<\kappa}\ \bigcap X \in \mathcal{U}$, that is, the intersection of fewer than κ members of $\mathcal{U}$ is also a member of $\mathcal{U}$. So countably complete is ω_1-complete with this terminology.

Exercise 30. Let $\mathcal{U}$ be a countably complete free ultrafilter over a set S. Prove that there is an uncountable cardinal $\kappa \leq |S|$ with a $\kappa-$complete free ultrafilter over κ. Hint: let κ be the least cardinal such that the intersection of κ many members of $\mathcal{U}$ is not a member of $\mathcal{U}$ and let this be witnessed by $\{A_\alpha : \alpha \in \kappa\}$. Let $B_\beta = \bigcap\{A_\alpha : \alpha < \beta\} \setminus A_\beta$ for all $\beta \in \kappa$.

An uncountable cardinal κ is said to be measurable whenever there exists a $\kappa-$complete free ultrafilter over κ. Thus, Exercise 30 says that the existance of a countably complete ultrafilter entails the existence of a measurable cardinal.

In some respects, a measurable cardinal acts like the cardinal ω.

Theorem 35. *Let kappa be a measurable cardinal. If* $P\colon [\kappa]^2 \to \{1,2\}$, *then there is* $H \subseteq \kappa$ *of size* κ *such that* P *is constant on* $[H]^2$.

Proof. The proof follows that of Ramsey's Theorem 32, replacing ω with κ. The $\kappa-$completeness of $\mathcal{U}$ will let us build the long sequence $\{\alpha_n : n \in \kappa\}$. $\square$

Theorem 36. *Every measurable cardinal is inaccessible.*

Proof. Let $\mathcal{U}$ be a $\kappa-$complete free ultrafilter over κ; then for each initial segment $\alpha \subseteq \kappa$ we have that $\kappa \setminus \alpha \in \mathcal{U}$. A cofinal subset of κ of size less than κ would immediately contradict the completeness of $\mathcal{U}$ and hence κ must be regular.

It remains to show that if $\lambda < \kappa$ then $|\mathcal{P}(\lambda)| < \kappa$ which we do by contradiction: suppose $\lambda < \kappa$ is the least cardinal such that $|\mathcal{P}(\lambda)| \geq \kappa$.

We impose a linear ordering on $\mathcal{P}(\lambda)$ by setting $A < B$ provided that there is some $\alpha < \lambda$ such that $A \cap \alpha = B \cap \alpha$ and $\alpha \in B \setminus A$. Checking that this is indeed a linear ordering is straightforward. Furthermore, between any $A < B$ in $\mathcal{P}(\lambda)$ there is some C with $A < C \leq B$ and

$$C \in \mathcal{D} = \{D \in \mathcal{P}(\lambda) : (\exists \alpha \in \lambda) D \subseteq \alpha\}.$$

By the choice of λ we have $|\mathcal{D}| < \kappa$. An argument similar to that used for the third part of Theorem 28 shows that this linear ordering on $\mathcal{P}(\lambda)$ can have no increasing or decreasing sequences of order type κ. We now use a variant of Sierpinki's partition from Theorem 34. Let $f\colon \kappa \to \mathcal{P}(\lambda)$ be an injection. Define P as follows: for $\alpha < \beta$, let

$$P(\{\alpha, \beta\}) = \begin{cases} 1, & \text{if } f(\alpha) < f(\beta); \\ 2, & \text{if } f(\alpha) > f(\beta). \end{cases} \tag{9.2}$$

We now apply Theorem 35 to get a contradiction. □

Let $\mathcal{U}$ be a free ultrafilter over ω. Form an equivalence (i.e. reflexive, symmetric and transitive) relation $\sim$ on ${}^{\omega}\mathbb{R}$ by the rule:

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

In order to verify that this is an equivalence relation we only need to know that $\mathcal{U}$ is a filter. 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 <u>ultrapower</u> of $\mathbb{R}$ with respect to $\mathcal{U}$ and is usually denoted by ${}^*\mathbb{R}$. The elements of ${}^*\mathbb{R}$ are called the <u>hyperreal numbers</u>.

There is a <u>natural embedding</u> 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 also define an ordering ${}^*\!\!<$ on ${}^*\mathbb{R}$ by the rule:

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

Exercise 31. Verify that ${}^*\!\!<$ is a linear ordering on ${}^*\mathbb{R}$ which extends the usual ordering of $\mathbb{R}$ and show where it is necessary to assume that $\mathcal{U}$ is an ultrafilter, not just a filter.

We usually omit the asterisk, writing $<$ for ${}^*\!\!<$.

Note that $\mathbb{R} \neq {}^*\mathbb{R}$; consider $a = [f]$, where $f(n) = 1/n$ for each $n > 0$. We have $a > 0$ but $a < r$ for each positive $r \in \mathbb{R}$; a member of ${}^*\mathbb{R}$ with this property is called a positive <u>infinitesimal</u>. Similarly, there are negative infinitesimals; 0 is also considered to be an infinitesimal.

Exercise 32. Prove that there is a subset of ${}^*\mathbb{R}$ of order type ω_1, Theorem 28 notwithstanding.

We can also extend addition and multiplication to ${}^*\mathbb{R}$. For example, for a, b and c in ${}^*\mathbb{R}$, $a + b = c$ means that

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

Multiplication of hyperreals is defined similarly. With the same techniques used for Exercise 31, it is now straightforward to show that ${}^*\mathbb{R}$ is an ordered field and the natural embedding places $\mathbb{R}$ as an ordered subfield of ${}^*\mathbb{R}$.

Since ${}^*\mathbb{R}$ is an ordered field, for any positive infinitesimal a we have that $1/a$ exists and $1/a > r$ for any real number r; this is an example of a positive infinite number.

A hyperreal number a is said to be finite whenever $-r < a < r$ 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,

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

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

$${}^*F\colon \; {}^*\mathbb{R} \; \to \; {}^*\mathbb{R}$$

given, for each $a \in {}^*\mathbb{R}$, by

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

where $s \colon \omega \to \mathbb{R}$ is some element of a. It is easily verified that indeed $^*F$ is a function which extends F. This allows us to write F for $^*F$ because omitting the asterisk will cause no confusion.

In order to do elementary calculus we consider a function $F \colon \mathbb{R} \to \mathbb{R}$ with $y = F(x)$. We define, as usual, the increment $\triangle y$ generated by $\triangle x$ as $F(x + \triangle x) - F(x)$.

We define the derivative, F', of F by setting $F'(x)$ to be

$$st\left(\frac{\triangle y}{\triangle x}\right)$$

provided this exists and is the same for each non-zero infinitesimal $\triangle x$. When F is differentiable at x, we define the differential dy of F at x generated by $\triangle x$ as $dy = F'(x)\triangle x$.

Assuming that $F'(x)$ exists it is straightforward to check that for each infinitesimal $\triangle x$ the quantities $\triangle y$ and dy are both infinitesimal. Furthermore, if we also denote $\triangle x$ by dx, the quantities $\triangle x$, $\triangle y$, dx and dy are all infinitesimal and are related by:

$$\frac{dy}{dx} = st(\frac{\triangle y}{\triangle x}) = F'(x)$$

Theorem 37. *(The Chain Rule)*

Suppose $y = F(x)$ and $x = G(t)$ are differentiable functions. Then $y = F(G(t))$ is differentiable and has derivative $\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt}$.

Proof. Let $\triangle t = dt$ be any non-zero infinitesimal. Since G is differentiable, we have that $\triangle x = G(t + \triangle t) - G(t)$ is infinitesimal. Now, in turn, we let $\triangle y = F(x + \triangle x) - F(x)$. We wish to calculate $st(\frac{\triangle y}{\triangle t})$, which will be the derivative of $y = F(G(t))$. We consider two cases.

Case 0: $\triangle x = 0$

We have $\triangle y = 0$, so $st(\frac{\triangle y}{\triangle t}) = 0$, and also $\frac{dx}{dt} = st(\frac{\triangle x}{\triangle t}) = 0$.
So $st(\frac{\triangle y}{\triangle t}) = \frac{dy}{dx} \cdot \frac{dx}{dt}$.

Case 1: $\triangle x \neq 0$

We have $\frac{\triangle y}{\triangle t} = \frac{\triangle y}{\triangle x} \cdot \frac{\triangle x}{\triangle t}$, so $st(\frac{\triangle y}{\triangle t}) = st(\frac{\triangle y}{\triangle x}) \cdot st(\frac{\triangle x}{\triangle t})$,
and again, $st(\frac{\triangle y}{\triangle t}) = \frac{dy}{dx} \cdot \frac{dx}{dt}$.

$\square$

The use of infinitesimals is now a valuable tool in both pure and applied mathematics, allowing us to follow the intuition of Leibniz and Euler, yet maintaining complete rigor. An important concept is the Leibniz Transfer Principle: any statement true for all real numbers must also be true for all hyperreal numbers. While we will not here make this more precise, it would, for example, entail that if

$$F\colon \mathbb{R}^n \to \mathbb{R} \text{ and } G\colon \mathbb{R}^n \to \mathbb{R}$$

then

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

Here ${}^*F\colon ({}^*\mathbb{R})^n \to {}^*\mathbb{R}$ is the natual extension of $F\colon \mathbb{R}^n \to \mathbb{R}$ 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 $j \in \omega$ we have $\vec{s}(j) = \langle s_0(j), \ldots, s_{n-1}(j)\rangle$ and $s_i \in a_i$ for each i.

But...let's not go there now.

Chapter 10

The Universe

Well, not *THE* universe — just the mathematical one. Nevertheless, $\mathbb{V}$ is quite complicated and it would help to have some way of dividing $\mathbb{V}$ into more manageable pieces, perhaps a hierarchy of collections of sets of greater and greater complexity. 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 procedure 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 demonstrate.

It is easy to see that a set X is transitive iff $\bigcup X \subseteq X$ and this motivates the following definition by recursion on $\mathbb{N}$:

$$U(X, 0) = X \text{ and } U(X, n+1) = \bigcup U(X, n) \text{ if } n \in \mathbb{N},$$

and using this we define the <u>transitive closure</u> of X as:

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

It is now straightforward to check that, for any set X, $trcl(X)$ is a transitive set and that it is, in fact, the smallest transitive set including X as a subset.

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

For each formula $\Phi(n, w_1, \ldots, w_k)$ of the language of set theory we have:

Theorem 38. Φ

For all $w_1, \ldots, w_k$, if

$$\forall n \in \mathbb{V}\ [(\forall m \in n\ \Phi(m)) \to \Phi(n)]$$

then

$$\forall n \in \mathbb{V}\ \Phi(n).$$

Proof. We will assume that the theorem is false and derive a contradiction. Take any fixed $w_1, \ldots, w_k$ and a fixed $l \in \mathbb{V}$ such that $\neg\Phi(l)$. Let t be any transitive set containing l, e.g. $t = trcl(\{l\})$.

The proof now proceeds verbatim as the proofs of Theorems 8 and 13.

□

The Recursion Principle for the Universe

We can also carry out recursive definitions over all of $\mathbb{V}$. This process is called $\underline{\in -\text{recursion}}$. For any formula $\Phi(x, f, y, w_1, \ldots, w_k)$ of the language of set theory, we denote by $REC(\Phi, \mathbb{V}, w_1, \ldots, w_k)$ the class

$$\bigcup\{f : (\exists n \in \mathbb{T})[f : n \to \mathbb{V} \wedge \forall m \in n\ \Phi(m, f|m, f(m))]\}.$$

where $\mathbb{T}$ denotes $\{n : n \text{ is transitive}\}$.

