

# Partial Universes and the Axioms of Set Theory

Author: Samuel Coskey  
Advisor: M. Scott Osborne

University of Washington

June, 2003

## **Abstract**

This paper is intended as a senior thesis. The topic is set theory. The material covered is not particularly advanced with respect to the field. However, this author has never had the opportunity to take an official class in set theory; it was in the course of preparing this document that he himself was introduced to set theory.

After introducing some fundamental notions from set theory and logic, a brief introduction to classes is presented. The discussion then turns to the Zermelo Hierarchy of partial universes of sets (the  $V_\alpha$  guys). In particular, the question of which Zermelo-Fraenkel axioms hold at each stage of the hierarchy is answered.

# 1 Introduction

Typically in an introduction, one gives an idea or an example to motivate the narrative. In a paper on set theory, the example is inevitably the Russell paradox, and regretfully this will be mine as well.

The contradictions furnished by the Russell “set”<sup>1</sup> necessitated a rigorous restriction on the class of allowable sets. To avoid many common paradoxes, one would like to disallow sets as members of themselves, and more generally, to demand that the relation of set membership be wellfounded.

What will be called the “Zermelo Hierarchy” is a concrete realization of this concept. It is a transfinite “sequence” (indexed by the full class of ordinal numbers) of sets  $\emptyset = V_0 \subset V_1 \subset \dots \subset V_\alpha \subset \dots$ , the union of which is taken to be the universe of sets. That is, for any set  $x$ , there is a least ordinal  $\alpha$  such that  $x \in V_\alpha$ . Furthermore, it is the case that  $x$  is composed, via the set theoretic operations, of elements of earlier sets in the sequence.

But first,

## 2 Some Foundations of Set Theory

For the most part, we will assume that the reader is familiar with the basics of naïve set theory, cardinality and cardinals, orderings and recursion. Nevertheless, we will develop most of the vocabulary and definitions which will be needed in our later analysis. We will present without proof some of the basic theorems as well.

### 2.1 First order languages

A first order language  $L$  is a collection of relational (true/false valued) symbols  $r, s, t, \dots$ , each of some arity (number of arguments), together with the following customary symbols.

- Variable symbols  $x, y, z \dots$
- Connectives  $\neg, \wedge$ , and the quantifier  $\forall$
- Grouping symbols  $( )$
- The symbol  $=$ , a symmetric relation of arity 2

The formulas of  $L$  may be constructed inductively via the connectives as follows.

- If  $x_1, \dots, x_n$  are variable symbols, and  $r$  is a relational symbol, then  $x_1 = x_2$  and  $rx_1 \dots x_n$  are atomic formulas.
- If  $p$  and  $q$  are formulas,  $\neg p$  and  $(p \wedge q)$  are formulas.
- If  $p$  is a formula and  $x$  is a variable symbol, then  $\forall x(p)$  is a formula.

---

<sup>1</sup>Let  $R = \{x : x \notin x\}$  and notice that  $R \in R$  if and only if  $R \notin R$ .

The symbols  $\neg$  and  $\wedge$  are to be interpreted as you would expect (not, and). The remaining connectives  $\vee$ ,  $\rightarrow$ , and  $\leftrightarrow$  are defined in terms of  $\neg$  and  $\wedge$  to make the DeMorgan and other laws of Boolean algebra work out. Similarly,  $\exists x$  is defined to abbreviate  $\neg\forall x\neg$ . We sometimes omit parentheses (when the meaning is unambiguous) or add parentheses (when there is cause for confusion).

The notion of proof is formalized with rules of inference, which define when some formula is a logical consequence of some others. More specifically, the relationship “ $p$  follows via the rules of inference from the set of formulas  $A$ ” is denoted  $A \vdash p$ . The rules of inference are generally defined in such a way that the connectives possess their usual meaning. For example, it is the case that  $A \cup \{q\} \vdash p$  if and only if  $A \vdash q \rightarrow p$ , so that  $\rightarrow$  possesses its usual meaning of implication.

If  $p$  is a formula and  $x$  is a symbol of  $p$ , then  $x$  is said to occur in  $p$ . This occurrence is either bound or free. It is bound if  $x$  occurs in the  $\dots$  in  $\forall x(\dots)$  and it is free otherwise. When we abuse the notation  $p(x)$ , we wish to tacitly imply that  $p$  is a formula and  $x$  occurs free in  $p$ .

A sentence is any formula which does not have any free occurrences of variables. A theory  $T$  is a collection of sentences. Such a theory may be given to you as a literal such collection, or defined to be the set of sentences which follow via the rules of inference from a given (smaller) set of sentences. A theory is said to be consistent if there is no sentence  $p$  in  $T$  such that  $T \vdash p$  and  $T \vdash \neg p$ .

The language of set theory is a first order language with just a single binary relation,  $\in$ . The variables are interpreted as sets. The formula  $\in xy$  is usually written  $x \in y$  and interpreted “the set  $x$  is a member of the set  $y$ .” All set theoretic concepts can be expressed in the language of set theory. For example, we write  $z \in x \rightarrow z \in y$  to mean simply that  $x \subset y$ . We will see many more examples in the next section.

We will very shortly discuss a rather popular collection of axioms, called ZFC. Set theory is the collection of formulas which follow from ZFC. There are several standard formulations of these Zermelo-Fraenkel axioms, of which I have chosen to present one. The axioms as given are not free of dependencies, and this will be discussed as the axioms are presented.

Furthermore, two of the axioms which will be heavily scrutinized are not properly axioms, but what would be termed axiom *schemas*<sup>2</sup>. In both cases, the formulation follows the template “For every formula  $p(x)$  in the language of set theory, the following is an axiom:  $\dots$ ”

## 2.2 The axioms of set theory

The way we have defined the language of set theory,  $=$  is an automatic binary relation. But we have not defined its meaning. How do we know when  $x$  and  $y$  are the same, so we can write  $x = y$ ? The intuitive solution is the answer:  $x = y$  when they contain exactly the same elements.

**Axiom 1 (Extensionality).**  $\forall xy(x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y))$

<sup>2</sup>rather than the perhaps more appropriate “schemæ”

Next, we need some sets to work with. We have a whole language for describing truths of set theory, but we haven't anywhere guaranteed that there are any sets. Of course, we could just demand that  $\exists x(x = x)$ , and this together with other axioms would actually be good enough. But it's traditional to be more specific.

**Axiom 2 (Empty Set).**  $\exists x \forall z (z \notin x)$

It follows from Extensionality that this empty set  $x$  is unique, so it is safe to give it its own symbol  $\emptyset$ . To be more precise, we let  $p(\emptyset)$ , the formula  $p$  with  $\emptyset$  substituted for some free variable  $x$ , abbreviate  $\forall x(\forall y(y \notin x) \rightarrow p(x))$ . Since we will be making many abbreviations, we will not be so precise in the future.

Now that we have a named set to work with, we want to form new sets from it. Perhaps the simplest way to create new sets is the following.

**Axiom 3 (Pairing).**  $\forall x \forall y \exists z \forall w (w \in z \leftrightarrow (w = x \vee w = y))$

As before, Extensionality implies that for fixed  $x, y$  the set  $z$  is unique. It is abbreviated  $\{x, y\}$ , and we might again write with an explicit formula what is meant by  $p(\{x, y\})$ . At least, the reader may certainly do so, but we don't recommend it. Applying the Axiom of Pairing twice, we may form the ordered pair  $\langle x, y \rangle = \{x, \{x, y\}\}$  from sets  $x$  and  $y$ . With this notion, we can say what we mean by a function.

**Definition 2.1.** Let  $\text{Fun}(f)$  denote the formula

$$(\forall z(z \in f \rightarrow \exists xy(z = \langle x, y \rangle))) \wedge (\forall xyw((\langle x, y \rangle \in f \wedge \langle x, w \rangle \in f) \rightarrow y = w))$$

That is,  $\text{Fun}$  expresses the fact that its argument is a set of ordered pairs, with a unique second coordinate for each first coordinate.

Suppose  $\text{Fun}(f)$  holds. We say that  $f$  is a function. If  $\langle x, y \rangle \in f$  we write  $y = f(x)$ . If  $a$  is any set, there is a function  $g \subset f$  such that  $y = g(x)$  if and only if  $y = f(x)$  and  $x \in a$ ; we write  $g = f|_a$ . If  $X$  is the collection of sets  $x$  such that some  $\langle x, y \rangle \in f$ , and  $Y$  is any collection of sets containing all  $y$  such that some  $\langle x, y \rangle \in f$ , then we write  $f : X \rightarrow Y$ .

