ALTERNATIVE SET THEORIES

M. Randall Holmes, Thomas Forster and Thierry Libert

CONTENTS

1 INTRODUCTION

The one thing that all alternative set theories have in common is the fact that they are alternatives to ZF or ZFC. They are not what the biologists call *monophyletic* nor do they form what the philosophers call a *natural kind*. Nor for that matter is there anything alternative about them in the vulgar sense: people who study alternative set theory are no more New Age or eco-friendly than any other kind of set theorist. What unites alternative set theories is what they are not. To this end the treatments of the alternative set theories will potter off into their little separate corners at the earliest opportunity.[1]

The diversity of set-theoretic systems that have been considered since the discovery of the paradoxes is rather unrecognized or disregarded. For one system has always overshadowed the others: Zermelo's and its variants, notably ZF, which is now the presumed referent of the phrase "axiomatic set theory". And this is perfectly understandable since—as well as providing a natural foundational framework for ordinary mathematics—no other system has proved to be more appropriate for investigating questions about infinity, which—after all—was the original motivation for Cantor's work.

Yet sets are *abstract* objects, and therefore subject to mathematical abstraction and generalization. Thus, as happened with the concept of number, it was to be expected that mathematical constructions, as algebraic, topological, or more generally categorical ones, would be used to extend the concept of set in one way or another. One often cites standard algebraic constructions in the theory of fields as inspiring the usual model-theoretic techniques for proving relative consistency and independence results in ZF. But a more significant example can be found in [Barwise and Moss, 1996], where the move from the integers to the rationals is used to motivate the enrichment of the ordinary well-founded universe of sets by so-called "hypersets", which provides solutions to certain reflexive set-theoretic equations. At least, the independence of the Axiom of Foundation and subsequently the consistency of anti-foundation axioms, discussed in [Aczel, 1988], have shown that the *iterative* conception of sets cannot be regarded as the sole motivation for ZF-like axiomatizations of set theory, and *a fortiori* as the sole method of avoidance of the paradoxes.[2]

Arguably, what characterizes ZF-like axiomatizations then is the principle of *specialized comprehension*, cited in [Church, 1974], according to which 'we seek axioms which are special cases of the general comprehension axiom and which promise to maintain consistency while at the same time being adequate for a large variety of mathematical purposes'. In the systems we are going to review the selection of admissible instances of comprehension is not guided by mathematical purposes, whatever they are, but is essentially of a *logico-syntactic* nature in order

[1] Part of this chapter was written in January 2008 while Thierry Libert was visiting Thomas Forster at the DPMMS in Cambridge, with the support of the Belgian 'Fonds National de la Recherche Scientifique'.

[2] But see discussion below of the realization of something like the cumulative hierarchy (and so of the iterative conception) in theories with anti-foundation axioms

to block the paradoxes.

One famous example of such systems is Quine's "New Foundations" [Quine, 1937]. This belongs to the respectable category of *type-theoretic* approaches, which make use of syntactical criteria to prohibit circularity in definitions – as typically in the instance of comprehension leading to Russell's paradox, i.e., $\{x \mid \neg(x \in x)\}$. A section on type theory is provided to support the discussion of New Foundations.

Another logico-syntactic device which can be used in a type-free setting to eradicate Russell's paradox is to proscribe negation in formulas defining sets, or at least to tamper with its use. Systems following that prescription will be referred to as *positive set theory*. A section below gives the reader an overview of such systems, in connection with the general topic of *topological set theory*. Although the subject would also encompass systems based on certain non-classical logics, these are not going to be discussed in this chapter. A historical account of positive set theory from that perspective can be found in [Libert, 2004].

Nor are we going to discuss set theories motivated by the application of intuitionistic logic. This is not because we do not consider these theories interesting, but because they introduce a new level of complexity which would make this chapter much longer.

We will briefly cover simple type theory [Wang, 1970] and Zermelo set theory [Zermelo, 1914], which are not strictly speaking alternative set theories (type theory is not precisely set theory and Zermelo set theory is not precisely "alternative", as it is the original version of ZFC: but we do present a version of Zermelo set theory which qualifies as an alternative set theory). We will discuss theories with proper classes: VGB set theory [Gödel, 1940] and Kelley-Morse set theory [Kelley, 1976] are very close to ZFC; Ackermann set theory [Levy, 1959] appears quite distant in its motivation from ZFC but turns out to be almost the same theory. We then discuss a further class of theories which are quite close to ZFC, in which the Axiom of Extensionality is weakened to allow atoms or in which the Axiom of Foundation is negated.

The next section takes us farther from the familiar. New Foundations [Quine, 1937] and the known-to-be consistent theories allied with it are motivated by simple type theory rather than Zermelo set theory and superficially look very different from the usual set theory. Some of the consistent fragments remain alien to ZFC on closer inspection, but the most mathematically fluent system of this kind, NFU (Jensen's modification of NF to allow urelements in [Jensen, 1969]) turns out to be quite closely related to the usual set theory in a way we will describe.

We then discuss positive set theories. The most mathematically fluent system of positive set theory is the system GPK^+ defined by Olivier Esser [Esser, 1999], which is based on earlier work on positive comprehension and "hyperuniverses" in which Marco Forti is prominent (see [Forti and Honsell, 1996b]). This system is elegant, superficially different from standard set theory but underneath deeply related to a powerful extension of ZFC, and surprisingly strong: its model theory takes us into the domain of large cardinals of moderate strength (this is also true of the higher reaches of the model theory of natural extensions of NFU, but in the case of Esser's theory one is involved with large cardinals from the outset). Esser's set theory and earlier related systems of Forti and others have topological motivation, and these theories might be called topological set theories.

We discuss a pair of set theories motivated by the surprising application of model theory to analysis subsumed under the term "nonstandard analysis". One of these, Nelson's IST [Nelson, 1977], is a deliberate adaptation of ZFC to support nonstandard analysis. The other, Vopěnka's "alternative set theory" or AST [Vopěnka, 1979], is more *sui generis*, and differs more fundamentally in its outlook from ordinary set theory: it is mathematically interesting though as a set theory it is rather weak.

Finally, we discuss two curiosities. The first one honestly exhibits a characteristic which has been ascribed by philosophers to all systems of set theory (incorrectly, we believe). The double extension set theory of Andrzej Kisielewicz [Kisielewicz, 1998] is a fascinating and bizarre system of set theory which is frankly an *ad hoc* solution to the paradoxes. Appropriately, we do not yet know whether this system is consistent. It is at least as strong as ZFC, for reasons which seem to resemble those which make the Ackermann set theory as strong as ZFC.

The other curiosity is the theory obtained by adjoining the existence of an elementary embedding $j : V \to V$ to Zermelo set theory with the Rank Axiom, which has been studied by Paul Corazza [Corazza, 2000]. The existence of such an embedding is inconsistent with ZFC (though it is not known to be inconsistent with ZF), and this extension of Zermelo set theory is not a weak theory but is in fact one of the strongest set theories ever proposed.

Zermelo set theory (first proposed in 1908 in [Zermelo, 1914]) was not in our view motivated by an *ad hoc* attempt to resolve the paradoxes. Zermelo appears to have catalogued constructions in set theory which had actual applications in mathematical reasoning and which were required to implement his proof that the Axiom of Choice implies the well-ordering theorem. The motivation of simple type theory seems to us to be similar (though more motivated by logical considerations than by mathematical practice): its history is enormously complicated (see [Wang, 1970]). New Foundations([Quine, 1937], hereinafter NF) does seem to have started out as an *ad hoc* attempt to further simplify "simple type theory" to eliminate the annoyance of type indices: but the simplification which motivates New Foundations is mathematically very appealing and the fact that almost any reasonable weakening of NF seems to give a demonstrably consistent theory suggests that Quine was onto something. Positive set theory appears to have an elegant topological motivation. A final point is that most of these set theories, different as they appear to be from one another, seem to converge on the same picture of the world of set theory, though there are some interesting byways.

2 THE BASICS: TYPE THEORY AND THE ORIGINAL THEORY OF ZERMELO

In this section we describe two systems which are not "alternative": they are standard theories. These are simple type theory (in a streamlined form—not the complex theory of *Principia Mathematica* [Whitehead and Russell, 1910-13] but the much simpler system whose evolution is described in [Wang, 1970]) and the original set theory of Zermelo [Zermelo, 1914], which differs in some ways from its final modern development.

2.1 Simple type theory

The simple type theory TST is a first-order many-sorted theory with equality and membership as primitive relations. The sorts, called "types" are indexed by the natural numbers $0, 1, 2 \ldots$. Atomic formulas are well-formed if they take one of the forms $x^n = y^n$ or $x^n \in y^{n+1}$, where the superscript indicates the type of the variable.

The intuition is that type 0 is inhabited by some objects of an unspecified nature called "individuals", type 1 is inhabited by sets of individuals, type 2 is inhabited by sets of sets of individuals, and so forth; in general type $n+1$ is inhabited by sets of type n objects.

There are two axiom schemes which make up the core of the theory and two further axiom schemes which are usually assumed. We advocate the use of variables without explicit type superscripts in this system, in situations where the types assigned can be deduced from the context or in which any types for the variables compatible with the syntax will work, but we will sometimes give explicit type indices for clarity.

Axiom of Extensionality: $(\forall AB.A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B))$. Objects of any positive type are equal iff they have the same elements.

Axiom of Comprehension: For any formula ϕ in which A does not appear free, $(\exists A.(\forall x.x \in A \leftrightarrow \phi))$. Any condition on objects of a given type defines a set of the next higher type.

Definition: For any formula ϕ in which A^{n+1} does not appear free, $\{x^n \mid \phi\}^{n+1}$ is defined as the unique A^{n+1} such that $(\forall x^n.x^n \in A^{n+1} \leftrightarrow \phi)$ (which exists by Comprehension and is unique by Extensionality).

Note that in the statement of the axiom schemes we do not assign types explicitly to the variables: any assignment of types which works will give an instance of the scheme.

We define V^{n+1} as $\{x^n \mid x^n = x^n\}^{n+1}$ and $\emptyset^{n+1}$ as $\{x^n \mid x^n \neq x^n\}^{n+1}$. We define $\{x^n, y^n\}^{n+1}$ as $\{z^n \mid z^n = x^n \vee z^n = y^n\}^{n+1}$. Type indices may be omitted.

Ordered pairs can be defined using the usual Kuratowski definition $(x, y) = \{\{x\}, \{x, y\}\}$, and relations and functions can then be defined as usual. It is unfortunate that in the notation $x\,R\,y$ the type of R is 3 higher than that of x or y, and similarly the type of f is three higher than that of x in the notation $f(x)$; there is a way to change this displacement to the more natural 1 (by developing an ordered pair which is of the same type as its projections) but that development would take us a bit far afield.

We define 0 as $\{\emptyset\}$: note that we can assign any type ≥ 2 to this object, and in fact we define 0 in each type at or above 2. For each set A, define $A+1$ as $\{a \cup \{x\} \mid a \in A \wedge x \notin a\}$. $A+1$ is the set of all objects obtained by adding a single new element to an element of A. Now $0+1$ is the set of all one-element sets, which we call 1, and $1+1$ is the set of all two-element sets, which we call 2, and so forth. We define $\mathbb{N}$, the set of natural numbers, as the intersection of all sets which contain 0 and are closed under the extended "successor" operation. Notice that a version of each natural number is defined in each type with index ≥ 2, and a set $\mathbb{N}$ is defined in each type with index ≥ 3. The elements of the natural numbers are the finite sets: we define the set $\mathbb{F}$ of finite sets as $\bigcup \mathbb{N}$.

We can now assert the remaining axioms usually assumed in TST.

Axiom of Infinity: $V^{n+1} \notin \mathbb{F}^{n+2}$. The universe (of objects of a given type) is not a finite set.

Axiom of Choice: Any pairwise disjoint collection of nonempty sets (of a given type) has a choice set (of the next lower type).

This theory figures in our discussion here primarily as the basis for the development of New Foundations and related theories, but there are other interesting relationships between TST and other alternative set theories.

This theory has been described carelessly as the simple type theory of Russell, but this is historically inaccurate. Russell had something like this in mind in the Appendix to *Principles of Mathematics* [Russell, 1903] though even there this is not clear. In the formalization of *Principia Mathematica* (hereinafter PM, [Whitehead and Russell, 1910-13]) Whitehead and Russell were hampered by the lack of a set-theoretical definition of ordered pair. They knew that relations were sets of ordered pairs, but they had no set-based notion of ordered pair, so in fact in PM they defined ordered pairs as a sort of relation! The type system of PM contains types of n-ary relations over each finite sequence of previously defined types, which makes it quite complex: it is further complicated by predicativity restrictions which motivate the further subdivision of types into "orders". Ramsey pointed out (in [Ramsey, 1926]) that the Axiom of Reducibility renders the orders redundant, and so they should simply be dropped. Norbert Wiener, in [Wiener, 1914], 1914, gave the first set-theoretical definition of ordered pair, defining (x, y) as $\{\{\{x\}, \emptyset\}, \{\{y\}\}\}$, a definition which has certain formal merits. He pointed out that this made it possible to collapse the complex type system of relations of PM to a simple linear hierarchy of set types. Hao Wang has an interesting discussion of the history of simple type theory [Wang, 1970] which suggests that this kind of theory was actually first described in detail around 1930.

Whitehead and Russell were aware of the ambiguity of the notation (what computer scientists call "polymorphism") which tempts us to omit explicit type indices: they called it "systematic ambiguity" in PM and took advantage of it in the same way to avoid giving explicit indications of type in his much more complex type system. (In fact, oddly, PM does not have a notation for types).

2.2 The original system of Zermelo

The axioms of Zermelo's original theory of 1908 [Zermelo, 1914] are presented. The content is presented (this is not a literal translation). Familiarity with common set-theoretical notation is assumed.

Axiom of Extensionality: If every element of a set A is also an element of B and *vice versa*, then $A = B$.

Axiom of Elementary Sets: $\emptyset$ exists; for any object x, $\{x\}$ exists; for any objects x and y, $\{x, y\}$ exists.

Axiom of Separation: If $P(x)$ is a proposition with a definite truth-value for each $x \in M$, $\{x \in M \mid P(x)\}$ exists.

Power Set Axiom: For every set A, the power set $\mathcal{P}(A)$, the set of all subsets of A, exists.

Union Axiom: For any set A, the union $\bigcup A$, the set of all elements of elements of A, exists.

Axiom of Choice: Any pairwise disjoint collection of nonempty sets has a choice set.

Axiom of Infinity: There is a set Z such that $\emptyset \in Z$ and for each $x \in Z$, $\{x\} \in Z$.

The Axiom of Extensionality is phrased in a way which allows non-set elements of the domain of discourse (objects with no elements distinct from one another and from the empty set, which are often called *atoms* or *urelements*). In modern theories, it is usually assumed that every object is a set.

The Axiom of Elementary Sets is usually replaced with the Axiom of Pairing, which asserts the existence of $\{x, y\}$ for any objects x and y. The existence of $\emptyset$ follows from Separation and the existence of any specific set (such as the one given in Infinity). It is a more modern observation that the existence of $\{x\} = \{x, x\}$ is a special case of pairing.

Zermelo did not have the modern formulation of the conditions in the Axiom of Separation in terms of formulas of the first-order language of set theory, and apparently when he saw this formulation he did not like it: he believed that it was too restrictive. The true formulation of Zermelo's intentions may be second-order.

The Axiom of Choice takes a modern form. It is interesting that Zermelo proves his theorem that the well-ordering principle follows from choice in the 1908 paper [Zermelo, 1914], without having the ability to express ordered pairs as sets (which was present in this theory of course, but first recognized by Norbert Wiener in [Wiener, 1914] in 1914).

The Axiom of Infinity differs from its usual modern form which asserts the existence of a set which contains $\emptyset$ and is closed under the von Neumann successor operation $x \mapsto x \cup \{x\}$. A different definition of the natural numbers is the reason: Zermelo defines the natural numbers as the iterated singletons of the empty set. It is an interesting fact that though the mathematical effect of either axiom of infinity is the same, these axioms are not equivalent in the presence of the other axioms of Zermelo set theory: neither one implies the other. They have the same consistency strength: each theory can interpret the other. The two axioms of infinity are equivalent in the presence of the very powerful Axiom of Replacement; they are also equivalent in the presence of a weaker axiom which asserts that each object belongs to some rank of the cumulative hierarchy.

The Axiom of Foundation is absent from Zermelo's original theory though it often appears in modern lists of axioms for Zermelo set theory.

Axiom of Foundation: For any set x, there is $y \in x$ such that $y \cap x = \emptyset$.

This axiom asserts that the membership relation restricted to any set is a well-founded relation. It rules out such things as self-membered sets. This is a consequence of the view that the world of set theory is constructed by the iterated application of the power set operation to an initial collection (the empty set or perhaps a set of atoms) through stages indexed by the ordinals. We will see below that variations of ZFC have been studied whose distinctive feature is that Foundation is omitted and replaced with principles which flatly contradict it.

This subsection introduces the original system of Zermelo, but does not aim to correct common usage: "Zermelo set theory" in the sequel will refer to the theory with the von Neumann form of the Axiom of Infinity and with the Axiom of Foundation.

For reference we include a statement of the final axiom scheme which turns Zermelo set theory (in either form) to the usual set theory ZFC:

Axiom of Replacement: If a is a set and ϕ is a formula such that we can prove $(\forall x.(\exists! y.\phi))$, then $\{y \mid (\exists x \in a.\phi)\}$ is a set.

2.3 The relationship between these systems. Mac Lane set theory

The simple theory of types and Zermelo set theory appear superficially to live at the same level of strength. Both of them can talk about an infinite set (type 0 for TST, the set provided by Infinity for Zermelo) and its iterated power sets. Each theory proves the existence of the cardinals $\aleph_n$ for each n but cannot prove the existence of $\aleph_\omega$. But Zermelo set theory is stronger: Kemeny proved in his PhD thesis [Kemeny, 1950] (which was never published) that Zermelo set theory proves the consistency of TST.

The source of strength in Zermelo set theory is the ability to quantify over the entire universe in instances of Separation. The variant of Zermelo set theory in which Separation is restricted to Δ_0 formulas (those in which every quantifier is bounded in a set) [with the von Neumann form of Infinity and the Axiom of Foundation] is known as Mac Lane set theory (because it was advocated as a foundational system by Saunders Mac Lane in [Mac Lane, 1986]), and Mac Lane set theory is demonstrably mutually interpretable with TST.

Adrian Mathias has written a beautiful survey of relations between Mac Lane set theory, Zermelo set theory, and some other systems, found in [Mathias, 2001a], [Mathias, 2001b].

2.4 Mac Lane or Zermelo set theory as an alternative set theory

As witnessed by the program of Mac Lane mentioned above, Zermelo set theory or variants of Zermelo set theory have been pressed into service themselves as alternative set theories, presumably by workers nervous about the high consistency strength of ZFC.

One is then faced with the problem that the implementations of cardinal and ordinal numbers traditional in ZFC do not work in Zermelo set theory. Cardinals are implemented as initial ordinals, so we need only consider the ordinals. And here we have a serious problem. Zermelo set theory with the modern form of Infinity proves the existence of the von Neumann ordinals $0, 1, 2, 3, \ldots \omega, \omega{+}1, \omega{+}2, \omega{+}3 \ldots$ but not of $\omega \cdot 2$. But these theories prove the existence of much longer well-orderings.

An elegant solution is obtained by adding the Axiom of Foundation (or restricting our attention to well-founded sets), and further stipulating that every set belongs to a rank of the cumulative hierarchy. It is then possible to define the cardinality of a set A as the set of all sets B which are equinumerous with A and of the lowest rank which contains such sets, and define the order type of a well-ordering $\leq$ as the set of all well-orderings similar to $\leq$ and of the lowest rank which contains such well-orderings (notice that we no longer identify cardinals with initial ordinals if we use these implementations). This is called “Scott’s trick”, introduced in [Scott, 1955].

It remains to explain how to stipulate that every set belongs to a rank of the cumulative hierarchy.

Definition: A set H is a *subhierarchy* iff H is well-ordered by inclusion, $\bigcup I \in H$ for each $I \subseteq H$, and for any $x \in H$ other than $\bigcup H$, the minimal element of H properly including x is $\mathcal{P}(x)$. A set r is a *rank* iff there is a subhierarchy H such that $r \in H$.

Axiom of Rank: For every set x there is a rank r such that $x \in r$.

Theorem: For any subhierarchies H_1 and H_2, either $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.

Proof: Omitted.

Corollary: Ranks are well-ordered by inclusion (all ranks, not just the ones in any given subhierarchy).

Definition: For any set x, the rank of x is the minimal rank which contains x as a subset.

Definition: For any set A, $|A| \equiv \{B \in r_A \mid B \sim A\}$ where r_A is the minimal rank which contains a set equinumerous with A. A *cardinal number* is a set $|A|$.

Definition: For any well-ordering W let r_W be the minimal rank which contains a well-ordering isomorphic to W. Define $ot(W)$ as $\{W' \in r_W \mid W \approx W'\}$. An *ordinal number* is a set of the form $ot(W)$.

Definition: For any ordinal α, V_α is the rank (if there is one) such that the order type of the inclusion order on the ranks properly included in α is α.

It is further interesting to observe that the following axioms suffice to present this extension of Zermelo (or Mac Lane) set theory.

Axiom of Extensionality: Sets with the same elements are the same.

Axiom of (Bounded) Separation: For any set A and (Δ_0) formula ϕ in which B is not free, $(\exists B.x \in B \leftrightarrow x \in A \wedge \phi)$.

Axiom of Power Set: For any set A, there is a set $\mathcal{P}(A)$ whose members are exactly the subsets of A.

Axiom of Rank: Every set belongs to some rank.

Axiom of Infinity: The von Neumann ordinal ω exists. [it is equivalent to assume the existence of the set of Zermelo natural numbers in this context.]

Axiom of Choice: Any pairwise disjoint collection of nonempty sets has a choice set.

The axioms of Pairing and Union in the original Zermelo axiom set turn out to be redundant. If x and y belong to ranks r_x and r_y, $\{x, y\} = \{z \in r_x \cup r_y \mid z = x \vee z = y\}$, where $r_x \cup r_y$ is a set because it is the larger of the two ranks. If x belongs to the rank r_x, $\bigcup x = \{y \in r_x \mid (\exists z.y \in z \wedge z \in x)\}$. The Axiom of Foundation follows quite directly from the Axiom of Rank as well.

A book-length development of set theory in this style is given by Michael Potter in [Potter, 2004]. The Axiom of Rank adds no essential strength to Zermelo or MacLane set theory: see [Mathias, 2001a].

For reference and comparison, we give the familiar definition of ordinal usual in ZFC, originally due to von Neumann.

Definition: A set A is *transitive* iff for all x, y, $x \in y \wedge y \in A$ implies $x \in A$. This admits the equivalent formulations $\bigcup A \subseteq A$ and $A \subseteq \mathcal{P}(A)$.

Definition: A (von Neumann) ordinal is a transitive set α which is strictly well-ordered by the membership relation.

Definition: The (von Neumann) order type of a well-ordering $\leq$ is the unique (von Neumann) ordinal α such that $\leq$ is isomorphic to the inclusion order on α.

The Axiom of Replacement suffices to prove that every well-ordering has a von Neumann order type. This is not a theorem of Zermelo set theory (or any of the related theories of similar strength that we have described).

3 THEORIES WITH CLASSES

3.1 General considerations

In any version of set theory, we find ourselves wanting to talk about collections which are not sets as if they were sets. It is easy to see that in a certain sense such talk is harmless: if (M, E) is a model of set theory, we could add $\mathcal{P}(M)$ (the true power set of M) as a new domain, and add the membership relation of elements of M in elements of $\mathcal{P}(M)$ to obtain a theory with signature

$$(M, \mathcal{P}(M), E, \in \cap (M \times \mathcal{P}(M))),$$

then finally identify each preimage of an element of M under E with that element of M and collapse E and $\in \cap (M \times \mathcal{P}(M)$ into a single relation appropriately. Of course this requires some strength in the metatheory, but not very much.

In such an extended theory, the domain of the original theory is definable as the collection of elements. In the family of theories with classes that we give first, general objects are called *classes* and objects which are elements of classes are called *sets*. The primitive relations of the theory are equality and membership as usual.