Analogous to ordinary recursion and transfinite recursion, this recursion is justified by a theorem scheme.

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

Theorem 39. Φ

For all $w_1, \ldots, w_k$ suppose that we have

$$(\forall x \in \mathbb{V})(\forall f)[(f\colon x \to \mathbb{V}) \to \exists! y\ \Phi(x, f, y, w_1, \ldots, w_k)].$$

Then, letting F denote the class $REC(\Phi, \mathbb{V}, w_1, \ldots, w_k)$, we have:

1. $F\colon \mathbb{V} \to \mathbb{V}$;

2. $\forall m \in \mathbb{V}\ \Phi(m, F|m, F(m), w_1, \ldots, w_k)$.

Proof. This proof, like that of Theorem 14 is almost identical to the proof of Theorem 9. Just replace $\mathbb{N}$ by $\mathbb{V}$ or $\mathbb{T}$ where appropriate. $\square$

The main application of $\in$-recursion is to define our first new measure of the complexity of a set, the rank function. This associates, to each set x, an ordinal $rank(x)$ by the following rule:

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

By Theorem 39 this is a recursive definition of a function:

$$rank\colon\ \mathbb{V}\ \to\ \mathbb{V}$$

and by $\in$-induction, Theorem 38, we have that

$$rank\colon\ \mathbb{V}\ \to\ \mathbb{ON}$$

Note that $rank(\emptyset) = 0$ and in fact, by transfinite induction we immediately see that $rank(\alpha) = \alpha$ for all $\alpha \in \mathbb{ON}$.

Since $x \in y$ implies that $rank(x) < rank(y)$, the $rank$ function is a global witness to the well founded nature of the membership relation in the sense of Theorem 18.

The cumulative hierarchy, an ordinal-gradation on $\mathbb{V}$ due to J. von Neumann, is defined by by recursion on $\mathbb{ON}$.

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

Sometimes we write $\mathbf{R}_\alpha$ or $\mathbb{V}_\alpha$ for $\mathbf{R}(\alpha)$. The next lemma gives two basic properties of the cumulative hierarchy. The proofs are straightforward transfinite inductions.

Lemma.

1. $\forall \alpha \in \mathbb{ON}\ \mathbf{R}(\alpha)$ is transitive.
2. $\forall \alpha \in \mathbb{ON}\ \forall \beta \in \mathbb{ON}\ \beta < \alpha \rightarrow \mathbf{R}(\beta) \subseteq \mathbf{R}(\alpha.$

Using the concept of rank we obtain this important theorem.

Theorem 40. $\mathbb{V} = \bigcup\{\boldsymbol{R}(\alpha) : \alpha \in \mathbb{ON}\}$*; i.e.,* $\forall x\ \exists \alpha \in \mathbb{ON}\ x \in \boldsymbol{R}(\alpha)$.

Proof. Since every set has a rank, the theorem will follow from this claim:

$$\{x : rank(x) < \alpha\} \subseteq \mathbf{R}(\alpha)$$

which we prove by transfinite induction.

If $\alpha = \beta + 1$ then, for any x with $rank(x) < \alpha$ we have:

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

Thus, $rank(y) < \beta$ for all $y \in x$, and by inductive hypothesis, $y \in \mathbf{R}(\beta)$. That is, $x \subseteq \mathbf{R}(\beta)$ and so $x \in \mathcal{P}(\mathbf{R}(\beta)) = \mathbf{R}(\alpha)$.

On the other hand, if α is a limit ordinal, then for any x with $rank(x) < \alpha$ we have $rank(x) < \beta$ for some $\beta < \alpha$ and so

$$x \in \bigcup\{\mathbf{R}(\beta) : \beta < \alpha\} = \mathbf{R}(\alpha).$$

□

Exercise 33. Verify these finer points about the rank hierarchy.

1. $\forall x\ \forall \alpha \in \mathbb{ON}\ (x \in \mathbf{R}(\alpha) \leftrightarrow \exists \beta \in \alpha\ x \subseteq \mathbf{R}(\beta))$.
2. $\forall x\ \forall \alpha \in \mathbb{ON}\ (x \in \mathbf{R}(\alpha + 1) \setminus \mathbf{R}(\alpha)\ \leftrightarrow\ rank(x) = \alpha)$.

We have discussed cardinality; it is also a way of measuring the complexity of a set. However, if the set is not transitive, cardinality does not tell the whole story since it cannot distinguish among the elements of the set. Some elements may have larger cardinality than others than others. For example, although $\mathbb{N} \in \{\mathbb{N}\}$, $|\mathbb{N}| = \aleph_0$ while $|\{\mathbb{N}\}| = 1$. However, in truth, if we look deeper, $\{\mathbb{N}\}$ is no less complicated than $\mathbb{N}$. We are led to 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,

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

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.

Theorem 41.

1. *For any infinite cardinal κ, $\exists z \ z = \boldsymbol{H}(\kappa)$.*
2. *$\mathbb{V} = \bigcup\{\boldsymbol{H}(\kappa) : \ \kappa$ is a cardinal$\}$.*

Proof. The first part follows from the exercise below. The second part follows from the fact that every set x has a transitive closure $trcl(x)$ and hence a hereditary cardinality $hcard(x) = \lambda$; so $x \in \mathbf{H}(\lambda^+)$. □

Exercise 34. Prove these three statements and show how they lead to a proof of the first part of Theorem 41.

1. For any infinite cardinal κ, $\mathbf{H}(\kappa)$ is transitive.
2. For any infinite cardinal κ, $\forall x \ hcard(x) < \kappa \rightarrow rank(x) < \kappa$.
3. For any infinite cardinal κ, $\mathbf{H}(\kappa) \subseteq \mathbf{R}(\kappa)$.

Theorem 42. *If either $\kappa = \omega$ or κ is an inaccessible cardinal, then $\boldsymbol{H}(\kappa) = \boldsymbol{R}(\kappa)$.*

Proof. From Exercise 34 we have $\mathbf{H}(\kappa) \subseteq \mathbf{R}(\kappa)$.

For the reverse inclusion, let $x \in \mathbf{R}(\kappa)$. Since κ is a limit ordinal there is some $\alpha < \kappa$ such that $x \in \mathbf{R}(\alpha)$. Since $\mathbf{R}(\alpha)$ is transitive $x \subseteq \mathbf{R}(\alpha)$ and so $trcl(x) \subseteq \mathbf{R}(\alpha)$.

So 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 the corollary to Theorem 25, observing that κ is regular. $\square$

It follows from this theorem that for each inaccessible cardinal κ, $\mathbf{H}(\kappa)$ is a transitive set closed under the operations of pairing, taking the power set and taking the union of a collection of elements of $\mathbf{H}(\kappa)$ which is indexed by a member of $\mathbf{H}(\kappa)$ and is therefore, by definition, a <u>Grothendieck universe</u>.

Exercise 35. Show that the following two statements (one from Set Theory, one from Algebraic Geometry) are equivalent.

1. $\forall\alpha \in \mathbb{ON}\ \exists\kappa > \alpha\ \kappa$ is an inaccessible cardinal.
2. $\forall x\ \exists U\ x \in U$ and U is a Grothendieck universe.

Chapter 11

Reflection

It is natural to wonder how close to $\mathbb{V}$ are the approximations given by rank and by hereditary cardinality. We ask: which statements true for $\mathbb{V}$ are also true for the various $\mathbf{H}(\kappa)$ and $\mathbf{R}(\kappa)$, i.e. which true statements are also true when relativised to these sets? We are particularly interested in the formulas of our $\mathcal{ZFC}$ axiom system which includes: Existence, Extensionality, Pairing, Union, Intersection, Foundation, Choice, Power Set, Infinity, and the Replacement Scheme. Which of these truths of $\mathbb{V}$ reflect down to an $\mathbf{H}(\kappa)$ or an $\mathbf{R}(\kappa)$?

We will rely upon the definition of relativisation given in Chapter 1. For reasons explained there, we will only speak about relativising to non-empty classes.

Whenever Φ is a formula with no free variables, we write $M \models \Phi$ for Φ^M and we say that M is a model of Φ or that Φ is true in M. More generally, if $\Phi(v_0, \ldots, v_k)$ does have have free variables and $\{m_0, \ldots, m_k\} \subseteq M$, we write

$$M \models \Phi(v_0|m_0, \ldots, v_k|m_k) \text{ or just } M \models \Phi(m_0, \ldots, m_k)$$

provided that

$$(\Phi(v_0|m_0, \ldots, v_k|m_k))^M \text{ holds.}$$

Notice that if the variable v_i does not occur free in the formula Φ then m_i does not play a role and can be omitted from the string $(m_0, \ldots, m_k)$. A formula of LOST with no free variables is called a sentence.

Since we are only relativising to non-empty M, each one is a model of the Axiom of Existence.

In the theorem schemes used to show $M \models \Phi$, we use a new proof technique, called *induction on complexity of a formula*. We prove something true for Φ using the fact that it is true for all subformulas of Φ. This is best first illustrated in a simple, straightforward situation.

For each sentence Φ of LOST we have:

Theorem 43. Φ

Let A and B be sets and $f\colon A \to B$ be an isomorphism (with respect to the membership relation). Then: $A \models \Phi$ iff $B \models \Phi$

Proof. Let $f\colon A \to B$ be fixed as in the statement of the theorem. We would like to prove this theorem by a kind of induction over the subformulas of Φ, the formulas used in the construction of the formula Φ. But we cannot use the statement of the theorem itself as any kind of an inductive hypothesis because the statement relates only to a formula with no free variables. Although Φ has no free variables, its subformulas certainly will have free variables; in fact all occurrences of variables in atomic formulas are always free.

So we must create a variation of the theorem which will be meaningful for any subformula Ψ of Φ and which will immediately give us the theorem when Ψ is Φ. This is the lemma scheme which follows. It is a *finite* lemma scheme — one lemma for each of the finitely many subformulas Ψ of our formula Φ. Each of these finitely many lemmas plays a role in the proof of the theorem for Φ. □

Let all the variables occurring in the formula Φ lie among $v_0, \ldots, v_k$. For each subformula Ψ of Φ we have

Lemma. Ψ

Let $f\colon A \to B$ be an isomorphism. For all $a_0, \ldots, a_k$ in A:

$$A \models \Psi(v_0|a_0, \ldots, v_k|a_k) \ \text{ iff } \ B \models \Psi(v_0|b_0, \ldots, v_k|b_k)$$

where each $b_0 = f(a_0), \ldots, b_k = f(a_k)$.

Proof. We use induction on the complexity of Ψ.

There are two main steps. The base step proves all those instances of the lemma scheme for which Ψ is an atomic subformula of Φ .

The inductive step proves the instances of the lemma scheme for each subformula Ψ which is built from other subformulas Θ and Ω by an application of a connective $\neg$, $\wedge$, $\vee$ or $\rightarrow$, or a quantifier $\forall$ or $\exists$. In this inductive step we are permitted to assume that the instances of the lemma scheme for Θ and for Ω are already known to be true; that is, the proof of the instance of lemma scheme for Ψ will rely upon the instances of the lemma scheme for Θ and for Ω.

We begin with the base step. There are two cases because there are two types of atomic formulas.