Working purely from the preceding axioms, we can only create sets like  $\emptyset$ ,  $\{\emptyset, \emptyset\} = \{\emptyset\}$ ,  $\{\emptyset, \{\emptyset\}\}$ , etc. Although these sets quickly become arbitrarily "deep," they can never have more than two elements. A more horizontal way to glue sets together from old ones is given by the next axiom.

**Axiom 4 (Union).**  $\forall x \exists y (z \in y \leftrightarrow \exists w (z \in w \wedge w \in x))$

The set  $y$  is often denoted  $\bigcup x$ , and we abbreviate  $\bigcup \{a, b\}$  by  $a \cup b$ . It is not immediately obvious, but we can actually create triples (and more!) from the axioms that we have so far. Repeatedly using pairing, one can create the set

$$x = \{\{\emptyset, \{\emptyset\}\}, \{\{\{\emptyset\}\}\}\}$$

and then notice that  $\bigcup x$  has three distinct elements.

Actually, the Union Axiom is not quite as powerful as it may at first seem. The reason is that to form the union of a collection of sets, those sets must already be contained in some set. But where do we get the container set? One common way to do this is with the next axiom.

**Axiom 5 (Power Set).**  $\forall x \exists y \forall z (z \in y \leftrightarrow z \subset x)$

The Power Set axiom gives us large sets from which we can extract interesting subsets. But what qualifies as an “interesting” subset? We’ve talked about so few sets so far that it might be hard to classify any of them as interesting. Nevertheless, the goal of axiomatizing set theory is to obtain a wide class of sets which don’t give rise to contradictions. We’ll take as our definition of interesting subset those that can be described by a formula. In other words, we can always put together all the elements of a set satisfying a common formula (sometimes called a property).

**Axiom 6 (Separation schema).** If  $p(z)$  is a formula with a free variable  $z$ , then the following is an axiom.  $\forall x \exists y (z \in y \leftrightarrow (z \in x \wedge p(z)))$ .

The usual abbreviation for a subset  $y \subset x$  which follows the formula  $p(z)$  is  $y = \{z \in x : p(z)\}$ .

The next axiom is such a ridiculously long formula that we will introduce some new notation just to state it. It will guarantee the existence of some sets with eclectic membership.

**Definition 2.2.** If  $x$  and  $y$  are sets, let  $p(z)$  denote  $z \in y$ . Denote by  $x \cap y$  the set such that  $z \in x \cap y \leftrightarrow (z \in x \wedge z \in y)$  guaranteed by the Separation Axiom. It is unique by the Axiom of Extensionality.

**Definition 2.3.** If  $p(x)$  is a formula with free variable  $x$ , then the formula of the form  $\exists! x (p(x))$  abbreviates  $\exists x (p(x) \wedge (p(y) \rightarrow y = x))$ .

**Axiom 7 (Choice).**  $\forall x ((\forall y z (y \in x \wedge z \in x \rightarrow y \cap z = \emptyset)) \rightarrow (\exists c \forall y (y \in x \rightarrow \exists! z (z \in y \cap c))))$

In other words, given a set of disjoint subsets, there is a set containing precisely one member from each subset. The Axiom of Choice is highly practical, and has many non-obvious consequences. We will discuss some of these shortly.

Any middle school student works with infinite sets. So far, we only have ways of constructing sets with finitely many elements. The power set of a finite set is finite, the finite union of finite sets is finite, etc. If we want our theory to have sets like  $\mathbb{R}$ , we need an axiom. Infinitude is rather difficult to write down in the first order language of set theory; here is a popular substitute.

**Axiom 8 (Infinity).**  $\exists x (\emptyset \in x \wedge (u \in x \rightarrow u \cup \{u\} \in x))$

As hoped, the the set  $x$  postulated by this axiom is infinite. We hold off on proving this until we have finished listing the axioms. The next interesting axiom also merits a definition.

**Definition 2.4.** A formula  $p(x, y)$  with free variables  $x$  and  $y$  will be called a functional relation if  $\forall x \exists! y(p(x, y))$ .

**Axiom 9 (Replacement schema).** Suppose  $p(w, z)$  is a functional relation and  $x$  is a set. Then the following is an axiom.  $\exists y(z \in y \leftrightarrow \exists w(w \in x \wedge p(w, z)))$ .

In other words, if  $x$  is any set then its image under a functional relation is also a set. This axiom vastly extends the scope of set theory, as will be seen in later sections.

The final axiom is at this moment rather unmotivated. We will discuss its consequences later, when we get to wellfounded relations. Essentially, this axiom prohibits things that just “don’t look quite right” from in fact being sets.

**Axiom 10 (Foundation).**  $\forall x(x \neq \emptyset \rightarrow \exists y(y \in x \wedge \nexists z(z \in y \wedge z \in x)))$

As an example, this axiom rules out the possibility  $x \in x$ . For if this were the case, consider  $y = \{x\}$ . Then the Axiom of Foundation requires that  $y$  be disjoint from  $x$ , but this is not the case since both sets contain  $x$ . This axiom also rules out the possibility of a set of all sets—such a set must be a member of itself. Further, it follows that the collection of sets which do not contain themselves encompasses all sets, and hence is not a set.

We’ll also use this axiom to show that the set  $x$  whose existence is guaranteed by the Axiom of Infinity is actually infinite. To see this, assume that  $x$  contains exactly  $n$  distinct elements. Then  $x$  contains an element (possibly not unique)  $u$  with the most members, say  $m$ . By the Axiom of Infinity,  $u \cup \{u\} \in x$  as well. By the Axiom of Foundation,  $u \notin u$  already, so  $u \cup \{u\}$  contains exactly  $m + 1$  elements, and hence must be distinct from the  $n$  elements of  $x$  previously discussed. Hence  $x$  contains  $n + 1$  elements.

Although set theory seems to require the strictest formalism, we will tend to state and prove theorems in the meta-language. For all of our convenience (but mostly mine), we will avoid the cumbersome notation of formulas and sentences and formal proofs, and speak instead in the usual dialect of mathematics.

## 2.3 Sets and classes

We now have a precise definition of what it means to be a set—a set is an object with certain properties that is guaranteed to exist by some axioms of set theory. The main lesson to be learned from Russell’s paradox is that even though the objects in the language of set theory are themselves “sets,” it is not the case that every collection of these objects is again a set.

However, it can be convenient and even enlightening to talk about arbitrary collections of sets. To do this, we need a way of specifying such collections, and fortunately we have the language of set theory to help us do so.

**Definition 2.5.** To each formula  $p(x)$  with free variable  $x$ , we associate a class  $C$ , whose members are precisely those sets  $x$  such that  $p(x)$  holds. We say  $x \in C$  if and only if  $p(x)$ .

For example, if  $p(x)$  is the formula  $x = x$ , then it holds for all  $x$  and the associated class contains all sets as members. We will denote this class  $V$ , the class of all sets. It also happens that every set is a class, for if  $x$  is a set we may define  $p(z)$  as  $z \in x$  to obtain the class consisting precisely of elements of  $x$ .

A surprising number of definitions and properties from set theory carry over to classes as well. For example, we may mimic Extensionality and define two classes  $C$  and  $D$  to be equal if and only if every element of  $C$  is also an element of  $D$  and conversely. Similarly, we may write  $C \subset D$  if and only if every element of  $C$  is also an element of  $D$ .

In addition to allowing us to talk about the class of all sets, this definition allows us to simplify our discussion of Zermelo-Fraenkel set theory. Whenever we are inclined to quantify formulas in the language of set theory, we may instead quantify classes. For example, the traditional Separation Axiom

- If  $x$  is a set and  $p(z)$  is any formula in the language of set theory with one free variable, then the set  $y = \{z \in x : p(z)\}$  consisting of elements of  $x$  satisfying a property  $p(z)$  is a set.

May be replaced with the much simpler

- Subclasses of sets are sets.

We have seen that every set is a class, and have hinted that the converse is false. (For example, the class  $V$  of all sets cannot be a set, for if it were we would have  $V \in V$ .) A class which is not a set is called a proper class. We will shortly see a more broad definition of classes.

## 2.4 Models and set-models

Let  $L$  be a first order language with relational symbols  $\{r_i\}$  (each of arity  $n(i)$ ). A model for  $L$  is a class  $A$  together with relations  $\{R_i\}$  corresponding to the  $r_i$ . More specifically, each  $R_i$  is a subset of  $A^{n(i)}$ .