We give the following

Definition: $set(x) \equiv (\exists Y.x \in Y)$: a *set* is a class which is an element of some class.

Convention: We use upper case letters for general variables and lower case letters for variables restricted to the sets.

Any theory of this kind has two characteristic axioms as preamble.

Axiom of Extensionality: $(\forall AB.A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B))$

Axiom of {Predicative} Class Comprehension: Where ϕ is a formula {in which all quantifiers are restricted to the domain of sets} in which A does not appear free, $(\exists A.(\forall x.x \in A \leftrightarrow \phi))$. Note the use of the variable case convention here: if we were to assume that variables ranged over all classes, this would be $(\exists A.(\forall x.x \in A \leftrightarrow set(x) \wedge \phi))$.

Definition: The object A such that $(\forall x.x \in A \leftrightarrow \phi)$ (where A is not free in ϕ) is denoted by $\{x \mid \phi\}$.

The differences between predicative and impredicative class comprehension will be brought out in the next subsection. The motivating discussion above suggests that any set theory can be extended to a theory of classes with extensionality and full class comprehension. Thus impredicative class comprehension as such adds no additional strength to a set theory, but it does add considerable expressive power.

3.2 Von Neumann-Gödel-Bernays and Kelley-Morse set theory

Two theories with sets and classes are given here. The theory of sets in each extends ZFC. The first theory, called VGB (von Neumann-Gödel-Bernays set theory – [Gödel, 1940] is a reference) is a conservative extension of ZFC: it proves nothing about sets that ZFC does not. This theory was originally proposed by von Neumann in 1925 [von Neumann, 1925] but as a theory of functions rather than sets. We give our own presentation of a theory with the same mathematical force.

Notice that if x and y are sets, the existence of $\{x, y\}$, $\bigcup x$, $\mathcal{P}(x)$ (understood as the collection of sub*sets* of x) follows from Class Comprehension: the additional axioms corresponding to the axioms of Pairing, Union, and Power Set assert that these classes are sets.

Axiom of Extensionality: as above

Axiom of Predicative Class Comprehension: as above

Axiom of Separation: Every subclass of a set is a set.

Axiom of Power Set: for any set x, $\mathcal{P}(x)$ is a set.

Axiom of Set Union: for any set x, $\bigcup x$ is a set.

Axiom of Infinity: There is a set which contains $\emptyset$ and contains $x \cup \{x\}$ if it contains x, for any set x.

Axiom of Limitation of Size: A class C is not a set iff there is a class bijection from the class V of all sets to C.

Axiom of Foundation: For any set x, there is an element y of x such that $x \cap y = \emptyset$.

Many of the axioms of ZFC (or even of Zermelo set theory) may seem to be missing here, but this is not the case.

The Axiom of Infinity implies that there are at least three distinct sets (say the von Neumann numerals $0, 1, 2$) so there cannot be a bijection between a class $\{x, y\}$ and the universe, so $\{x, y\}$ must be a set so Pairing holds. Since Pairing holds, ordered pairs of sets are sets, and so every logical relation on sets is realized by a class of ordered pairs.

The Axiom of Choice (in a strong form) follows from Limitation of Size (surprise!). The class of von Neumann ordinals is not a set, so there is a class bijection from the class of von Neumann ordinals to the class of all sets, which determines a well-ordering of the universe in the obvious way. From any element of a collection of pairwise disjoint sets, select the least element with respect to this global well-ordering: collect these least elements to build a choice set.

The Axiom of Replacement follows from Limitation of Size. Suppose a is a set and $(\forall x \in a.(\exists! y.\phi))$. The class $B = \{y \mid (\exists x \in a.\phi)\}$ exists as a class by Class Comprehension. Note that $B = F``a$ where $F = \{(x, y) \mid \phi\}$. Define $F^{-1}(b)$ as $\{a \in A \mid F(a) = b\}$: notice that F^{-1} is an injection from B into $\mathcal{P}(a)$. $\{F^{-1}(b) \mid b \in B\}$ is a set because it is a subclass of $\mathcal{P}(a)$: if there were a class bijection from V to B, there would be a class bijection from V to this set, which contradicts Limitation of Size. Thus B is a set and Replacement holds.

If one wishes to avoid proving Choice, one can replace Limitation of Size with the following weaker axiom:

Weak Limitation of Size: For any set a and class B, if there is a bijection from a to B then B is a set.

It is straightforward to check that the derivation of Replacement above still succeeds. In the presence of the other axioms and weak Limitation of Size, we can deduce full Limitation of Size from a strong version of Choice (such as "there is a class well-ordering of the universe of sets") but not from weaker forms such as "every pairwise disjoint set of sets has a choice set".

The role of Limitation of Size in encapsulating the powerful axioms of Replacement and Choice is one of the striking features of VGB class theory. The other striking feature, found in all the original presentations, and obscured in our presentation so far, is that VGB class theory is finitely axiomatizable, because the axiom scheme of Predicative Class Comprehension can be replaced by finitely many of its instances.

We present a reduction of Predicative Class Comprehension to finitely many axioms. This proceeds by induction on the structure of formulas.

To handle classes in which the top level logical operator in the defining formula is propositional, note that

$$\{x \mid \phi \wedge \psi\} = \{x \mid \phi\} \cap \{x \mid \psi\}$$

and

$$\{x \mid \neg\phi\} = V - \{x \mid \phi\},$$

and every propositional operator can be defined in terms of conjunction and negation. So we stipulate that if A and B are classes, $A \cap B$ and $V \setminus A$ are classes.

We stipulate that for any sets x and y, $\{x, y\}$ is a set and so $(x, y) = \{\{x\}, \{x, y\}\}$ is a set. We define $(a_1, a_2, \ldots, a_n)$ as $(a_1, (a_2, \ldots, a_n))$. We stipulate that $(x_1) = x_1$: a 1-tuple is its sole item.

We define π_1 as $\{((x, y), x) \mid x = x\}$ and π_2 as $\{((x, y), y) \mid x = x\}$. We define $A|B$ for A and B classes as $\{(x, y) \mid (\exists z.(x, z) \in A \wedge (z, y) \in B)\}$. Note that if f and g are functions in the usual sense, $f \circ g = g|f$. We stipulate that π_1 is a class, π_2 is a class, and if A and B are classes, so is $A|B$. Note that each of the sets $\pi_i^n = \{((a_1, \ldots, a_i, \ldots, a_n), a_i) \mid a_1 = a_1\}$, the function sending an n-tuple to its ith element, is a class because constructible from the projection functions and relative product, except for $\pi_1^1 = \{(x, x) \mid x = x\}$, which we stipulate is a class.

For any class R we stipulate that $\mathrm{dom}(R) \equiv \{x \mid (\exists y.(x, y) \in R)\}$ is a class and $R^{-1} \equiv \{(y, x) \mid (x, y) \in R\}$ is a class.

Define $R \otimes S$ as $(R|\pi_1^{-1}) \cap (S|\pi_2^{-1})$. Define $R_1 \otimes R_2 \otimes \ldots \otimes R_n$ as $R_1 \otimes (R_2 \otimes \ldots \otimes R_n)$ (same grouping as for n-tuples).

Our strategy is first to fix a finite sequence of variables $x_1, \ldots, x_n$ and show how to define $\{(x_1, \ldots, x_n) \mid \phi\}$ for any formula in which no variable other than the x_i's appears free.

We provide $\{(x_1, x_2) \mid x_1 = x_2\}$ and $\{(x_1, x_2) \mid x_1 \in x_2\}$ as basic classes (actually, we have already provided the former).

If we have already shown $X_\psi = \{(b_1, \ldots b_p) \mid \psi\}$ to be a class, where all free variables in ψ appear among the b_i's, then $\{(a_1, \ldots, a_n) \mid \psi(a_{s_1}, \ldots, a_{s_p})\}$ can be shown to exist using the constructions we have already defined and one additional construction. It is constructible as $\mathrm{dom}((\pi_n^{s_1} \otimes \ldots \otimes \pi_n^{s_p}) \cap V \times X_\psi)$, so we stipulate that for any class X, $V \times X$ is a class (and for any classes X and Y, $X \times Y = (V \times X)^{-1} \cap (V \times Y)$ is thus a class).

If we have already shown $X_\phi = \{(a_1, \ldots, a_n) \mid \phi\}$ to be a class, we construct $\{(a_1, \ldots, a_n) \mid (\exists a_i.\phi)\}$ as $\mathrm{dom}(\pi_n^1 \otimes \ldots \otimes \pi_n^{i-1} \otimes (V \times V) \otimes \pi_n^{i+1} \otimes \ldots \otimes \pi_n^n) \cap (V \times X_\phi)$. Of course there are typographical variations of this if $i = 1$, $i = n$.

Now we show how to construct an arbitrary $\{x \mid \phi\}$. Let $x, x_2, x_3, \ldots, x_n$ be all the variables free or bound in ϕ. Construct $\{(x, x_2, \ldots, x_n) \mid \phi\}$. For each of the x_i's which is free in ϕ, intersect this class with $(\pi_n^i)^{-1}``\{x_i\}$ ($R^{-1}``A \equiv \mathrm{dom}(R^{-1} \cap V \times A)$; pairing implies the existence of singletons as sets and so as classes). Finally, the domain of this intersection is $\{x \mid \phi\}$.

We present the finite list of axioms which we have just shown to imply predicative class comprehension.

boolean intersection: For any classes A and B, $A \cap B = \{x \mid x \in A \wedge x \in B\}$ is a class.

complement: For any class A, $V \setminus A = \{x \mid x \notin A\}$ is a class.

pairing: For any sets x and y, $\{x, y\} = \{z \mid z = x \vee z = y\}$ is a set.

definition of ordered pair: (x, y) is defined as $\{\{x\}, \{x, y\}\}$.

projections: $\pi_1 = \{((x, y), x) \mid x = x\}$ is a class. $\pi_2 = \{((x, y), y) \mid x = x\}$ is a class.

relative product: For any classes R and S, $R|S = \{(x, y) \mid (\exists z.(x, z) \in R \wedge (z, y) \in S)\}$ is a class.

converse: For any class R, $R^{-1} = \{(y, x) \mid (x, y) \in R\}$ is a class.

domain: For any class R, $\mathrm{dom}(R) = \{x \mid (\exists y.(x, y) \in R)\}$ is a class.

primitive relations: $\{(x, x) \mid x = x\}$ and $\{(x, y) \mid x \in y\}$ are classes.

cartesian product: for any classes A and B, $A \times B = \{(a, b) \mid a \in A \wedge b \in B\}$ is a class (we only used $V \times A$ in the construction above, but one might as well provide the general construction).

Various finite sets of constructions have been given for this purpose. The primitives here are basic operations of the algebras of sets and relations, and the approach is inspired by the work of Givant and Tarski in [Tarski and Givant, 1987], in which they show that standard axioms for relation algebra augmented with basic properties of projection operations are sufficient to interpret first-order logic.

One can note here that our elimination of the Axiom of Pairing from the axioms of VGB is undone when we "unpack" the predicative comprehension scheme, but we are still correct that it is not needed in the presence of that scheme: what we have apparently revealed is that it is a component of that scheme. It is unique among those components in being a set existence principle rather than a class existence principle.

Kelley-Morse set theory (see the appendix to Kelley's [Kelley, 1976]) differs from VGB simply in having the full scheme of class comprehension instead of the predicative scheme. We can quantify over all classes in constructions of both classes and sets in Kelley-Morse set theory. We remarked above that adding classes which satisfy impredicative class comprehension to a set theory adds no real strength to the theory, but impredicative class comprehension can be used here to prove the existence of sets as well as classes, since any subclass of a set is a set and any class which can be placed in a class one-to-one correspondence with a set is a set.

Kelley-Morse set theory is stronger than VGB in a very marked way: Kelley-Morse set theory proves the consistency of VGB. Further, the class comprehension scheme of Kelley-Morse set theory cannot be finitely axiomatized.

The significant nature of the difference between VGB and Kelley-Morse set theory is not at first obvious. One way of seeing it is to observe that the classes of VGB can be identified with formulas of the language of ZFC; we will explain in what sense this is true and why it is not true of Kelley-Morse set theory.

If (M, E) is a model of ZFC, we show how to construct a model of VGB (with the weak version of Limitation of Size) which satisfies exactly the same theorems. A *preclass* is a formula with no variable other than x free in the language of ZFC augmented with constants for each element of M. We define $\phi \sim \psi$ as holding iff

$(\forall x.\phi \leftrightarrow \psi)$ holds in (M, E). We define a *class* as an equivalence class of preclasses under $\sim$. We define $[\phi]\, E^*\, [\psi]$ as holding iff $(\exists y.(\forall x.x \in y \leftrightarrow \phi) \wedge \psi[y/x])$. It is straightforward to establish that the structure (C, E^*), where C is the set of classes, satisfies VGB (with the weak form of Limitation of Size; if a global well-ordering of the universe is added, we can get full Limitation of Size) and moreover satisfies exactly the same sentences about sets as (M, E). This establishes that VGB (at least the version with weak Limitation of Size) proves nothing about sets that is not provable in ZFC.

We pause to observe that we can define ordered pairs of classes as classes. Let σ be the function which sends each natural number to its successor and fixes each other object. For classes A and B, define (A, B) as $\sigma``A \cup \{0\} \times \sigma``B \cup \{0\}$. From a cartesian product of nonempty classes one can of course extract the first and second projections. This means that we can represent formulas with arbitrary class parameters as classes in either VGB or Kelley-Morse set theory by a suitable coding scheme. What we can do in Kelley-Morse set theory that we cannot do in VGB is define satisfaction for formulas with no quantifiers over classes. In fact, we can define the structure defined in the previous paragraph over the universe of sets and code every element of it (set or preclass) as a set, so we obtain a class structure which is a model of VGB and we can prove the consistency of VGB. This shows that Kelley-Morse set theory is essentially stronger than VGB.

3.3 Ackermann set theory

The set theory of Ackermann (see [Levy, 1959]) has sets and classes but not in the same sense as VGB or Kelley-Morse. In the former theories, classes which are not sets are not elements; in Ackermann set theory some non-set classes are elements of other classes. The notion of set in Ackermann set theory is an independent primitive notion not definable in terms of the usual primitives of equality and membership: in fact, the whole thing works precisely because sethood is not definable.

Ackermann set theory is a first-order theory. General objects are called *classes*. Primitive relations are equality and membership and there is a primitive unary predicate of sethood.

Here are the five axioms of the theory.

Extensionality: Classes with the same elements are equal.

Class Comprehension: For any formula ϕ in which the variable A is not free, $(\exists A.(\forall x.x \in A \leftrightarrow set(x) \wedge \phi))$.

Elements: Any element of a set is a set.

Subclasses: Any subclass of a set is a set.

Set Comprehension: Let ϕ be a formula in which the variable A is not free, in which the predicate *set* does not occur, and in which all free variables other than x denote sets, and suppose $(\forall x.\phi \rightarrow set(x))$. Then $(\exists A.set(A)\wedge(\forall x.x \in A \leftrightarrow \phi))$.

We give the flavor of this theory by doing some basic proofs.

Universe: There is a class V which contains exactly the sets as elements (this is obvious from Class Comprehension).

Empty Set: $\emptyset$ is a set. The formula $x \neq x$ satisfies the conditions of Set Comprehension vacuously.

Pairing: Suppose a and b are sets. Then $\{a, b\}$ is a set.

Proof: The formula $x = a \vee x = b$, for any sets a and b, satisfies the conditions of Set Comprehension: it does not mention set, the parameters a and b are sets, and if it is true of x then x is a set.

Union: Suppose a is a set. Then $(\exists y.x \in y \wedge y \in a)$ satisfies the conditions of Set Comprehension. Note that if $x \in y$ and $y \in a$, then y must be a set by the Axiom of Elements (it belongs to the set a) and so x must also be a set (as it belongs to the set y).

Power Set: Suppose a is a set. Then $(\forall y.y \in x \rightarrow y \in a)$ satisfies the conditions of Set Comprehension. If x satisfies this condition it is a set by the Axiom of Subclasses.

Infinity: There is a set which contains $\emptyset$ as an element and is closed under the operation $x \mapsto x \cup \{x\}$.

Proof: "x belongs to every class C such that $\emptyset \in C$ and $(\forall X.X \in C \rightarrow X \cup \{X\} \in C)$" is a formula which satisfies the conditions of Set Comprehension. Note that $\emptyset$ is a set and for any set x, $x \cup \{x\}$ is a set (by applications of Pairing and Union proved above). This means that the class V of all sets is a class with the desired closure conditions. This verifies that if x satisfies this condition then $x \in V$ so x is a set! Note that the condition here quantifies over all *classes* (in fact we cannot quantify over all sets in an instance of Set Comprehension, because this would require us to mention the set predicate).

Separation: If a is a set, $\{x \in a \mid \phi\}$ is a set for any formula ϕ at all, by the Axiom of Subclasses.

Transitive Closure: The transitive closure $TC(A)$ of a set A, the intersection of all transitive sets which contain A as a subset, is a set.

Proof: Let A be a set. Consider the formula "x belongs to every class C such that $A \in C$ and for any $x \in C$, $\bigcup x \in C$". Note that V has the closure property in question, so this formula satisfied the requirements for Set Comprehension and defines a set. The union of this set is a transitive set containing A as a subset (and in fact is the minimal such set).

The proofs of Union and Power Set seem to be "set up" in advance by the axioms of Elements and Subclasses, but the fact that we can prove Infinity is a surprise.

At this point we have shown that Ackermann set theory has the strength of Zermelo set theory and so is adequate for classical mathematics.

We pause before considering how much stronger this theory might be to wonder what the motivation of the theory might be. Suppose that sets are actual (completed, definite) collections and classes are merely potential collections. The collection of all sets is a merely potential collection (however many actual sets we have constructed, we can construct some more). The class of all sets with a given property is a potential collection (as we built more and more sets we may add more objects to this class). An element of an actual collection is actual; a subclass of an actual collection is actual. The tricky part (as always) is getting one's mind around the interpretation of Set Comprehension. A collection defined mentioning only actual parameters (and in particular not mentioning the merely potential collection V of all sets) all of whose members turn out to be actual. . .is actual. This is not to our minds entirely convincing, but it may be suggestive.

We now develop an argument for the proposition that the well-founded sets of Ackermann set theory satisfy ZFC.

Definition: A *hierarchy* is a class which is well-ordered by inclusion, contains the unions of all its subclasses as elements, and in which the immediate successor of each element in the inclusion order (if it has one) is its power class. A *rank* is a class which belongs to a hierarchy. (Note that this definition is similar to the definition of *subhierarchy* in section 3.4, but note also that there is a distinction between set and class here which is not drawn in the earlier context).

Comments: A linear order is a well-ordering iff each subclass of its domain has a least element; since our class comprehension principles are apparently limited to classes of sets this might mean less than it appears to mean once we get past set ranks: we would not necessarily be able to carry out inductions on non-set ranks because not all definable subcollections of arbitrary classes are necessarily classes. The power class of a class is the class of all subclasses of that class, if such a class exists.

Theorem: For any two hierarchies H_1 and H_2 which are classes of sets, either $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.

Comment: We omit the proof of this result as we did with the similar result stated in section 3.4: we note that the restriction to hierarchies of sets is necessary because our comprehension principles are restricted.

Corollary: The set ranks are well-ordered by inclusion.

Definition: A *good rank* is a rank which is either included in or includes each other rank, and such that the ranks included in it are well-ordered by inclusion.

Theorem: The class of all ranked sets is a good rank.

Proof: Any hierarchy which contains all the set ranks will contain the class of all ranked sets as an element: then every element of such a hierarchy (and so every rank) will either be a set rank (and so be included in the class of all ranked sets) or include the class of all ranked sets. A hierarchy which does not contain all ranked sets will have its union a set rank, which is a subset of the class of all ranked sets.

Theorem: The class of all ranked sets is not the maximal good rank in the inclusion order.

Proof: If it were, it would be a set by set comprehension, as it would be definable without reference to the sethood predicate (as the maximal good rank in the inclusion order) and all of its elements are sets (since it is the class of all ranked sets). If the class of all ranked sets were a set, its power class would also be a set, so it would itself be a ranked set, and an element of itself. It is straightforward to establish by induction on inclusion that all set ranks are non-self-membered.

Corollary: There are non-sets which are elements: any rank which properly includes the class of all ranked sets has the class of all ranked sets itself as an element!

Theorem: Each of the axioms remains true if the term “ranked set” replaces the term “set”.

Proof: Extensionality does not mention sethood at all. Any definable subclass of the ranked sets is a class by Class Comprehension, so the modified version of Class Comprehension holds. Any element of a ranked set is a ranked set: one proves by induction on the inclusion order that all set ranks are transitive. So the modified version of the Axiom of Elements holds. Any subclass of a ranked set is a ranked set: if x is included in a set rank r, the power set of x is included in the power set of r. So the modified version of the Axiom of Subclasses holds. A class x which is defined without reference to sethood, with no parameters but ranked set parameters, and all of whose members are ranked sets is a set by Set Comprehension. x is included as a subset in a rank (the class of all ranked sets will serve). The first rank which includes x must be a set, because we just defined it using only a set parameter, and all of its elements are sets. Thus the modified version of Set Comprehension holds.

Comment: We have just shown that we can without affecting the strength of the theory simply add the assertion that every set belongs to a rank as an axiom.

Theorem: Let $\phi(x, y)$ be a formula containing only ranked set parameters which does not mention the sethood predicate and which is functional (for every ranked set x there is exactly one ranked set y such that $\phi(x, y)$). Then for any ranked set a, $\{y \mid (\exists x \in a.\phi(x, y))\}$ is a set.

Proof: This set exists by set comprehension, by inspection of its definition. There is a rank which includes all of its elements, namely the class of all ranked sets, and the minimal rank in the inclusion order which includes this set is also a set by set comprehension, so it follows that this set is a ranked set.

Comment: This assertion looks like the Axiom of Replacement, but some delicacy is required to show that the Axiom of Replacement holds for all formulas. The difficulty is that the formula ϕ in an instance of Replacement as directly translated from ZFC may contain quantifiers over all sets, which would be quantifiers over the class of sets in Ackermann set theory, to which set comprehension could not be applied. Nonetheless Replacement does hold. The idea is to show that $(\exists x \in V.\phi(a_1, \ldots, a_n))$ can be translated to a formula not mentioning nonset parameters or the sethood predicate if ϕ is equivalent to a formula not mentioning nonset parameters or the sethood predicate. Define $r(a_1, \ldots, a_n)$ as the least rank in the inclusion order which includes an x such that $\phi(a_1, \ldots, a_n)$, if there is one, and otherwise the empty rank. Note that this rank will be a set in any case. $(\exists x \in V.\phi(a_1, \ldots, a_n))$ is equivalent to $(\exists x \in r(a_1, \ldots, a_n).\phi(a_1, \ldots, a_n))$. By this technique we can systematically eliminate quantifiers over the domain of all sets from formulas that meet the appropriate syntactical restrictions. We should now have given at least a strong indication why the final theorem holds. Full details of the equivalence of Ackermann and ZF are found in [Reinhardt, 1970].

Theorem: The interpretation of the Axiom of Replacement (and, as we have already seen, of all the other axioms of ZFC) holds on the domain of ranked sets of Ackermann set theory.

The consistency of Ackermann set theory also follows from the consistency of ZFC, in a way which reveals something about the nature of the domain of "sets" in Ackermann set theory.

Augment the language of ZFC with an additional symbol M. Provide axioms asserting that M is transitive and contains all subsets of its elements. For each formula ϕ and variables x, y add an axiom to the effect that $(\forall x.(\exists y.\phi)) \leftrightarrow (\forall x \in M.(\exists y \in M.\phi))$ (for all values of any parameters). Note that a corollary of this is that $(\forall x.\phi) \leftrightarrow (\forall x \in M.\phi)$ will be an axiom for any ϕ (use a dummy variable y that does not appear in ϕ). It is straightforward to show that any finite collection of such axioms is consistent with ZFC. If $(\forall x.(\exists y.\phi))$ holds, we arrange for M to be closed under a suitable operation (applying replacement); if $(\forall x.(\exists y.\phi))$ does not hold, we arrange for M to include a counterexample (if ϕ contains parameters, this is still a closure operation, as we need counterexamples for all appropriate values of the parameters).