1. When Ψ is the atomic formula $v_i = v_j$ where $i \leq k$ and $j \leq k$ we need to show that for all $a_0, \ldots, a_k$ in A

$$A \models \Psi(a_0, \ldots, a_k) \ \text{ iff } \ B \models \Psi(b_0, \ldots, b_k).$$

   But this is just

$$a_i = a_j \ \text{ iff } \ b_i = b_j,$$

   which, in turn, is just

$$a_i = a_j \ \text{ iff } \ f(a_i) = f(a_j),$$

   and this is true since f is an injection.

2. When Ψ is the atomic formula $v_i \in v_j$ where $i \leq k$ and $j \leq k$ we need to show that for all $a_0, \ldots, a_k$ in A

$$A \models \Psi(a_0, \ldots, a_k) \ \text{ iff } \ B \models \Psi(b_0, \ldots, b_k).$$

   But this is just

$$a_i \in a_j \ \text{ iff } \ b_i \in b_j,$$

which, in turn, is just

$$a_i \in a_j \text{ iff } f(a_i) \in f(a_j),$$

and this is true since f is an isomorphism.

We now turn to the inductive step. There are six cases, one for each of the four connectives $\neg$, $\wedge$, $\vee$ and $\rightarrow$, as well as one for each of the two quantifiers $\forall$ and $\exists$. We will assume, as an inductive hypothesis, that for all $a_0, \ldots, a_k$ in A:

$$A \models \Theta(v_0|a_0, \ldots, v_k|a_k) \text{ iff } B \models \Theta(v_0|b_0, \ldots, v_k|b_k)$$

and

$$A \models \Omega(v_0|a_0, \ldots, v_k|a_k) \text{ iff } B \models \Omega(v_0|b_0, \ldots, v_k|b_k)$$

where each $b_0 = f(a_0), \ldots, b_k = f(a_k)$.

1. In the case that Ψ is $(\neg\Theta)$, we prove

$$A \models \Psi(a_0, \ldots, a_k) \text{ iff } B \models \Psi(b_0, \ldots, b_k)$$

by a chain of equivalences. We begin with: for all $a_0, \ldots, a_k$ in A

$$A \models (\neg\Theta)(v_0|a_0, \ldots, v_k|a_k)$$

using the definition of relativisation in Chapter 1, we obtain

$$\text{iff it is not true that } A \models \Theta(v_0|a_0, \ldots, v_k|a_k))$$

and now using the inductive hypothesis applied to Θ, we have

$$\text{iff it is not true that } B \models \Theta(v_0|b_0, \ldots, v_k|b_k))$$

again using the definition of relativisation, we finish with

$$\text{iff } B \models (\neg\Theta)(v_0|b_0, \ldots, v_k|b_k).$$

2. In the case that Ψ is $(\Theta \wedge \Omega)$, we use a chain of equivalences as in the previous case. We begin with: for all $a_0, \ldots, a_k$ in A

$$A \models (\Theta \wedge \Omega)(v_0|a_0, \ldots, v_k|a_k)$$

using the definition of relativisation, we get

$$\text{iff} \quad \text{both } A \models \Theta(v_0|a_0, \ldots, v_k|a_k) \text{ and } A \models \Omega(v_0|a_0, \ldots, v_k|a_k)$$

by the inductive hypothesis applied to Θ and to Ω, we get

$$\text{iff} \quad \text{both } B \models \Theta(v_0|b_0, \ldots, v_k|b_k) \text{ and } B \models \Omega(v_0|b_0, \ldots, v_k|b_k)$$

again using the definition of relativisation, we finish with

$$\text{iff} \quad B \models (\Theta \wedge \Omega)(v_0|b_0, \ldots, v_k|b_k).$$

3. The case that Ψ is $(\Theta \vee \Omega)$ is similar to the previous case.

4. The case that Ψ is $(\Theta \rightarrow \Omega)$ is similar to the previous cases.

5. In the case that Ψ is $(\forall v_j)\Theta$, we once more use a chain of equivalences. We begin with: for all $a_0, \ldots, a_k$ in A

$$A \models (\forall v_j)\Theta(v_0|a_0, \ldots, v_k|a_k).$$

But since v_j does not occur free in $(\forall v_j)\Theta$, we have

$$\text{iff} \quad A \models (\forall v_j)\Theta(v_0|a_0, \ldots, v_{j-1}|a_{j-1}, v_{j+1}|a_{j+1}, \ldots v_k|a_k).$$

Now by the definition of relativisation and the simple logical equivalence obtained by replacing the bound variable v_j by the new variable a, we obtain

$$\text{iff} \quad \text{for all } a \text{ in } A: \ A \models \Theta(v_0|a_0, \ldots, v_j|a, \ldots, v_k|a_k)$$

We now invoke the inductive hypothesis applied to Θ. Notice that here we use the fact that f is a surjection.

$$\text{iff} \quad \text{for all } b \text{ in } B: \ B \models \Theta(v_0|b_0, \ldots, v_j|b, \ldots, v_k|b_k)$$

Again by a simple logical equivalence and the definition of relativisation we get

$$\text{iff} \quad B \models (\forall v_j)\Theta(v_0|b_0, \ldots, v_{j-1}|b_{j-1}, v_{j+1}|b_{j+1}, \ldots v_k|b_k)$$

and since v_j does not occur free in $(\forall v_j)\Theta$ we finish with

$$\text{iff} \quad B \models (\forall v_j)\Theta(v_0|b_0, \ldots, v_k|b_k)$$

6. The case that Ψ is $(\exists v_j)\Theta$ is similar to the previous case and so is left for the reader.

$\square$

We say that a formula $\Phi(v_0, \ldots, v_k)$ is absolute for M provided that for all $m_0, \ldots, m_k$ in M we have:

$$M \models \Phi(m_0, \ldots, m_k) \text{ iff } \Phi(m_0, \ldots, m_k).$$

By the definition of relativisation, every atomic formula is absolute for any term M; in fact, any formula without quantifiers is absolute for any M. It will become valuable for us to know which formulas Φ are absolute for which models M, especially transitive sets M.

Let's look at a detailed example: a non-empty transitive set M and the formula "$v_0 = \emptyset$". Formally, to show that the formula is absolute for M we need to verify that for any m in M:

$$M \models (v_0 = \emptyset)(v_0|m) \text{ iff } (v_0 = \emptyset)(v_0|m)$$

which is just a long way of writing:

$$\text{for all } m \in M \quad M \models m = \emptyset \text{ iff } m = \emptyset$$

To prove this, we must unravel the abbreviations of LOST. Let $m \in M$:

$$M \models m = \emptyset \text{ iff } M \models m = \{x : x \neq x\}$$

$$\text{iff } M \models (\forall x)(x \in m \leftrightarrow x \neq x) \text{ iff } \forall x \in M \; M \models (x \in m \leftrightarrow x \neq x)$$

$$\text{iff } \forall x \in M \; [M \models x \in m \Leftrightarrow M \models x \neq x] \text{ iff } \forall x \in M \; [x \in m \Leftrightarrow x \neq x]$$

$$\text{iff } \forall x \in M \; x \notin m \text{ iff } M \cap m = \emptyset \text{ iff } m = \emptyset.$$

The assumption that M is transitive only gets used for the last step.

As another example, let's check that "being disjoint" is absolute for transitive models. Let M be a transitive set containing a and b.

$$M \models a \cap b = \emptyset \text{ iff } M \models \{u : (u \in a) \wedge (u \in b)\} = \{u : u \neq u\}$$

$$\text{iff} \quad M \models (\forall u)((u \in a \ \wedge u \ \in b) \ \leftrightarrow \ u \neq u)$$
$$\text{iff} \quad \forall u \in M \ [(M \models u \in a \text{ and } M \models u \in b) \ \Leftrightarrow \ M \models u \neq u]$$
$$\text{iff} \quad \forall u \in M \ [(u \in a \text{ and } u \in b) \Leftrightarrow u \neq u]$$
$$\text{iff} \quad \forall u \in M \ [u \notin a \cap b] \ \text{ iff } \ a \cap b \cap M = \emptyset \ \text{ iff } \ a \cap b = \emptyset$$

where, again, transitivity is only invoked for the last step.

The Axiom of Extensionality and the Axiom of Foundation are absolute for transitive models. Since they are each true (by assumption) all we need to verify is the following lemma.

Theorem 44. *Any transitive set M is a model of the Axiom of Extentionality and the Axiom of Foundation.*

Proof. M is a model of the Axiom of Extensionality because the membership relation is extensional on any transitive set. We address the Axiom of Foundation.

$$M \models \forall x \ [x \neq \emptyset \ \rightarrow \ (\exists y \in x)(x \cap y = \emptyset)]$$
$$\text{iff} \quad \forall x \in M \ [M \models x = \emptyset \text{ or } M \models (\exists y \in x)(x \cap y = \emptyset)].$$

We use the above examples to help us get:

$$\text{iff} \quad \forall x \in M \ [x = \emptyset \text{ or } (\exists y \in x \cap M)(x \cap y = \emptyset)].$$

Now since M is transitive, we get:

$$\text{iff} \quad \forall x \in M[x = \emptyset \text{ or } (\exists y \in x)(x \cap y = \emptyset)],$$

and this obviously follows from the Axiom of Foundation. □

Before continuing to investigate models, we pause to extend our analysis of the formulas of LOST; this will make our later work a lot easier.

A <u>bounded formula</u> (also called a $\triangle_0$ formula or a restricted formula) is one which is built up as usual with respect to atomic formulas and connectives, but where the $(\exists v_i)\Phi$ clause is replaced by $(\exists v_i \in v_j)\Phi$, and the $(\forall v_i)\Phi$ clause is replaced by $(\forall v_i \in v_j)\Phi$. Thus each *bound* variable is *bounded* by another variable.

Bounded formulas are absolute for transitive models. We prove this formally as a theorem scheme. For each bounded formula $\Phi(v_0, \ldots, v_k)$ we have:

Theorem 45. Φ *If M is a (non-empty) transitive set and $\{m_0, \ldots, m_k\} \subseteq M$, then*

$$M \models \Phi(m_0, \ldots, m_k) \text{ iff } \Phi(m_0, \ldots, m_k)$$

Proof. We again use the technique of induction on complexity of the formula. This time the atomic formula step is trivial. The connective cases in the inductive step slide by with no problem. Only the quantifier cases of the inductive step require some discussion. In fact, these two cases are similar to each other, so we will only address the existential quantifier case.

Since Φ is a bounded formula we can suppose that Φ is $(\exists v_i \in v_j)\Theta$ where v_i and v_j are among $v_0, \ldots, v_k$. Without loss of generosity we shall assume that v_i and v_j are different variables.

Let M and $\{m_0, \ldots, m_k\}$ be as in the statement of the theorem. We assume as inductive hypothesis that

$$M \models \Theta(m_0, \ldots, m_k) \text{ iff } \Theta(m_0, \ldots, m_k)$$

and carry out the proof through a chain of equivalences.