To explain exactly what a model is for, we need some new notation. Recall the construction of formulas over a language  $L$  in Section 2.1. Suppose that  $p(x_1, \dots, x_n)$  is a formula over  $L$ , and suppose that every free variable occurring in  $p$  is among the  $x_1, \dots, x_n$ . Let  $a_1, \dots, a_n \in A$ . Then  $p(a_1, \dots, a_n)$  is a new string obtained by substituting each free occurrence of  $x_i$  in  $p$  with  $a_i$ .

Recall the construction of formulas over a language  $L$  in Section 2.1. We construct a relation  $\models$  as follows.

- If  $a_1 = a_2$ , then  $A \models a_1 = a_2$ .  
If  $R_i a_1 \dots a_n$ , then  $A \models r_i a_1 \dots a_n$ .
- If  $A \models p(a_1, \dots, a_n)$  and  $A \models q(a_1, \dots, a_n)$  then  $A \models (p \wedge q)(a_1, \dots, a_n)$ .  
If  $A \not\models p(a_1, \dots, a_n)$ , then  $A \models \neg p(a_1, \dots, a_n)$ .
- If for every  $a \in A$ ,  $A \models p(a, a_2, \dots, a_n)$ , then  $A \models \forall x(p(x, a_2, \dots, a_n))$ .

Note that any formula without free variables is both a formula over  $L$  and over  $A$ . If  $T$  is a collection of formulas with no free variables, and for each  $p$  in  $T$   $A \models p$ , then we write  $A \models T$ , and we say that  $A$  is a model for  $T$ .

Intuitively speaking, a model for a theory  $T$  is a structure in which all of the formulas in  $T$  are true. In particular, if  $T$  is contradictory, then it should have no models. In fact, we have following theorem.

**Theorem 2.6 (Completeness).** A theory  $T$  over a language  $L$  has a model if and only if  $T$  is consistent. Equivalently,  $p \vdash q$  if and only if  $A \models p$  implies that  $A \models q$  for every model  $A$  of  $L$ .

A model for the language of set theory is a class  $A$  together with a binary relation  $E$ , corresponding to the usual relation  $\in$ .  $A$  is a model for set theory if  $A \models \text{ZFC}$ , where here ZFC represents the set of axioms introduced earlier. Constructing models for all of ZFC is a rather difficult task. Here are some examples of models for just some of the axioms of ZFC.

- Let  $a, b, c$  be any distinct sets. Let  $A = \{a, b, c\}$ , with  $aEb$  and  $aEc$ . Then  $A$  has a unique empty set,  $a$ . But  $A$  does not satisfy Extensionality because  $b = \{a\} = c$ , yet  $c \neq b$ .
- Again let  $A = \{a, b, c\}$ , but this time  $aEb$ ,  $bEc$ , and  $cEa$ . Then  $A$  satisfies Extensionality—no two sets are equal and no two sets have equal members. On the other hand,  $A$  does not have an empty set.

Due to the nature of set theory, it is possible to have models  $X$  for the language of set theory where the relation for  $\in$  is  $\in$  itself. If  $X$  is in fact a set (and not a proper class), then  $X$  will be called a set-model.

The classes inside a set-model  $X$  are collections of elements of  $X$  satisfying some property  $p(x)$ . By the usual Separation Axiom, these are all subsets of  $X$ . However, not every subset of  $X$  need be definable by a formula  $p(x)$ . This only holds in the so-called constructible universe of Gödel, in which many other ridiculous properties such as the Generalized Continuum Hypothesis hold as well.

## 2.5 Other languages

First order languages are not the only way of doing things. There are simpler systems which do not involve quantifiers, such as the propositional calculus or boolean algebra. Such systems have proven themselves both instructive, and useful in applications.

On the other end of the spectrum, one might allow more powerful tools. A common class of examples are the so-called systems of second order logic, which employ more general quantifiers. Recall that a first order language has a sequence of variable symbols which stand for the objects of the language. A system of second order logic will additionally possess a new sequence of variable symbols to stand either for collections of the objects, or functions on the objects, or both.

Another second order tool is to allow quantification over formulas. This construction implicitly allows one to quantify over collections of sets (rather than just sets), since given any property  $p(x)$  there is a collection of elements  $x$  satisfying  $p(x)$ . Under our earlier definitions, this is equivalent to quantifying over classes.

## 2.6 Von-Neumann Bernays class theory

The Von-Neumann Bernays (VNB) theory is one of the more mild extensions of first order logic. It extends the language by allowing variable symbols called class variables. These symbols are typically denoted by capital letters  $A, B, \dots$ . As the usual variables are meant to stand for sets, the class variables are meant to stand for collections of sets. The  $\in$  symbol is extended to allow usual objects (sets) to be elements of the new objects (classes). More specifically, the formulas of VNB are constructed as before from the atomic formulas  $A = B$ ,  $x = y$ ,  $x \in y$ , and  $x \in A$ .

We define the Axiom of Extensionality for classes in the expected way;

**Axiom 1' (Extensionality).**  $\forall AB(A = B \leftrightarrow \forall z(z \in A \leftrightarrow z \in B))$

An extension of VNB theory due to Gödel is as follows. If, given a class  $A$ , there is a set  $x$  such that  $z \in x \leftrightarrow z \in A$ , we say  $x = A$  and  $A = x$  and identify the two objects. With this, we define all symbols  $A \subset B$ ,  $x \subset A$ ,  $A \subset x$ ,  $A \cap B$ , etc., as one would expect.

Most of the axioms of ZFC also show up in VNB. We will highlight some of the exceptions and additions here. Other than Extensionality, there are two axioms that tell us something new about classes. First, the Axiom of Foundation is replaced by an identical statement for all classes. Second, we have an Axiom Schema of Separation for Classes, which allows us to construct all of the classes that we'll need.

**Axiom 11 (Separation schema for Classes).** Suppose  $p(z)$  is a formula of VNB which contains no quantifiers over classes, that is, no part of the form  $\forall A(\dots)$ . Then the following is an axiom.  $\exists A \forall z(z \in A \leftrightarrow p(z))$ .

In fact, there exist five sentences which, when taken together as axioms, are equivalent to the axiom schema just stated. The construction is somewhat technical; the axiom given here is far more useful than five somewhat arcane axioms. However, it is an important property of VNB that it can be expressed by a finite list of axioms.

It follows from the Axiom Schema of Separation for Classes that every set is a class. To see this, let  $x$  be a set and let  $p(z)$  stand for  $z \in x$ . Then the axiom gives a class  $A$  with exactly the same members as  $x$ , in other words,  $x = A$  is a class.

Using the powerful notion of classes, we may reduce the axiom schemas of ZFC to single sentence axioms in the language of VNB. We first simplify the Axiom Schema of Separation. By the discussion in the section on classes above,

the the clumsy “for every formula  $p(z)$ ” can be replaced by “for every class  $A$ ,” a valid quantification in VNB.

**Axiom 6' (Separation).**  $\forall A(\forall x\exists y(z \in y \leftrightarrow z \in x \cap A))$

It is easy to see that this is equivalent to the old Axiom 6. It is nearly as easy to replace the Axiom Schema of Replacement by a single sentence in the language of VNB, as follows.

**Axiom 9' (Replacement).**  $\forall F(\text{Fun}(F) \rightarrow \forall x\exists y\forall z(z \in y \leftrightarrow \exists w(w \in x \wedge z = F(w))))$

Technically, by  $\text{Fun}(F)$  we mean the expression shown in Definition 2.1, with a class variable  $F$  substituted for the argument  $f$ . Thus, in this context,  $\text{Fun}(F)$  is a sentence of VNB, and not of ZFC.

We have already mentioned that there are finitely many axioms of VNB. A second fact which makes the use of VNB so appealing is that it proves precisely the same theorems about sets as ZFC. That is, if  $p$  is a formula with no class variables, then  $\text{ZFC} \vdash p$  if and only if  $\text{VNB} \vdash p$ .

## 3 Important Definitions and Theorems

### 3.1 Relations and orderings

We temporarily consider a first order language with a single binary relation, denoted  $<$ . We will abbreviate  $x < y \vee x = y$  with  $x \leq y$ .

**Definition 3.1.** A strict partial order satisfies the following axioms.

- (areflexivity)  $\forall x(x \not< x)$
- (transitivity)  $\forall xyz(x < y \wedge y < z \rightarrow x < z)$

**Definition 3.2.** A strict total order is a strict partial order satisfying the following additional axiom.

- (trichotomy)  $\forall xy(x < y \vee x = y \vee y < x)$ , and  $\forall xy(x < y \rightarrow x \neq y \wedge y \not< x)$

**Definition 3.3.** If  $(x, <)$  is a strict partial order, a chain  $y \subset x$  is any subset of  $x$  which is a strict total order under  $<$ .