This extension of ZFC has a model if ZFC has a model. We claim that this is a model of Ackermann set theory, with the sethood predicate taken as meaning "is an element of M". The only axiom which requires any attention is Set Comprehension. Let ϕ be a predicate not mentioning M with all parameters other than x belonging to M. We may suppose without loss of generality that there is only one parameter $a \in M$. So we express the predicate as $\phi(x,a)$ and further suppose that for any $a \in M$, $\phi(x,a)$ is in M. Note that in our original model of ZFC it will be the case for every $a \in M$ that $\{x \mid \phi(x,a)\}$ exists as a set in ZFC (it will be in the power set of M), and so for every $a \in M$ there is a (dummy) $b \in M$ such that $\{x \mid \phi(x,a)\}$ exists, and so for every a whatsoever, $\{x \mid \phi(x,a)\}$ exists. From this it follows by a more natural application of the special axioms for M that for each $a \in M$, $\{x \mid \phi(x,a)\} \in M$, and this establishes that the interpretation of Ackermann set theory's Set Comprehension Axiom holds in the extended model of ZFC.

It is worth noting that though externally M is a model of ZFC, the extended model of ZFC does not necessarily think that M is a model of ZFC (after all, the original model may not see any models of ZFC, in which case the extended model will not see any either). In Ackermann set theory, the class V of all sets cannot be shown to be a model of ZFC, though in some external sense it is, for similar reasons.

3.4 A pocket set theory

This final theory with classes is an expansion by one of the authors (Holmes) of a suggestion of Rudy Rucker. It is the weakest set theory discussed so far in this chapter. It is also very funny.

It is a folk observation that there are two sizes of infinite sets which occur in nature (that is, which occur naturally in mathematics outside of set theory). These are $\aleph_0$ and c, the cardinalities of the set of natural numbers and the set of real numbers respectively. The theory under consideration here asserts that these are the *only* infinite cardinalities (the alternative set theory of Vopěnka discussed below also has this consequence). The very amusing aspect is that with carefully crafted axiomatics the assertion that these are the only two infinite cardinalities turns out to be almost the entire theory, and the theory is strong enough to support most classical mathematics.

Pocket set theory is a theory with sets and classes. We have the usual axioms and definitions: this is a first-order single sorted theory with equality and membership as primitive relations.

Extensionality: $(\forall AB.A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B))$

Definition: $set(x) \equiv (\exists Y.x \in Y)$

Class Comprehension: For any formula ϕ and variable A not free in ϕ, $(\exists A.(\forall x.x \in A \leftrightarrow set(x) \wedge \phi))$.

Definition: We define $\{x \mid \phi\}$ as the unique A (A not free in ϕ) such that $(\forall x.x \in A \leftrightarrow set(x) \wedge \phi)$.

Definition: For any sets a and b, we define $\{a, b\}$ as $\{x \mid x = a \vee x = b\}$. We define $\{a, a\}$ as $\{a\}$. We define (a, b) as $\{\{a\}, \{a, b\}\}$. Note that these are strictly definitions: we do not as yet know that there are any sets at all, nor have we assumed that an unordered pair of sets is itself a set.

Definition: A *relation* is a class of ordered pairs. For any relation R, we define R^{-1} as $\{(y, x) \mid (x, y) \in R\}$ (if this class exists, which requires that (y, x) be a set for each $(x, y) \in R$). A relation is a function f such that $(\forall xyz.(x, y) \in f \wedge (x, z) \in f \rightarrow y = z)$. A function f is a bijection iff f^{-1} exists and is a function. We define $\mathrm{dom}(R)$ for any relation R as $\{x \mid (\exists y.(x, y) \in R)\}$. We say that f is a bijection from A to B if f is a bijection, $\mathrm{dom}(f) = A$ and $\mathrm{dom}(f^{-1}) = B$. Sets A and B are said to be the same size iff there is a bijection from A to B.

Now we give the axioms which provide the specific content of this theory (so far we have just given the generic theory of sets and classes and some definitions).

Definition: A class C is *infinite* iff there is a bijection from C to some proper subclass of C. A class C is *proper* if it is not a set.

Axiom of Infinite Sets: There is an infinite set. Any two infinite sets are the same size.

Axiom of Proper Classes: Any two proper classes are the same size, and any class the same size as a proper class is proper.

These axioms assert that there are just two infinite cardinalities for classes, the cardinality of the infinite sets and the cardinality of the proper classes.

This may not seem like a mathematically adequate set of axioms. But it is.

Definition: We fix an infinite set I. We define R as the class $\{x \mid x \notin x\}$, the Russell class, which is demonstrably proper.

Theorem: The empty class is a set.

Proof: Note that $\emptyset = \{x \mid x \neq x\}$ does exist as a class by Class Comprehension. Suppose that $\emptyset$ is not a set and so is proper. It follows that $\emptyset$ is the same size as R, so R is empty from which it follows that every set is self-membered. It follows further that $\{I\}$ is a set (because it is certainly not the same size as $\emptyset$) but also $\{I\} \notin \{I\}$, because $I \neq \{I\}$, because $\{I\}$ clearly is not the same size as any of its proper subclasses, so it cannot be the infinite I, so $\{I\} \in R$, which is a contradiction.

Theorem: For any set x, $\{x\}$ is a set.

Proof: Suppose otherwise. Then $\{x\}$ is proper and must be the same size as R, so R has just one element, which must be $\emptyset$ as obviously $\emptyset \in R$. From this it follows that $\{I, \emptyset\}$ is a set (because both I and $\emptyset$ are sets and there cannot be a bijection from $\{I, \emptyset\}$ to $\{x\}$), but neither I nor $\emptyset$ can have exactly two elements, so $\{I, \emptyset\} \in R$, which is a contradiction.

Theorem: For any sets x and y, $\{x, y\}$ is a set.

Proof: R contains $\emptyset$, $\{\emptyset\}$ and $\{\{\emptyset\}\}$, so $\{x, y\}$ cannot be the same size as R.

Corollary: (x, y) is a set for any sets x and y.

Theorem: For any formula $\phi(x, y)$, there is a class $\{(x, y) \mid \phi(x, y)\}$ which implements it. This makes it possible to demonstrate the existence of definable bijections as classes (note that until this point we have never appealed to the existence of specific bijections but only to the logical impossibility of the existence of bijections between concrete finite sets of specific sizes).

Definition: A von Neumann ordinal is a transitive set which is strictly well-ordered by the membership relation (where this means that every sub*class* of a von Neumann ordinal has a membership-minimal element).

Theorem: The class of all von Neumann ordinals is strictly well-ordered by the membership relation. The proof of this is standard, involved, and omitted.

Theorem: The class of all von Neumann ordinals is not a set.

Proof: This class is transitive and it is strictly well-ordered by membership. If it were a set it would be a von Neumann ordinal and so self-membered and so not strictly well-ordered by membership, which is a contradiction.

Theorem: If two classes are each the same size as a subclass of the other, then they are the same size. This is the Cantor-Schröder-Bernstein theorem and the standard proof is omitted.

Theorem: The Axiom of Choice holds.

Proof: By Cantor-Schröder-Bernstein, the universe V of all sets and R are the same size (R is the same size as $R \subseteq V$ and V is the same size as

$$\{\{\{x\}, \emptyset\} \mid x \in V\} \subseteq R.)$$

R is the same size as the class of von Neumann ordinals. Thus the universe is the same size as the class of von Neumann ordinals and a specific bijection from V to the ordinals gives an ordinal indexing of the universe which induces a well-ordering in the obvious way. A global well-ordering of the universe obviously implies Choice: choose the least element in the well-ordering of each element of a pairwise disjoint family of nonempty sets in order to get a choice set.

Theorem: Each set is the same size as a von Neumann ordinal.

Proof: Each set is the same size as a set of von Neumann ordinals, which can be mapped by an obvious recursion to an initial segment of the von Neumann ordinals, which is itself a von Neumann ordinal, which will be a set because it is in bijection with a set.

Theorem: There is an infinite von Neumann ordinal.

Proof: The von Neumann ordinal which is the same size as I is infinite.

Theorem: All infinite von Neumann ordinals are the same size as the first infinite von Neumann ordinal ω (that is, all infinite ordinals are countably infinite).

Proof: All infinite sets are the same size.

Theorem: The class of all subsets of ω is not a set.

Proof: The standard argument for Cantor's Theorem shows that ω is not the same size as the class of its subsets. Since the class of all subsets of ω is clearly infinite, it cannot be a set, and so must be a proper class.

Construction: The usual constructions of integers, rational numbers, and real numbers as Dedekind cuts in the rationals can be carried out. The set of reals is the same size as the class of all sets of natural numbers for standard reasons, and so is a proper class. Individual reals are countable collections of rationals, thus sets.

Theorem: Each infinite subclass of the class of real numbers is either the size of the class of real numbers or countable. $c = \aleph_1$.

Proof: The real numbers make up a proper class. An infinite subclass of the reals is either a set, in which case it is the same size as the infinite set ω, or it is a proper class, in which case it is the same size as the proper class of all reals. The real numbers can be placed in one-to-one correspondence with the countable ordinals, as both collections are proper classes.

Comment: The principle of Limitation of Size of the original class theories, which appears here as the Axiom of Proper Classes, in this context allows us to prove not only the "Axiom" of Choice (as in the original theories) but also the Continuum Hypothesis! This theory is much weaker than the theory of types (all objects it constructs appear by type 5 on the most uncharitable interpretation of the capabilities of TST), but every mathematical object needed in physics is constructible here. The collection of *all* functions from the reals to the reals is too large, but notice that the collection of continuous functions from the reals to the reals is of size c and can be represented in fairly natural ways, and in general the constructions actually needed in mathematical physics (or any mathematics short of set theory and shorn of excessive levels of abstraction) do not transcend the cardinality c. Points of Hilbert space are countable sequences of real numbers (thus sets) and continuous functions on Hilbert space are representable just as continuous functions on the reals are representable, and so forth.

4 THEORIES WITH ATOMS AND THEORIES WITH ANTI-FOUNDATION AXIOMS

4.1 Weak extensionality and ZFA

Most modern set theories have sets (or more generally classes) as the only objects. This is odd because originally sets were conceived as collections of other, previously given objects. It has seemed less odd (at least to those who have completed their indoctrination in foundations of mathematics) since we have adopted the view that all mathematical objects are sets. It also affords a logical simplification of the theories.

The usual form of the axiom of extensionality is

Axiom of Extensionality: $(\forall AB.A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B))$

A more natural axiom from a naïve standpoint would be

Weak Axiom of Extensionality: $(\forall AB.class(A) \wedge class(B) \rightarrow (A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B)))$

Here we restrict the scope of extensionality to classes (we say "class" here to avoid collision with the notion of "sethood" found in the theories with sets and classes above; below in NFU we will use "set" for this concept).

Further, it seems natural to assert that classes are the only things which have elements:

Axiom of Classhood: $(\forall Ax.x \in A \rightarrow class(A))$

The non-classes have no elements. These objects are generally called *atoms* or *urelements*. There is (in most theories) an elementless class $\emptyset$ as well (and it is the only one). This observation about $\emptyset$ indicates that, though we can formalize theories with weak extensionality using a primitive notion of classhood there is no need to do this.

Definition: $class(x) \equiv x = \emptyset \vee (\exists y.y \in x)$

This treatment applies equally well to theories with or without proper classes. If no distinction is drawn between sets and classes, then the predicate *class* could equally well be written *set*. In either case it is definable if we are willing to take $\emptyset$ as a primitive notion instead of the classhood (or sethood) predicate.

In most cases it is mathematically preferable to suppose that everything is a set and adopt the strong form of extensionality. In the context of New Foundations the adoption of strong extensionality has very strong consequences: the theory NFU with weak extensionality is known to be consistent where the original theory NF of Quine is not. This will be discussed below. In Kripke-Platek set theory (see [Barwise, 1975]) it is found useful to have a variant KPU with urelements: this we do not discuss at all here.

The theory (or theories) ZFA obtained by replacing Extensionality with Weak Extensionality in the usual set theory ZFC have a notable practical application which make them worthy of note as an alternative set theory. These theories support a relatively straightforward proof of the independence of the Axiom of Choice from the other axioms of set theory [Fraenkel, 1922]. It is trickier to prove the relative consistency of $\neg$AC with extensional ZF: this was not done until forcing was developed by Cohen (and the proof owes something to the technique based in ZFA).

We note that there are two different sorts of ZFA, depending on whether the atoms make up a set or a proper class. A very strong denial of Extensionality would stipulate that every set is the same size as some set of atoms; a weaker nonextensional theory would provide that the universe consisted of the union of all the iterated power sets of a set A of atoms.

If we fix an infinite set of atoms A, we can consider the action of permutations on A on all sets: if π is a permutation of A, there is a uniquely determined permutation π^* acting on all sets such that $\pi^*(a) = \pi(a)$ for $a \in A$ and $\pi^*(B) = \pi^*``B$ for all sets B (the existence of this uniquely determined permutation depends on Foundation).

We can then say that a set B has support $S \subseteq A$ iff every permutation π^* fixing each element of $A \setminus S$ also fixes B. The punchline is that the class of all sets with finite support in A satisfies all the axioms of ZFA except Choice: for certainly A itself has finite support in A and no linear order of A has finite support in A (much less any well-ordering of A). The demonstration that the other axioms of ZFA hold in the domain of sets with finite support in A is technically involved and beyond the scope of this chapter.

While there might seem to be philosophical advantages to providing many non-classes with no elements (because it is a very sophisticated ontological perspective indeed that would lead us to decide that everything is a set), the mathematical advantage generally seems to be with stipulating that everything is a set, and where we do find mathematical applications for atoms they do not seem to have anything to do with the pre-set-theoretical reasons for supposing that there are such objects.

4.2 Aczel's AFA

Nowadays when people use the phrase 'anti-foundation axiom' it is almost always to denote specifically the anti-foundation axiom of Forti and Honsell [Forti and Honsell, 1983] rather than any of the other anti-foundation axioms that have appeared from time to time. Broadly the common noun 'anti-foundation axiom' is used to refer to an axiom (scheme) which can be added to the axioms of ZF(C) once the Axiom of Foundation has been dropped. Thus anti-foundation axioms are alternative in the sense of being alternative to the axiom scheme of Foundation *once the other axioms of ZF have been agreed on.* This is a different sense of 'alternative' from the way in which NF or GPK^+_∞ are alternative.

Forti-Honsell's axiom (which is called AFA by Peter Aczel in his [Aczel, 1988]) is the most interesting and gives rise to the clearest mathematics, and this is probably because of all the anti-foundation axioms it is that one which most clearly arises from a sensible conception of what sets are. It arises from the idea that sets are things denoted by set pictures.

Definition: The *field* of a relation R is defined as the union of its domain and range. A pair (R, r) is a *set picture* iff R is a relation, r is an element of the field of R, and there is a map f with domain the field of R with the property that $f(x) = \{f(y) \mid y\, R\, x\}$ for all x in the field of R. We say further that (R, r) is a *set picture* of $f(r)$. There is no immediate reason that a set picture should be a set picture of just one set (but this is the case in the first theory we consider).

In the usual set theory ZFC (or for that matter in ZF: choice is not an issue), the question of what relations are set pictures is fully settled in a very simple way.

Theorem (ZFC): A pair (R, r) with r in the field of R is a set picture iff the relation R is well-founded. Further it is the picture of a uniquely determined set.

The Axiom of Foundation implies immediately that any set picture must be well-founded, and the Mostowski Collapsing Lemma establishes that any well-founded relation is actually a set picture of a uniquely determined set.

Now we change our attention to the theory ZFC– which is obtained by omitting the Axiom of Foundation from ZFC. It now becomes conceivable that more general relations R may be set pictures.

Consider the very simple example of a reflexive relation R with field $\{r\}$. If (R, r) is a set picture, there is a set which is its own sole element. Note that nothing tells us that there cannot be many such sets if the possibility of one is admitted: the same relation may be the picture of many different sets.

It is a theorem of ZFC– that for each well-founded set picture (R, r) the set of which it is a picture is uniquely determined, but this depends strongly on the well-foundedness of R.

The axiom of Forti and Honsell which is called AFA by Aczel asserts this:

Axiom of Anti-Foundation: Each pair (R, r) with R a relation and r in the field of R is a set picture and the set picture of a uniquely determined set.

There are two things going on here: that every pair (R, r) is a set picture says that the Axiom of Foundation fails in very elaborate ways (though not in all possible ways as we will see); that each relation R is a set picture in *only one way* amounts to a strong form of extensionality. There is a set which its own sole element under AFA, but also there is exactly one such set.

At least one of the authors (Holmes) can testify that the appearance of Aczel's book which popularized this anti-foundation axiom caused a deal of philosophical excitement, as it seemed that the possibility of self-membered sets somehow transcended the iterative conception of set. This excitement was ill-founded (the author begs forgiveness for the pun): the relation of ZFC– + AFA to ZFC is intimate, and even in its own terms it reveals the stamp of the iterative conception of set. The theory does have technical uses in avoiding certain inconveniences of the universe of well-founded sets, but its philosophical differences from ZFC have been overstated.

ZFC– + AFA is readily interpreted in ZFC. If we have two relations R and S in ZFC we say that a relation b from the field of R to the field of S is a *bisimulation* iff $x\, b\, y$ iff for each $z\, R\, x$ there is $w\, S\, y$ such that $z\, b\, w$ and for each $w\, S\, y$ there is $z\, R\, x$ such that $z\, b\, w$. The objects of our interpretation of ZFA will be pairs (R, r) where R is a relation and r is an element of the field of R. Two pairs (R, r) and (S, s) are equivalent iff there is a bisimulation from the field of R to the field of an end extension of S which relates r to s (an end extension of S is a relation including S under which elements of the field of S have the same preimages they have under S). We say that $(R, r)\, E\, (S, s)$ iff (R, r) is equivalent to some (S, t) where $t\, S\, s$. E is the membership relation of our interpretation of ZFC– + AFA. We will not present a proof that this works in the limited space of this chapter.

In terms of ZFC– + AFA, a version of the cumulative hierarchy is readily recovered in spite of the ill-founded nature of the theory. The stages of this hierarchy are indexed by the ordinals as is the usual hierarchy of ranks. $H_0 = \{0, 1\}$ (it will clear in a moment why we cannot start with an empty rank). Once H_α is defined, we define $H_{\alpha+1}$ as the collection of all sets which have set pictures R with field a subset of H_α. At limit stages, we take unions as in the usual hierarchy. Observe for example that H_1 contains three sets, the empty set and $\{\emptyset\}$ (pictured by the usual order relation on $\{0, 1\}$), and the solitary object which is its own sole element (pictured in any reflexive relation). It is straightforward to show that the power set of H_α is included in $H_{\alpha+2}$ (and in $H_{\alpha+1}$ if α is infinite) but of course each new level of the hierarchy includes new sets defined in non-well-founded ways. However, the novelty is quite constrained: H_α is the same size as V_α for each infinite ordinal α.

4.3 Boffa's axiom

An anti-foundation axiom with effects rather different those of AFA was proposed by Boffa [Boffa, 1972]. Aczel's axiom allows relatively few non-well-founded objects because of its strong extensionality attributes (for example, it allows only one self-singleton). Boffa's axiom allows (speaking roughly) as many non-well-founded sets as possible. For example, under ZFC– with Boffa's anti-foundation axiom there is a proper class of self-singletons.

Definition: A relation R is *extensional* iff each element x of the range of R is uniquely determined by its preimage $R^{-1}``\{x\}$.

Definition: Let R be an extensional relation. A *set-labelling* of R is an injection s with domain a subset of the field of R with the property that if x is in the domain of s then $R^{-1}``\{x\}$ is included in the domain of s and

$$s(x) = \{s(y) \mid y\,R\,x\}.$$

Boffa's anti-foundation axiom: Any set-labelling of an extensional relation R can be extended to a set-labelling of R (not necessarily unique) whose domain is the entire field of R.

So, for example, if we define a relation R on the domain $\mathbb{N}$ so that $m\,R\,n$ iff either $n = 0$ or $m = n > 0$, we find that under Aczel's axiom this is a set picture of the unique self-singleton, whereas under Boffa's axiom we can extend the empty set-labelling of this set to a total set-labelling of its range and discover that each positive integer is sent to a self-singleton and 0 is sent to a countably infinite set of self-singletons. The same technique can be used to get as many distinct self-singletons as we might want.

Boffa's axiom is similarly motivated by the idea of sets as derived from set pictures, but with leeway for many sets to be represented by the same picture, and ZFC– with Boffa's axiom admits a fairly straightforward relative consistency proof from ZFC.

5 NEW FOUNDATIONS AND RELATED SYSTEMS

In this section we review the set theory "New Foundations" (NF) proposed by W. v. O. Quine in 1937 [Quine, 1937] and related systems. This system is not so far known to be consistent, and it was shown by Specker in [Specker, 1953] that it disproves the Axiom of Choice (and so it proves Infinity). However, there are several known subsystems of *NF* which are known to be consistent (none of which reproduce Specker's disproof of Choice, though one of them does allow use of Specker's argument to prove Infinity, as we will see below).

5.1 Stratified comprehension

New Foundations is a variant of TST rather than of Zermelo set theory, in spite of the fact that it is an untyped set theory. The starting point for the development of NF is the observation that TST enjoys a great deal of what Russell called "systematic ambiguity", modern set theorists call "typical ambiguity", and computer scientists call “polymorphism”. For any sentence ϕ in the language of TST, define ϕ^+ as the sentence obtained by raising the type of each variable in ϕ by one. Note that for any axiom ϕ, ϕ^+ is also an axiom, and that for any rule of inference allowing ψ to be deduced from χ, ψ^+ can be deduced from χ^+ using the same rule. It follows from these considerations that for any formula ϕ, if ϕ is a theorem so is ϕ^+. If we associate with each variable x a specific variable x^+ of the next higher type, we can extend the definition of ϕ^+ to all formulas. Note that each object $\{x \mid \phi\}$ defined by a set abstract has a precise analogue $\{x^+ \mid \phi^+\}$ in the next higher type: for example the Frege natural number 3 defined in type 2 (the set of all type 1 sets with three elements) has a precise analogue in type 3 (the set of all type 2 sets with three type 1 elements) and indeed in each higher type.

Quine’s suggestion is that this is an indication that the types need not actually be different. He proposed that all the types might be the same domain: the sentence ϕ says the same thing as the sentence ϕ^+ and the object $\{x \mid \phi\}$ is the same object as $\{x^+ \mid \phi^+\}$: in our example, there is exactly one Frege natural number 3 which is simply the set of all sets with three elements.

The axioms of Quine’s theory “New Foundations” (so called from the name of the paper [Quine, 1937]) are the extensionality and comprehension axioms of TST with all indications of type removed. The details follow.

NF is a first-order single-sorted theory with equality and membership as primitive relations.

Axiom of Extensionality: $(\forall AB.A = B \leftrightarrow (\forall x.x \in A \leftrightarrow x \in B))$.

Axiom of Comprehension: For any formula ϕ in which A does not appear free, and which can be converted to a well-formed formula of TST by an assignment of types to variables, $(\exists A.(\forall x.x \in A \leftrightarrow \phi))$ is an axiom.

Definition: For any formula ϕ in which A does not appear free, $\{x \mid \phi\}$ is defined as the unique A (if there is one) such that $(\forall x.x \in A \leftrightarrow \phi)$ (this exists by Comprehension if ϕ can be obtained by dropping types from a formula of TST and is unique if it exists by Extensionality).

This presentation of the axioms should make it clear why dropping types from the axioms of TST does *not* give us the inconsistent Axiom of Comprehension of naïve set theory. The Russell class $\{x \mid x \notin x\}$ is not provided by the Axiom of Comprehension of NF because there is no way to assign types to the variables in $x \notin x$ which gives a well-formed formula of TST.

It is usual to rephrase the comprehension axiom of NF in a way which hides the apparent dependence on the language of another theory.

Definition: A formula ϕ is said to be *stratified* iff there is a function σ from the set of all variables to the natural numbers (equivalently, to the integers) such that for each atomic subformula $x = y$ of ϕ we have $\sigma(x) = \sigma(y)$ and for each atomic subformula $x \in y$ of ϕ we have $\sigma(x)+1 = \sigma(y)$. Such a function σ is called a *stratification* of ϕ.