Since v_i does not occur free in Φ

$$M \models \Phi(m_0, \ldots, m_k) \text{ iff } M \models \Phi(m_0, \ldots, m_{i-1}, m_{i+1}, \ldots, m_k)$$

and when the abbreviation is unraveled we obtain

$$M \models (\exists v_i)((v_i \in v_j) \wedge \Theta)(m_0, \ldots, m_{i-1}, m_{i+1}, \ldots, m_k).$$

We now use the definition of relativisation and change the bound variable v_i to the new variable m to get

$$\exists m \in M \quad M \models (m \in v_j) \wedge \Theta)(m_0, \ldots, m_{i-1}, m, m_{i+1}, \ldots, m_k).$$

Again using the definition of relativisation gives

$$\exists m \in M \quad M \models (m \in m_j) \text{ and } M \models \Theta(m_0, \ldots, m_{i-1}, m, m_{i+1}, \ldots, m_k).$$

Using the inductive hypothesis, this simplifies to

$$\exists m \in M \ m \in m_j \text{ and } \Theta(m_0, \ldots, m_{i-1}, m, m_{i+1}, \ldots, m_k).$$

We now use that M is transitive; $m_j \subseteq M$, so we get the equivalence

$$\exists m \in m_j \ \Theta(m_0, \ldots, m_{i-1}, m, m_{i+1}, \ldots, m_k).$$

This immediately gives us

$$(\exists v_i \in v_j)\Theta(m_0, \ldots, m_{i-1}, m_{i+1}, \ldots, m_k),$$

and since v_i does not occur free, we finish with

$$(\exists v_i \in v_j)\Theta(m_0, \ldots, m_k).$$

□

Many mathematical concepts can be expressed by bounded formulas. Many more concepts have definitions which are formulas that are logically equivalent to bounded formulas. That is, the original formula Φ can be proved to be equivalent to a bounded formula Φ^* by purely logical means without invoking any of the axioms of $\mathfrak{ZFC}$. According to the discussion at the end of Chapter 1, this equivalence will persist to the relativised formulas, so that $\Phi^M \iff (\Phi^*)^M$ and we have:

$$\Phi \iff \Phi^* \iff M \models \Phi^* \iff M \models \Phi$$

which shows that Φ is absolute for M. As a matter of fact, for transitive models M, according Theorem 44, we would be permitted to invoke the Axioms of Extensionality and Foundation in the proof of equivalence.

For example $x = \emptyset$ is formally an abbreviation for a formula which is not a bounded formula:

$$(\forall v)(((v \in x) \to (v \neq v)) \ \wedge \ ((v \neq v) \to (v \in x))).$$

But $v \neq v$ is false (by logical assumption) so that $((v \neq v) \to (v \in x))$ is logically true. Hence $x = \emptyset$ is logically equivalent to $(\forall v)((v \in x) \to (v \neq v))$ which is the bounded formula $(\forall v \in x)(v \neq v)$.

The formula $y = succ(x)$ is $(\forall v)[(v \in y) \leftrightarrow (v \in x \cup \{x\})$ and this is logically equivalent to:

$$[x \in y] \ \wedge \ [(\forall v \in x)(v \in y)] \ \wedge \ (\forall v \in y)[(v \in x) \vee (v = x)]$$

which is a bounded formula. As a consequence we see that the formula "n is a natural number" is also logically equivalent to a bounded formula. Other examples of formulas logically equivalent to bounded formulas include:

$$\begin{array}{rcl} x \text{ is transitive} & \text{iff} & (\forall y \in x)(\forall z \in y)(z \in x) \\ x \text{ is an ordinal} & \text{iff} & (x \text{ is transitive}) \wedge (\forall y \in x)(y \text{ is transitive}) \\ z = \{x, y\} & \text{iff} & (x \in z) \wedge (y \in z) \wedge (\forall u \in z)(u = x \vee u = y) \end{array}$$

Of course, applying the connectives ($\neg$, $\wedge$, $\vee$ and $\rightarrow$) to bounded formulas results in another bounded formula. In particular, the formula $x \in \mathbb{ON} \setminus \mathbb{N}$ is, by Theorem 45 and the discussion above, absolute for any transitive model. We use this observation below to show that a transitive set is a model of the Axiom Of Infinity exactly when it contains ω.

Theorem 46. *Let M be transitive. $\omega \in M$ iff $M \models \mathbb{N} \neq \mathbb{ON}$.*

Proof. Suppose $\omega \in M$. Then $(\exists x \in M)(x \in \mathbb{ON} \setminus \mathbb{N})$. By absoluteness

$$(\exists x \in M) M \models (x \in \mathbb{ON} \setminus \mathbb{N})$$

and so $M \models (\exists x)(x \in \mathbb{ON} \setminus \mathbb{N})$ so $M \models \mathbb{N} \neq \mathbb{ON}$.

Conversely, suppose that

$$M \models (\exists x)(x \in \mathbb{ON} \setminus \mathbb{N}) \ \vee \ (\exists x)(x \in \mathbb{N} \setminus \mathbb{ON})$$

By absoluteness, we get:

$$(\exists x \in M)(x \in \mathbb{ON} \setminus \mathbb{N}) \text{ or } (\exists x \in M)(x \in \mathbb{N} \setminus \mathbb{ON}).$$

Since $\mathbb{N} \setminus \mathbb{ON} = \emptyset$, we must have the first statement holding. This means that M contains an ordinal $\alpha \geq \omega$. Since M is transitive $\alpha \subseteq M$ and so $\omega \in M$. ∎

The formula $p = \langle x, y \rangle$, that is

$$(\forall z)[(z = \{x\} \vee z = \{x, y\}) \rightarrow z \in p] \wedge (\forall z \in p)(z = \{x\} \vee z = \{x, y\})$$

is not absolute for all transitive models M. In fact it is is not absolute for the transitive set $M = 1 = \{\emptyset\}$. This means that the formula cannot be logically equivalent to a bounded formula, even using the axioms of Extensionality and Foundation. Nevertheless, using the Axiom of Pairing the formula is equivalent to

$$(\exists z \in p)(z = \{x\}) \wedge (\exists z \in p)(z = \{x, y\}) \wedge (\forall z \in p)(z = \{x\} \vee z = \{x, y\})$$

and is therefore absolute for transitive models of the Axiom of Pairing.

Exercise 36. Verify the statements in the last paragragh.

This immediately gives more formulas which are logically equivalent to bounded formulas assuming the Axioms of Extensionality, Foundation and Pairing.

$$\begin{array}{rcl} p \text{ is an ordered pair} & \text{iff} & (\exists u \in p)(\exists x \in u)(\exists v \in p)(\exists y \in v)(p = \langle x, y\rangle) \\ f \text{ is a function} & \text{iff} & (\forall p \in f)(p \text{ is an ordered pair}) \text{ and} \end{array}$$

$$(\forall p \in f)(\forall q \in f)[(\exists u \in p)((u \in q) \wedge (\exists x \in u)(u = \{x\})) \to (p = q)]$$

$$\begin{array}{rcl} x \in dom(f) & \text{iff} & (\exists p \in f)[(p \text{ is an ordered pair}) \wedge (\exists u \in p)(u = \{x\}) \\ X \subseteq dom(f) & \text{iff} & (\forall x \in X)(x \in dom(f)) \\ dom(f) \subseteq X & \text{iff} & (\forall p \in f)(\forall u \in p)(\forall x \in u)[x \in dom(f) \to x \in X] \\ f\colon X \to \mathbb{ON} & \text{iff} & f \text{ is a function and } dom(f) \subseteq X \text{ and } X \subseteq dom(f) \text{ and} \end{array}$$

$$(\forall x \in X)(\exists p \in f)(\exists v \in p)(\exists y \in v)[(p = \langle x, y\rangle) \wedge (y \text{ is an ordinal})]$$

A transitive model M of the Axiom of Pairing is simply a transitive set such that

$$\forall x \in M\ \forall y \in M\ \exists z \in M\ z = \{x, y\}$$

which means that M is closed under pairing, the formation of pairs. So these formulas above are absolute for transitive models closed under pairing, for example $\mathbf{H}(\theta)$ for an infinite cardinal θ or $\mathbf{R}(\lambda)$ for a limit ordinal λ.

The following observation will be helpful. For any formula $\Phi(v, w_1, \ldots, w_k)$ of LOST, we have:

Lemma. Φ

If M is a transitive set and $z \in M$, then for all $m_1, \ldots, m_k$ in M we have:

$$M \models z = \{v : \Phi(v, m_1, \ldots, m_k)\} \text{ iff } z = \{v \in M : \Phi^M(v.m_1, \ldots, m_k)\}.$$

Exercise 37. Prove this lemma scheme and use it to show that whenever λ is a limit ordinal $\mathbf{R}(\lambda)$ is a model of the Intersection Axiom, the Union Axiom, the Axiom of Choice and the Power Set Axiom.

Let us now investigate whether a transitive set M is a model of an instance of the Axiom of Replacement scheme for a formula $\Phi(x, u, v, w_1, \ldots, w_k)$. To this end, let $w_1, \ldots, w_k$ and x be in M and assume that

$$M \models (\forall u \in x)(\exists! v)\Phi$$

while we attempt to show that

$$M \models \exists z \; z = \{v : \exists u \in x \; \Phi\}.$$

From our assumption, after a little unravelling we get:

$$(\forall u \in M \cap x)(\exists v \in M) M \models \Phi \wedge (\forall t)(\Phi(v|t) \rightarrow v = t)$$

which, since M is transitive, is equivalent to:

$$(\forall u \in x)(\exists v)[v \in M \;\wedge\; \Phi^M \;\wedge\; (\forall t)(t \in M \;\wedge\; \Phi^M(v|t) \rightarrow v = t)].$$

Notice that this does *not* give us the right to invoke Replacement for the formula Φ^M because we are not guaranteed a unique v for each u; v will only be unique among the elements of M, not necessarily among all sets.

However, we are not at a standstill. Let Ψ be the formula $(v \in M) \wedge \Phi^M$. We can rewrite the assumption as

$$(\forall u \in x)(\exists v)[\Psi \;\wedge\; (\forall t)(\Psi(v|t) \rightarrow v = t)]$$

which is $(\forall u \in x)(\exists! v)\Psi$ and this enables us to use the instance of Replacement for the formula Ψ. Doing this, we obtain:

$$\exists z \; z = \{v : (\exists u \in x)\Psi\}$$

which when put in terms of our original formula Φ is:

$$\exists z\ z = \{v : (\exists u \in x)\ (v \in M) \wedge \Phi^M\}.$$

This is logically equivalent to

$$\exists z\ z = \{v \in M : (\exists u \in x)\Phi^M\}.$$

which, since M is transitive and $x \subseteq M$ gives:

$$\exists z\ z = \{v \in M : ((\exists u \in x)\Phi)^M\}.$$

Now we are close; if there were such a z in M, by the previous lemma we would immediately get:

$$M \models \exists z\ z = \{v : \exists u \in x\ \Phi\}.$$

which is exactly what we are seeking. But of course, we are not always guaranteed that there is such a z in M. For example, $\mathbf{R}(\omega + \omega)$ does not contain the set

$$\{v \in \mathbf{R}(\omega + \omega) : (\exists u \in \omega)(v = \omega + u)^{\mathbf{R}(\omega+\omega)}\}$$

and so $\mathbf{R}(\omega + \omega)$ is not a model of this instance of the Replacement scheme. Nevertheless, to show that the instance of Replacement for the formula Φ holds in a transitive model M we only need to verify that for all $x \in M$:

$$\{v \in M : (\exists u \in x)\Phi^M\} \in M.$$

For each axiom Φ of $\mathcal{ZFC}$ except for the Axiom of Infinity and the Power Set Axiom, we have the following.

Theorem 47. Φ

For each regular cardinal κ $\boldsymbol{H}(\kappa) \models \Phi$.

Exercise 38. Verify this theorem scheme.