If  $y \subset x$  is a chain, an upper bound for the chain is an element  $u$  of  $x$  such that for all  $z \in y$ ,  $z \leq u$ .

The most useful models for orderings are classes  $X$  with some relation  $<$ , which can be thought of as a subclass of  $X \times X$ . From now on, we will consider these.

**Definition 3.4.** Let  $X$  be a class, and let  $(X, <)$  be a partial order. Then  $(X, <)$  is called wellfounded if it satisfies the following additional condition.

- For each nonempty  $Y \subset X$ ,  $Y$  has a least element  $a$ . That is,  $\forall b \in Y (a < b \vee a = b)$ .

If  $(X, <)$  is simultaneously wellfounded and strictly totally ordered, then we say it is strictly well ordered. In this case, if  $x \in X$ , we denote by  $x^+$  the unique least element greater than  $x$ .

We now turn to the most important theorem about wellfounded relations on classes. First, we narrow slightly the classes that we'll work with.

**Definition 3.5.** A class  $X$  with a wellfounded relation  $<$  is called set-like if for every  $a \in X$ , the subclass  $\text{Init}_X(a)$  consisting of all  $x \in X$  with  $x < a$  is in fact a set.

**Theorem 3.6 (Recursive Definitions).** Let  $D$  be a well ordered set-like class, and  $R$  any class. Suppose  $G : \mathcal{F} \rightarrow R$  is a class function with domain  $\mathcal{F} \subset V \times D$  a class satisfying the following conditions.

- $\mathcal{F}$  consists of pairs  $(f, x)$  such that  $f : \text{Init}_D(x) \rightarrow R$ .
- Let  $(f, x) \in \mathcal{F}$ , and define  $g : \text{Init}_D(x^+) \rightarrow R$  by  $g|_{\text{Init}_D(x)} = f$ ,  $g(x) = G(f, x)$ . Then  $(g, x^+) \in \mathcal{F}$ .
- If  $x$  is limit,  $f : \text{Init}_D(x) \rightarrow R$ , and  $(f|_{\text{Init}_D(y)}, y) \in \mathcal{F}$  for all  $y \in \text{Init}_D(x)$ , then  $(f, x) \in \mathcal{F}$ .

Then there exists a unique functional class  $F : D \rightarrow R$  such that for all  $x \in X$ ,  $F(x) = G(F|_{\text{Init}_D(x)}, x)$ . This theorem is thus a generalization of Theorem 3.6.

When the classes  $D$  and hence  $G, \mathcal{F}$  are in fact sets, the theorem is well known. We will assume the theorem holds in this case, and prove the stated generalization to classes.

*Proof.* For each  $d \in D$ , let  $\mathcal{F}_d = \{(f, x) \in \mathcal{F} : x \in \text{Init}_D(d)\}$  and  $G_d = G|_{\mathcal{F}_d}$ . Then since  $D$  is set-like,  $\text{Init}_D(d)$  is a set, and the hypotheses of the set version of the theorem are trivially satisfied. Hence there is a unique  $F_d : \text{Init}_D(d) \rightarrow R$  such that for all  $x \in \text{Init}_D(d)$ ,  $F_d(x) = G_d(F_d|_{\text{Init}_D(x)}, x)$ . From the fact that each  $F_d$  is unique, it follows that if  $d_1 < d_2$ , then  $F_{d_2}|_{\text{Init}_D(d_1)} = F_{d_1}$ .

Now, for each  $x \in D$ , define  $F(x) = F_{x^+}(x)$ . Then for all  $d \in D, x < d$ ,

$$F|_{\text{Init}_D(d)}(x) = F(x) = F_{x^+}(x) = G_{x^+}(F_{x^+}|_{\text{Init}_D(x)}, x) = G_d(F_d|_{\text{Init}_D(x)}, x) = F_d(x)$$

In other words, for each  $d \in D$ ,  $F|_{\text{Init}_D(d)} = F_d$ . Now in particular,

$$F(x) = F_{x^+}(x) = G_{x^+}(F_{x^+}|_{\text{Init}_D(x)}, x) = G(F_x, x) = G(F|_{\text{Init}_D(x)}, x)$$

So that  $F$  satisfies the desired property.  $\square$

## 3.2 Ordinals

**Definition 3.7.** An ordinal is a set  $\alpha$  which is transitive and which is a well order under the relation  $\in$ .

**Proposition 3.8.** The ordinals themselves are strictly well ordered by  $\in$  and we will often denote this ordering  $<$ . It follows that every nonempty collection of ordinals has a minimal element.

If  $\alpha$  is the least ordinal greater than some other ordinal  $\lambda$ , then we write  $\alpha = \lambda + 1$  and say that  $\alpha$  is the successor of  $\lambda$ . If there is no such  $\lambda$ , we say  $\alpha$  is limit.

The ordinals follow the following recipe:

- The smallest ordinal is  $\emptyset$
- $\alpha = \lambda + 1$  if and only if  $\alpha = \lambda \cup \{\lambda\}$
- $\alpha$  is limit if and only if it is not  $\lambda + 1$  for any  $\lambda$
- In both cases,  $\alpha = \bigcup_{\beta < \alpha} \beta$

For each natural  $n$ , there is a unique ordinal with  $n$  elements, and we always identify the natural  $n$  with this ordinal. The first infinite ordinal is  $\omega$ , the union of all finite ordinals. If  $\alpha$  is an ordinal, we denote by  $\alpha + n$  the  $n^{\text{th}}$  least ordinal greater than  $\alpha$ .

Finally, the relation  $\in$  is a strictly well ordered relation on the class of ordinals.

We now have our first example of a non-ridiculous collection of sets which is not itself a set.

**Proposition 3.9.** The collection of ordinals is not a set.

*Proof.* Suppose  $O$  were the set of ordinals. Then by the above remarks,  $O$  is well ordered and it is clearly transitive, so it is an ordinal. Hence  $O \in O$ , contradicting the Axiom of Foundation.  $\square$

It is convenient to use an ordinal as an index set. If  $\alpha$  is used to index a set  $x = \{x_\beta : \beta < \alpha\}$ , we say  $x$  is an  $\alpha$ -sequence.

**Definition 3.10.** If  $x$  is any set of ordinals, then  $\sup x$  is defined to be the least ordinal which is greater than every element of  $x$ .

If  $x = \{x_\beta : \beta < \alpha\}$  is an  $\alpha$ -sequence, then we sometimes write  $\sup_{\beta < \alpha} x_\beta$  for  $\sup x$ .

**Definition 3.11.** If  $\lambda$  is any limit ordinal, define the cofinality of  $\lambda$ ,  $\text{cf}(\lambda)$  to be the least ordinal  $\alpha$  such that there is an  $\alpha$ -sequence  $x$  with  $\sup x = \lambda$ .

**Definition 3.12.** Call a limit ordinal  $\lambda$  regular if  $\text{cf}(\lambda) = \lambda$ .

We now return to recursion in the context of ordinals. Consider Theorem 3.6 in the special case that  $X$  is the class of ordinals, with the wellfounded relation  $\in$ . We always have that  $\text{Init}_X(\alpha) = \alpha$ . This special case is called Ordinal Recursion, and allows us to define sequences indexed by the class of ordinals. We have the following application of this theorem, together with the axioms of Choice and Replacement.

**Lemma 3.13 (Zorn's Lemma).** If  $(x, <)$  is a strict partial order such that every chain in  $x$  has an upper bound, then  $x$  contains a maximal element.

*Proof.* Suppose that each chain in  $x$  has an upper bound, but  $x$  contains no maximal element. We will recursively construct a map  $F$  on the ordinals.

Let  $\mathcal{F}$  be the class of pairs  $(f, \alpha)$  such that  $f : \alpha \rightarrow x$  is increasing. If  $f$  is such a function, then the image of  $f$  is a chain, and we may choose an upper bound  $b$  for this chain. Since  $b$  is not maximal, choose  $c \in x$  with  $b < c$  and define  $G(f, \alpha) = c$ . Then the function  $f' : \alpha + 1 \rightarrow x$  is such that  $f'|_\alpha = f$  and  $f'(\alpha) = c$  is increasing. The three points that  $\mathcal{F}$  must satisfy are now readily verified.

By Ordinal Recursion, there is an increasing function  $F$  on the ordinals such that  $F(\alpha) = G(F|_\alpha, \alpha)$  for all  $\alpha$ . The whole point of this construction is that  $F$  itself is now increasing.