Axiom of Stratified Comprehension: For any stratified formula ϕ, $(\exists A.(\forall x.x \in A \leftrightarrow \phi))$ is an axiom.

A stratification of ϕ of course codes an assignment of types to the variables in ϕ which converts ϕ to a well-formed formula of TST.

The dependence of the axiomatization of NF on an understanding of TST really is just apparent as we said above. The axiom scheme of Stratified Comprehension is equivalent to a finite collection of its instances, and so can be given as a list of particular comprehension axioms with no reference to types or stratification at all. The original reference for this is [Hailperin, 1944], but the implementation given there is terrible. The construction of such a finite axiomatization is similar to our development of the finite axiomatization of predicative class comprehension given above.

5.2 New Foundations with urelements

In this section we will explore the development of set theory with stratified comprehension, but we will begin with a critique of the theory history presents to us.

Motivation of NFU

Quine claimed in [Quine, 1937] that the choice of strong extensionality over weak extensionality is purely a matter of convenience. He suggested that any urelements (objects with no elements distinct from the empty set and from one another) could be reinterpreted as self-singletons $x = \{x\}$, thus restoring strong extensionality. In Zermelo-style set theory this suggestion makes sense. One can define a (proper class) map f which sends each urelement and iterated singleton of an urelement to its singleton and redefine membership to the relation $x \in_f y \equiv_{def} x \in f(y)$. This will preserve the axioms of Zermelo-style set theory and convert all urelements to self-singletons. But this procedure makes no sense in a set theory with stratified comprehension, because the map f has an unstratified definition, so sets defined in terms of $\in_f$ cannot be expected to exist in general.

This suggests that the theory we should be considering is the theory NFU whose primitive notions are equality, membership, and the empty set, and whose axioms are as follows:

Axiom of the Empty Set: $(\forall x.x \notin \emptyset)$

Axiom of Weak Extensionality: $(\forall ABx.x \in A \rightarrow (A = B \leftrightarrow (\forall y.y \in A \leftrightarrow y \in B)))$

Axiom of Stratified Comprehension: For any stratified formula ϕ, $(\exists A.(\forall x.x \in A \leftrightarrow \phi))$ is an axiom.

Definition: $set(x) \equiv_{def} x = \emptyset \vee (\exists y.y \in x)$. Notice that $(\forall AB.set(A) \wedge set(B) \rightarrow (A = B \leftrightarrow (\forall y.y \in A \leftrightarrow y \in B)))$

Definition: For any stratified formula ϕ, we define $\{x \mid \phi\}$ as the unique A such that $set(x)$ and for all x, $x \in A \leftrightarrow \phi$. This exists by stratified comprehension and is unique by weak extensionality and the definition of the empty set.

And indeed history suggests that this is the theory we should be thinking about. The consistency of NF remains an open question after more than 70 years, but R. B. Jensen proved the consistency of NFU on quite weak assumptions in [Jensen, 1969] in 1969. Most of the standard development of mathematics in NF (which looks quite odd from the Zermelo-trained standpoint) works perfectly well in NFU, which has well-understood models which can be examined to see what is going on.

For this reason, we will follow the plan of giving the general presentation of mathematics with stratified comprehension in NFU, then following with a section in which distinctive features of mathematics in NF (notably Specker's disproof of the Axiom of Choice in [Specker, 1953]) are described.

A model construction for NFU

We first give a model construction for NFU. This is not the original consistency proof due to Jensen, but a similarly motivated model construction due to Maurice Boffa [Boffa, 1988].

We consider a model of Mac Lane set theory without the Axiom of Infinity or the Axiom of Choice (though both could be adjoined and usually will be) with an external automorphism j which moves a rank V_α of the cumulative hierarchy downward. The existence of such models is a standard result of model theory. The domain of the model of NFU to be constructed is the extension of the nonstandard rank V_α. The membership relation of the model is the relation $x \in_{NFU} y \equiv_{def} j(x) \in y \wedge y \in V_{j(\alpha)+1}$. Each element x of $V_{j(\alpha)+1}$ is assigned the extension of $j^{-1}(x)$ in the original model (notice in particular that $V_{j(\alpha)}$ is assigned the extension of V_α, the entire universe of the model!), and each element of $V_\alpha \setminus V_{j(\alpha)+1}$ is treated as an urelement. This should make it clear that weak extensionality is satisfied. Of course the empty set of the model of NFU is the empty set of the original model of set theory.

We argue that stratified comprehension is satisfied in the model. Let ϕ be a stratified formula in the language of NFU with a stratification σ. Let N be a natural number larger than the value of σ at any variable in ϕ. This formula has a translation ϕ_1 into the language of the model of Mac Lane in which the model construction is carried out. Each reference to $\emptyset$ in NFU is replaced with a reference to $\emptyset$ in the language of Mac Lane set theory, each atomic formula $u = v$ is replaced with $u = v$, each atomic formula $u \in v$ is replaced with $j(u) \in v \wedge v \in V_{j(\alpha)+1}$, and each quantifier is bounded in V_α. It might seem that we just need to construct the set $\{x \in V_\alpha \mid \phi_1\}$ and note that $j(\{x \in V_\alpha \mid \phi_1\})$ is the set which we assign this extension in the model of NFU. That is what we do in the end, but it requires justification. The problem is that the formula ϕ_1 contains references to the automorphism j, and the Axiom of Separation of Mac Lane set theory does not apply to formulas containing j, unless all references to j are in parameters (terms not containing bound variables). We show how to effect a transformation of ϕ_1 to an equivalent formula in which j occurs only in parameters. We begin by replacing each atomic formula $u = v$ with $j^{N-\sigma(v)}(u) = j^{N-\sigma(v)}(v)$. This is equivalent because j is an automorphism. Notice that $j^{N-\sigma(v)}(u) = j^{N-\sigma(u)}(u)$ because σ is a stratification. We replace each formula $j(u) \in v$ with $j^{N-\sigma(v)}(j(u)) \in j^{N-\sigma(v)}(v)$. This is equivalent because j is an automorphism. Notice that $j^{N-\sigma(v)}(j(u)) = j^{N-\sigma(u)}(u)$ because σ is a stratification. We replace formulas of the shape $v \in V_{j(\alpha)+1}$ with the equivalent $j^{N-\sigma(v)}(v) = j^{N-\sigma(v)}(V_{j(\alpha)+1})$. These substitutions produce a formula ϕ_2. In ϕ_2, every variable u occurs with exactly $N - \sigma(u)$ applications of j. An equivalent formula is produced by replacing each $j^{N-\sigma(u)}(u)$ such that u is bound by a quantifier (restricted to V_α) with the variable u, replacing the bounding set with $j^{N-\sigma(u)}(V_\alpha)$. These substitutions produce a formula ϕ_3. The variable x occurs in ϕ_3 only in the context $j^{N-\sigma(x)}(x)$. ϕ_4 is obtained by replacing this term with x. Note that ϕ_4 contains no occurrences of j except in constants. $j^{-(N-\sigma(x))}(\{x \in j^{N-\sigma(x)}(V_\alpha) \mid \phi_4\})$ is the set $\{x \in V_\alpha \mid \phi_1\}$ of the original model of Mac Lane set theory, and the image of this set under j is the set $\{x \mid \phi\}$ of the model of NFU.

If we let the ordinal α above be a nonstandard natural number, we obtain a model of NFU in which the universe is finite! This witnesses an error in [Quine, 1937]: Quine says there that the existence of $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, etc. shows that Infinity holds in NF. Infinity *does* hold in NF but not for this reason. This sequence of sets witnesses the fact that any model of NFU (including a model of NFU + "the universe is finite") must actually be externally infinite, but it does not witness the Axiom of Infinity because there is no reason to expect that it is a set (its definition is not stratified). In a model of NFU + "the universe is finite" it is *not* a set. It is interesting to observe that in fact this collection *can* be a set but this would take us far afield.

If we assume that Infinity and Choice hold in the underlying model of Mac Lane, Infinity and Choice hold in the model of NFU. We can obtain strong axioms of infinity by assuming them in the underlying model of Mac Lane set theory, but more satisfying results can be achieved by making strong assumptions about the automorphism j (which are generally equivalent to strong axioms of infinity in Mac Lane or *ZFC*, but give nicer characteristics to the model of NFU).

Mathematics in NFU

In this section, we outline the development of the foundations of mathematics in NFU. What we do here will also work in NF, as long as we do not assume Choice. Infinity, which we must add as an assumption to NFU for the purposes of this development, is a theorem of NF, as we will see in the next section.

We develop the natural numbers (following Frege's definition) exactly as we did above in TST. We define 0 as $\{\emptyset\}$. For each set A, define $A + 1$ as $\{a \cup \{x\} \mid a \in A \wedge x \notin a\}$. $A + 1$ is the set of all objects obtained by adding a single new element to an element of A. Now $0 + 1$ is the set of all one-element sets, which we call

1, and $1+1$ is the set of all two-element sets, which we call 2, and so forth. We define $\mathbb{N}$, the set of natural numbers, as the intersection of all sets which contain 0 and are closed under the extended "successor" operation. The elements of the natural numbers are the finite sets: we define the set $\mathbb{F}$ of finite sets as $\bigcup \mathbb{N}$.

Notice that NFU witnesses the coherence of Frege's implementation of the natural number n as the set of all sets with n elements.

We are not quite done. There is a bad possibility which we have avoided discussing in the development above. One can prove by mathematical induction that no natural number contains a proper subset of one of its elements. This implies that if the universal set V is finite, the natural number which contains it as an element is $\{V\}$. It then follows that the successor of V is $\emptyset$, and of course the successor of $\emptyset$ is also $\emptyset$. This gives an exception to Peano's fourth axiom (all the others are easily seen to hold for this implementation of the natural numbers).

We rule out this badness by adopting the

Axiom of Infinity: $V \notin \mathbb{F}$

We further usually adopt the

Axiom of Choice: Each pairwise disjoint collection of nonempty sets has a choice set.

though in this context it should be noted that the Axiom of Choice is false in NF.

We discuss the extension of the definition of stratification to support term constructions. Where ψ is a stratified formula, let $(\iota x.\psi)$ represent the unique object x such that ψ (if there is one) and the empty set otherwise. We define $(\exists! x.\psi)$ as $(\exists x.\psi) \wedge (\forall xy.\psi \wedge \psi[y/x] \rightarrow x = y)$ (y being a fresh variable). Notice that this is stratified iff ψ was stratified. Now note that $\phi[(\iota x.\psi)/x]$ is equivalent to $((\exists! x.\psi) \wedge \phi) \vee (\neg(\exists! x.\psi) \wedge \phi[\emptyset/x])$. The correct notion of stratification for general terms containing definite description terms $(\iota x.\psi)$ is as follows. A stratification of a formula ϕ is a function σ from terms to natural numbers (or integers) such that for any atomic subformula $t = u$ we have $\sigma(t) = \sigma(u)$, for any atomic subformula $t \in u$ we have $\sigma(t) + 1 = \sigma(u)$ [note that t and u may be complex terms rather than variables], and $\sigma((\iota x.\psi)) = \sigma(x)$: each definite description is assigned the same type as the variable bound in it. To see that this works, check that in the transformations above a stratification of a formula is the same as the stratification of the transformed formula. We do not intend to explicitly use definite description terms in what follows, but any defined term construction introduced can be thought of as being implemented in this way.

As we noted above under TST, the usual Kuratowski definition $(x, y) \equiv_{def} \{\{x\}, \{x, y\}\}$ of the ordered pair is inconvenient in type theory because the pair thus defined is two types higher than its projections.

If we assume Infinity, it is possible to define an ordered pair on $\mathcal{P}^2(V)$ (the collection of sets of sets) which is the same type as its projections. In NF, since

$\mathcal{P}^2(V) = V$, this is a definable type level ordered pair on the universe. In NFU, we develop a modified interpretation of NFU in which there is a type level pair on the universe. It should also be noticed that it is provable in NFU + Infinity + Choice that there is a type level pair (this follows from the theorem $\kappa^2 = \kappa$ for infinite cardinals κ, in the special case $\kappa = |V|$, but this line of development requires that a lot of mathematics be developed using the Kuratowski pair first).

Definition: We define $\sigma(x)$ as $x + 1$ if x is a natural number and as x otherwise.

Definition: We define $\sigma_1(x)$ as $\{\sigma(y) \mid y \in x\}$ and $\sigma_2(x)$ as $\sigma_1(x) \cup \{0\}$, for any $x, y \in \mathcal{P}(V)$.

Observations: Notice that $\sigma_i(x) = \sigma_i(y) \rightarrow x = y$ for $i = 1, 2$, and note also that $\sigma_1(x) \neq \sigma_2(y)$ for any x, y.

Definition: We define σ_1"x as $\{\sigma_1(y) \mid y \in x\}$, σ_2"x as $\{\sigma_2(y) \mid y \in x\}$, and (x, y) (the Quine ordered pair of x and y) as σ_1"$x \cup \sigma_2$"y, for any $x, y \in \mathcal{P}^2(V)$.

Definition: For any $w \in \mathcal{P}^2(V)$, we define $\pi_1(w)$ as $\{x \mid \sigma_1(x) \in w\}$ and $\pi_2(w)$ as $\{x \mid \sigma_2(x) \in w\}$.

Observation: $\pi_1(x, y) = x; \pi_2(x, y) = y; (\pi_1(x), \pi_2(x)) = x$, for all $x, y \in \mathcal{P}^2(V)$.

This gives an ordered pair on sets of sets which is assigned the same value as its projections by any stratification (due to Quine in [Quine, 1945] and thus called the Quine ordered pair) . Now for a trick. We define a new interpretation of NFU whose domain is the extension of $\mathcal{P}^2(V)$ in the model we start with. The empty set of the new interpretation is the empty set of the old model. The equality relation of the new interpretation is the equality relation of the old interpretation restricted to the new domain. The membership relation $x \in_{new} y \equiv_{def} x \in y \wedge y \in \mathcal{P}^3(V)$. In other words, each set of sets of sets retains its original extension (so all subsets of the domain of the new interpretation are realized) and sets of sets which are not sets of sets of sets are treated as urelements. The verification that the new interpretation satisfies the axioms of NFU is straightforward. The type level ordered pair on sets of sets of the old interpretation serves as a type level pair in the new interpretation. If we replace natural numbers n with their restrictions $n \cap \mathcal{P}^2(V)$ to sets of sets in the definition of the Quine pair, and use the modified Quine pair to carry out the construction of the new interpretation, then the type level pair of the new interpretation will coincide with the original (unmodified) Quine pair of the new interpretation on sets of sets. The reason for this is that the restricted natural number $n \cap \mathcal{P}^2(V)$ becomes a natural number in the new interpretation.

The upshot of this discussion is that we might as well adjoin relations π_1 and π_2 with the same stratification conditions as equality to our language, define (x, y) as $(\iota z.z\pi_1 x \wedge z\pi_2 y)$, define $\pi_1(x)$ as $(\iota y.x\pi_1 y)$ and $\pi_2(x)$ as $(\iota y.x\pi_2 y)$, then adjoin an

Axiom of Ordered Pairs: $\pi_1(x,y) = x; \pi_2(x,y) = y; (\pi_1(x), \pi_2(x)) = x$.

We could but do not need to stipulate that the type level ordered pair coincides with the Quine pair on sets of sets. The introduction of type level ordered pairs by axiom in NFU is a proposal of one of the authors (Holmes) in his [Holmes, 1998].

The Axiom of Ordered Pairs is actually an inessential strengthening of the Axiom of Infinity (in the presence of Choice, they are equivalent); we have just shown that NFU + Infinity interprets NFU + Ordered Pairs. The function sending x to $(x, 0)$ has a stratified definition and maps the universe V to its proper subset $V \times \{0\}$ which is sufficient to show that Ordered Pairs implies Infinity.

Whether we introduce a primitive type level pair as we have just done or use the Kuratowski pair, the definitions of relations and functions and the basic properties of and operations on relations and functions are similar. There is a precise correspondence between the functions and relations implementable with a type level pair and those implementable with the Kuratowski pair. One has to bear in mind with the Kuratowski pair that the pair is two types higher than its projections: this causes some inconveniences. A fundamental difference between the Quine pair and the Kuratowski pair is that the projection operations of the Quine pair (or of a primitive type level pair) are set functions, while the projection functions of the Kuratowski pair are not. If $\langle x, y\rangle$ is the Kuratowski pair, there is a set function π_1^* such that $\pi_1^*(\langle x, y\rangle) = \{\{x\}\}$ but there is no set function π_1^* such that $\pi_1^*(\langle x, y\rangle) = x$. Notice that the double application of the singleton operation ensures that the relative types of the two occurrences of x are the same.

We define cardinal number. For any set A, we define the relation $A \sim B$ as holding iff there is a bijection f from A onto B. This is proved to be an equivalence relation in a standard way. We define $|A|$ as $\{B \mid B \sim A\}$. It is useful to observe that $|A|$ is exactly the same set whether we use the type level pair or the Kuratowski pair (because every function implemented with the type level pair has an exact analogue implemented with the Kuratowski pair and *vice versa*). $|A|$ is one type higher than A for purposes of stratification. Note that the natural numbers defined above are cardinal numbers of finite sets. This can be proved by mathematical induction. Objects which are $|A|$ for some set A are called *cardinal numbers*.

We define isomorphism types of relations. $R \approx S$ is defined as holding iff there is a bijection f from the field of R onto the field of S such that $x \, R \, y \leftrightarrow f(x) \, S \, f(y)$. Isomorphism is an equivalence relation on relations for standard reasons. The isomorphism type of a relation R is defined as $\{S \mid S \approx R\}$. The identity of the isomorphism type of R does depend on the pair used, and so does the type differential, because the isomorphism type is a set of functions and so a set of sets of ordered pairs. If we use the type level pair, the isomorphism type of R is one type higher than R for purposes of stratification; if we use the Kuratowski pair it is three types higher than R.

Isomorphism types of well-orderings are of special interest to us. If $\leq$ is a well-ordering, we define $ot(\leq)$ as the isomorphism type of $\leq$. An isomorphism type of a well-ordering is called the *order type* of a well-ordering, and an object which is the order type of some well-ordering is called an *ordinal number*.

These collections are very big. The analogous classes in Zermelo-style set theory are proper. Further, we have the universal set $V = \{x \mid x = x\}$, the very standard of bigness. We can contemplate $|V|$, the cardinality of the universe. We can prove that there is a natural well-ordering on the ordinal numbers in a quite standard way, then consider the order type Ω of the natural order on the ordinal numbers. We seem to be very close to paradox here, as $|V|$ and Ω are precisely the objects which instigate the classic paradoxes of Cantor and Burali-Forti!

However, we will now discover that NFU has its own characteristic way of averting paradox. We begin with the Cantor paradox of the largest cardinal. The Cantor theorem in TST takes the form $|\mathcal{P}_1(A)| < |\mathcal{P}(A)|$, where $\mathcal{P}_1(A)$ is the set of one-element subsets of A; the usual form $|A| < |\mathcal{P}(A)|$ would be ill-typed. The proof of the Cantor theorem of TST is inherited by NFU, so we have $|\mathcal{P}_1(A)| < |\mathcal{P}(A)|$ for any set A and so in particular $|\mathcal{P}_1(V)| < |\mathcal{P}(V)|$. Now of course we can add in $|\mathcal{P}_1(V)| < |\mathcal{P}(V)| \leq |V|$, so we get the odd result $|\mathcal{P}_1(V)| < |V|$. This is odd because we can "see" the one-to-one correspondence between these two sets defined by the singleton map. But there is no reason to believe that the singleton map is a set function, and we have just shown that it can't be.

Since we have a counterexample to the intuitively plausible proposition $|\mathcal{P}_1(A)| = |A|$ $(A = V)$ and since it is easy to show that $|A| = |B| \leftrightarrow |\mathcal{P}_1(A)| = |\mathcal{P}_1(B)|$, we define an operation on cardinals with defining equation $T(|A|) = |\mathcal{P}_1(A)|$. We have just shown that $T(|V|) < |V|$, so this is a nontrivial operation. There are sets A such that $|\mathcal{P}_1(A)| = |A|$: standard finite sets have this property, for example. Such sets are called *cantorian* sets and their cardinals are called cantorian cardinals. An even stronger property (motivated by the apparent witness our incorrect intuition that $|\mathcal{P}_1(A)| = |A|$ for all A, and also holding of concrete finite sets) is this: a set A (and its cardinality) is said to be *strongly cantorian* (s.c). iff $\iota\lceil A$, the restriction of the singleton map to A, is a set. Clearly strongly cantorian sets are cantorian; the converse is a strong axiom of infinity.

The Burali-Forti paradox is resolved in an even more interesting way. One can prove in a quite standard way that for any two well-orderings, either one is similar to an initial segment of the other, or the two are similar (the three alternatives being mutually exclusive). This determines a natural linear order on the ordinal numbers (*qua* isomorphism types of well-orderings) which is itself a well-ordering and so has an isomorphism type Ω. The paradox depends on the "obvious" observation that the order type of the restriction of this order to the ordinals less than α is α. The resolution depends on the discovery that this "obvious" observation is false.

For any relation R, we can define a relation $R^\iota = \{(\{x\}, \{y\}) \mid x\,R\,y\}$. If the isomorphism type of R is ρ, we define $T(\rho)$ as the isomorphism type of R^ι. It is straightforward to show that the definition of $T(\rho)$ does not depend on the choice of the relation R. Note that R^ι is one type higher than R for purposes of stratification and $T(\rho)$ is one type higher than ρ. We have just seen that the singleton map is not a function, so we have no reason to believe that $\rho = T(\rho)$ in general. A relation type ρ is cantorian iff $T(\rho) = \rho$ and strongly cantorian iff $T(\rho) = \rho$ is witnessed for each $R \in \rho$ by a set function equal to the restriction of the singleton map to the field of ρ.

Suppose that $\leq$ is a well-ordering belonging to an ordinal number α. For each x in the field of $\leq$, we can define α_x as the order type of the restriction of $\leq$ to $\{y \mid y \leq x\}$. If $\leq$ is also used for the order on ordinal numbers, we have $x \leq y \leftrightarrow \alpha_x \leq \alpha_y$. This would establish the desired isomorphism between the natural order on the ordinals less than α and the element $\leq$ of α if the map $x \mapsto \alpha_x$ were a function. But this cannot be shown to be the case, for α_x is two types higher than x for purposes of stratification (four types if we were using the Kuratowski pair). What *is* a function is $\{\{x\}\} \mapsto \alpha_x$; the application of the singleton operation repairs the failure of stratification. So what we can actually show is that $(\leq^\iota)^\iota$ is similar to the natural order on the ordinals less than α, so the order type of the ordinals less than α is $T^2(\alpha)$.

Now the paradox resolves itself. For the order type of the ordinals less than Ω is $T^2(\Omega)$ by the preceding discussion, and it then follows that $T^2(\Omega) < \Omega$. This is not altogether appetizing, because it is easy to show that for any ordinals α and β, $T(\alpha) < T(\beta) \leftrightarrow \alpha < \beta$, so $T^2(\alpha) < T^2(\beta) \leftrightarrow \alpha < \beta$, so $\ldots T^6(\Omega) < T^4(\Omega) < T^2(\Omega) < \Omega$. This does not give an infinite descending sequence of ordinals in NFU, because T is not a function and the sequence just revealed cannot be shown to be a set (fortunately!). This does imply that any set model of NFU has ordinals whose elements are not well-orderings from an external standpoint. If our working set theory were NFU, this would not lead to contradiction, because though the domain V of the "model" of NFU in which we would work is a set, the membership relation $\in$ of the "model" is not a set.

Notice that the T operation appears to be an external endomorphism of the ordinals. In our model construction for NFU, we used an external automorphism of our model of set theory which moved ordinals. This similarity is not an accident. The T operation on isomorphism types is closely related to the automorphism j of the underlying model of set theory if we are working in a Boffa model (the same is true of the T operation on cardinals defined above).

The considerations so far give us some assurance that the original system NF does not easily fall prey to the paradoxes. We *know* that NFU does not (unless a quite weak fragment of the usual set theory is inconsistent).