Since $\mathbf{H}(\omega) = \mathbf{R}(\omega)$ a consequence of all this work is that the hereditarily finite sets, $\mathbf{H}(\omega)$ is a model of each axiom of $\mathcal{ZFC}$ except for the Axiom of

Infinity. This allows us to draw some conclusions about the $\mathcal{ZFC}$ axioms which cannot be expressed as formulas of LOST. These will be statements *about* Set Theory rather than statements *of* Set Theory.

Gödel's Second Incompleteness Theorem tells us that even if the $\mathcal{ZFC}$ axioms are consistent we cannot prove this from these same $\mathcal{ZFC}$ axioms alone. Nevertheless, assuming only the $\mathcal{ZFC}$ axioms we can infer the consistency of the collection of all of $\mathcal{ZFC}$ except the Axiom of Infinity. We reason as follows. If this collection of axioms were inconsistent they would entail a proof of the formula $(0 = 1)$. This proof could involve only finitely many axioms from the collection and we form the formula Θ as the conjunction of these finitely many axioms. According to the discussion at the end of Chapter 1, the relativisation of Θ to a non-empty set would entail the relativisation of $(0 = 1)$ to that same set. In particular, we would get:

$$\mathbf{H}(\omega) \models \Theta \;\Rightarrow\; \mathbf{H}(\omega) \models (0 = 1)$$

But we have that $\mathbf{H}(\omega) \models \Theta$ because $\mathbf{H}(\omega)$ is a model of each of the finitely many formulas in the conjunction. This gives a contradiction since $\mathbf{H}(\omega)$ is definitely not a model of $(0 = 1)$.

In a similar way, we can also show that the Axiom of Infinity is not redundant — it does not follow from the remainder of the $\mathcal{ZFC}$ axioms. If it did, then as above, we would get a formula Θ, a conjunction of finitely many of these axioms and which logically implies the Axiom of Infinity. But, as above, since $\mathbf{H}(\omega) \models \Theta$ we would get that $\mathbf{H}(\omega)$ would be a model of the Axiom of Infinity. Since $\omega \notin \mathbf{H}(\omega)$ this would contradict Theorem 46.

We also have $\mathbf{H}(\kappa) = \mathbf{R}(\kappa)$ whenever κ is an inaccessible cardinal. Using $\mathbf{H}(\kappa)$ instead of $\mathbf{H}(\omega)$ in the arguments above gives us that with the addition of the Axiom of Inaccessibles to $\mathcal{ZFC}$ we can show the consistency of $\mathcal{ZFC}$. Furthermore, we can use the result of the following exercise to show that the Axiom of Inacessibles does not follow from $\mathcal{ZFC}$.

Exercise 39. Prove that if κ is the least inaccessible cardinal then

$$\mathbf{H}(\kappa) \models \text{“there are no inaccessible cardinals”}$$

and use this to argue that we cannot prove the Axiom of Inaccessibles from $\mathcal{ZFC}$.

The hereditarily countable sets also give us something interesting.

Exercise 40. Prove that $\mathbf{H}(\omega_1) \models (\forall x)(x \text{ is countable})$. Use this to argue that we cannot prove the existence of uncountable sets if we remove the Power Set Axiom from $\mathcal{ZFC}$.

The proof of this next theorem shows how to obtain absoluteness results for some formulas which are not bounded.

Theorem 48.

For any infinite cardinal κ with $X \in \boldsymbol{H}(\kappa)$ and $R \in \boldsymbol{H}(\kappa)$ we have:

$$R \text{ is a well founded relation on } X \Leftrightarrow$$
$$\boldsymbol{H}(\kappa) \models \text{“}R \text{ is a well founded relation on } X\text{”}$$

Proof. The reader may check that that the definition of “R is a well founded relation on X” is equivalent to a formula of the form $\forall Y \Phi(Y, X, R)$ where Φ is a bounded formula and the Axiom of Pairing is used for the equivalence. Furthermore, the last characterisation of well foundedness in Theorem 18 can similarly be written in the form $\exists f \Psi(f, X, R)$ where Ψ is another bounded formula. Since the proof of Theorem 18 did not use the Power Set Axiom, we have that $\mathbf{H}(\kappa) \models \forall Y \Phi(Y, X, R) \leftrightarrow \exists f \Psi(f, X, R)$. To finish the proof we use Theorem 45 in the following chain of implications.

$$\begin{array}{lc}
 & R \text{ is a well founded relation on } X \\
\Rightarrow & \forall Y \Phi(Y, X, R) \\
\Rightarrow & \forall Y \in \mathbf{H}(\kappa)\ \Phi(Y, X, R) \\
\Rightarrow & \forall Y \in \mathbf{H}(\kappa)\ \ \mathbf{H}(\kappa) \models \Phi(Y, X, R) \\
\Rightarrow & \mathbf{H}(\kappa) \models \forall Y \Phi(Y, X, R) \\
\Rightarrow & \mathbf{H}(\kappa) \models \text{“}R \text{ is a well founded relation on } X\text{”} \\
\Rightarrow & \mathbf{H}(\kappa) \models \exists f \Psi(f, X, R) \\
\Rightarrow & \exists f \in \mathbf{H}(\kappa)\ \ \mathbf{H}(\kappa) \models \Psi(f, X, R) \\
\Rightarrow & \exists f \in \mathbf{H}(\kappa)\ \Psi(f, X, R) \\
\Rightarrow & \exists f \Psi(f, X, R) \\
\Rightarrow & R \text{ is a well founded relation on } X
\end{array}$$

□

From the previous exercise we see that no $\mathbf{H}(\kappa^+)$ can be expected to be a model of the Power Set Axiom. Our first use of that axiom was in the proof of the Well Ordering Principle. It is remarkable that nevertheless this latter result still holds in each $\mathbf{H}(\kappa)$.

Exercise 41. Use Theorem 48 to prove that Theorem 21 holds in $\mathbf{H}(\kappa)$ for any infinite cardinal κ. Moreover, show that:

$$\mathbf{H}(\kappa) \models \text{every set has a cardinality.}$$

This tells us that most of what we have proven from the $\mathfrak{ZFC}$ axioms will actually hold in $\mathbf{H}(\kappa)$ for some κ large enough so that $\mathbf{H}(\kappa)$ contains all the relevant parameters.

The next theorem scheme is called the Levy Reflection Principle. It shows that the global truth of any sentence in $\mathbb{V}$ is reflected locally to some $\mathbf{R}(\beta)$ in the rank hierarchy. More poetically: the truth will eventually be revealed.

For each formula $\Phi(v_0, \dots, v_k)$ of the language of set theory, we have:

Theorem 49. Φ

$$\forall \alpha \in \mathbb{ON}\ \exists \beta \in \mathbb{ON}\ [\beta \geq \alpha \text{ and } \Phi \text{ is absolute for } \boldsymbol{R}(\beta)].$$

Proof. We will use the method of induction on complexity of formulas to verify the absoluteness, but first, for a fixed formula Φ and ordinal α we must find an appropriate corresponding ordinal β.

For each subformula Θ of Φ, for each variable v_i with $i \leq k$, for each ordinal γ and for each sequence $\vec{s}\colon (k+1) \to \mathbb{V}$ of length $k+1$, we let $E(\Theta, v_i, \gamma, \vec{s}) = \gamma$ if there is no ordinal $\delta \geq \gamma$ with the property that for some $a \in \mathbf{R}(\delta)$:

$$\Theta(v_0|\vec{s}(0), \dots, v_{i-1}|\vec{s}(i-1), v_i|a, v_{i+1}|\vec{s}(i+1), \dots, v_k|\vec{s}(k))$$

holds. On the other hand, if such an ordinal δ does exist, then we let $E(\Theta, v_i, \gamma, \vec{s})$ be the least such ordinal.

In an analogous manner, for each subformula Θ of Φ, for each variable v_i with $i \leq k$, for each ordinal γ and for each sequence $\vec{s}\colon (k+1) \to \mathbb{V}$ of length

$k+1$, we let $U(\Theta, v_i, \gamma, \vec{s}) = \gamma$ if there is no ordinal $\delta \geq \gamma$ with the property that for some $a \in \mathbf{R}(\delta)$:

$$\Theta(v_0|\vec{s}(0), \ldots, v_{i-1}|\vec{s}(i-1), v_i|a, v_{i+1}|\vec{s}(i+1), \ldots, v_k|\vec{s}(k))$$

fails to hold. On the other hand, if such an ordinal δ does exist, then we let $U(\Theta, v_i, \gamma, \vec{s})$ be the least such ordinal.

For each ordinal γ and each sequence $\vec{s}\colon (k+1) \to \mathbb{V}$, let $B(\gamma, \vec{s})$ be the maximum of the finitely many ordinals $E(\Theta, v_i, \gamma, \vec{s})$ and $U(\Theta, v_i, \gamma, \vec{s})$ where Θ is a subformula of Φ and v_i is one of the variables $v_0, \ldots, v_k$.

We now use recursion on $\mathbb{N}$ to define $F\colon \mathbb{N} \to \mathbb{ON}$ and β.

$$\begin{aligned} f(0) &= \alpha \\ f(n+1) &= sup\{B(f(n), \vec{s}) : \ \ \vec{s}\colon (k+1) \to \mathbf{R}(f(n))\} \\ \beta &= sup\{f(n) : \ n \in \mathbb{N}\} \end{aligned}$$

Clearly $\beta \geq \alpha$. In order to show that β satisfies the statement of the theorem, it suffices to prove that for each subformula Ψ of Φ we have the following.

$$\text{For all } \{a_0, \ldots, a_k\} \subseteq \mathbf{R}(\beta) : \quad \Phi(a_0, \ldots, a_k) \ \Leftrightarrow\ \mathbf{R}(\beta) \models \Phi(a_0, \ldots, a_k).$$

We prove this by induction on complexity. As usual, the atomic formula step and the connective cases of the inductive step are straightforward. We proceed to the case of the existential quantifier. Suppose Ψ is $\exists v_i \Theta$; as inductive hypothesis we have:

$$\text{For all } \{a_0, \ldots, a_k\} \subseteq \mathbf{R}(\beta) : \quad \Theta(a_0, \ldots, a_k) \ \Leftrightarrow\ \mathbf{R}(\beta) \models \Theta(a_0, \ldots, a_k).$$

Fix $\{a_0, \ldots, a_k\} \subseteq \mathbf{R}(\beta)$; there is some $n \in \mathbb{N}$ with $\{a_0, \ldots, a_k\} \subseteq \mathbf{R}(f(n))$. Since v_i is not free in $\exists v_i \Theta$:

$$\exists v_i\ \Theta(a_0, \ldots, a_k) \ \Leftrightarrow\ \exists a\ \Theta(a_0, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_k)$$

by the definition of $E(\Theta, v_i, f(n), \vec{s})$ where $\vec{s}(j) = a_j$ for each $j \leq k$, we get:

$$\Leftrightarrow\ \exists a \in \mathbf{R}(\beta)\ \Theta(a_0, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_k)$$

by inductive hypothesis

$$\Leftrightarrow\ \exists a \in \mathbf{R}(\beta)\ \mathbf{R}(\beta) \models \Theta(a_0, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_k)$$

and since v_i is not free

$$\Leftrightarrow \ \mathbf{R}(\beta) \models \exists v_i \Theta(a_0, \ldots, a_k).$$

With the same inductive hypothesis we take up the case of the universal quantifier; Ψ is $\forall v_i \Theta$. Again fix $\{a_0, \ldots, a_k\} \subseteq \mathbf{R}(\beta)$; there is again some $n \in \mathbb{N}$ with $\{a_0, \ldots, a_k\} \subseteq \mathbf{R}(f(n))$. Since v_i is not free in $\forall v_i \Theta$:

$$\forall v_i \Theta(a_0, \ldots, a_k) \ \Leftrightarrow \ \forall a \Theta(a_0, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_k)$$

by the definition of $U(\Theta, v_i, f(n), \vec{s})$ where $\vec{s}(j) = a_j$ for each $j \leq k$, we get:

$$\Leftrightarrow \ \forall a \in \mathbf{R}(\beta) \ \Theta(a_0, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_k)$$

by inductive hypothesis

$$\Leftrightarrow \ \forall a \in \mathbf{R}(\beta) \ \mathbf{R}(\beta) \models \Theta(a_0, \ldots, a_{i-1}, a, a_{i+1}, \ldots, a_k)$$

and since v_i is not free

$$\Leftrightarrow \ \mathbf{R}(\beta) \models \forall v_i \Theta(a_0, \ldots, a_k).$$

$\square$

The astute reader will have already noticed that we actually proved that not only is Φ absolute for $\mathbf{R}(\beta)$, but so also is each subformula of Φ. We state this as a corollary scheme; each fixed instance of the corollary scheme can be formalised in LOST.