Now  $F$  is a map from the ordinals into  $x$ . Let  $y \subset x$  be the range of  $F$ , a set by Separation. Since  $F$  is strictly increasing, it is in particular injective. Hence there is a well-defined, surjective inverse map from  $y$  into the ordinals. Replacement then implies that the collection of ordinals is a set, a contradiction.  $\square$

### 3.3 Cardinals

Cardinals are a special subclass of the ordinals. In this section a lower-case Greek letter will most frequently denote a cardinal rather than an ordinal. In future sections the greek letters will be used to refer to ordinals and cardinals.

**Definition 3.14.** A cardinal is an ordinal  $\alpha$  which is not in bijection with any ordinal  $\beta$  such that  $\beta < \alpha$ .

Since one can easily construct a bijection between any infinite  $\alpha$  and  $\alpha + 1$ , it is immediate that all infinite cardinals are limit ordinals. We now present (without proofs) some basic facts about cardinals which will be needed in later sections.

**Proposition 3.15.** Each nonempty set  $x$  is in bijection with a unique cardinal, called the cardinality of  $x$  and denoted  $|x|$ . For consistency,  $|\emptyset| = \emptyset$ .

For  $x, y \neq \emptyset$ ,  $|x| \leq |y|$  if and only if there is an injective function  $x \rightarrow y$ ;  $|x| \geq |y|$  if and only if there is a surjective function  $x \rightarrow y$ ; and  $|x| = |y|$  if and only if there is a bijection  $x \rightarrow y$ .

**Definition 3.16.** If  $\alpha$  and  $\beta$  are cardinals, define  $\alpha \cdot \beta = |\alpha \times \beta|$ , and  $2^\alpha = |\mathcal{P}(\alpha)|$ .

**Proposition 3.17.** If  $\alpha$  and  $\beta$  are infinite cardinals, then  $\alpha \cdot \beta = \max(\alpha, \beta)$ .

**Corollary 3.18.** The union of at most  $\kappa$  sets, each of cardinality at most  $\kappa$  in turn has cardinality at most  $\kappa$ .

**Proposition 3.19 (Cantor).** If  $x$  is any set, then  $|x| < 2^{|x|}$ .

*Proof.* Clearly,  $|x| \leq |\mathcal{P}(x)| = 2^{|x|}$  by the injection  $y \mapsto \{y\}$ . On the other hand, suppose  $|x| = |\mathcal{P}(x)|$ . Then there is a bijection  $f : x \rightarrow \mathcal{P}(x)$ . Let  $d = \{z \in x : z \notin f(z)\} \subset x$ . Since  $f$  is a bijection, there is a  $c \in x$  such that  $f : c \mapsto d$ . One then sees from the definition of  $d$  that  $c \in f(c) = d$  if and only if  $c \notin f(c) = d$ , a contradiction.  $\square$

**Proposition 3.20.** We may recursively construct a hierarchy of cardinals. We write  $\omega_0 = \omega$  is the least infinite cardinal. If  $\omega_\alpha$  is a cardinal, there is a least cardinal greater than it, denoted  $\omega_{\alpha+1}$ . If  $\lambda$  is a limit ordinal, then  $\sup_{\beta < \lambda} \omega_\beta$  is again a cardinal denoted  $\omega_\lambda$ .

In fact,  $\sup_{\beta < \lambda} \omega_\beta = \bigcup_{\beta < \lambda} \omega_\beta$ , but to prove that this is even a set, the Axiom of Replacement is required. It is this fact that will cause us very much grief towards the end of the paper.

**Definition 3.21.** A cardinal  $\alpha$  is called a strong limit if for every  $\beta < \alpha$ ,  $2^\beta < \alpha$ .

**Definition 3.22.** A cardinal  $\alpha$  is called inaccessible if it is regular and a strong limit.

Note that any ordinal satisfying the strong limit or inaccessible properties must in fact be a cardinal.

## 4 The Zermelo Hierarchy

**Definition 4.1.** For each ordinal  $\alpha$ , define the set  $V_\alpha$  as follows:

- $V_0 = \emptyset$
- $V_{\alpha+1} = \mathcal{P}(V_\alpha)$
- $V_\alpha = \bigcup_{\beta < \alpha} V_\beta$  if  $\alpha$  is a limit ordinal

The collection of  $V_\alpha$  is called the Zermelo Hierarchy. The class  $V = \bigcup V_\alpha$  (where the union symbol means that  $V$  contains every element which is in some  $V_\alpha$ ) is the collection of sets for Zermelo-Fraenkel set theory, as we'll see later.

Each  $V_\alpha$  is called a partial universe. We will show shortly that together the  $V_\alpha$  comprise all sets.

**Definition 4.2.** For each set  $x$ , define the ordinal  $\text{rank}(x)$  to be the least ordinal  $\alpha$  such that  $x \in V_{\alpha+1}$ .

It should be noted that in our terms, the classes of a partial universe  $V_\alpha$  are the subsets  $C \subset V_\alpha$  which are definable by a formula  $p(x)$ , in which all free variables are understood to refer only to elements of  $V_\alpha$ . The collections of sets of  $V_\alpha$  are just subsets of  $V_\alpha$ .

#### 4.1 Some eclectic facts about the hierarchy

Here, we state and prove a few little properties of the Zermelo hierarchy that will turn out indispensable in later discussion.

**Lemma 4.3.** For every  $\alpha$ ,  $x \subset V_\alpha$  if and only if  $x \in V_{\alpha+1}$ .

*Proof.* This follows from the second part of the recursive definition in 4.1. If  $x \subset V_\alpha$ , then by the very definition of the power set,  $x \in \mathcal{P}(V_\alpha) = V_{\alpha+1}$ , and conversely.  $\square$

**Proposition 4.4.** If  $x \in V_\alpha$ , then there is a  $\beta < \alpha$  with  $x \subset V_\beta$ .

*Proof.* We proceed by induction on  $\alpha$ . The base case is vacuous. Lemma 4.3 proves this in the case that  $\alpha$  is successor. If  $\alpha$  is limit, then  $V_\alpha = \bigcup_{\gamma < \alpha} V_\gamma$ , so there is a  $\gamma < \alpha$  such that  $x \in V_\gamma$ . By the inductive hypothesis, there is a  $\beta < \gamma$  such that  $x \subset V_\beta$ .  $\square$

**Proposition 4.5.** If  $\beta < \alpha$ , then  $V_\beta \subset V_\alpha$ .

*Proof.* We proceed by induction on  $\alpha$ . The base case is trivial. Assume that for every  $\gamma < \alpha$ , if  $\beta < \gamma$  then  $V_\beta \subset V_\gamma$ . Let  $\beta < \alpha$ . If  $\alpha$  is limit, then clearly  $V_\beta \subset \bigcup_{\gamma < \alpha} V_\gamma = V_\alpha$ , by the definition. If  $\alpha = \lambda + 1$ , then the inductive hypothesis gives that  $V_\beta \subset V_\lambda$ . Let  $x \in V_\lambda$ . Then by Proposition 4.4, there is a  $\gamma < \lambda$  with  $x \subset V_\gamma$ , and again by the inductive hypothesis  $V_\gamma \subset V_\lambda$  so that  $x \subset V_\lambda$ . Hence, by Lemma 4.3,  $x \in V_{\lambda+1} = V_\alpha$ .  $\square$

**Corollary 4.6.** If  $y \subset x$  and  $x \in V_\alpha$ , then  $y \in V_\alpha$ .

*Proof.* By Proposition 4.4, there is a  $\beta < \alpha$  with  $y \subset x \subset V_\beta$ . Hence  $y \in V_{\beta+1} \subset V_\alpha$ .  $\square$

**Definition 4.7.** A set  $A$  is transitive if each element of  $A$  is also a subset of  $A$ .

**Corollary 4.8.** Each  $V_\alpha$  is transitive.

*Proof.* By Propositions 4.4 and 4.5, if  $x \in V_\alpha$  then there is a  $\beta < \alpha$  such that  $x \subset V_\beta \subset V_\alpha$ .  $\square$

**Theorem 4.9.** For any set  $x$ , there is an  $\alpha$  with  $x \in V_\alpha$ .

This is the most important result about the hierarchy, and is the true reason for studying it at all.