We now turn our attention to the implementation of mathematics in *NF(U)*. We have already described the implementations of natural numbers, ordered pairs, relations, functions, and cardinal and ordinal numbers. Operations on cardinal and ordinal numbers admit quite usual definitions (hereafter in this subsection we assume without comment that the pair is type level). For example, $|A| + |B| = |(A \times \{0\}) \cup (B \times \{1\})|$ and $|A| \cdot |B| = |A \times B|$ are quite familiar-looking definitions of cardinal addition and multiplication. If we define B^A as the collection of functions from A to B, we might want to define $|B|^{|A|}$ as $|B^A|$, but we do not for reasons of relative type. The type of $|B^A|$ is one higher than that of A or B: we correct this by defining $|B|^{|A|}$ as $T^{-1}(|B^A|)$. This is not a total operation: there are cardinals κ such that there is no λ such that $T(\kappa) = \lambda$, such as $\kappa = |V|$. A special case of this definition is $2^{|A|} = T^{-1}(|\{0,1\}^A|) = T^{-1}(|\mathcal{P}(A)|)$ (the last equation uses the one-to-one correspondence between sets and their characteristic functions). The argument for Cantor's theorem above shows that $|A| < 2^{|A|}$ for all A for which $2^{|A|}$ is defined: we have $|\mathcal{P}_1(A)| < |\mathcal{P}(A)|$: applying T^{-1} to both sides of the inequality gives $|A| < T^{-1}(|\mathcal{P}(A)|)$, whenever the latter is defined, and so $|A| < 2^{|A|}$. It is straightforward to establish that the T operation distributes over all these operations and relations in the obvious sense, and so does T^{-1} where it is defined. Very similar considerations apply to operations on ordinal numbers, though all the usual operations on ordinals turn out to be total. One should note that the identification between cardinal numbers and initial ordinals does not hold here, so one needs to make use from time to time of explicit operations $card(\alpha)$ = the cardinality of the field of an element of the ordinal α and $init(\kappa)$ = the smallest order type of a well-ordering of a set of size κ.

All of the mathematical constructions we have done so far can be clarified by looking at how one would do the same mathematics in TST. The appearance of T and T^{-1} is unmysterious there: a set of size $T(\kappa)$ is "the same size" as a set of size κ but appears one type higher. A set of size $T^{-1}(\kappa)$ is a set the same size as a set of size κ but one type lower, and it is not surprising that this does not work for all κ because higher types are larger (being "power sets" of lower types).

Some mathematical constructions are awkward in NFU because they are awkward in TST. A general class of constructions for which this is true is "indexed families of sets". The difficulties are best illustrated by giving an example. We are going to develop the definition of the product of a family of cardinals, $\Pi_{i\in I}\kappa_i$. The product is the size of the generalized cartesian product of a collection of sets A_i each of cardinality κ_i. The generalized cartesian product $\times_{i\in I}A_i$ is the collection of all functions f such that $f(i) \in A_i$ for each $i \in I$. The problem that now arises is how to read the various indexed expressions here. The difficulty is that the sets A are one type lower than the cardinals κ, so if we read A_i as $A(i)$ (letting A be a function from I to sets) then we need to read κ_i as $\kappa(\{i\})$, so κ can be construed as a function from $\mathcal{P}_1(I)$ to cardinals. The set $\times_{i\in I}A_i$ can then be defined explicitly as the set of all functions f such that $f(i) \in \mathcal{P}_1(A(i))$ for each $i \in I$. We cannot have $f(i) \in A_i$ because once again the types are wrong: the type of i is the same as the type of $A(i)$ and so is one higher than the type of an element of A_i: we fix this by having f send each $A(i)$ to the singleton of one of its elements. We want $\Pi_{i\in I}\kappa_i$ to be of the same type as each of the cardinals κ_i. $\times_{i\in I}A_i$ is a set of functions which are each one type higher than elements of I, so it is itself two types higher than an element of I, and its cardinality is three types higher than that of an element of I. We want $\Pi_{i\in I}\kappa_i$ to be of the same type as the $\kappa(i)$'s. A $\kappa_i = \kappa(\{i\})$ is one type higher than that of an element of i. So the full definition of $\Pi_{i\in I}\kappa_i$ is $T^{-2}(|\times_{i\in I} A_i|)$, where $A(i) \in \kappa(\{i\})$ for each $i \in I$ (the existence of such an A and the well-definedness of this cardinal depends on Choice; this leaves open the possibility that products of indexed families of cardinals may be undefined because T^{-2} is partial, and indeed they may). There is no particular mystery to do with NF(U) in this construction: it is a construction in TST imported into NFU. The awkwardness could be reduced in special cases if we had index sets made up of self-singletons (if $i = \{i\}$ for each $i \in I$). One can prove the consistency of the existence of such index sets, but it is also provable that they are all small sets (such an index set would certainly be strongly cantorian, and it is consistent with NFU that every s.c. set is the same size as a set of self-singletons).

There are some mathematical techniques appropriate to NFU which take advantage of the fact that it is a type-free theory and cannot be replicated in TST. These generally require "reasonable" assumptions about the behavior of T operations on cardinals and ordinals which are not provable in NFU (and are in effect strong axioms of infinity).

We defined the notion of *strongly cantorian set* above. "Strongly cantorian" is a notion of smallness (every subset of a strongly cantorian set is strongly cantorian). A variable x restricted to a strongly cantorian set in a set definition does not need to have a type assigned to it for purposes of stratification. The reason is that its type can be freely manipulated: if $x \in A$ and A is s.c., then $x = \bigcup(\iota\lceil A(x)) = (\iota\lceil A)^{-1}(\{x\})$: in one of these expressions the type of x is raised by one and in one it is lowered by one, and these operations can be iterated.

The first special axiom we might consider is Rosser's Axiom of Counting (proposed by Rosser in [Rosser, 1953]), which asserts that the set of natural numbers is strongly cantorian. Rosser's original formulation was the assertion that $\{1, \ldots, n\} \in n$ for each natural number n: the connection of this to our usual counting procedure should be clear and the theorem provable in NF(U) is $\{1, \ldots, n\} \in T^2(n)$ It is clear that Rosser's axiom implies and is implied by the assertion that each natural number n is cantorian ($T(n) = n$). It is only very slightly less obvious that this is equivalent to the assertion that the set of natural numbers is s.c. It is provable in NFU + Infinity that $\mathbb{N}$ is cantorian, and the natural inductively defined bijection from $\mathbb{N}$ to $\mathcal{P}_1(\mathbb{N})$ sends $T(n)$ to $\{n\}$ for each n, so will witness that $\mathbb{N}$ is s.c. if and only if every natural number is cantorian.

The practical effect of the use of Rosser's Axiom of Counting is that one does not need to assign types to variables with natural number values. In fact, the same becomes true of many familiar classes of mathematical objects, as the class of s.c. sets is closed under cartesian product, power set, and the formation of function spaces, so for example the set of real numbers or the set of points in Hilbert space will be s.c. sets.

The metamathematical effect of the Axiom of Counting is surprising. It does not actually add practical strength on the level of arithmetic (even arithmetic of higher order): we simply avoid notational complications. The truly substantial effects are in set theory. NFU + Infinity does not prove the existence of $\beth_\omega$; NFU + Counting proves the existence of $\beth_{init(\beth_n)}$ for each natural number n. NFU + Counting has models $\beth_{init(\beth_\omega)}$ is needed in the construction) and NF + Counting is known to prove the consistency of NF (such independence results are rare for

NF).

The strong principle of mathematical induction (induction for unstratified as well as stratified formulas) is stronger than the Axiom of Counting (which it easily implies). Its precise strength is not known, but it is weaker than ZFC, as an ω-model of NFU can be constructed on fairly weak assumptions (existence of $\beth_{init(\beth_{\omega_1})}$ is enough). Again, its effects are not seen in higher-order arithmetic (which is adequately handled by stratified induction) but in set theory.

An assumption considered by C. Ward Henson in [Henson, 1973] (in a slightly more restricted form) is "every cantorian set is strongly cantorian", which we call CS (short for "the Axiom of Cantorian Sets"). Notice that Infinity and CS together imply Counting. The strength of NFU + Infinity + CS is rather shocking: it is equiconsistent with ZFC + "there is an n-Mahlo cardinal" for each concrete natural number n (see [Holmes, 2001]). Rieger-Bernays permutation methods can be used to convert a model of this theory to one in which the hereditarily strongly cantorian sets make up a model of ZFC with an n-Mahlo cardinal for each concrete n.

There are still stronger assumptions which have been investigated and shown to correspond in strength to further strong axioms of infinity, but this is enough to show the pattern: natural axioms formulated in NFU which regularize the behavior of strongly cantorian sets have surprisingly strong effects on the consistency strength of the extended version of NFU which results. We give one more example: the Axiom of Small Ordinals (introduced in [Holmes, 1998] and [Holmes, 2001]) in its weakest form asserts that any definable class of strongly cantorian ordinals is the intersection of the class of strongly cantorian ordinals (which is not a set) with some set. NFU + Infinity + Small Ordinals has the same strength as Kelley-Morse set theory + "the proper class ordinal is weakly compact" or ZFC - Power Set + "there is a weakly compact cardinal" (Solovay showed this in [Solovay, preprint]).

5.3 Peculiarities of NF

We now address the oddities of New Foundations itself. These are twofold. On the negative side, no consistency proof for NF is known, nor is there any proof that NF is any stronger than TST with the Axiom of Infinity (or equivalently Mac Lane set theory with Infinity) which is somewhat weaker than Zermelo set theory. Any suspicion that NF must be strong because it allows "big" sets such as the universe should be dispelled by the observation that NFU, which proves the existence of the same big sets for the same reasons, is weaker than Peano arithmetic, and NFU + Infinity (which is a more reasonable set theory) is exactly as strong as TST + Infinity or Mac Lane set theory with Infinity. There is a useful distinction to be drawn between "big" objects such as the universe V or the ordinal Ω and "large" objects such as inaccessible or measurable cardinals.

On the positive side, NF is known to prove the Axiom of Infinity and disprove the Axiom of Choice [Specker, 1953]. This means that NFU + Choice (which is known to be consistent) proves the existence of urelements! Specker gave separate proofs of these two results, but of course the proof that Choice is false also proves

Infinity: if the universe were finite, it could be well-ordered, which would allow the proof of Choice in a familiar way (it is worth noting here that all the standard equivalences between forms of Choice hold in NF(U), as long as one is careful to state them in stratified forms).

We present Specker's proof that the Axiom of Choice is false.

Theorem: (NF, due to Specker) $\neg AC$

Proof: Suppose otherwise. Then the natural order on the cardinal numbers would be a well-ordering (this is equivalent to AC for quite standard reasons).

We know from above that the exponential map $\kappa \mapsto 2^\kappa$ is partial. We define for each cardinal κ a set called the exp-closure of κ (this is a nonce notion): the exp-closure of κ is the smallest set which contains κ and contains 2^λ whenever it contains λ and 2^λ exists. We define the set SM (German *Speckermenge*, "the Specker set", a coinage of Thomas Forster) as the set of all cardinals whose exp-closure is finite. This set is nonempty: for example, $|V|$ belongs to SM.

Now comes a move which is peculiar to NF. We have $2^{|\mathcal{P}_1(V)|} = |\mathcal{P}(V)|$ by the definition of exponentiation. But this means that $2^{T(|V|)} = |V|$ (this requires $|\mathcal{P}(V)| = |V|$, which is a consequence of strong extensionality). Properties of the T operation then tell us that $2^{T^{n+1}(|V|)} = T^n(|V|)$ for each concrete natural number n, so all the cardinals $T^n(|V|)$ (which do not make up any kind of set sequence!) are in SM.

We argue that for any cardinal κ in SM, we also have $T^n\kappa \in SM$ and $T^{-1}(\kappa) \in SM$, if the latter cardinal exists. Suppose $\kappa \in SM$. We give the name exp to the map $\kappa \mapsto 2^\kappa$. If λ is an element of the exp-closure of κ (so of the form $exp^n(\kappa)$ for some natural number n) then $T(\lambda)$ is in the exp-closure of $T(\kappa)$, because it is equal to $exp^{T(n)}(T(\kappa))$ (this proof is considerably simpler if the Axiom of Counting is assumed, which would give $T(n) = n$!). Now consider the largest element ν of the exp-closure of κ. It must be greater than $T(|V|)$, or else its image under exp would be defined (notice that linear ordering of cardinals by the natural order is used here). $exp(T(\nu))$ will be defined, since $T(\nu) \leq T(|V|)$: since $\nu > T(|V|)$, we have $T(\nu) > T^2(|V|)$ and so $exp(T(\nu)) \geq exp(T^2(|V|)) = T(|V|)$. From this it follows that $exp^2(T(\nu))$ either fails to exist or is equal to $|V|$. So if the cardinality of the exp-closure of κ is n, the cardinality of the exp-closure of $T(\kappa)$ will be either $T(n)+1$ or $T(n)+2$, and in any case finite, so $T(\kappa) \in SM$. Now suppose that $\kappa = T(\lambda)$, and that $\lambda \notin SM$. This would imply that $exp^n(\lambda)$ was defined for every n, and it would follow that $T(exp^n(\lambda)) = exp^{T(n)}(\kappa)$ was defined for every n, and though the T operation might fail to be the identity on $\mathbb{N}$ it is onto: it follows that $\kappa \notin SM$ contrary to hypothesis. So if $\kappa \in SM$ and $T^{-1}(\kappa)$ exists, it follows that $T^{-1}(\kappa) \in SM$.

Because SM is a set of cardinals, it has a smallest element μ. The results above imply that $T(\mu) = \mu$. For otherwise we would have to have $T(\mu) > \mu$, whence $T^{-1}(\mu)$ would exist, belong to SM, and be less than μ, contrary to choice of μ (because the T operation preserves order on cardinals and its domain of definition is downward closed). Now let n be the finite number of elements in the exp-closure of μ. We see from the previous paragraph that the number of elements in the exp-closure of $T(\mu)$ must be $T(n)+1$ or $T(n)+2$, and of course this must be the same number. But $n = T(n) \mathtt{~mod~} 3$ is easy to show, so $n = T(n)+1$ or $n = T(n)+2$ is impossible.

As we have already remarked, there is a nicer proof of this (in detail; the basic idea is the same) if the Axiom of Counting is assumed, found in [Holmes, 1998] (mod 3 arithmetic does not come into play), but we will shortly indicate reasons why we do not regard the assumption of Counting as harmless.

For any cardinal κ, we define the *Specker tree* of κ as the smallest set which contains κ and includes the preimage under exp of each of its elements. This is the set of all cardinals λ such that $exp^n(\lambda) = \kappa$ for some $n \in \mathbb{N}$. If the Axiom of Choice is assumed, it is straightforward to show that the Specker tree of any cardinal has finite depth (there is an n such that for no λ is $exp^n(\lambda) = \kappa$). It is a theorem of ZF that the Specker tree of any cardinal is well-founded, due to Forster in [Forster, 1976] (using Sierpinski's result that $\aleph(\kappa) < exp^3(\kappa)$, where $\aleph$ is the Hartogs aleph function: $\aleph(\kappa)$ is the first ordinal which is not the order type of a well-ordering of a subset of a set of size κ [this definition would need to be modified slightly to be stratified]). Thus in ZF the Specker tree of a cardinal must have an ordinal rank in an obvious sense, which if AC is assumed must be finite. It is an open question in ZF whether it is possible for a cardinal to have a Specker tree of infinite rank (and so have inverse images under exp of every index); it is a theorem of NF + Counting that the Specker tree of the cardinality of the universe has infinite rank: so here we have a combinatorial situation arising in what should not be a terribly strong extension of NF whose possibility has not been established in the usual Choice-free mathematics.

The only obvious mathematical advantage of NF over NFU which we see is that the Quine pair is definable in NF. There may be other elegant features of mathematical development in NF which follow from the absence of urelements. The use of Choice in mathematics is pervasive, which is a strong apparent *disadvantage* of NF over NFU. Now it might be said that the form of Choice most often used in classical mathematics is Dependent Choices, and so far as anyone knows DC is not inconsistent with NF. Unfortunately, no one has any idea how to provide a relative consistency proof for DC relative to NF.

The only method for relative consistency and independence proofs which has been used extensively with NF is the Rieger-Bernays permutation method, which can prove various amusing results, but is limited by the fact that stratified sentences (i.e, anything which makes sense in TST) are invariant in permutation models. Orey showed that NF + Counting is essentially stronger than NF using metamathematical techniques. It is possible to emulate work done for NFU to show that NF + Infinity + CS proves the relative consistency of the existence of n-Mahlos, but we do not know that NF + Counting does not prove the existence of

n-Mahlos. Similar considerations apply to other strong extensions of NFU whose consistency strength is known. Recently, one of the authors has shown that forcing arguments can be carried out in NF, but without some Choice forcing is not useful. A typical result which can be shown is that if NF + DC is consistent then NF + DC + "the continuum is well-ordered" is consistent.

The appeal that NF has over NFU for some is that it is a theory of pure sets. In NF (as in ZF) everything without exception is a set. This appeal should be resisted or at least viewed with care. Notice that in ZF the idea that something is a pure set can be expressed and used to define subsets using the Separation Axiom. In NFU, the predicate "is a pure set" has no stratified definition, and so cannot be used in a set definition (if there were such a definition, then the collection of pure sets in any model of NFU would be a model of NF, and this is not the case in known models). These considerations are intimately related to the fact that there are easy ways to interpret ZFA in ZF and *vice versa*, whereas all efforts to interpret NF in NFU have failed.

5.4 Extensional fragments of New Foundations

NFU, which is known to be consistent, differs from NF "only" in that extensionality is weakened to allow urelements; it has the same comprehension scheme. There are other fragments of NF which are known to be consistent which include the strong Axiom of Extensionality and perforce involve limitations on Stratified Comprehension. We describe two of them.

In the same year (1969) that Jensen showed the consistency of NFU, Grishin [Grishin, 1969] showed the consistency of NF_3, the fragment of New Foundations whose axioms are the strong Extensionality Axiom and all those comprehension axioms which admit a stratification whose range has no more than three elements: that is, all those axioms which can be typed in TST using types 0,1,2. He also showed that NF_4 is the same theory as NF, and that NF_3 + "$\{\{\{x\}, y\} \mid x \in y\}$ exists" is the same theory as NF.

To put this result in context, we cite a model-theoretic result of Specker. He showed that NF is equiconsistent with TST plus the scheme "$\phi \leftrightarrow \phi^+$" (for each formula ϕ). This scheme is more briefly called Amb (the ambiguity scheme). This result is a verification of Quine's motivation for *NF*. The technique of his proof generalizes to relate fragments of NF to fragments of TST. For example, NFU is equiconsistent with TSTU, the version of TST in which urelements are allowed in each positive type. Grishin's system NF_3 is equiconsistent with TST_3 + Amb, where TST_3 is the system of type theory with only three types (0,1,2), and the notation ϕ^+ of course only applies to formulas not mentioning type 2. It turns out that *all* infinite models of TST_3 satisfy Amb, so the consistency of NF_3 and indeed of powerful extensions of NF_3 is easy to establish, and it appears that NF_3 is a much more natural theory than NF. On the other hand, NF_3 is not an environment in which it is easy to do mathematics (essentially because TST_3 is rather restrictive). Henrard, in unpublished work, has shown that NF_3 does admit a theory of cardinal number. Pabion has shown that NF_3 + Infinity is precisely as strong as second-order arithmetic.

In [Crabbé, 1982], 1983, Marcel Crabbé defined a system NFP (for "predicative NF") and a stronger system NFI (which Holmes has called "mildly impredicative NF). NFI consists of the strong Axiom of Extensionality and those axioms $\{x \mid \phi\}$ exists" which admit a stratification in which the type assigned to $\{x \mid \phi\}$ (one higher than the type assigned to x) is the highest type in the range. NFP further restricts to those comprehension axioms in which no variable of the type of $\{x \mid \phi\}$ is bound in ϕ. Crabbé showed that NFI (and so of course NFP) is consistent. NFI is exactly as strong as second order arithmetic and NFP is weaker than Peano arithmetic.

A striking feature of NFP is that one can make use of the Specker argument for the failure of Choice to prove Infinity. All the known consistent fragments (including NFP and NFI are consistent with Choice. The argument goes like this. Use $\iota(x)$ as a nonce notation for $\{x\}$. For any formula ϕ, for sufficiently large n, "$\{\iota^n(x) \mid \phi\}$ exists" will be a comprehension axiom of NFP (make n large enough, and the variable representing $\iota^n(x)$ will have type higher than any of the variables in ϕ). From this it follows that NFP + Union = NF. Argue in NFP that Infinity holds as follows: if Union holds then Infinity follows by Specker's argument in NF; if Union does not hold, then Infinity holds because finite sets have unions. There is of course a bit of detail to this argument hidden in this sketch. NFP has interesting relationships to weak subsystems of arithmetic, and Holmes has shown that it is precisely equiconsistent with the ramified theory of types of Whitehead and Russell's *Principia Mathematica* without the Axiom of Reducibility (and it is much simpler!) NFI admits just enough impredicativity to implement impredicative arithmetic, and is as we observed above precisely as strong as second-order arithmetic.

5.5 Reflections on New Foundations and ZF

Perhaps the single most important point to make about the Quine systems is that—properly understood—they do not contradict ZF at all. This point is in any case important, being both fundamental and correct, but it is important also in the sense that there is a continuing and pressing need to make it. Let us explain.

ZF and ZF-like theories are the endeavour to axiomatise our understanding of the cumulative hierarchy of well-founded sets. NF casts its net a little wider, in that it is an attempt to not only do that but to reason about some other sets as well. It is possible to define a predicate "is a well-founded set" in NF, and this predicate[3] enables us to reason about well-founded sets in NF. *No known NF theorem about well-founded sets contradicts anything provable in ZF or any large cardinal upgrade thereof.* Models of strong extensions of NFU can be constructed

[3] which of course is unstratified; were it stratified, its extension would be a set and we would have a paradox.

in which the well-founded sets make up a model of (a strong extension of) ZFC.[4] If you are interested in facts about well-founded sets then you will find nothing in NF to offend you. If you wish to be offended by NF then you will have to believe in the proposition that all sets are well-founded. Of course a lot of people do believe this proposition, but one of the most striking things about the set-theoretic literature is the complete absence of any arguments for it. Not only does one find any arguments, one is hard put to find even an acknowledgement that one might come in handy. The closest thing I can find is Boolos [Boolos, 1971] "There does not seem to be any argument that is guaranteed to persuade someone who really does not see the peculiarity of a set's belonging to itself, or to one of its members *etc.*, that these states of affairs are peculiar". NF's peculiar charm is that it attempts to axiomatise the behaviour of large collections while taking care not to say anything that will contradict what we know or believe about well-founded sets. It's the lover that the spouse can't object to. That being so, one would expect NF to be the toast of Society. But there are several respects in which NF gets what in popular parlance would be called a *bum rap*. Admittedly NF does not prove mathematical induction for unstratified properties of natural numbers, and it disappoints in other ways by somehow neglecting to establish that the sets that are conceptually small have all the nice properties that one might expect them to have. But this shows merely that we haven't yet determined what axioms should be added to NF; that's our fault not NF's[5] I think the most illuminating parallel for NF in this context is with Zermelo set theory, Z. Z is clearly a theory about the cumulative hierarchy (or at least it has been so interpreted in retrospect)—and although no one is suggesting that it makes allegations about the cumulative hierarchy that are *false*—it does stand accused of not achieving absolutely everything that was asked of it. NF is an attempt to axiomatise not only the cumulative hierarchy but some sets beyond it. One shouldn't expect one's first attempt in this to be successful any more than one would expect that Z should be not only our first word on the axiomatisation of the cumulative hierarchy but our last as well.[6] Clearly we are going to have to add axioms to NF to ensure that whenever $ZF \vdash \phi$ then $NF \vdash \phi^{WF}$ (where ϕ^{WF} is the relativisation of ϕ to the well-founded sets); the fact that this task has not been completed is no cause to belittle the NF project: it simply means that there are axioms that remain to be found. One could of course simply add this aforementioned principle by main force as an axiom scheme, but the idiomatic and gentler way to achieve this result would be to add judiciously designed axioms concerning big sets. In NFU the Axiom of Cantorian Sets ensures that this is true in a certain permutation model. In [Forster, 2006] it is shown how

[4]It is true that NF refutes the Axiom of Choice, but the only known failures of AC involve the sets that in NF studies we call **big** (following a suggestion of Thomas Forster) as opposed to *large*, as in *large* cardinals in ZF. These are the collections like the universal set, and the set of all cardinals and the set of all ordinals: collections denoted by expressions which in ZF-like theories will pick out proper classes. Of course in NFU there is no such problem.