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

Corollary. Φ

For any ordinal α there is an ordinal $\beta \geq \alpha$ such that each subformula of Φ is absolute for $\mathbf{R}(\beta)$.

The extra information given by this corollary to the Levy Reflection Principle is important for the hypothesis of this next result.

Exercise 42. Fix a formula Φ of LOST and a set $S \subseteq \mathbb{ON}$ such that for each $\alpha \in S$, each subformula of Φ is absolute for $\mathbf{R}(\alpha)$. Prove, using the technique of induction on the complexity of the suformulas of Φ that Φ is absolute for $\mathbf{R}(\delta)$ where $\delta = supS$.

We can now demonstrate that $\mathcal{ZFC}$ cannot be finitely axiomatised. That is, there is no finite collection of sentences, each of which is implied by our $\mathcal{ZFC}$ axioms and which in turn, together imply all the axioms of $\mathcal{ZFC}$.

If such a finite collection of sentences would exist, let Φ be their conjunction. Let Ω be the formula $(\exists\alpha \in \mathbb{ON})(\mathbf{R}(\alpha) \models \Phi)$. This is a formula of LOST which follows from the instance of the Levy Reflection Principle for the sentence Φ. Since this instance of the Levy Reflection Principle follows from $\mathcal{ZFC}$, it must follow from Φ. That is, Ω is provable from Φ using only our basic logical assumptions. By the discussion at the end of Chapter 1,

$$\mathbf{R}(\beta) \models \Phi \;\Rightarrow\; \mathbf{R}(\beta) \models \Omega$$

for all ordinals $\beta > 0$, in particular for the least β (given by the Levy Reflection Principle) such that $\mathbf{R}(\beta) \models \Phi$. For this β we have $\mathbf{R}(\beta) \models \Omega$, that is:

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

So

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

Since $\beta = \mathbf{R}(\beta) \cap \mathbb{ON}$ we get:

$$(\exists\alpha \in \beta)\ \mathbf{R}(\beta) \models (\mathbf{R}(\alpha) \models \Phi)$$

and by the definition of relativisation this becomes:

$$(\exists\alpha \in \beta)\ (\mathbf{R}(\beta) \cap \mathbf{R}(\alpha)) \models \Phi.$$

Since $\mathbf{R}(\alpha) \subseteq \mathbf{R}(\beta)$ we get that $(\exists\alpha \in \beta)\ \mathbf{R}(\alpha) \models \Phi$ and this contradicts the minimality of β.

Exercise 43. Enumerate the (countably many) axioms of $\mathcal{ZFC}$ as $\Psi_n : n \in \mathbb{N}$ in such a way that each axiom appears infinitely often in the enumeration. Using the corollary to Theorem 49, recursively construct a strictly increasing sequence of ordinals $\{\beta_n : n \in \mathbb{N}\}$ such that each subformula of Ψ_n is absolute for $\mathbf{R}(\beta_n)$. Let $\delta_0 = sup\{\beta_n : n \in \mathbb{N}\}$. According to Exercise 42 we have that $\mathbf{R}(\delta_0) \models \Psi$ for each axiom Ψ of $\mathcal{ZFC}$, so we have proven that:

$$\exists \delta \in \mathbb{ON}\ \mathbf{R}(\delta) \models \mathcal{ZFC}.$$

Now, since each axiom of $\mathcal{ZFC}$ holds in $\mathbf{R}(\delta_0)$ so too does their logical consequence, that is:

$$(\exists \delta \in \mathbb{ON}\ \mathbf{R}(\delta) \models \mathcal{ZFC})^{\mathbf{R}(\delta_0)}.$$

So there is some $\delta_1 \in \delta_0$ such that $\mathbf{R}(\delta_1) \models \mathcal{ZFC}$. Again, since each axiom of $\mathcal{ZFC}$ holds in $\mathbf{R}(\delta_1)$ so too does their logical consequence:

$$(\exists \delta \in \mathbb{ON}\ \mathbf{R}(\delta) \models \mathcal{ZFC})^{\mathbf{R}(\delta_1)}.$$

We obtain $\delta_2 \in \delta_1$ such that $\mathbf{R}(\delta_2) \models \mathcal{ZFC}$. We can continue in the same manner, obtaining $\{\delta_n : n \in \mathbb{N}\}$. This is an infinite decreasing sequence of ordinals, so something is wrong here. What?

Chapter 12

Elementary Submodels

In order to introduce some elementary set operations, we use the ordered triple notation: $\langle x, y, z\rangle$ is $\langle\langle x, y\rangle, z\rangle$ and the ordered quadruple notation: $\langle w, x, y, z\rangle$ is $\langle\langle w, x, y\rangle, z\rangle$.

$$\begin{aligned}
G_0(A, B) &= \{\langle u, v\rangle : u \in A \wedge v \in B\};\ \text{i.e., } A \times B \\
G_1(A, B) &= \{\langle v, u\rangle : u \in A \wedge v \in A \wedge \langle u, v\rangle \in B\} \\
G_2(A, B) &= \{\langle u, \langle v, w\rangle\rangle : \{u, v, w\} \subseteq A \wedge \langle u, v, w\rangle \in B\} \\
G_3(A, B) &= \{\langle u, v, w\rangle : \{u, v, w\} \subseteq A \wedge \langle u, \langle v, w\rangle\rangle \in B\} \\
G_4(A, B) &= \{\langle v, u, w\rangle : \{u, v, w\} \subseteq A \wedge \langle u, v, w\rangle \in B\} \\
G_5(A, B) &= \{\langle t, v, u, w\rangle : \{t, u, v, w\} \subseteq A \wedge \langle t, u, v, w\rangle \in B\} \\
G_6(A, B) &= \{u : u \in A \wedge u = B\};\ \text{i.e., } \{B\} \text{ or } \emptyset \\
G_7(A, B) &= \{u : u \in A \wedge u \in B\};\ i.e., A \cap B \\
G_8(A, B) &= \{u : u \in A \wedge B \in u\} \\
G_9(A) &= \{\langle u, v\rangle : u \in A \wedge v \in A \wedge u = v\} \\
G_{10}(A) &= \{\langle u, v\rangle : u \in A \wedge v \in A \wedge u \in v\} \\
G_{11}(A, B) &= \{u : u \in A \wedge u \notin B\};\ i.e., A \setminus B \\
G_{12}(A, B, C) &= \{u : u \in A \wedge \exists v \in B\ \langle u, v\rangle \in C\} \\
G_{13}(A, B, C) &= \{u : u \in A \wedge \forall v \in B\ \langle u, v\rangle \in C\}
\end{aligned}$$

Elementary set operations were first proposed by K. Gödel; he had nine elementary operations, but we shall enlarge the number to fourteen Gödel Operations in order to make the proofs run smoothly.

It is natural to extend the ordered pair notation past triples and quadruples to ordered n-tuples by recursion on $\mathbb{N}$:

$$\begin{aligned} Tuples(0, X) &= \emptyset \\ Tuples(1, X) &= X \\ Tuples(2, X) &= X \times X \\ Tuples(n+1, X) &= Tuples(n, X) \times X. \end{aligned}$$

We denote a typical member of $Tuples(m, X)$ by $\langle x_1, \ldots, x_m \rangle$. Notice that for $m \in \mathbb{N}$, $Tuples(m, X)$ is the $(m-1)-$ fold composition of G_0:

$$Tuples(m, X) = G_0(G_0 \ldots (G_0(X, X), \ldots X), X).$$

Shuffling the coordinates of an m-tuple can be done with Gödel operations via the following Shuffle Lemma.

Lemma. For any $m \in \mathbb{N} \setminus \{0, 1\}$ and any permutation

$$\sigma \colon \{1, \ldots, m\} \to \{1, \ldots, m\}$$

there is a composition F_σ of the operations G_1, G_2, G_3, G_4 and G_5 such that for any X and any $S \subseteq Tuples(m, X)$,

$$F_\sigma(X, S) = \{\langle x_{\sigma(1)}, \ldots, x_{\sigma(m)} \rangle \in Tuples(m, X) : \langle x_1, \ldots, x_m \rangle \in S\}.$$

Proof. Fix $m \in \mathbb{N}$, X and S. For ease of reading, let's denote the function of one variable $G_1(X, \cdot)$ by $F_1(\cdot)$ and similarly for G_2, G_3, G_4 and G_5. It may be of independent interest that each of our F_σ's will be compositions of F_1, F_2, F_3, F_4 and F_5 only.

Binary exchanges generate the symmetric group. So, noting that the identity permutation is given by $F_1 \circ F_1$, it suffices to consider only those σ such that for some $1 \leq l < m \geq 2$:

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

Letting $F_1^{(n)}$ denote the n-fold composition of F_1, etc., we address all cases.

$$\begin{array}{rl}
\text{if } m = 2 & F_\sigma = F_1 \\
\text{if } m = 3, \text{ and } l = 1 & F_\sigma = F_4 \\
\text{if } m \geq 3, \text{ and } l = m-1 & F_\sigma = F_3 \circ F_1 \circ F_4 \circ F_1 \circ F_2 \\
\text{if } m \geq 4, \text{ and } l = 1 & F_\sigma = F_3^{(m-3)} \circ F_4 \circ F_2^{(m-3)} \\
\text{if } m \geq 4, \text{ and } 2 \leq l \leq m-2 & F_\sigma = F_3^{(m-l-2)} \circ F_5 \circ F_2^{(m-l-2)}
\end{array}$$

□

It is remarkable that many sets can be realised as the result of the fourteen Gödel operations. For each formula Φ of the language of set theory with free variables x, $w_0, \ldots, w_k$ we have:

Theorem 50. Φ

For all X and for all $w_0, \ldots, w_k$ in X, $\{x \in X : \Phi^X(x, w_0, \ldots, w_k)\}$ is the result of a finite composition of Gödel operations on $X, w_0, \ldots w_k$.

Proof. Fix the formula Φ. Without loss of generosity we may assume that the free variables x, $w_0, \ldots, w_k$ do not also occur bound in Φ. Let x be x_1 and let $x_2, \ldots, x_m$ be the bound variables of Φ.