*Proof.* Suppose  $x_0$  is a set such that  $x_0 \notin V_\alpha$  for all  $\alpha$ . Inductively choose  $x_i \in x_{i-1}$  such that  $x_i \notin V_\alpha$  for all  $\alpha$ . Then  $x_i$  must contain an element  $x_{i+1}$  such that  $x_{i+1} \notin V_\alpha$  for all  $\alpha$ . For if otherwise, that is, if for all  $y \in x_i$ ,  $y \in V_{\alpha_y}$ , let  $\alpha = \bigcup \alpha_y$ . Then  $y \in V_{\alpha_y} \subset V_\alpha$  for each  $y \in x_i$ , hence  $x_i \in \mathcal{P}(V_\alpha) = V_{\alpha+1}$ , a contradiction. We have thus constructed an infinite decreasing chain

$$\dots \in x_i \in x_{i-1} \in \dots \in x_0$$

This contradicts that  $\in$  is wellfounded, as the set  $\{x_i\}$  has no  $\in$ -minimal element.  $\square$

**Proposition 4.10.** Definition 4.1 may be formulated in a slightly different fashion. That is, if we define

- $V'_0 = \emptyset$
- $V'_\alpha = \bigcup_{\beta < \alpha} \mathcal{P}(V'_\beta)$

then for all  $\alpha$ ,  $V'_\alpha = V_\alpha$ .

*Proof.* We proceed by induction on  $\alpha$ . Clearly,  $V_0 = \emptyset = V'_0$ . Suppose  $V_\beta = V'_\beta$  for all  $\beta < \alpha$ . Now, if  $\beta < \alpha$ , then  $\mathcal{P}(V'_\beta) = V_{\beta+1}$ . From the definitions, we have

$$V'_\alpha = \bigcup_{\beta < \alpha} \mathcal{P}(V'_\beta) = \bigcup_{\beta < \alpha} \mathcal{P}(V_\beta) = \bigcup_{\beta < \alpha} V_{\beta+1}$$

In the case that  $\alpha$  is limit, the right-hand side is just  $\bigcup_{\beta < \alpha} V_\beta = V_\alpha$ . On the other hand, if  $\alpha = \lambda + 1$ , then Proposition 4.5 implies that the right-hand side just becomes  $\bigcup_{\beta \leq \lambda} V_{\beta+1} = V_{\lambda+1} = V_\alpha$ .  $\square$

**Proposition 4.11.** For every  $\alpha$ , we have  $\alpha \notin V_\alpha$  but  $\alpha \in V_{\alpha+1}$ . In other words,  $\text{rank}(\alpha) = \alpha$ .

*Proof.* Again by induction on  $\alpha$ . Suppose  $\alpha = \lambda + 1$ . Then if  $\alpha \in V_\alpha$ , Lemma 4.3 gives that  $\lambda \in \alpha \subset V_\lambda$ , a contradiction. On the other hand,  $\alpha \subset \mathcal{P}(\lambda) \subset \mathcal{P}(V_\lambda) = V_\alpha$ , and hence  $\alpha \in V_{\alpha+1}$ .

On the other hand, suppose  $\alpha$  is limit. Then if  $\alpha \in V_\alpha$ , there is a  $\beta < \alpha$  with  $\alpha \in V_\beta$ . By Corollary 4.8,  $\beta + 1 \in \alpha \subset V_\beta$ , contrary to the inductive hypothesis. For the second statement, the inductive hypothesis together with Proposition 4.5 gives that each member of  $\alpha$  is in  $V_\alpha$ , so that  $\alpha \in \mathcal{P}(V_\alpha) = V_{\alpha+1}$ .  $\square$

**Proposition 4.12.** If  $\alpha$  is limit,  $x, y \in V_\alpha$ , and  $f : x \rightarrow y$  is any function, then  $f \in V_\alpha$ .

*Proof.* Choose by Proposition 4.4 an ordinal  $\beta < \alpha$  such that  $x, y \subset V_\beta$ , and hence  $x, y \in V_{\beta+1}$ . Each element of  $f$  is of the form  $\langle a, b \rangle = \{\{a\}, \{a, b\}\}$  where  $a \in x$ ,  $b \in y$ . This pair is evidently an element of  $\mathcal{P}(\mathcal{P}(V_{\beta+1})) = V_{\beta+3}$ . Since  $f$  consists solely of such elements, it follows that  $f \in \mathcal{P}(V_{\beta+3}) = V_{\beta+4} \subset V_\alpha$ .  $\square$

**Proposition 4.13.** If  $\alpha$  is limit,  $x, y \subset V_\alpha$ , and  $f : x \rightarrow y$  is any function, then  $f \subset V_\alpha$ .

*Proof.* Let  $\langle a, b \rangle$  be an arbitrary element of  $f$ . Then  $a, b \in V_\alpha$ . Since  $\alpha$  is limit, there is a  $\beta < \alpha$  such that  $a, b \in V_\beta$ . As before, it is evident that  $\langle a, b \rangle = \{\{a\}, \{a, b\}\} \in V_{\beta+3} \subset V_\alpha$ . Since this is true of an arbitrary element of  $f$ ,  $f \subset V_\alpha$ .  $\square$

## 5 Zermelo Hierarchy and the Axioms of ZFC

Since the universe of sets arises from the Zermelo hierarchy, it is natural to expect that each partial universe  $V_\alpha$  provides a partial model for ZFC (that is, a model for some of the axioms of ZFC). In fact, since the  $V_\alpha$  are themselves sets, they provide set-models for set theory.<sup>3</sup> It is thus useful to know precisely which axioms of the theory will actually hold at each  $V_\alpha$ , and this is the primary purpose of this section.

The presentation will be as plain as possible—we no longer use the symbols from the language of set theory and revert to ordinary mathematical notation. We will also need to review the axioms of ZFC, this time they will be presented in plain english.

### 5.1 Axioms that always hold

**Axiom 1 (Extensionality).** If  $x, y$  are sets, then  $x = y$  if and only if the following holds. For all sets  $z$ ,  $z \in x$  if and only if  $z \in y$ .

**Theorem 5.1.** The Axiom of Extensionality holds in all  $V_\alpha$ .

*Proof.* If there is a  $z \in V_\alpha$  (that is,  $z$  is a set in the universe  $V_\alpha$ ) such that  $z \in x - y$  or  $z \in y - x$ , then obviously  $x \neq y$ .

For the converse, note that if  $x, y$  are sets in the universe  $V_\alpha$ , then  $x, y \in V_\alpha$ , so that by Proposition 4.8,  $x, y \subset V_\alpha$ , whence each element of  $x$  and of  $y$  is a set. Thus, if for all sets  $z$ ,  $z \in x$  if and only if  $z \in y$ , then in particular every element of  $x$  is an element of  $y$  and vice versa. That is,  $x = y$ .  $\square$

**Axiom 10 (Foundation).** If  $x$  is a set, then there is a  $z \in x$  such that  $z \cap x = \emptyset$ .

**Theorem 5.2.** The Axiom of Foundation holds in all  $V_\alpha$ .

*Proof.* If  $x \in V_\alpha$ , then there is a  $z \in x$  such that  $z \cap x = \emptyset$ . Since  $V_\alpha$  is transitive, in fact  $z \in V_\alpha$  as well.  $\square$

**Axiom 4 (Union).** If  $x$  is a set, then there is a set, denoted  $\bigcup x$ , whose elements are precisely those which are elements of some element of  $x$ .

<sup>3</sup>No, it's not possible to say the word "set" too many times

**Lemma 5.3.** If  $x \in V_\alpha$ , then  $\bigcup x \in V_\alpha$ .

*Proof.* If  $x \in V_\alpha$ , by Proposition 4.4 there is a  $\beta < \alpha$  such that  $x \subset V_\beta$ . In other words, for each  $y \in x$ ,  $y \in V_\beta$ , and hence the transitivity of  $V_\beta$  gives that for each  $z \in y \in x$ , we have  $z \in V_\beta$  as well. Hence, if  $w = \bigcup x$  (the usual union), then  $w \in V_{\beta+1} \subset V_\alpha$ .  $\square$

**Theorem 5.4.** The Union Axiom holds in all  $V_\alpha$ .

*Proof.* Let  $w = \bigcup x \in V_\alpha$  (the usual union). It remains only to check that  $w = \bigcup x$  in the model  $V_\alpha$  as well. To do this, let  $z \in V_\alpha$ . If  $z \in w$ , then by definition of  $w$  there is a  $y \in x$  such that  $z \in y$ . Since  $V_\alpha$  is transitive,  $y \in V_\alpha$  as well. Hence  $w$  satisfies the definition of  $\bigcup x$  in  $V_\alpha$ .  $\square$

**Corollary 5.5.** If  $\alpha$  is limit,  $a \subset \alpha$ , and  $\sup a = \alpha$ , then  $a \notin V_\alpha$ .

*Proof.* In this case, we have that  $\sup a = \alpha = \bigcup a$ . To see the second equality, note that for every  $\beta < \alpha$ , there is a  $b \in a$  such that  $\beta < b$ . In other words,  $\beta \in \alpha$  implies there is a  $b$  with  $\beta \in b \in a$ . This shows that  $\alpha \subset \bigcup a$ .