[5]for NFU, the question of what axioms to add has been extensively investigated, and these are probably similar to what would be added to NF

[6]Why is it that I always find my spectacles in the *last* place I look for them not the *first*?

axioms postulating good behaviour on the part of the T function systematically correspond to axioms giving the existence of well-founded sets of high rank. Clearly this is the way in which we have to go. The reason why nobody has explored this in the detail which will eventually be needed is that the whole project is stalled—and will remain stalled until nerves around the consistency question for NF have been calmed: adding new axioms to a theory not yet known to be consistent invites a wealth of disparaging metaphors—*running before one can walk* or *building castles in the air* ... (of course such remarks do not apply to NFU). However, any reason for disparaging the project-to-find-sensible-extensions-for-NF is a reason for trying to prove NF inconsistent[7] and *that* project seems to have no takers.

In coming to grips with NF the hard part is to fully understand stratification, to know how to recognise a stratified formula when you see one. There is an easy rule of thumb with formulas that are in primitive notation, for one can just ask oneself whether the formula could become a wff of type theory by adding type indices. It's harder when one has formulas no longer in primitive notation, and the reader encounters these difficulties very early on, since even something as basic as the ordered pair is not a set-theoretic notion. How does one determine whether or not a formula is stratified when it contains subformulas like $y = f(x)$? The technical/notational difficulty here lands on top of—as so often—a conceptual difficulty. The answer is that of course one has to fix an implementation of ordered pair and stick to it. Does that mean that—for formulas involving ordered pairs—whether or not the given formula is stratified depends on how one implements ordered pairs? The answer is 'yes' but the situation is not as grave as this suggests. Let us consider again the formula $y = f(x)$. This is of course a molecular formula, and how we stratify it will depend on what formula it turns out to be in primitive notation once we have settled on an implementation of ordered pairs. If we use Kuratowski ordered pairs then the formula we abbreviate to $y = f(x)$ is stratified with x and y having the same type, and that type is three types lower than the type of f. If we use Quine ordered pairs then the formula we abbreviate to $y = f(x)$ is stratified with x and y having the same type, and that type is one type lower than the type of f. There are yet other implementations of ordered pair under which the formula we abbreviate to $y = f(x)$ is stratified with x and y having the same type, and that type is one or more types lower than the type of f.

The point is that our choice among the possible implementations will affect the difference in level between x (and y) and f but will not change the formula from a stratified one to an unstratified one. This is subject to two important provisos:

1. We restrict ourselves to ordered pair implementations that ensure that in $x = \langle y, z\rangle$ y and z are given the same type.

2. We do not admit self-application: $(f(f))$.[8]

[7]Thus Forster, but Holmes comments that some (not he!) will raise an objection of the form "NF does not satisfy Choice and so is irrelevant to mathematics as she is Done", which does not necessarily involve doubts about the *consistency* of NF.

[8]There is a further important distinction between type differential 0 and all other cases: the

Thes two provisos are of course related. The second will seem reasonable to anyone who thinks that mathematics is strongly typed. (The typing system in NF interacts quite well with the endogenous strong typing system of mathematics; this is a striking and deep fact that has not so far attracted the attention it should). If we consider expressions like $x = \langle x, y\rangle$ we see that their truth-value depends on how we implement ordered pairs. There is a noncontroversial sense (entirely transparent in the theoretical CS tradition) in which expressions of this kind are not part of mathematics—in contrast to expressions like $x = \langle y, z\rangle$ which are. The only formulas whose stratification status are implementation-sensitive in this way are formulas that are not in this sense part of mathematics.

The second one is a bit harder to understand: why should we not have an implementation that compels y and z to be given different types in a stratification of $x = \langle y, z\rangle$—or even make the whole formula unstratified?

If we make $x = \langle y, z\rangle$ into something unstratified then we cannot be sure that $X \times Y$ exists, nor that compositions of relations (that are sets) are sets; converses of relations might fail to exist; and we will not really be able to do any mathematics. After all, $X \times Y$ is $\{z \mid (\exists x \in X)(\exists y \in Y)(z = \langle x, y\rangle)\}$ and if $z = \langle x, y\rangle$ is not stratified then the set abstraction expression might not denote a set.

However, even if we muck things up only to the extent of allowing $x = \langle y, z\rangle$ to be stratified with y and z of different types then we will find not only that some compositions of relations (that are sets) are not sets but also that, for some big sets X, the identity relation 1_X is not a set. For example 1_V would not be a set. Let's look into this last point a bit more closely. Suppse "$x = \langle y, z\rangle$" is stratified but with y and z being given different types. Then $X \times Y$ is $\{z \mid (\exists x \in X)(\exists y \in Y)(z = \langle x, y\rangle)\}$ which this time is stratified, so $X \times Y$ is a set. However if $R \subseteq X \times Y$ and $S \subseteq Y \times Z$ then $R \circ S$ is

$$\{w \mid (\exists x \in X)(\exists y \in Y)(\exists z \in Z)(\langle x, y\rangle \in R \wedge \langle y, z\rangle \in S \wedge w = \langle x, z\rangle)\}.$$

This is not stratified. If the difference between the types of the two components of an ordered pair is n, then x and y have types differing by n, and y and z too have types differing by n, and x and z have types differing by $2n$, so although we can stratify $\langle x, y\rangle \in R \wedge \langle y, z\rangle \in S$ we will not be able to stratify $\langle x, y\rangle \in R \wedge \langle y, z\rangle \in S \wedge w = \langle x, z\rangle$.

The problem with 1_X arises because $(\exists x \in X)(y = \langle x, x\rangle)$ is not stratified, so its extension is not certain to be a set. By the same token no permutation of a set can be relied upon to be a set. The (graph of the) relation of equipollence might fail to be reflexive, or symmetrical, or transitive.

The conclusion is that if we want our implementation of mathematical concepts into set theory to be tractable from the NF point of view, and to deliver the

projection maps are set functions, which allows further constructions which could be construed as subverting the type security of the pair, though they are mathematically convenient. For example, in NFU the existence of a type-level pair implies Infinity. In either theory, the definition of cardinal multiplication is much simpler with a type level pair, but the definition clearly involves a subversion of the security of the cartesian product type.

routine banalities about relational algebra that we take for granted (existence of 1_X, existence of compositions of relations, existence of transitive closures of relations whose graphs are sets, and so on) then we want a pairing/unpairing function that interprets $x = \langle y, z\rangle$ as a stratified formula with y and z having the same type. One such ordered pair is the Kuratowski ordered pair that we all know and love. As it happens, in NF we usually use the Quine ordered pair which we defined in section 6.2.3.

Does the difference between Quine pairs and Kuratowski pairs matter? Much less than you might think. In some deep sense it doesn't matter at all.[9] Both of them make the formula "$\langle x, y\rangle = z$" stratified and give the variables x and y the same type; z takes a higher type in most cases (never lower). *How* much higher depends on the version of ordered pair being used, but there are very few formulas that come out stratified on one version of ordered pair but unstratified on another, and they are all pathological in ways reminscent of the paradoxes. The best way to illustrate this is by considering ordinals in NF. (In NF we implement ordinals as isomorphism classes of well-orderings). For any ordinal α the order type of the set (and it is a set) of the ordinals below α is well-ordered. In ZF one can prove that the well-ordering of the ordinals below α is of length α. In NF one cannot prove this equation for arbitrary α since the formula in the set abstract whose extension is the graph of the requisite isomorphism is not stratified for any implementation of ordered pair. Now any well-ordering R of a set A to length α gives rise to a well-ordering of $\{\{\{a\}\} \mid a \in A\}$, and if instead one tries to prove (in NF) that the ordinals below α are isomorphic to the well-ordering of length α decorated with curly brackets, one finds that the very assertion that there is an isomorphism between these two well-orderings comes out stratified or unstratified depending on one's choice of implementation of ordered pair! This is because, in some sense, the applications of the pairing function are *two* deep in well-ordering of the ordinals below α, but only *one* deep in the well-ordering of the set of double singletons. If we use Quine ordered pairs, the assertion is stratified—and indeed provable. If one uses Kuratowski ordered pairs (or Wiener ordered pairs) then the assertion is unstratified and refutable. However if one uses Kuratowki ordered pairs there is instead the assertion that the ordinals below α are isomorphic to the obvious corresponding well-ordering of $\{\{\{\{\{a\}\}\}\} \mid a \in A\}$, which comes out stratified (and provable). In general for each implementation of ordered pair there is a depth of nesting of curly brackets which will make a version of this equality come out stratified and true. This does not work with deviant implementations of ordered pair under which "$\langle x, y\rangle = z$" is unstratified or even with those which are stratified but give the variables x and y different types.

Let us try to prove Cantor's theorem. The key step in showing there is no surjection $f : X \twoheadrightarrow \mathcal{P}(X)$ by *reductio ad absurdum* is the construction of the diagonal set $\{x \in X \mid x \notin f(x)\}$. The proof relies on this object being a set, which it will be if "$x \in X \wedge x \notin f(x) \wedge f : X \to \mathcal{P}(X)$" is stratified. This in turn depends on "$(\exists y)(y \in \mathcal{P}(X) \wedge \langle y, x\rangle \in f \wedge f : X \to \mathcal{P}(X))$" being stratified. And it *isn't*

[9] as long as one does not care whether the projection maps are set functions.

stratified, because "$\langle y, x\rangle \in f$" compels '$x$' and '$y$' to be given the same type, while "$f : X \to \mathcal{P}(X)$" will compel 'y' to be given a type one higher than 'x'. This is because we have subformulas '$x \in X$' and '$y \subseteq x$'. Notice that we can draw this melancholy conclusion without knowing whether the type of 'f' is one higher than that type of its argument, or two, or three ....

However if we try instead to prove that $\{\{x\} \mid x \in X\}$ is not the same size as $\mathcal{P}(X)$ (the form of Cantor's theorem appropriate in TST) we find that the diagonal set is defined by a stratified condition and exists, so the proof succeeds. This tells us that we cannot prove that $|X| = |\{\{x\} \mid x \in X\}|$ for arbitrary X: graphs of restrictions of the singleton function tend not to exist. (If they did, we would be able to prove Cantor's theorem in full generality). This gives rise to an endomorphism T on cardinals, where $T(|X|) =: |\{\{x\} \mid x \in X\}|$ which we have seen earlier. T misbehaves in connection with *big* sets. If $|X| = |\{\{x\} \mid x \in X\}|$ we say that X is **cantorian**. If the singleton function restricted to X exists, we say that X is **strongly cantorian**. Sets whose sizes are concrete natural numbers are strongly cantorian. $\mathbb{N}$ (the set of Frege natural numbers) is cantorian, but the assertion that it is strongly cantorian (which is the same as the assertion that every inductively finite set is cantorian) implies the consistency of NF.

NF has various natural notions of 'small'. *Cantorian* and *strongly cantorian* are two obvious examples. *Strongly Cantorian* turns out to be mathematically more natural. It has nicer closure properties and does more to ensure nice behaviour on the part of its bearers than *cantorian* does. Another natural notion is *well-founded*. We have recently established [Bowler and Forster, 2009] that every well-founded set is smaller than $T^n(|V|)$ for every concrete n, so there are no *big* well-founded sets. *Well-Ordered* sounds like a notion of smallness too, but there is a theorem of Hartogs' that says that there are in some sense arbitrarily large well-ordered sets. Hartogs' theorem in the exact form in which it is usually stated in ZF is not provable in NF but the form of Hartogs's theorem appropriate to TST is provable. There is no analogue for well-ordered sets of the result alluded to above for well-founded sets: there are well-ordered sets whose size is $\not\leq T(|V|)$.

INDUCTIVE DEFINITIONS IN NF

Mathematics is replete with "inductively" defined sets—sets defined by a recursion. Any set theory that claims to provide a foundation for mathematics must prove that these sets exist, and the more transparent the proof is the happier we will all be. The paradigmatic inductively defined set is $\mathbb{N}$, the set of natural numbers. It is the $\subseteq$-least set containing 0 and closed under successor. That is to say it is the intersection of all sets that contain 0 and are closed under successor. When can we prove the existence of intersections? Well (as long as Y is nonempty) $\bigcap Y$ is a set because it is a subset of $\bigcup Y$—and *that* will be a set as long as Y is, because we have the Axiom of Union. That is the situation in ZF. However we should steel ourselves for trouble because *prima facie* there is no reason to suppose that the collection of all sets-containing-0-and-closed-under-successor is a set. However if there is even one set X that contains 0 and is closed under successor then we know that the set $\mathbb{N}$ that we are trying to define will be a subset of X. If Y is a set-containing-0-and-closed-under-successor then $X \cap Y$ will be another such. In these circumstances the intersection of *all* sets-containing-0-and-closed-under-successor is going to be the same as the intersection of *all-subsets-of-X*-containing-0-and-closed-under-successor. This last object is certainly going to be a set because it is a subset of X.

The usual formulation of the Axiom of Infinity in ZF is geared precisely to the problem of proving that $\mathbb{N}$—indeed a particular implementation of $\mathbb{N}$—exists. It says that there is a set which contains $\emptyset$ and is closed under $\lambda x.x \cup \{x\}$. In the von Neumann implementation of arithmetic $\emptyset$ is the implementation of 0 and $\lambda x.x \cup \{x\}$ is the implementation of successor. Thus it hands us on a plate an X containing-0-and-closed-under-successor.

The reader may feel that this is rather *ad hoc*. It is, and in two ways. For one thing this form of the Axiom of Infinity is unnecessarily specific and informative. It would be equally satisfactory to take the axiom in a form that says merely that there is an infinite set (one the same size as a proper subset of itself). It is *ad hoc* also in the sense that it solves only one case of the problem of proving the existence/sethood of inductively defined collections. The general problem remains unsolved: in ZF we cannot in general define inductively defined sets "top-down" as the intersection of a suitably closed family of sets; we cannot rely on there automatically being a set that contains the founders and is closed under the operations in question.[10]

The predicament can be illustrated by considering two examples of inductively defined families. In any topological space, the family of Borel sets is the smallest collection of sets containing all open and closed sets and closed under countable unions and complementation. In this case we have a set—namely the power set of the space in question—that contains all open and closed sets and is closed under complementation and countable union, so we can obtain the family of Borel sets as a subset of it. In contrast, if we try to prove the sethood of the collection of hereditarily countable sets (which is the $\subseteq$-least set containing all its countable subsets) there is no obvious natural set, visible in the light of day, which contains all its countable subsets.[11] What can we do? We can approach the desired collection "from below" by starting with the set of all the founder objects (which in the case of HC is empty) and iteratively closing under the operations of interest. In the case of HC this means adding at each stage—to the stage-in-hand—all countable subsets of that stage. Then we hope to reach a fixed point. In fact we do reach one—and after fewer than ω_2 steps—but establishing this by purely combinatorial means requires considerable ingenuity.

[10] More accurately: we cannot do it as easily. Replacement is a great help but the process is complex and a known problem for beginners. In Zermelo set theory it is in general impossible, even with foundation or the Axiom of Rank.

[11] It is true that V_{ω_1} can be proved to contain all its countable subsets if we use AC_ω but this is a red herring, since the existence of HC can be proved without using AC_ω.

Typically these iterative "from below" constructions are done by iteration along the ordinals. It doesn't much matter how we implement ordinals, and in principle any sufficiently long well-ordering will do. (This point is not generally appreciated. The fact that von Neumann ordinals from $\omega \cdot 2$ onwards cannot be proved to exist in Zermelo doesn't mean that we cannot in Zermelo prove termination of iterative processes of arbitrary countable length: we most certainly can). There's the rub: how do we know that there always *is* a sufficiently long well-ordering? That's where Hartogs' theorem comes in. Suppose our desired object D can be obtained as the union of a well-ordered collection of stages that we construct by recursion over a (any) well-ordering. Hartogs' theorem says that for any set X there is a well-ordering too big to be embeddable into X. Thus if our desired object D really does exist, then a well-ordering too big to be embeddable into D will be a well-ordering long enough for us to construct the sequence of stages by recursion over it. Thus Hartogs' theorem tells us that if a recursive definition crashes, it won't be because we have run out of ordinals.

Of course we should not expect to be able to prove the sethood of arbitrary inductively defined collections: some such collections are paradoxical. But what one would like is to have smooth proofs of the existence/sethood of the nonparadoxical collections, and this we do not really have. (A good illustration of this is the difficulty—alluded to above—that we have in proving that the collection of hereditarily countable sets is a set).

In NF the existence of big sets restores the possibility of direct top-down definitions of inductively defined sets: any inductively defined set that can be defined at all can be given a direct "top-down" definition. (This is for the gratifyingly simple reason that—whatever your founders and operations—the universal set contains all founders and is closed under all operations, so when we take the intersection of the set of all sets containing the founders and closed under the operations we are not taking the intersection of the empty set). Thus we obtain the effect of Hartogs' theorem without actually having the theorem itself (as it is usually phrased in ZF).

That is not to say that every inductively defined collection is a set. An inductively defined collection (least thing containing this and closed under that) will be a set as long as the property of being-a-set-that-contains-this-and-is-closed-under-that is stratified. If that is a stratified property then the collection of all sets that contain-this-and-are-closed-under-that is a set and its intersection (which is the inductively defined collection we want) will be a set. This proviso is important, because we need somehow to block the existence of arbitrary inductively defined collections. The trouble is that this stratification proviso seems to block too much. The property of containing-all-your-countable-subsets is not stratified, so we cannot use NF axioms to prove the sethood of the intersection of all sets with it. The existence of HC (the set of well-founded hereditarily countable sets) is not provable in NF if NF is consistent. One should emphasise here that NF does not appear to have anything like a proof that HC cannot be a set. There is a suite of sensible-looking axioms that one can add to NF to get the well-founded sets to exhibit all the behaviour that ZF says they should. These axioms take the form of saying that the collections of cantorian ("nice") cardinals and ordinals are closed under certain natural operations.

The stratification constraint means that we have to be very careful when implementing these operations. An important example is cardinal exponentiation. Since (as we have seen) A and $\{\{y\} \mid y \in A\}$ are not reliably the same size, the two definitions of 2^α—(i) as $|\mathcal{P}(A)|$ where $|A| = \alpha$ or (ii) as $|\mathcal{P}(A)|$ where $|\{\{y\} \mid y \in A\}| = \alpha$—are not equivalent as they are in ZF. It makes a difference which one we choose. If we choose the second it turns out (easy to check) that the property of containing a given cardinal α and being closed under exponentiation is stratified. This means that the collection $\{\alpha, 2^\alpha, 2^{2^\alpha} \ldots\}$ is a set for all α. Were we to use the other definition it wouldn't be.

However, although such inductive constructions as can be executed at all can be executed in the direct top-down fashion, it is still possible to import ordinals into a description of this activity. Suppose our inductive construction starts from a set X with a stratified definition (so it is $\{x \mid \phi\}$ for some stratified formula ϕ with one free variable) and we want to obtain the least superset of X closed under some infinitary homogeneous operation. Examples would be: union of countable subsets; or $F(X) =: \{y \mid (\exists f : y \twoheadrightarrow X)($ is countable-to-one$\}$. F is the operation taking us from each stage to the next stage. The collection of F-stages is the least set containing X, and closed under F and unions of chains. It is of course a set, and it is—for the usual reasons—well-ordered by $\subseteq$. Therefore one can associate an ordinal with every F-stage. (As usual there are several ways of doing it: (i) the set of stages and the set of ordinals are alike wordered so there is a canonical map between them; (ii) each stage bounds an initial segment which has an ordinal for its length. (ii) is guaranteed to work even though (i) isn't). Notice that in this treatment we do not use Hartogs' theorem.

Now we are in a position to find an echo of the ZF way of doing things. The closure ordinal α is in a weak sense well-behaved. Let f be the map that sends the ordinal α to the αth stage in the construction. f has a stratified definition without parameters, so the expression

$$(\forall\alpha)(\forall\beta)(f(\alpha) = f(\beta) \longleftrightarrow f(T(\alpha)) = f(T(\beta)))$$

is stratified (fully stratified: it has no parameters and only the one free variable) and so it can be proved by induction on ordinals. This means that if α is the closure ordinal (that is to say, the least β such that $f(\beta) = f(\beta + 1)$) then so is $T(\alpha)$.

It would close the circle very nicely if we knew that every closure ordinal of a stratified recursion were strongly cantorian, but I see no proof. Perhaps it's a very strong assumption.

Armed as we now are with a clearer understanding of recursive definitions in NF, we can return to an earlier topic. There does seem to be a faint chance that the big sets could yet tell us something about well-founded sets. If we define exponentiation of cardinals in NF so that 2^α is $|\mathcal{P}(A)|$ where α is $|\{\{x\} : x \in A\}|$—

as we have in fact resolved to do—we find that for some cardinals α the cardinal 2^α cannot be defined. A brief look at the definition will reassure us that if α is a cardinal of the form $|\{\{x\} \mid x \in A\}|$ then 2^α is defined, but that 2^α is not defined if $\alpha \not\leq T(|V|)$. So, for a cardinal α, it may well happen that the sequence

$$\alpha, 2^\alpha, 2^{2^\alpha} \ldots$$

is finite. We now know that this object will always be a set in NF, since the property of containing α and being closed under exponentiation is stratified. In the NF literature this set is called $\phi(\alpha)$. Now let A be suitable big set (this is only ever going to give us interesting results if A is big) and α its cardinal and consider the sequence

$$|\phi(\alpha)|, |\phi(T(\alpha))|, |\phi(T^2(\alpha))| \ldots$$

of natural numbers. We can prove that $|\phi(T(\alpha))| \geq T(|\phi(\alpha)|)+1$ always; we know that for some α it can happen that for n sufficiently large $|\phi(T^n(\alpha))|$ is $\aleph_0$ but it is known that, for at least some big α (such as $|V|$) $|\phi(T^n(\alpha))|$ is finite for at least all concrete n. If $|\phi(T^n(\alpha))|$ is finite for all natural numbers n of the model then the model has a class of natural numbers definable with a big cardinal α as a parameter. *Prima facie* it is a proper class rather than a set because its definition is highly unstratified, but—again *prima facie*—there is no obvious reason why the axiom saying that all such classes are sets should be inconsistent. Nor is there any obvious reason why sets defined in this way with big parameters should be definable without them, so we just might have here a way in which big sets can tell us something about rather more familiar and mundane sets like $\mathbb{N}$ and $\Re$.

6 POSITIVE SET THEORY

6.1 Positive set theory from the Fregean notion of set

The Fregean notion of set is about sets as *extensions* of propositional expressions, as presented in Frege's *Grundgesetze der Arithmetik*. More properly, given that Frege takes the notion of function as primitive in developing his system, and particularly in his analysis of propositional expressions, the Fregean notion of set is about sets as extensions of propositional *functions*. Anyway, whatever it is exactly, in its original presentation, it is inconsistent.

This was also true of the first systems proposed subsequently by the founders of the λ-calculus and related combinatory logics, which, like the Fregean paradise of type-free functions, were aimed at incorporating propositional notions, and which were thus exposed to paradoxes akin to Russell's, such as Curry's paradox.[12]

Nevertheless, logicians have always been trying to get back into the Fregean paradise in one way or another. And naturally, since the pure functional part of

[12] We would here refer the reader to [Seldin, 2009] for a precise account of the history of the λ-calculus and combinatory logic.

Frege's system, which is essentially the pure λ-calculus, was proved to be consistent, some have been tempted to locate the flaw in Frege's interpretation of the propositional component.

Such a way out is taken in [Aczel, 1980] and followed through in [Flagg and Myhill, 1987a; Flagg and Myhill, 1987b; Flagg and Myhill, 1989] where consistent systems of illative λ-calculus are proposed. As noticed in [Flagg and Myhill, 1987a], that line of research is in fact closely related to much earlier work by Fitch, as [Fitch, 1948] and papers therein cited. Additionally, a related approach that makes explicit reference to Fitch's work as source of inspiration is [Scott, 1975].