We prove the theorem by induction on the complexity of the subformulas of Φ, using the finite lemma scheme below; there is one lemma for each subformula Ψ of Φ. The theorem follows from the lemma in which Ψ is Φ after an $(m-1)-$fold application of the Gödel operation $G_{12}(X, \cdot)$ to $\{\langle x_1, \ldots, x_m \rangle \in Tuples(m, X) : \Phi^X\}$ □

Continuing the notation of the previous theorem, for each subformula Ψ of Φ we denote $\{\langle x_1, \ldots, x_m \rangle \in Tuples(m, X) : \Psi^X\}$ by A_Ψ and we have this lemma.

Lemma. Ψ

A_Ψ is the result of a finite composition of Gödel operations on $X, w_0, \ldots w_k$.

Proof. There are two steps in a proof by induction on complexity. This time neither the base step nor the inductive step is immediate.

For the base step we consider atomic formulas. Since we have treated the variables x_i different from the variables w_i we have nine different types of atomic formulas. We must address each case separately.

1. In the case that Ψ is $w_i = w_j$ for $i, j \leq k$ the set A_Ψ is either $Tuples(m, X)$ or $\emptyset = G_{11}(X, X)$ depending simply upon whether or not w_i is actually equal to w_j or not.

2. The case where Ψ is $w_i \in w_j$ for $i, j \leq k$ is similar to the previous case.

3. In the case that Ψ is $w_i = x_j$ for $i \leq k$ and $1 \leq j \leq m$ the Shuffle Lemma allows us to reduce to the situation in which $j = 1$ so that:
$$A_\Psi = G_0(G_0 \ldots (G_0(G_6(X, w_i), X) \ldots X), X),$$
where the composition is $(m-1)-$fold.

4. In the case that Ψ is $x_j \in w_i$ for $i \leq k$ and $1 \leq j \leq m$ the Shuffle Lemma again allows us to reduce to the situation in which $j = 1$, so that:
$$A_\Psi = G_0(G_0 \ldots (G_0(G_7(X, w_i)) \ldots X), X),$$
where the composition is $(m-1)-$fold.

5. In the case that Ψ is $w_i \in x_j$ for $i \leq k$ and $1 \leq j \leq m$ the Shuffle Lemma again allows us to reduce to the situation in which $j = 1$ and
$$A_\Psi = G_0(G_0 \ldots (G_0(G_8(X, w_i), X) \ldots X), X),$$
where the composition is $(m-1)-$fold.

6. In the case that Ψ is $x_j = x_j$ for $1 \leq j \leq m$ we simply have $A_\Psi = Tuples(m, X)$.

7. In the case that Ψ is $x_i = x_j$ for $1 \leq i, j \leq m$ and $i \neq j$ the Shuffle Lemma allows us to reduce to the situation in which $i = 1$ and $j = 2$.
$$A_\Psi = G_0(G_0 \ldots (G_0(G_9(X), X) \ldots X), X),$$
where the composition is $(m-2)-$fold.

8. In the case that Ψ is $x_j \in x_j$ for $1 \leq j \leq m$ we simply have $A_\Psi = \emptyset = G_{11}(X, X)$.

9. In the case that Ψ is $x_i \in x_j$ for $1 \leq i, j \leq m$ and $i \neq j$ the Shuffle Lemma allows us to reduce to the situation in which $i = 1$ and $j = 2$.
$$A_\Psi = G_0(G_0 \ldots (G_0(G_{10}(X), X) \ldots X), X),$$
where the composition is $(m-2)-$fold.

In the six cases of the inductive step, for the inductive hypothesis, we assume that

$$\begin{aligned} \text{both } A_\Theta &= \{\langle x_1, \ldots, x_m\rangle \in \mathit{Tuples}(m, X) : \Theta^X\} \\ \text{and } A_\Omega &= \{\langle x_1, \ldots, x_m\rangle \in \mathit{Tuples}(m, X) : \Omega^X\} \end{aligned}$$

are the result of a finite composition of Gödel operations on $X, w_0, \ldots w_k$.

1. In the case that Ψ is $\neg\Theta$ we have $A_\Psi = \mathit{Tuples}(m, X) \setminus A_\Theta$ and so $A_\Phi = G_{11}(\mathit{Tuples}(m, X), A_\Theta)$

2. In the case that Ψ is $\Theta \wedge \Omega$ we have $A_\Psi = A_\Theta \cap A_\Omega$. Noticing that $A \cap B = A \setminus (A \setminus B)$, this case finishes in a manner similar to the previous one.

3. In the case that Ψ is $\Theta \vee \Omega$ we have $A_\Psi = A_\Theta \cup A_\Omega$
$$= \mathit{Tuples}(m, X) \setminus [(\mathit{Tuples}(m, X) \setminus A_\Theta) \cap (\mathit{Tuples}(m, X) \setminus A_\Omega)]$$
and so this case finishes like the previous two.

4. In the case that Ψ is $\Theta \to \Omega$ we have $A_\Psi = (\mathit{Tuples}(m, X) \setminus A_\Theta) \cup A_\Omega$ so that this case is also similar.

5. In the case that Ψ is $\exists x_j \Theta$ we use the Shuffle Lemma to reduce to the situation in which $j = m$. Now, since the variable x_m occurs bound in Ψ we have
$$A_\Psi = \{\langle x_1, \ldots, x_{m-1}\rangle \in \mathit{Tuples}(m-1, X) : (\exists x_m \Theta)^X\} \times X$$
so that
$$A_\Psi = G_0(G_{12}(\mathit{Tuples}(m-1, X), X, A_\Theta), X).$$

6. In the case that Ψ is $\forall x_j \Theta$ we again use the Shuffle Lemma to reduce to the situation in which $j = m$. In a manner similar to the previous case, we have:

$$A_\Psi = G_0(G_{13}(Tuples(m-1, X), X, A_\Theta), X).$$

$\square$

It is evident that not all possible compositions of Gödel operations were used in this proof. In fact $G_0(A, B)$ was only used in the form $G_0(\cdot, X)$. We have already noted that G_1 through G_5 were only used in the form $G(X, \cdot)$. Moreover G_6, G_7 and G_8 were only used in the form $G(X, w)$ where w was a parameter; G_9 and G_{10} were only used in the form $G(X)$. G_{12} and G_{13} were only used in the form $G(Tuples(n, X), X.\cdot)$. Only G_{11} was fully used, yet all of its inputs were subsets of $Tuples(n, X)$ for some $n \in \mathbb{N}$. Furthermore, all outputs of all compositions were subsets of $Tuples(n, X)$ for some $n \in \mathbb{N}$.

We will enumerate all compostions of Gödel operations actually used in the proof of Theorem Scheme 50. However, one complicating factor is that although a composition of functions of one variable is a function of one variable, the composition of functions of two variables can be a function of more than two variables; in general the input to a finite composition of functions of two variables is a finite sequence $\vec{s}$.

We define a master operation $\mathbf{G} \colon \mathbb{V} \times \mathbb{V} \times \mathbb{N} \to \mathbb{V}$ such that for each X, each finite sequence $\vec{w}$ and each $X \in \mathbb{N}$ the value $\mathbf{G}(X, \vec{w}, n)$ is the result of a composition of Gödel operations on X and the range of $\vec{w}$. Furthermore, each composition of Gödel operations used in proving any instance of Theorem Scheme 50 involving a fixed set X and parameters $w_0, \ldots, w_k$ is $\mathbf{G}(X, \vec{w}, n)$ for some $n \in \mathbb{N}$ where $\{w_0, \ldots, w_k\}$ is the range of $\vec{w}$.

More precisely, for each X and each $\vec{s}$ we define $\mathbf{G}(X, \vec{s}, \cdot) \colon \mathbb{N} \to \mathbb{V}$ by recursion on $\mathbb{N}$ as follows:

$$\mathbf{G}(X,\vec{s},n) = \begin{cases} X & \text{if } n = 1; \\ G_0(\mathbf{G}(X,\vec{s},i),X) & \text{if } n = 3^i \text{ and } i > 0; \\ G_1(X,\mathbf{G}(X,\vec{s},i)), & \text{if } n = 2\cdot 3^i; \\ G_2(X,\mathbf{G}(X,\vec{s},i)), & \text{if } n = 2^2\cdot 3^i; \\ G_3(X,\mathbf{G}(X,\vec{s},i)), & \text{if } n = 2^3\cdot 3^i; \\ G_4(X,\mathbf{G}(X,\vec{s},i)), & \text{if } n = 2^4\cdot 3^i; \\ G_5(X,\mathbf{G}(X,\vec{s},i)), & \text{if } n = 2^5\cdot 3^i; \\ G_6(X,\vec{s}(i)) & \text{if } n = 2^6\cdot 3^i; \\ G_7(X,\vec{s}(i)) & \text{if } n = 2^7\cdot 3^i; \\ G_8(X,\vec{s}(i)) & \text{if } n = 2^8\cdot 3^i; \\ G_9(X,X) & \text{if } n = 2^9; \\ G_{10}(X,X) & \text{if } n = 2^{10}; \\ G_{11}(\mathbf{G}(X,\vec{s},i),\mathbf{G}(X,\vec{s},j)) & \text{if } n = 2^{11}\cdot 3^i\cdot 5^j; \\ G_{12}(\mathbf{G}(X,\vec{s},i),X,\mathbf{G}(X,\vec{s},j)) & \text{if } n = 2^{12}\cdot 3^i\cdot 5^j; \\ G_{13}(\mathbf{G}(X,\vec{s},i),X,\mathbf{G}(X,\vec{s},j)) & \text{if } n = 2^{13}\cdot 3^i\cdot 5^j; \\ \emptyset, & \text{otherwise.} \end{cases}$$

Although we will be mainly interested in situations for which $\vec{s}\colon l \to X$ for some $l \in \mathbb{N}$, we can nevertheless show by a straightforward induction on $\mathbb{N}$ that for any X and $\vec{s}$, each $\mathbf{G}(X,\vec{s},n) \subseteq \mathit{Tuples}(m,X)$ for some m. Moreover, we obtain an immediate corollary to (the proof of) Theorem 50.

Corollary. Φ

For all X and for all $w_0,\ldots,w_k$ in X there is an $n \in \mathbb{N}$ such that

$$\{x \in X : \Phi^X(x,w_0,\ldots,w_k)\} = \mathbf{G}(X,\vec{w},n))$$

where $\vec{w}(j) = w_j$ for each $j \in \mathit{dom}(\vec{w})$.

We are now prepared for the most important definition of this chapter.

M is said to be an elementary submodel of N and we write $M \prec N$ whenever $M \subseteq N$ and for all $k \in \mathbb{N}$, for all $\vec{w} \in {}^kM$ and for all $n \in \mathbb{N}$

$$\mathbf{G}(N,\vec{w},n) \cap N \neq \emptyset \Leftrightarrow \mathbf{G}(N,\vec{w},n) \cap M \neq \emptyset.$$

Two theorems help us obtain elementary submodels. The first is sometimes called the Löwenheim-Skolem Theorem.

Theorem 51. *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:

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

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

$$\begin{aligned} X_0 &= X \\ X_{m+1} &= X_m \cup F^{\to}(\omega \times X_m) \end{aligned}$$

Let $M = \bigcup_{m\in\omega} X_m$. As such, (2) and (3) are clearly satisfied. To check (1) let $n \in \mathbb{N}$ and $\vec{w} \in {}^kM$ such that $\mathbf{G}(N,\vec{w},n) \cap N \neq \emptyset$. Then $\vec{w} \in {}^kX_m$ for some $m \in \omega$ so that $F(n,\vec{w}) \in \mathbf{G}(N,\vec{w},n) \cap X_{m+1}$ and $\mathbf{G}(N,\vec{w},n) \cap M \neq \emptyset$.