Now, if it were the case that  $a \in V_\alpha$ , since  $\alpha$  is limit there must be a  $\beta < \alpha$  such that  $a \in V_\beta$ . By the theorem,  $\alpha = \sup a = \bigcup a \in V_\beta$ , contradicting Proposition 4.11.  $\square$

**Axiom 6 (Separation).** If  $x$  is a set, and  $y$  is a class such that  $y \subset x$ , then  $y$  is a set.

**Theorem 5.6.** The Separation axiom schema holds in all  $V_\alpha$ .

*Proof.* Suppose  $x$  is an element of  $V_\alpha$ . If  $y$  is any class at all which is a subset of  $x$ , then we have  $y \subset x \in V_\alpha$  and so by Corollary 4.6,  $y \in V_\alpha$  as well.  $\square$

**Axiom 7 (Choice).** If  $x$  is set of nonempty, pairwise disjoint sets, then there is a choice set  $z$  consisting of precisely one element from each  $y \in x$ .

**Theorem 5.7.** The Axiom of Choice holds in all  $V_\alpha$ .

*Proof.* Suppose such an  $x \in V_\alpha$  is given. If the elements of  $x$  are pairwise disjoint in  $V_\alpha$ , they are trivially pairwise disjoint. With the Axiom of Choice one chooses a choice set  $z$  for  $x$ . Then by an argument parallel to that given in Lemma 5.3,  $z \in V_\alpha$ .

Once again, it remains only to check that  $z$  is a choice set for  $x$  in the model  $V_\alpha$ . For each  $y \in x$ , we know  $z \cap y$  is a singleton, say  $\{u\}$ . Since  $u \in z$ , it follows that  $u \in V_\alpha$ .  $\square$

## 5.2 Existence axioms

Not surprisingly, this type of axiom holds true when the set whose existence it postulates first appears in the hierarchy. By Proposition 4.5, the axiom will hold for all larger ordinals as well.

**Axiom 2 (Empty Set).** There is a set, denoted  $\emptyset$ , consisting of no elements. That is, for every set  $x$ ,  $x \notin \emptyset$ .

**Theorem 5.8.** The Empty Set axiom holds in  $V_\alpha$  if and only if  $\alpha > 0$ .

*Proof.*  $V_0 = \emptyset$ , so nothing postulating the existence of a set holds. On the other hand,  $\emptyset \in \mathcal{P}(\emptyset) = V_1$ .  $\square$

**Axiom 8 (Infinity).** There is a nonempty set  $x$  such that  $u \in x$  implies  $u \cup \{u\} \in x$ .

**Lemma 5.9.**  $V_\omega$  contains no infinite set.

*Proof.* By a similar induction, for each natural  $n$ ,  $V_n$  consists only of finite sets. Hence,  $V_\omega = \bigcup_{n \in \omega} V_n$  consists only of finite sets as well, whence  $V_\omega$  cannot contain an infinite set.  $\square$

It follows from elementary properties of the ordinals that if  $\omega$  is a set, then  $x$  in the Axiom of Infinity may be taken to be  $\omega$ . As it turns out,  $\omega$  is the earliest (rank-wise) such  $x$  one can choose.

**Theorem 5.10.** The Axiom of Infinity holds in  $V_\alpha$  if and only if  $\alpha > \omega$ .

*Proof.* Lemma 5.9 shows that the Infinity cannot hold if  $\alpha < \omega$ . Proposition 4.11 shows that  $\omega \in V_{\omega+1}$ , and hence Infinity does hold if  $\alpha > \omega + 1$ .  $\square$

### 5.3 Axioms holding at limit ordinals

These are the set construction axioms. Not surprisingly, since they create larger sets from old, they require a little bit of wiggle room. Limit ordinals furnish this wiggle room quite nicely.

**Axiom 5 (Power Set).** If  $x$  is a set, then there is a set, denoted  $\mathcal{P}(x)$ , such that  $y \subset x$  if and only if  $y \in \mathcal{P}(x)$ .

**Lemma 5.11.** If  $\alpha$  is limit, then for each  $x \in V_\alpha$ ,  $\mathcal{P}(x) \in V_\alpha$ .

*Proof.* If  $x \in V_\alpha$ , there is a  $\beta < \alpha$  with  $x \in V_\beta$ . If  $z \in \mathcal{P}(x)$ , then  $z \subset x \subset V_\beta$ , so  $z \in V_{\beta+1}$ . Hence  $\mathcal{P}(x) \subset V_{\beta+1}$ . Since  $\alpha$  is limit,  $\mathcal{P}(x) \in V_{\beta+2} \subset V_\alpha$ .  $\square$

**Theorem 5.12.** If  $\alpha$  is limit, then the Power Set Axiom holds in  $V_\alpha$ .

*Proof.* If  $x \in V_\alpha$ , let  $w = \mathcal{P}(x) \in V_\alpha$ . We must check that  $w = \mathcal{P}(x)$  in the model  $V_\alpha$ . First, note that  $z \subset x$  holds if and only if  $z \subset x$  holds in  $V_\alpha$ . To see this, suppose  $z \subset x$ . Then trivially  $y \in z$  and  $y \in V_\alpha$  implies  $y \in x$ . Conversely, suppose  $z \subset x$  in  $V_\alpha$ , and  $y \in z$ . Then since  $z \in V_\alpha$  and  $V_\alpha$  is transitive,  $y \in V_\alpha$ . Hence  $y \in x$ .

Now, let  $z \in V_\alpha$ . If  $z \in w$ , then by definition of the power set,  $z \subset x$  and hence  $z \subset x$  in  $V_\alpha$ . Conversely, if  $z \subset x$  in  $V_\alpha$ , then  $z \subset x$  so  $z \in w$ . Hence  $w$  serves as a valid  $\mathcal{P}(x)$  in  $V_\alpha$ .  $\square$

If  $\alpha = \lambda + 1$ , then one may easily find counterexamples to the Power Set Axiom. To see this, consider  $\lambda \in V_\alpha$ . If we had Power Sets, then  $\alpha = \lambda + 1 \subset \mathcal{P}(\lambda) \in V_\alpha$  as well, contradicting Proposition 4.11. Thus, the statement of the previous theorem may be rephrased as “if and only if.”

**Axiom 3 (Pairing).** If  $x, y$  are sets, then there is a set, denoted  $\{x, y\}$ , consisting of just  $x$  and  $y$ .

**Theorem 5.13.** If  $\alpha$  is limit, the Pairing Axiom holds in  $V_\alpha$ .

*Proof.* Let  $x, y \in V_\alpha$ . Since  $V_\alpha = \bigcup_{\beta < \alpha} V_\beta$ , there is a  $\beta < \alpha$  such that  $x, y \in V_\beta$ . The usual set  $\{x, y\}$  is thus a member of  $V_{\beta+1}$ , which in turn is a subset of  $V_\alpha$ .  $\square$

As with Union, one may find counterexamples to the Pairing Axiom in the case that  $\alpha = \lambda + 1$ . We know that  $\lambda \in V_\alpha$ , but if we could form the set  $\{\lambda, \lambda\} = \{\lambda\} \in V_\alpha$ , then by Proposition 4.4,  $\{\lambda\} \subset V_\lambda$ . Hence  $\lambda \in V_\lambda$ , contradicting Proposition 4.11. Once again, the converse of the preceding theorem holds.

## 5.4 One final kick in the teeth

This axiom is perhaps the most powerful, in that it can be used together with just a couple other axioms to derive the rest. It also provides the most difficult step in my analysis.

**Axiom 9 (Replacement).** If  $f$  is a class function whose domain is a set, then its range is a set.

Not surprisingly, Replacement does not hold in most of our familiar  $V_\kappa$ 's. For example, if  $\kappa = \omega_1$ , then  $\mathcal{P}(\omega) \in V_{\omega+2} \subset V_\kappa$ , but  $\mathcal{P}(\omega)$  has larger cardinality than  $\kappa$  and so naturally surjects onto it. Hence Replacement would imply that  $\kappa \in V_\kappa$ , a contradiction. Replacement does hold in some special  $V_\kappa$ , though, as we'll see shortly.

**Lemma 5.14.**  $\kappa$  is inaccessible if and only if  $\kappa$  is regular and  $|V_\kappa| = \kappa$ .

*Proof.* If  $\kappa$  is inaccessible, it is automatically regular. For any ordinal  $\beta$ , we have  $\beta \subset V_\beta$ , so in particular we have  $\kappa \leq |V_\kappa|$ .