All of these proposals apply the 'method of iterated truth definitions', which been used successfully (in one variant or another) for consistency proofs of various systems related to positive set theory, as in [Gilmore, 1974], and whose origin can thus be traced back to Fitch. However, we shall not elaborate on them herein as the logics of the corresponding systems are essentially non-classical (excluded middle fails). That is the price of *unrestricted* comprehension, after all, to which one should add the tax of intensionality here, for another drawback common to those approaches is the loss of the basic principle of extensionality for sets – which is inherent in the Fregean notion of set.

But was Frege's original interpretation of the propositional component flawed? In fact, no, as lately shown in [Libert, 2008b], provided one is willing to restrict comprehension, of course. And the restriction just consists in proscribing the use in abstraction of those logical connectives that invalidate the fixed-point theorem, such as negation and implication, which respectively give rise to Russell's and Curry's paradoxes. Drastically, but remarkably enough, this indeed suffices to block the paradoxes. The fragment of Frege's system so obtained is presented as a system of illative λ-calculus in [Libert, 2008b]. Its pure set-theoretic part, which is extensional, will provide us with our first example of system of (pure) positive set theory. Here is its axiomatization.

Consider the language of set theory given by the following abstract syntax:

$$\begin{array}{rccl} \mathrm{var_a}: & \upsilon & ::= & x, y, z, ... \\ \mathrm{term_a}: & \tau & ::= & \upsilon \mid \{\upsilon \mid \varphi\} \\ \mathrm{form_a}: & \varphi & ::= & \bot \mid \top \mid \tau_1 \in \tau_2 \mid \\ & & & \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \forall \upsilon \varphi \mid \exists \upsilon \varphi \mid \\ & & & \neg \varphi \mid \varphi_1 \to \varphi_2 \mid \varphi_1 \leftrightarrow \varphi_2 \end{array}$$

Note that the equality relation is not taken as primitive here. As usual, in any extensional framework, $x = y$ can be defined by $\forall z(z \in x \leftrightarrow z \in y)$, provided the following formulation of the Axiom of Extensionality is adopted:

$$\mathrm{ext}: \quad \forall z(z \in x \leftrightarrow z \in y) \to \forall z(x \in z \leftrightarrow y \in z)$$

Positive formulas, which we denote by adding the superscript '+' to metavariables, are those formulas obtained without using the last line of the formation rules.

Naturally, Positive Comprehension in this language is defined as the Axiom of Comprehension restricted to the set $\mathrm{form}_\mathrm{a}^+$ of positive formulas, i.e.:

$$\{\mathrm{form}_\mathrm{a}^+\}: \quad \forall x(x \in \{x \mid \varphi^+\} \leftrightarrow \varphi^+)$$

And the first system of positive set theory on which we shall now focus is:

$$\mathrm{PST}_\mathrm{a} :\equiv \mathrm{ext} + \{\mathrm{form}_\mathrm{a}^+\}$$

The consistency of PST_a was established in [Hinnion and Libert, 2003], without any reference to Frege's system. Note that the consistency of $\{\mathrm{form}_\mathrm{a}^+\}$ can easily be proved by a *term model* construction, using the method of iterated truth definitions, as in [Gilmore, 1974]; but that *extensional* term models can indeed be obtained by that method requires a subtle analysis of it, as shown in [Hinnion and Libert, 2003]. A historical account of consistency problems for positive comprehension principles can be found in [Libert, 2004], where these are traced to Skolem's late work, e.g. [Skolem, 1960] which already provides some insight into another consistency proof of PST_a, as follows.

The appropriate way of looking at sets described by PST_a is as extensions of propositional *functions*, indeed. This was emphasized in [Libert, 2008a] where natural *topological* models are considered. And these are obtained in exactly the same way as the topological models of the pure λ-calculus: they come up from solutions to the reflexive equation $U \cong [U \to 2]$ in the category of Scott spaces, where $[U \to 2]$ is the set of Scott-continuous functions from U to the Sierpinski space 2 of truth values[13] – whereas we recall that Scott topological models of the pure λ-calculus appear as solutions to $U \cong [U \to U]$. Connections with the λ-calculus being disclosed, it was not long before PST_a could finally be related in [Libert, 2008b] to some fragment of Frege's system.

The condition on a propositional function for its extension to exist here is *Scott-continuity*, which can thus be regarded as a safety property for avoidance of the paradoxes. Given that $[U \to 2]$ is naturally isomorphic to the space of Scott *open* subsets of U, one may equally look at PST_a as the pure set-theoretic system associated with the limitation of comprehension to so-called *observable* predicates. Moreover, the following first-order axiom scheme, whose semantical interpretation is clear, is then satisfied in any topological model of PST_a described above:

$$(\Box): \quad \exists y \forall z(z \subseteq y \leftrightarrow \forall x(x \in z \to \varphi))$$

So a natural extension of PST_a as formal system is $\mathrm{PST}_\mathrm{a}^\Box :\equiv \mathrm{PST}_\mathrm{a} + (\Box)$, which is our first example of a system related to topological set theory.

Any extension of PST_a will however be characterized by a rather peculiar feature: the non-existence of singletons. For the reader may have noticed that the extensional equality relation is not positive, so it cannot be used in comprehension. And this seems to be unavoidable, as the consideration of $\{x \mid \{y \mid x \in x\} = \{y \mid$

[13] I.e., $2 := \{1, 0\}$, with 1 as the true, 0 as the false, and where $\{1\}$ is open, but not $\{0\}$.

$\perp\}\}$ will show.[14] But this can be related to another feature of the formalization of $\mathrm{PST_a}$: the use in the object language of the abstraction operator '$\{\cdot \mid -\}$', a remnant of the development of that system from systems with functional abstraction (λ-calculi).

Although it is related to the Fregean notion of set, such a system as $\mathrm{PST}^{\Box}_{\mathrm{a}}$ is likely to disconcert anyone who attempts to resurrect Frege's logicist program through it. Anyway, as we shall now see, there are more *natural* versions of positive set theory which are more appropriate for mathematical purposes.

6.2 Positive set theory seen from the Cantorian point of view

The Cantorian notion of set is sets-as-used-in-mathematics. Despite the emergence of paradoxes in its naïve formulation, the general understanding was that the mathematical notion of set should be coherent. This notion is now usually regarded as being based on an intuitively safe conception of sets as abstract collections, which we now describe.

In any instance of the abstraction process of set formation, it seems natural to assume that we must have available all the objects that are to be members of the set to be constructed, and that this always provides us with a *new* abstract object, which therefore cannot be involved in that particular instance of application of the abstraction process, nor in any previous one, but can be involved in subsequent ones.

That view is supported by the observation that sets occurring in every classical field of mathematics are indeed obtainable by *iterated* application of that "well-founded" abstraction process from some given collection of primitive objects (e.g. numbers). This gives rise to the intuitive *cumulative hierarchy* picture: a universe of sets built up in stages, which appears as the core of any set-theoretic universe described by what we called '*ZF*-like axiomatizations' – the underlying principle of specialized comprehension essentially guaranteeing that that construction can be carried out.

Let us now have a closer look at sets that can be so obtained *ex nihilo* in a finite number of steps (this will avoid questions about infinity). The collection of those sets, each of which is obviously finite, is denoted by V_ω. It is called the universe of *hereditarily finite well-founded sets*, which is the first limit step of von Neumann's implementation of the cumulative hierarchy, i.e.:

$$V_0 := \emptyset, \qquad V_{n+1} := \mathcal{P}(V_n), \qquad V_\omega := \bigcup\{V_n \mid n < \omega\},$$

where $\mathcal{P}(\cdot)$ is the powerset operation available in any *ZF*-like axiomatization.

As widely known, the set-theoretic structure $\langle V_\omega; \in\rangle$ is an inner model of *ZF* minus infinity. Note also that $\langle V_\omega; \in\rangle$ is somehow captured by the family of (finite) set-theoretic structures $\langle V_n; \in\rangle$, $n > 0$, seeing that $x \in y$ for $x, y \in V_\omega$ iff there

[14] Let $R := \{x \mid \{y \mid x \in x\} = \{y \mid \perp\}\}$. We have $R \in R$ iff $\{y \mid R \in R\} = \{y \mid \perp\}$ iff $R \in R \leftrightarrow \perp$ iff $\neg(R \in R)$.

is some $n > 0$ such that $x, y \in V_n$ and $x \in y$. In categorical terms, we say that $\langle V_\omega; \in\rangle$ is the *inductive* (or *direct*) limit of the $\langle V_n; \in\rangle$'s. Let us formalize things carefully, if perhaps pedantically, in order to make clearer the comparison with the construction we are going to consider next.

For each $n > 0$, let $i_n : V_n \hookrightarrow V_{n+1} : x \mapsto x$ be the canonical inclusion map, and let $\langle V_n; \in_n^i\rangle$ be the set-theoretic structure defined by $x \in_n^i y$ iff $x \in i_n(y)$, that is, $x \in y$. So $\langle V_n, \in_n^i\rangle$ is nothing but $\langle V_n; \in\rangle$. We note that each i_n is obviously an injective $\in$-morphism of $\langle V_n; \in_n^i\rangle$ into $\langle V_{n+1}; \in_{n+1}^i\rangle$, i.e., $x \in_n^i y$ implies $i_n(x) \in_{n+1}^i i_n(y)$. Thus we have properly defined a *direct* system $(\langle V_n; \in_n^i\rangle, i_n)_{0<n<\omega}$ of set-theoretic structures, whose so-called *direct* or *inductive* limit is just $\langle V_\omega; \in\rangle$. Indeed, $\varinjlim_i V_n = \bigcup\{V_n \mid 0 < n < \omega\} = V_\omega$, by definition, and $x \in_\omega^i y$ iff there exists $n > 0$ such that $x \in_n^i y$, i.e., iff $x \in y$.

Given that categorical construction, and the familiar set-theoretic structure that comes with it, it is natural to look at the *dual* one, as follows.

Let $p_1 : V_2 \twoheadrightarrow V_1 : x \mapsto \emptyset$, and then let $p_{n+1} : V_{n+2} \twoheadrightarrow V_{n+1}$, $n > 0$, be inductively defined by $p_{n+1}(x) := \{p_n(y) \mid y \in x\}$. It is easily seen by induction that $p_n \circ i_n = \mathrm{id}_{V_n}$, i.e., the restriction of p_n on V_n is id_{V_n}, so p_n is indeed a surjection. Now, for each $n > 1$, let $\langle V_n; \in_n^p\rangle$ be the set-theoretic structure defined by $x \in_n^p y$ iff $p_{n-1}(x) \in y$; we conveniently define $\in_1^p$ as $V_1 \times V_1$. It is easy to prove that each p_n is a surjective $\in$-morphism of $\langle V_{n+1}; \in_{n+1}^p\rangle$ onto $\langle V_n; \in_n^p\rangle$, i.e., $x \in_{n+1}^p y$ implies $p_n(x) \in_n^p p_n(y)$. Thus we have defined an *inverse* system $(\langle V_n; \in_n^p\rangle, p_n)_{0<n<\omega}$ of set-theoretic structures, which is naturally associated with the direct system above. Let $\langle N_\omega; \in_n^p\rangle$ be the so-called *inverse* or *projective* limit of that system, i.e.,

$$N_\omega := \varprojlim_p V_n = \{\xi \in \prod\{V_n \mid 0 < n < \omega\} \mid \forall n > 0, p_n(\xi_{n+1}) = \xi_n\},$$

and $\xi \in_\omega^p \zeta$ iff for all $n > 0$, $\xi_n \in_n^p \zeta_n$. What is that toy, yet natural, set-theoretic structure a model of?

The first thing to note is that $\langle N_\omega; \in_\omega^p\rangle$ is extensional. A proof of this would make use of the following 'extension lemma': given $x \in V_\omega, \zeta \in N_\omega$, if $x \in_n^p \zeta_n$ for some $n > 0$, there exists $\xi \in N_\omega$ such that $\xi_n = x$ and $\xi \in_\omega^p \zeta$. We would also invite the reader to convince himself that $\langle N_\omega; \in_\omega^p\rangle$ is an *end-extension* of $\langle V_\omega; \in\rangle$. More precisely, for any $x \in V_\omega$, let $\rho(x)$ be the rank of x, that is, the least $n < \omega$ such that $x \in V_{n+1}$. Then, one can prove that the map $e : V_\omega \to N_\omega$ defined by $e(x) := (\ldots, p_{\rho(x)-1}(p_{\rho(x)}(x)), p_{\rho(x)}(x), x, x, \ldots)$, is an embedding of $\langle V_\omega; \in\rangle$ as *initial part* of $\langle N_\omega; \in_\omega^p\rangle$, i.e.: e is an injective $\in$-morphism and if $\xi \in_\omega^p e(y)$ for some $y \in V_\omega$, there exists $x \in V_\omega$ such that $\xi = e(x)$ and $x \in y$. Thereupon we note that $\langle N_\omega; \in_\omega^p\rangle$ is really a proper extension of $\langle V_\omega; \in\rangle$. For instance, it is easy to see that $p_n(V_n) = V_{n-1}$ for all $n > 0$, so $\nu := (V_{n-1})_{0<n<\omega} \in N_\omega$, and then that for all ξ in N_ω, $\xi \in_\omega^p \nu$. Therefore, there is a universal set in $\langle N_\omega; \in_\omega^p\rangle$, namely ν, whereas there is no such set in $\langle V_\omega; \in\rangle$. There are by far many other unusual sets in $\langle N_\omega; \in_\omega^p\rangle$, simply because N_ω has the power of the continuum, whereas V_ω is countable!

As a matter of fact, N_ω could have been presented as a *topological* completion of V_ω, in the same way as the real numbers can be presented as the Cauchy completion of the rational numbers. The notion of convergence in N_ω is defined so that for all ξ in N_ω, $\xi = \lim_{n\to\infty} e(\xi_n)$. In other words, sets in N_ω can be described as limits of hereditarily finite well-founded sets. Other presentations of N_ω as a mathematical structure can be found in [Abramsky, 1988], where $\langle N_\omega; \in^p_\omega\rangle$ is depicted as the universe of *finitary non-well-founded sets.*

One precise characterization of N_ω is as the solution to a reflexive equation of the form $U \cong \mathcal{P}_{cl}(U)$ in the category of compact metric spaces, where $\mathcal{P}_{cl}(U)$ is the set of *closed* subsets of U. This reminds us of the topological models of $\mathrm{PST_a}$ discussed earlier, which appear as solutions to the equation $U \cong \mathcal{P}_{op}(U)$ in the category of Scott spaces, where $\mathcal{P}_{op}(U)$ is the set of *open* subsets. So this suggests that set-theoretic universes as $\langle N_\omega; \in^p_\omega\rangle$ could be related to positive set theory as well. Actually, such set-theoretic structures turn out to be the topological models associated with another natural version of positive set theory, which can be axiomatized as follows.

Consider the language of set theory whose formulation rules are:

$$\begin{array}{rcll} \mathrm{var_b}: & \upsilon & ::= & x, y, z, ... \\ \mathrm{form_b}: & \varphi & ::= & \bot \mid \top \mid \upsilon_1 \in \upsilon_2 \mid \\ & & & \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid (\forall \upsilon_1 \in \upsilon_2)\varphi \mid (\exists \upsilon_1 \in \upsilon_2)\varphi \mid \\ & & & \neg\varphi \mid \varphi_1 \to \varphi_2 \mid \varphi_1 \leftrightarrow \varphi_2 \end{array}$$

Again, we call *positive* those formulas obtained without using the last line of the formation rules, which we indicate by the superscript '+' on metavariables. We stress the presence of bounded quantifiers as primitive *positive* operators here, but the absence of the abstractor '$\{\cdot \mid -\}$' as primitive term constructor. Positive comprehension, that is, the restriction of the Axiom of Comprehension to the set $\mathrm{form}^+_\mathrm{b}$ of positive formulas, is then expressed in this language by:

$$\{\mathrm{form}^+_\mathrm{b}\}: \quad \exists y \forall x (x \in y \leftrightarrow \varphi^+)$$

And the corresponding system of positive set theory in which we are interested now is:

$$\mathrm{PST_b} :\equiv \mathrm{ext} + \{\mathrm{form}^+_\mathrm{b}\}$$

Note immediately that the equality relation, which is not primitive here, can however be defined by a *positive* formula, i.e.,

$$x = y :\equiv (\forall z \in y) z \in x \wedge (\forall z \in x) z \in y.$$

So singletons are innocent definable objects in $\mathrm{PST_b}$, in contrast with $\mathrm{PST_a}$. Thanks to bounded quantification, other set-theoretic operations that were not available in $\mathrm{PST_a}$ are now legitimate, as notably the power-set operation, $\mathcal{P}(y) := \{x \mid (\forall z \in x) z \in y\}$, which thus can be defined positively according to this criterion. Note also that the usual 'unbounded' quantifiers can be recovered from the primitive bounded ones, by defining $\forall v\varphi$ as $(\forall v \in \mathrm{V})\varphi$ and $\exists v\varphi$ as $(\exists v \in \mathrm{V})\varphi$, where V is the universal set, whose existence follows from $\mathrm{PST_b}$.

As said above, the consistency of $\mathrm{PST_b}$ can be established by means of N_ω, and more generally of topological models that appear as solutions to the reflexive equation $U \cong \mathcal{P}_{cl}(U)$ in some category of compact Hausdorff spaces. These topological models clearly satisfy the *dual* of the axiom scheme $(\Box)$ in Section 2, namely:

$$(\Diamond): \quad \exists y \forall z(y \subseteq z \leftrightarrow \forall x(\varphi \to x \in z))$$

So the natural extension of $\mathrm{PST_b}$ as formal system is $\mathrm{PST_b^\Diamond} :\equiv \mathrm{PST_b} + (\Diamond)$.

Interestingly, it was shown in [Esser, 1997][15] that *ZF* minus infinity can be interpreted within $\mathrm{PST_b^\Diamond}$, which definitely distinguishes that system from $\mathrm{PST_a^\Box}$. In fact, *ZF* (with infinity) is naturally interpretable in $\mathrm{PST_b^\Diamond}$ augmented with von Neumann's Axiom of Infinity.[16] The resulting system, called GPK^+_∞ in [Esser, 1997], is then proved to be mutually interpretable with a large cardinal extension of the Kelley-Morse set theory. So there is undoubtedly a natural version of positive set theory which is at least as suitable as ZF for mathematical purposes.

The relative consistency of GPK^+_∞ in [Esser, 1997] was essentially established in [Forti and Hinnion, 1989], which provides a general construction of a hierarchy of N_α's, such that for any weakly compact cardinal κ, N_κ is a topological model of GPK^+_∞ containing V_κ (the von Neumann hierarchy up to κ). Such structures, subsequently called *hyperuniverses*, have been extensively studied since then, e.g. in [Forti and Honsell, 1996a; Forti and Honsell, 1996b]. As is explained in [Forti and Hinnion, 1989], they were independently investigated in [Weydert, 1989] for the purpose of topological set theory, to which subject we shall now turn. Note incidentally that the constructions given in [Forti and Hinnion, 1989] and [Weydert, 1989] were both inspired by the pioneering work of Malitz [Malitz, 1976], which can also be classed as a contribution to topological set theory.

6.3 Topological set theory

We have met two variants of positive set theory, namely $\mathrm{PST_a}$ and $\mathrm{PST_b}$, which can respectively be extended by the (first-order) *topological* axioms schemes $(\Box)$ and $(\Diamond)$.[17] These were suggested by the corresponding topological models, whose interest is manifest: in such topological set-theoretic structures, even though not all definable subsets of the domain can be abstracted (because of Russell's paradox), each of these can at least be optimally *approximated* by the largest [resp. the smallest] set contained in [resp. containing] it. So each of the axiom schemes $(\Box)$

[15]Esser's thesis was subsequently published in [Esser, 2004]; see also [Esser, 1999] for a related paper.

[16]In the form: 'there exists an infinite von Neumann ordinal'.

[17]By the way, we mention that $\mathrm{PST_a}$ could equally be extended by $(\Diamond)$, by duality, that is, $\mathrm{PST_a^\Diamond}$ is consistent (see [Libert, 2008a] for details). On the other hand, $\mathrm{PST_b^\Box}$ is inconsistent merely because $(\Box)$ is incompatible with the existence of all singletons (see [Libert and Esser, 2005]).

and $(\Diamond)$ provides on its own an approximate version of the naïve comprehension scheme.

Such an idea of approximation of naïve comprehension – the *essence* of topological set theory – originated in a paper by Skala [Skala, 1974], which was subsequently refined by Manakos [Manakos, 1984]. Although Skala's paper is cited in the references of [Weydert, 1989] and [Malitz, 1976], and so might have been a source of inspiration in those works, the proposal she made is clearly distinguishable from the others in that her system is based upon *both* $(\Box)$ and $(\Diamond)$, *together with* an axiom for the existence of complements, which we denote by (C).[18] So it is not at all a system of positive set theory! Besides, Skala's proposal is purely axiomatic, with no attempt at giving any topological characterization of the models. As shown in [Libert and Esser, 2005], the model theory of Skala's system is quite simple though.

The natural models of $(\Box)+(\Diamond)+(C)$ appear as solutions to a reflexive equation of the form $U \simeq \mathcal{P}_{clop}(U)$ in the category of quasi-discrete spaces, where $\mathcal{P}_{clop}(U)$ is the set of *clopen* subsets of U, which precisely coincides for quasi-discrete spaces with both $\mathcal{P}_{cl}(U)$ and $\mathcal{P}_{op}(U)$ – the reason why both $(\Box)$ and $(\Diamond)$ hold in these models. Thereupon we recall that $\mathcal{P}_{clop}(U)$ is a complete Boolean algebra, so there is actually no restriction in Skala's set theory on the use of logical connectives and quantifiers in comprehension, which is in contrast with the situation in positive set theory. The restriction, for there should be one, is relegated to the atomic level and is rather of a syntactical nature here, as follows.

Consider the simple language of set theory whose formation rules are:

$$
\begin{array}{rcll}
\mathrm{var_c}: & \upsilon & ::= & x, y, z, \dots \\
\mathrm{form_c}: & \varphi & ::= & \bot \mid \top \mid \upsilon_1 \in \upsilon_2 \mid \\
 & & & \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \forall \upsilon \varphi \mid \exists \upsilon \varphi \mid \\
 & & & \neg \varphi \mid \varphi_1 \rightarrow \varphi_2 \mid \varphi_1 \leftrightarrow \varphi_2
\end{array}
$$

Let $\varphi[\upsilon[$ denote any formula of $\mathrm{form_c}$ in which the variable υ does not appear on the *right-hand* side of $\in$, and let $\mathrm{form_c}[\upsilon[$ be the set of those formulas. Given a concrete variable, say x, we then naturally define the restriction of the axiom scheme of Comprehension to $\mathrm{form_c}[x[$ by:

$$\{\mathrm{form_c}[x[\} : \quad \exists y \forall x (x \in y \leftrightarrow \varphi[x[)$$

It is proved in [Libert and Esser, 2005] that $\mathrm{ext} + (\Box) + (\Diamond) + (C)$ is equivalent to $\mathrm{ext} + \{\mathrm{form_c}[x[\}$. So Skala's set theory can equally be presented as a type-free system based on a very simple syntactic criterion for avoidance of the paradoxes, which, in view of Russell's paradox, remains the only option if one wants a *Boolean* set-theoretic universe in a type-free setting.[19] But this has drastic consequences.

[18] Manakos in [Manakos, 1984] dropped this last assumption in order to consider other possible extensions of the basic system $(\Box) + (\Diamond)$; see [Libert and Esser, 2005] for details.

[19] Clearly, the restriction on the abstracted variable to appear on *left-hand* side of $\in$, which is the other option, would seem even less sensible.

Thus we observe that $x = y$, which was defined by $\forall z(z \in x \leftrightarrow z \in y)$, is neither in $\mathrm{form}_c[x[$ nor in $\mathrm{form}_c[y[$. Hence, the extensional equality relation cannot be used in instances of comprehension. However, what is called the *indiscernibility relation*, which is defined by $x \doteq y :\equiv \forall z(x \in z \leftrightarrow y \in z)$, can be used instead. Furthermore, *quasi-singletons*, i.e., sets of the form $\{x \mid x \doteq y\}$, appear as atomic objects in Skala's set theory, for any set is provably equal to the union of the quasi-singletons of its members. That the naïve conception of set can somehow be salvaged when indiscernibility plays the role of equality may not be devoid of philosophical interest. This was taken seriously in [Apostoli and Kanda, to appear; Apostoli and Kanda, 2000], where elaborated quasi-discrete solutions to $U \simeq \mathcal{P}_{clop}(U)$ are constructed, without any reference to Skala set theory, but in connection with *rough set theory*.