$\square$

The second way of obtaining elementary submodels is through a version of the Elementary Chain Theorem:

Theorem 52. *Suppose that δ is a limit ordinal and $\{M_\alpha : \alpha < \delta\}$ is a set of elementary submodels of N such that*

$$\forall\alpha\ \forall\alpha'\ (\alpha < \alpha' < \delta \to M_\alpha \subseteq M_{\alpha'}).$$

Let

$$M_\delta = \bigcup\{M_\alpha : \alpha < \delta\}.$$

Then $M_\delta \prec N$.

Proof. Let $k \in \omega$, let $\vec{w} \in {}^kM_\delta$, and let $n \in \omega$. We need to show that

$$\mathbf{G}(N, \vec{w}, n) \cap N \neq \emptyset \Rightarrow \mathbf{G}(N, \vec{w}, n) \cap M_\delta \neq \emptyset.$$

But this is easy since $\vec{w} \in {}^kM_\alpha$ for some $\alpha < \delta$.

$\square$

The power of elementary submodels arises from the following theorem scheme. For each formula $\Phi(v_0, \ldots, v_k)$ of LOST we have:

Theorem 53. Φ

Suppose $M \prec N$. Then for all $m_0, \ldots, m_k$ in M we have:

$$M \models \Phi(m_0, \ldots, m_k) \text{ iff } N \models \Phi(m_0, \ldots, m_k).$$

Proof. Fix a formula Φ with all of its variables lying among $v_0, \ldots, v_k$. We will prove the theorem using the technique of induction on the complexity of the subformulas Ψ of Φ.

Both cases of the atomic subformula step and the four connective cases of the inductive step follow immediately from the definition of relativisation. This allows us to directly proceed with the two quantifier cases of the inductive step. In these cases, the existential quantifier case in which Ψ is $(\exists v_i)\Theta$ and the universal quantifier case in which Ψ is $(\forall v_i)\Theta$, we have the following inductive hypothesis.

For all $\{m_0, \ldots, m_k\} \subseteq M$: $M \models \Theta(m_0, \ldots, m_k)$ iff $N \models \Theta(m_0, \ldots, m_k)$.

We first take up the existential quantifier case. Let $m_0, \ldots, m_k$ be in M. We begin a chain of equivalences with

$$M \models (\exists v_i)\Theta(m_0, \ldots, m_k).$$

By the definition of relativisation, since v_i is not free, this is equivalent to

$$\text{for some } x \in M \ M \models \Theta(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k).$$

By inductive hypothesis this is equivalent to

$$\text{for some } x \in M \ N \models \Theta(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k).$$

Since $M \subseteq N$, this is equivalent to

$$M \cap \{x \in N : \Theta^N(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k)\} \text{ is nonempty.}$$

By Theorem 50 there is $n \in \omega$ such that this latter set is $\mathbf{G}(N, \vec{m}, n)$ where $\vec{m}$ is a sequence with $\vec{m}(j) = m_j$ for all $0 \leq j \leq k$. So we get the equivalence:

$$M \cap \mathbf{G}(N, \vec{m}, n) \neq \emptyset.$$

Since $M \prec N$ this is equivalent to

$$N \cap \mathbf{G}(N, \vec{m}, n) \neq \emptyset,$$

that is:

$$N \cap \{x \in N : \Theta^N(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k)\} \text{ is nonempty.}$$

This is equivalent to

$$\text{for some } x \in N \ N \models \Theta(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k)$$

which, since v_i is not free in $(\exists v_i)\Theta$, is also equivalent to

$$N \models (\exists v_i)\Theta(m_0, \ldots, m_k).$$

We now turn our attention to the universal quantifier case, beginning with

$$M \models (\forall v_i)\Theta(m_0, \ldots, m_k).$$

By the definition of relativisation, since v_i is not free, this is equivalent to

$$\text{for all } x \in M \ M \models \Theta(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k).$$

By inductive hypothesis this is equivalent to

$$\text{for all } x \in M \ N \models \Theta(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k).$$

This is equivalent to

$$\text{for no } x \in M \text{ does } N \models (\neg\Theta)(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k).$$

Since $M \subseteq N$, this is equivalent to

$$M \cap \{x \in N : (\neg\Theta)^N(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k)\} \text{ is empty}$$

By Theorem 50 there is $n' \in \omega$ such that this latter set is $\mathbf{G}(N, \vec{m}, n')$ where $\vec{m}$ is a sequence with $\vec{m}(j) = m_j$ for all $0 \leq j \leq k$. So we get the equivalence:

$$M \cap \mathbf{G}(N, \vec{m}, n') = \emptyset.$$

Since $M \prec N$ this is equivalent to

$$N \cap \mathbf{G}(N, \vec{m}, n') = \emptyset,$$

that is:

$$N \cap \{x \in N : (\neg\Theta)^N(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k)\} = \emptyset.$$

This is equivalent to

$$\text{for all } x \in N \ N \models \Theta(m_0, \ldots, m_{i-1}, x, m_{i+1}, \ldots m_k)$$

which, since v_i is not free in $(\forall v_i)\Theta$, is also equivalent to

$$N \models (\forall v_i)\Theta(m_0, \ldots, m_k).$$

□

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

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

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 $q \in M$.

□

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, $\alpha \subseteq M$.

□

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

Theorem 54. *Erdős-Dushnik-Miller*

If $P : [\omega_1]^2 \ \to\ \{1, 2\}$, *then either*

1. *there is an infinite* $H \subseteq \omega_1$ *such that* $P(\{\alpha, \beta\}) = 1$ *for all distinct* α *and* β *in* H*, or*

2. *there is an uncountable* $H \subseteq \omega_1$ *such that* $P(\{\alpha, \beta\}) = 2$ *for all distinct* α *and* β *in* H*.*

Theorem 55. *(Pressing Down Lemma)*

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

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

Theorem 56. *(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. $\forall D_1, D_2 \in \mathcal{D}\ D_1 \cap D_2 = R$*.*

Proof of Pressing Down Lemma

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 $\mathcal{D} = \{D(\alpha) : \alpha < \omega_1\}$.

□

Chapter 13

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

$$\mathbb{L} : \mathbb{ON} \to \mathbb{V}$$

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

$$\begin{gathered}[(\forall u \in x \ \forall v \in u \ v \in x) \wedge (\forall u \in x \ \forall v \in x \ (u \in v \ \vee \ v \in u \ \vee \ u = v)) \\ \wedge(\forall u \in x \ \forall v \in x \ \forall w \in x \ (u \in v \ \wedge \ v \in w \to u \in w))].\end{gathered}$$

   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:

$$\begin{aligned}\beta = \mathbf{L}(\beta) \cap \mathbb{ON} &= \{x \in \mathbf{L}(\beta) : x \text{ is an ordinal}\} \\ &= \{x \in \mathbf{L}(\beta) : \Phi(x)\} \\ &= \{x \in \mathbf{L}(\beta) : \Phi^{\mathbf{L}(\beta+1)}(\alpha)\} \\ &\in cl(\mathbf{L}(\beta)) \text{ using Theorem 50} \\ &= \mathbf{L}(\beta + 1).\end{aligned}$$

□

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 50 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 50 we can construct an injection $cl(X) \to \omega \times \bigcup\{{}^k X : k \in \omega\}$. Hence, $|X| \leq |cl(X)| \leq \max\,(\aleph_0, |\{{}^k X : k \in \omega\}|) = |X|$.

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,

$$\begin{aligned} |\mathbf{L}(\omega)| &= |\bigcup\{\mathbf{L}(n) : n \in \omega\}| \\ &= \max\ (\aleph_0, \sup\ \{|\mathbf{L}(n)| : n \in \omega\}) \\ &= \max\ (\aleph_0, \sup\ \{|R(n)| : n \in \omega\}) \\ &= \aleph_0. \end{aligned}$$

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

$$\begin{aligned} |\mathbf{L}(\delta)| &= |\bigcup\{\mathbf{L}(\beta) : \beta \in \delta\}| \\ &= \max\ (|\delta|, \sup\ \{|\mathbf{L}(\beta)| : \beta \in \delta\}) \\ &= \max\ (|\delta|, \sup\ \{|(\beta)| : \beta \in \delta\}) \text{ by inductive hypothesis} \\ &= |\delta|. \end{aligned}$$

□

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 $\mathcal{ZFC}$ we have:

Theorem 57. 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 50, 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 57 Φ and motivates calling "$\mathbb{V} = \mathbb{L}$" the Axiom of Constructibility.

□

Remark. $\mathbb{V} = \mathbb{L}$ is consistent with $\mathcal{ZFC}$ in the sense that no finite subcollection of $\mathcal{ZFC}$ can possibly prove $\mathbb{V} \neq \mathbb{L}$; To see this, suppose

$$\{\Psi_0, \ldots, \Psi_n\} \vdash \mathbb{V} \neq \mathbb{L}.$$

Then

$$\Psi_0^{\mathbb{L}}, \ldots, \Psi_n^{\mathbb{L}}\} \vdash (\mathbb{V} \neq \mathbb{L})^{\mathbb{L}}.$$

by Theorem 57. 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 $\mathcal{ZFC}$ 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 58. *(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 $\underline{\kappa\text{−complete}}$ if $\forall X \in [\mathcal{U}]^{<\kappa}\ \bigcap X \in \mathcal{U}$.

An uncountable cardinal κ is said to be $\underline{\text{measurable}}$ whenever there exists a $\kappa-$complete free ultrafilter over κ.

Lemma. If $\mathcal{U}$ is a countably complete ultrafilter (in particular if $\mathcal{U}$ is a $\mu-$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.

□

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 <u>elementary embedding</u> 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.

□

Theorem 59. *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 $X \cap \{I\}$, which is either empty or a singleton, contradicting that $\mathcal{U}$ is a free filter.

□

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 60. *(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 sentence; 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.

$\square$

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 supercompact 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 14

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 Extensionality**

$$\forall x \; \forall y \; [x = y \leftrightarrow \forall u \; (u \in x \leftrightarrow u \in y)]$$

2. **Axiom of Existence**

$$\exists z \; z = z$$

3. **Axiom of Pairing**

$$\forall x \; \forall y \; \exists z \; z = \{x, y\}$$

4. **Union Axiom**

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

5. **Intersection Axiom**

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

6. **Axiom of Foundation**

$$\forall x\ [x \neq \emptyset \to (\exists y \in x)(x \cap y = \emptyset)]$$

7. **Replacement Axiom Scheme**

   For each formula $\Phi(x, u, v, w_1, \ldots, w_k)$ of the language of set theory,

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

8. **Axiom of Choice**

$$\forall X\ [(\forall x \in X\ \forall y \in X\ (x = y \leftrightarrow x \cap y \neq \emptyset))$$

$$\to \exists z\ (\forall x \in X\ \exists! y\ y \in x \cap z)]$$

9. **Power Set Axiom**

$$\forall x\ \exists z\ z = \{y : y \subseteq x\}$$

10. **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**

$$\exists \kappa\ \kappa \text{ is an inaccessible cardinal}$$

2. **Continuum Hypothesis**

$$|\mathcal{P}(\omega)| = \omega_1$$

3. **Generalised Continuum Hypothesis**

$$\forall\kappa\ [\kappa \text{ is a cardinal} \to |\mathcal{P}(\kappa)| = \kappa^+]$$

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**

$$\mathbb{V} = \mathbb{L}$$