For the opposite inequality, we first show by induction that for all  $\beta < \kappa$ ,  $|V_\beta| < \kappa$ . If  $\beta = \lambda + 1$  and  $|V_\lambda| < \kappa$ , then since  $\kappa$  is inaccessible,  $|V_\beta| = |\mathcal{P}(V_\lambda)| = 2^{|V_\lambda|} < \kappa$ . If  $\beta$  is limit and for each  $\lambda < \beta$  we have  $|V_\lambda| < \kappa$ , then  $|V_\beta| = |\bigcup_{\lambda < \beta} V_\lambda| < \kappa$  follows from the fact that  $\kappa$  is regular. Finally, this gives

$$|V_\kappa| \leq \left| \bigcup_{\beta < \kappa} V_\beta \right| \leq \kappa \cdot \kappa = \kappa$$

Conversely, suppose that  $\kappa$  is regular (hence limit) and  $|V_\kappa| = \kappa$ . We must show  $\kappa$  is a strong limit. Let  $\beta < \kappa$  be a cardinal. Then since  $\kappa$  is limit,  $\beta + 2 < \kappa$  as well. Recall also that  $\beta \subset V_\beta$ , so that  $\beta \leq |V_\beta|$ . Therefore,

$$2^\beta \leq 2^{|V_\beta|} = |V_{\beta+1}| < |\mathcal{P}(V_{\beta+1})| = |V_{\beta+2}| \leq |V_\kappa| = \kappa$$

so that  $\kappa$  is a strong limit.  $\square$

**Theorem 5.15.** If  $\kappa$  is inaccessible, the Axiom Schema of Replacement holds in  $V_\kappa$ .

*Proof.* Suppose  $\kappa$  is inaccessible, and  $x \in V_\kappa$ . If  $f$  is a class function on  $x$  in  $V_\kappa$  (i.e. a map  $x \rightarrow V_\kappa$ ), we must show that  $\text{Rng } f \in V_\kappa$ . To start, by Proposition 4.4 there is a  $\lambda < \kappa$  so that  $x \subset V_\lambda$ . Then  $|x| \leq |V_\lambda| < |V_\kappa| = \kappa$ , by the previous lemma. Hence, let  $\iota = |x| < \kappa$ .

Next, choose a bijection  $\iota \rightarrow x$ , and write  $x = \{x_i : i < \iota\}$ . For each  $i$  let  $r_i = \text{rank } f(x_i)$ . Note that since  $f(x_i) \in V_\kappa$ , we must have  $\text{rank } f(x_i) < \kappa$ . Thus  $r_i$  is an  $\iota$ -sequence in  $\kappa$ . Since  $\kappa$  is regular and  $\iota < \kappa$ , we must have  $\beta = \sup r_i < \kappa$ . Hence, for each  $i < \iota$ ,  $\text{rank } f(x_i) \leq \beta$  so that  $f(x_i) \in V_{\beta+1}$ . This means that  $\text{Rng } f \subset V_{\beta+1}$  and hence  $\text{Rng } f \in V_\kappa$ .  $\square$

**Lemma 5.16.** If  $\beta \in V_\alpha$ , then  $\beta$  is a cardinal if and only if it is a cardinal in  $V_\alpha$ .

*Proof.* Recall that  $\beta$  is an ordinal if and only if it is transitive and strictly well ordered by  $\epsilon$ . Since  $V_\alpha$  is transitive, it is clear that  $\beta$  is an ordinal if and only if it is an ordinal in  $V_\alpha$ .

Now suppose that  $\beta$  is an ordinal. If there is a  $\gamma < \beta$  and a bijective function  $f : \beta \leftrightarrow \gamma$ , then since the domain and range of  $f$  are elements of  $V_\alpha$ , we have  $f \in V_\alpha$  by Proposition 4.12. Conversely, if  $\beta$  and  $\gamma$  are in bijection in  $V_\alpha$ , then clearly  $\beta$  and  $\gamma$  are in bijection (period).  $\square$

**Theorem 5.17.** If the Axiom Schema of Replacement holds in  $V_\kappa$ , and  $\kappa$  is infinite, then  $\kappa$  is a strong limit.

*Proof.* We first show that  $\kappa$  is limit. If  $\kappa = \lambda + 1$ , then  $\lambda \in V_\kappa$  and we may define a function  $f : \lambda \rightarrow V_\kappa$  as follows. Let  $f(0) = \lambda$  and  $f(\alpha) = \alpha$  otherwise. Then  $\text{Rng } f \cup \{0\} = \lambda + 1 = \kappa$  is not a set in  $V_\kappa$ , so by the Union Axiom, neither is  $\text{Rng } f$ , a contradiction.

Now, let  $\beta < \kappa$ , so that  $\beta \in V_\kappa$ . We have just seen  $\kappa$  is limit, hence  $V_\kappa$  satisfies every axiom of ZFC. In particular, Lemma 5.11 shows that  $\mathcal{P}(\beta) \in V_\kappa$ . Next, Proposition 3.15 gives that  $\mathcal{P}(\beta)$  stands in bijection with a cardinal, which by uniqueness and the previous lemma, must be  $2^\beta$ . Hence, by Replacement,  $2^\beta \in V_\kappa$ .  $\square$

It is also the case that  $V_\kappa$  satisfies replacement when  $\kappa$  is either 0 or 1—in both cases there are no nonempty elements, and hence no possible nontrivial domains for functions. For any other finite  $\kappa$ , there are nontrivial functions and the argument of the preceding theorem goes through.

Note that it is impossible to show that if the Axiom Schema of Replacement holds in  $V_\kappa$  that  $\kappa$  is regular. It is, however, true with a strong hypothesis on the model  $V_\kappa$ .

**Definition 5.18.** Suppose that  $A$  is a set-model of ZFC. Then  $A$  satisfies Strong Replacement if the following holds.

- Let  $F \subset A$  such that  $\text{Fun}(F)$ . Then for all  $x \in A$ , there is a  $y \in A$  which is the image of  $x$  under  $F$ .

**Theorem 5.19.** Suppose some  $V_\kappa$  satisfies Strong Replacement (and hence Replacement). Then  $\kappa$  is regular.

*Proof.* Suppose that  $\kappa$  is not regular, in other words, there is an ordinal  $\alpha < \kappa$  and a  $\beta$ -sequence  $A = \{a_\alpha : \alpha < \beta\} \subset \kappa$  such that  $\sup A = \kappa$ . We may define a bijective map  $f : \beta \rightarrow A$  by  $\alpha \mapsto a_\alpha$ . By Proposition 4.13,  $f \subset V_\kappa$ . Hence by Strong Replacement,  $A$  is a set, which is to say  $A \in V_\kappa$ . But since  $\sup A = \kappa$ , this contradicts Corollary 5.5.  $\square$

## 5.5 Models for subsets of ZFC

From the above discussion, it is clear that  $V_\alpha$  is a model for all of ZFC if  $\alpha$  is an inaccessible cardinal which is larger than  $\omega$ . Since all of the arguments of the previous section are immune to the pass to the Axiom of Second Order Replacement, this statement can in fact be strengthened to “if and only if” when this axiom of assumed.

However, smaller models are available for certain subsets of ZFC. For example, we have that if Replacement is removed from our axiom list, then  $\alpha$  is required only to be limit and greater than  $\omega$ . The smallest such choice is  $\alpha = \omega + \omega = \bigcup_{n < \omega} (\omega + n)$ .

Alternatively, we might remove the Axiom of Infinity, and ask for a model containing only finite sets. In this case, we remove the restriction that  $\alpha$  is uncountable, so that  $\alpha$  need only be an inaccessible cardinal. The smallest such cardinal is  $\alpha = \omega$ .

## References

- [1] Cohn, Henry. Math 583CA Lecture at the University of Washington. May 30, 2003.
- [2] Devlin, Keith. The Joy of Sets. Springer-Verlag. 1979.
- [3] Jech, Thomas J. Set Theory. Academic Press. 1978.
- [4] Just, Winfred and Martin Weese. Discovering Modern Set Theory I: The Basics. American Mathematical Society. 1991.
- [5] Kanamori, Akihiro. The Higher Infinite. Springer-Verlag. 1991.
- [6] Lyndon, Roger C. Notes on Logic. Van Nostrand. 1966.
- [7] Müller, Gert H., ed. Sets and Classes. “The Role of Classes in Set Theory” by Azriel Lévy. 1976. Pages 173-215.