On the other hand, the non-definability of the usual operation of forming singletons does not entail their non-existence. As a matter of fact, it is proved in [Libert and Esser, 2005] that *any* given set-theoretic universe in which all singletons exist (e.g. a model of *ZF*) can consistently be embedded as the class of objects admitting singletons in some model of Skala's set theory. In other words, although it is intrinsically weak, Skala's system is very adaptable.

In any case, such an axiomatic set theory should not be regarded as (and then compared to) a *classical* axiom system, that is, one for which we have in mind some *standard* interpretation – which is the case of *ZF* for instance. Rather, it appears as a *general* axiom system, as the one of the theory of Boolean algebras, to which it is closely related. After all, thinking of sets as extensions of predicates, one might expect that some set-theoretic universe or another could reflect the Boolean structure of classical logic.

Hopefully the various set-theoretic systems presented in this survey will have at least convinced the reader of what mathematicians have discovered through their investigations, namely that set theory is definitely *not* unique.

6.4 A development of mathematics in GPK^+_∞

In this section we give a brief self-contained presentation of Oliver Esser's system GPK^+_∞ of [Esser, 1999], the most mathematically effective system of positive set theory.

GPK^+_∞ is a first-order theory with equality and membership as primitive relations.

Extensionality: $(\forall AB.(\forall x.x \in A \leftrightarrow x \in B) \rightarrow A = B)$

Definition: Let the class of (bounded) positive formulas be the smallest class containing the formula $x \neq x$ (useful because uniformly false), all atomic formulas, and closed under conjunction, disjunction, bounded universal quantification $(\forall x \in A.\phi)$ and existential quantification $(\exists x \in A.\phi)$.

Positive Comprehension: For any positive formula ϕ in which A does not appear, $(\exists A.(\forall x.x \in A \leftrightarrow \phi))$.

Closure: For any formula ϕ, there is a set C such that $(\forall D.C \subseteq D \leftrightarrow (\forall x.\phi \rightarrow x \in D))$ (a set minimal in the inclusion order among all sets which include the possibly proper class $\{x \mid \phi\}$). This is what was referred to above as $(\diamond)$ and it should be noted that this is equivalent to the assertion that any definable class of sets has set intersection.

Definition: The von Neumann ordinal ω is the intersection of all sets which contain $\emptyset$ and are closed under the von Neumann successor operation $x \mapsto x \cup \{x\}$. By Closure, ω is a set.

Definition: A set x is *isolated* iff $\{x\}^c = \{y \mid y \neq x\}$ is a set.

Infinity: Each element of ω is isolated. (We have stated this axiom in a form specifically designed to make our brief development more efficient).

We start working toward an interpretation of ZF in this theory.

Observation: $\{x, y\}$ is a set ($\{z \mid z = x \vee z = y\}$) Moreover $z = \{x, y\}$ is equivalent to a positive formula $x \in z \wedge y \in z \wedge (\forall w \in z.z = x \vee z = y)$. It is important in a development of this theory to know what term constructions can appear freely in positive formulas, as we have just shown the unordered pair (and so the usual ordered pair) can. Similarly, $x \cup y$ and $x \cap y$ can be mentioned freely in positive formulas. For any set x, $\bigcup x = \{y \mid (\exists z \in x.y \in z)\}$ and $\mathcal{P}(x) = \{y \mid (\forall z \in y.z \in x)\}$. Note that the axioms of Pairing, Union, and Power Set of the usual set theory thus hold. $y = \bigcup x$ is known to be equivalent to a positive formula; $y = \mathcal{P}(x)$ is not known to expressible in positive form. We cannot expect the Axiom of Separation to hold, as the formula ϕ determining an instance of the axiom is not necessarily positive.

Observation: Let A be a set all of whose elements are isolated. Then $\{x \in A \mid \phi\}$ exists for *every* formula ϕ by an application of Closure (take the intersection of all the sets $A \setminus \{y\}$ such that $y \in A$ and $\neg\phi[y/x]$.

Similarly, if $\phi(x, y)$ is a functional formula $((\forall x.(\exists y!.\phi(x, y))))$ and A is a set all of whose elements are isolated, then $\{y \mid (\exists x \in A.\phi(x, y))\}$ is a set. Let F^* be the closure of the class of all (x, y) such that $\phi(x, y)$. There is a set $\{(x, y) \mid x \in A \wedge (x, y) \in F^*\}$. Let $F(x)$ denote the unique y such that $\phi(x, y)$. For each $x \in A$, there is a set $\{(u, v) \mid x \neq u \vee v = F(x)\}$. From the existence of these sets, it follows that the intersection of all sets containing the class $\{y \mid (\exists x \in A.\phi(x, y))\}$ is exactly that class, and so the class is a set.

Observation: If all elements of A are isolated, then A is isolated. The function d such that $d(x) = V \setminus \{x\}$ for each $x \in A$ is a set by the second half of the previous observation. We want to show that A^c is a set. $x \neq A$ is equivalent to $(\exists y \in A.y \neq x) \vee (\exists y \in x.y \notin A)$. This is in turn equivalent to $(\exists y \in A.y \in d(x)) \vee (\exists y \in x.(\forall z \in A.y \in d(z)))$. Finally, $u \in d(v)$ is equivalent to $(u, v) \in d$, and we know that pairs can be mentioned freely in positive formulas. Since $x \neq A$ is equivalent to a positive formula, A^c is a set as desired.

The observations above are sufficient to suggest that the class of hereditarily isolated sets satisfies the axioms of ZF, though some additional details would need to be attended to in a full presentation.

One can then develop an adequate theory of proper classes of hereditarily isolated sets (closures of arbitrary classes of hereditarily isolated sets). This theory will satisfy Kelley-Morse set theory, with one further surprise: one can prove that the proper class ordinal in the interpreted Kelley-Morse set theory is weakly compact.

Conversely, there is a construction of a model of GPK^+_∞ in ZFC with a weakly compact, and in fact one can show that GPK^+_∞ is precisely as strong as Kelley-Morse set theory with the proper class ordinal weakly compact. This is considerably stronger than ZFC.

An interesting weaker theory is obtained if one drops the Axiom of Infinity (which was referred to above as $\mathrm{PST}^\diamond_\flat$, after the description of its natural model N_ω). This allows ω to have a non-isolated element, namely ω itself (ω becomes the topological limit point of the finite von Neumann ordinals). This theory has the same strength as second-order arithmetic.

7 SYSTEMS MOTIVATED BY NONSTANDARD ANALYSIS

In this section we discuss systems of set theory which can be regarded as motivated by the nonstandard analysis of Abraham Robinson [Robinson, 1966]. These theories can be characterized by the presence of what Vopěnka calls *proper semisets*: subcollections of sets which are proper classes. Of course, any set theory with a universal set must have proper semisets (the Russell class being a subclass of the universe) but in the theories here we postulate that "small" familiar sets such as $\mathbb{N}$ have proper subclasses.

7.1 Nonstandard analysis

Consider the first-order theory of the real numbers (the precise details of this theory are not important here). The collection of sentences $\{c > 0, c < 1, c < \frac{1}{2}, \ldots, c < \frac{1}{n}, \ldots\}$, where c is a new constant, is consistent (any finite subcollection is readily satisfied by taking c to be positive and sufficiently small), so by the compactness theorem the theory of the real numbers has models in which this collection of sentences is satisfied.

In such a model, we can think of c as a positive infinitesimal, for c is smaller than any of the familiar real numbers, because it is smaller than all the reciprocals of the familiar natural numbers. But note that of course there are "natural numbers" N of the model such that $\frac{1}{N} < c$.

If we define an "infinitesimal" as any real (of our nonstandard model) which is smaller than all standard $\frac{1}{n}$ in absolute value (a notion which cannot be defined in terms of the theory of the reals with which we started, because the notion of a standard natural number does not make sense there), then we can attempt the program of 17th century analysis.

For example, if f is a function from the reals to the reals definable in our original theory of real numbers, with a derivative $f'(x)$, it is straightforward to show that $\frac{f(x+dx)-f(x)}{dx}$ differs infinitesimally from $f'(x)$. So we can *define* $f'(x)$ for each standard real x as the unique standard real y which differs infinitesimally from $\frac{f(x+dx)-f(x)}{dx}$ for each infinitesimal dx. A very appealing development of analysis in this style is possible.

Robinson presented nonstandard analysis in the framework of model theory. In his approach, one considers a model $\mathbb{R}^*$ of the theory of the real numbers, containing all the usual reals (which are called the "standard" reals) and having the property that any real of the model which is bounded in absolute value by a standard real differs infinitesimally from one of the standard reals.

There is an alternative approach, which is to "augment" our set theory with axioms which ensure that infinite sets such as $\mathbb{N}$ or $\mathbb{R}$ have the properties we expect and at the same time have proper subclasses (proper semisets) which can play the role of the collections of "standard" reals or natural numbers in the discussion above.

7.2 Nelson's IST

Without further ado, we present Nelson's IST (internal set theory, defined in [Nelson, 1977]) which is motivated exactly by the desire to streamline the development of nonstandard analysis.

IST is a first-order theory with the familiar predicates of equality and membership and an additional predicate $st(x)$ of "standardness". It is convenient to abbreviate $(\forall x.st(x) \rightarrow \phi)$ as $(\forall^{st}x.\phi)$ and $(\exists x.st(x) \wedge \phi)$ as $(\exists^{st}x.\phi)$.

The axioms of IST are of two sorts. First we have the axioms of standard set theory. Nelson uses the axioms of ZFC, but one could equally well use a weaker theory such as Mac Lane set theory. The axioms of Separation and Replacement of IST are modified in that the formulas mentioned in the schemes may not mention the predicate *st*: sets may not be defined using the predicate *st*.

There are three additional axioms in Nelson's theory which directly support the program of standard analysis.

Idealization: Let $\phi(x, y)$ be a formula not mentioning st or the variable f (it may have nonstandard parameters). "for every standard finite set f, $(\exists x.(\forall y \in f.\phi(x, y)))$" is equivalent to "$(\exists x.(\forall^{st}y.\phi(x, y)))$".

Corollary: If ϕ is taken to be the formula "x is finite and $y \in x$", we observe that the first formula is true and by Idealization equivalent to the second, and so we assert that there is a finite set which contains every standard object.

Standardization: For every standard set A, and for any formula ϕ whatsoever (this formula may mention st and nonstandard parameters freely), $(\exists^{st} B.(\forall^{st} x.x \in B \leftrightarrow x \in A \wedge \phi))$: the standard elements x of B are exactly the standard elements x of A such that ϕ.

Transfer: For any formula ϕ not mentioning the predicate st and containing only standard parameters, $(\forall x.\phi) \leftrightarrow (\forall^{st} x.\phi)$.

The consistency of these axioms is readily established by standard techniques of model theory. The advantage for the nonstandard analyst is that they can then use IST as their set-theoretic framework and not have to directly discuss nonstandard models of (say) the reals at all.

For the set theorist, it is interesting that the axioms of IST are not recognizably about analysis: they are set-theoretical in character, though certainly what they assert is rather odd.

We review how these axioms support the program of nonstandard analysis. Suppose that the reals are implemented as left sets of Dedekind cuts in the rationals (which can in turn be supposed implemented in set theory in a quuite usual way).

Idealization gives us a positive infinitesimal real number. We define an infinitesimal as a real number which is less in absolute value than any standard real number. Obviously 0 is infinitesimal. But we can define a positive infinitesimal: let ϵ be the intersection of all positive reals (considered as left sets of Dedekind cuts in the rationals) which belong to a specific finite set F containing all standard sets as elements (the existence of such a set being a corollary of Idealization). Clearly ϵ is a positive real (it is the minimum of a finite set of reals) but it is also less than every positive standard real (because all of them belong to F).

Standardization gives us a unique standard real number infinitesimally close to any real number which is bounded in absolute value by some standard real. Let r be a real number which is bounded above in absolute value by some standard real. By Standardization, there is a standard set r^* which contains the same standard elements (r being a subset of the standard set of rationals). It is straightforward to use Transfer to verify that r^* is also a real number (since all standard elements of r^* are rational numbers, all elements of r^* without exception are rational numbers; since the standard elements of r^* make up a downward closed subset of the standard reals without a largest element, r^* itself is a downward closed subset of the reals without a largest element: it is important here than r is bounded above in absolute value by some standard real, as otherwise r^* could be either empty or the entire set of rationals and so of course not a real). If r^* differed from r by a non-infinitesimal amount, there would be a standard real in their symmetric difference, contrary to the choice of r^*.

Transfer (and the restriction of Separation and Replacement to formulas not mentioning the new predicate st) ensures that the underlying set theory (and so the implementation of the reals in the underlying set theory) works exactly as expected, in spite of the obvious weirdness caused by the presence of proper subclasses in every infinite set (strictly speaking, we do not have a theory of classes in IST, but we will in the theory considered next, and it would be straightforward to add a theory of classes to IST itself)

7.3 Vopěnka's alternative set theory

The alternative set theory of Vopěnka (see [Vopěnka, 1979]) is also usable for nonstandard analysis, but it is conceptually more different from the usual set theories. It is also much weaker, but mathematically rather interesting.

Vopěnka's theory has sets and classes. We have the usual axioms of Extensionality and Class Comprehension. We follow the convention that capital letters represent classes and lower-case letters represent sets.

Definition: "x is a set" means $(\exists A.x \in A)$ as usual.

Extensionality: $(\forall AB.(\forall x.x \in A \leftrightarrow x \in B) \rightarrow A = B)$. Classes with the same elements are the same.

Class Comprehension: For any formula ϕ in which A is not free, the universal closure of $(\exists A.(\forall x.x \in A \leftrightarrow \phi))$ is an axiom.

The axioms for sets are quite surprising. The theory asserted of sets is equivalent to ZFC with the Axiom of Infinity replaced by its negation, though the form presented has independent interest as a way to present the theory of hereditarily finite sets.

Empty set: The empty class $\emptyset$ (which exists by class comprehension) is a set.

Successor: For any sets x and y, $x \cup \{y\} = \{z \mid z \in x \vee z = y\}$, which exists by class comprehension, is a set.

Induction: For any formula ϕ in which all parameters are sets and all quantifiers are bounded in the class V of all sets, $\phi[\emptyset/z] \wedge (\forall xy \in V.\phi[x/z] \wedge \phi[y/z] \rightarrow \phi[x \cup \{y\}/z]) \rightarrow (\forall z \in V.\phi)$. This axiom says in effect that all sets are built up by the iterated application of the successor operation starting with the empty set.

The Axiom of Foundation for sets is usually listed as an axiom but follows from Induction.

Now come the peculiarities obviously related to nonstandard analysis.

Definition: A *semiset* is a subclass of a set. A *proper semiset* is a semiset which is not a set.

Semisets: There is a proper semiset.

Definition: A set is *finite* iff all of its subclasses are sets.

Definition: An order of type ω is a well-ordering whose domain is not finite but all initial segments of which have finite domain. A class is countable if there is a class bijection between it and the domain of an order of type ω. The class of finite von Neumann ordinals is provably an order of type ω.

Prolongation: Each countable function F can be extended to a set function.

The Axiom of Semisets has the effect of the Idealization Axiom in IST (it ensures that there are nonstandard sets, in this case nonstandard hereditarily finite sets). The notion of *finite* here can be understood as (roughly) capturing the notion of a set of standard finite size. The function of the Axiom of Prolongation is similar to that of standardization in IST. It is a distinct oddity of the alternative set theory as an implementation of nonstandard analysis that there is no analogue of the Axiom of Transfer.

We review the construction of number systems. There are two flavors of natural numbers: the system of all the von Neumann ordinals (nonstandard naturals) and the system of finite von Neumann ordinals (standard naturals). The real numbers can be defined as Dedekind cuts in the *standard* rationals (ratios of finite natural numbers and their additive inverses): these are not sets, but prolongation can be used to show that any real can be approximated infinitesimally by nonstandard rationals, and an interesting nonstandard analysis (not really the same as Robinson's) can be developed.

There are two more axioms in Vopěnka's system.

Vopěnka considers representations of superclasses of classes using relations on sets. A class relation R on a class A is said to code the superclass of inverse images of elements of A under R. A class relation R on a class A is said to extensionally code this superclass if distinct elements of A have distinct preimages. He "tidies up" the theory of such codings by adopting the very technical

Axiom of Extensional Coding: Every collection of classes which is codable is extensionally codable.

The final axiom tidies up the theory of infinite cardinality.

Axiom of Cardinalities: Any two uncountable classes are the same size (i.e, there is a class bijection between them).

Extensional Coding and Cardinalities together can prove Choice (any class can be well-ordered). Since Choice easily implies Extensional Coding, the former axiom might seem more natural. As in pocket set theory, there are only two infinite cardinalities, $\aleph_0$ and c. $\aleph_0$ is here the cardinality of the collection of standard natural numbers, and c is the cardinality of the collection of all natural numbers (?!).

A model of the alternative set theory in the usual set theory is a nonstandard model of V_ω of size ω_1 in which every countable external function extends to a function in the model. It might be best to suppose that this model is constructed inside L (the constructible universe) so that the Axiom of Cardinalities will be satisfied. The Axiom of Extensional Coding follows from Choice in the ambient set theory.

8 CURIOSITIES

In this section we introduce two eccentric theories. One of them has the property which is usually ascribed to New Foundations (we believe not entirely fairly) of being motivated by a syntactical trick without any semantic motivation. The double extension set theory of Andrzej Kisielewicz definitely has this property (and it is not known whether it is consistent).

The other is one of the strongest set theories ever proposed. It is known that it is inconsistent with ZFC that there is a proper elementary embedding from the universe to itself. But this proof does not work in Zermelo set theory: Zermelo set theory with an axiom asserting that there is an elementary embedding from V to V is of the order of consistency strength of the strongest extensions of ZFC that have been proposed.

8.1 Double extension set theory

Andrzej Kisielewicz has proposed in two papers (describing three systems: the not-known-to-be-inconsistent one is described in [Kisielewicz, 1998])) that the paradoxes of set theory can be averted by providing two different membership relations, and allowing extensions for each relation to be defined using formulas in the other relation. His strongest systems have been shown to be inconsistent by Holmes (in [Holmes, 2004]). Holmes also showed that the ordinals in his weaker system have startling properties (in [Holmes, 2005]).

We describe the version of Kisielwicz's theory which is not known to be inconsistent. This is a first-order theory with equality and two primitive relations $\in$ and ϵ, both of which are to be thought of as membership. For any formula ϕ, the formula ϕ^* is obtained by replacing $\in$ with ϵ and *vice versa* in ϕ.

Definition: A set x is said to be *regular* iff $(\forall y, y \in x \leftrightarrow y \,\epsilon\, x)$ (i.e., if it has the same extension with respect to both membership relations).

Axiom of Comprehension: For each formula ϕ which does not mention ϵ, in which any parameters are regular, and in which the variable A is not free, $(\exists A.(\forall x.x \,\epsilon\, A \leftrightarrow \phi) \wedge (\forall x.x \in A \leftrightarrow \phi^*))$. The object A is denoted by $\{x \mid \phi\}$.

Axiom of Mixed Extensionality: $(\forall AB.(\forall x.x \in A \leftrightarrow x \,\epsilon\, B) \rightarrow A = B)$ Note that any sets shown to be equal by mixed extensionality are necessarily regular.

Definition: We say that x is *partially contained* in y iff $(\forall z.z \in x \rightarrow z \in y) \vee (\forall z.z \,\epsilon\, x \rightarrow z \,\epsilon\, y)$. We say that a set *has regular elements* iff all of its elements in either sense are regular.

Axiom of Regular Containment: Any set partially contained in a set with regular elements is regular.

This set avoids Russell's paradox in the following curious way. There is a set $R = \{x \mid x \in x\}$ by Comprehension, but we have to recall what this means: $x \,\epsilon\, R \leftrightarrow x \in x$ and $x \in R \leftrightarrow x \,\epsilon\, x$. Thus we have $R \in R \leftrightarrow \neg R \,\epsilon\, R$: we conclude that R is not regular.

This theory is at least as strong as ZF (the class of hereditarily regular sets satisfies ZF), but the presentation of the proof is hampered by the extremely counterintuitive character of the reasoning involved: Holmes has commented that computer proof checking really recommends itself when reasoning with this system, because of the curious interchanges of the two membership relations in applications of comprehension. An important observation is that the proof that double extension set theory is strong turns out to be driven by a not immediately obvious formal resemblance of this theory to the Ackermann set theory discussed above.

The most recent result about this theory (whose consistency remains an entirely open question) is the result of Holmes that the formal symmetry between the two membership relations is broken: the sequence of von Neumann ordinals in terms of one of the relations is a proper initial segment of the sequence of von Neumann ordinals defined in terms of the other.

This theory would have some real interest if it could be showed to be consistent with respect to ZFC or some generally accepted extension thereof, because it is remarkably economical in its notions and axioms. But it is in practical terms rather hard to reason in.

8.2 Zermelo set theory with an elementary embedding

It is a theorem of ZFC due to Kunen [Kunen, 1971] that there is no nontrivial elementary embedding from V to V. The status of ZF with an axiom scheme asserting the existence of such an embedding is unclear (this extension of ZF would be extremely strong).

Kunen's proof depends on examination of the limit of the $j^n(\kappa)$'s, where j is the elementary embedding and κ is the first ordinal moved by j. This suggests that it might be profitable to consider Zermelo set theory with an elementary embedding (which would have a model if it were consistent with ZFC that there is an elementary embedding from a limit rank onto itself, which is among the strongest "axioms of infinity" ever proposed). The Axiom of Rank is useful here, as one wants an elementary embedding which moves a rank, not one which merely moves an ordinal (the latter being not obviously strong in Zermelo set theory).

We present a set of axioms for a first-order theory with primitives equality, membership, and a function j. An equivalent theory was introduced by Paul Corazza in private discussions with one of the authors (Holmes): we present it in a slightly different way. The theory presented in [Corazza, 2000] appears to be essentially the same, but the fact that it is naturally viewed as a version of Zermelo set theory rather than ZFC seems not to be mentioned.

Extensionality: $(\forall AB.(\forall x.x \in A \leftrightarrow x \in B) \rightarrow A = B)$

Separation: For any formula ϕ in which the variable A is not free, $(\forall B.\exists A.(\forall x.x \in A \leftrightarrow x \in B \wedge \phi))$. A is called $\{x \in B \mid \phi\}$ as usual.

Power Set: $(\forall x.(\exists y.(\forall z.z \in y \leftrightarrow z \subseteq x)))$

Rank: Every set is included in a rank (referring back to the definition of hierarchy and rank in our discussion of Zermelo set theory and Mac Lane set theory above).

Elementarity: For any formula ϕ in which j is not mentioned and in which no variable other than x appears free, $(\forall x.\phi \leftrightarrow \phi[j(x)/x])$.

Nontriviality: There is a rank r such that $j(r) \neq r$.

Choice: Every nonempty collection of disjoint sets has a choice set.

We have enjoyed stating these axioms minimally. It should be noted that Pairing and Union can be deduced from the axioms given and the Axiom of Rank just as in Zermelo or Mac Lane set theory with Rank. Infinity is not needed because it is straightforward to show that a rank moved by j must be infinite.

Further, it is straightforward to prove that the index κ of the first rank moved by j is inaccessible, Mahlo, weakly compact, measurable, etc. It follows further that the Axiom of Replacement holds for all formulas not mentioning j: the apparent weakness of this system is illusory. It could further be augmented with classes (with the embedding j extending to classes, and classes freely definable using j just as j can be freely used in instances of Separation in the version given here).

9 CONCLUSIONS

We believe that there are two conclusions to be drawn from this survey, one positive and one negative, for a program of alternative set theory.

The positive result is that there are mathematically fluent and useful systems of alternative set theory. The most unequivocal examples of this are quite close to standard set theory: Nelson's IST and the theory ZFC– + AFA are modifications of ZFC specifically designed to facilitate certain kinds of mathematics, and they are fit for their respective purposes. Further, there are at least two systems which at least appear to be conceptually fundamentally different from ZFC in which it is clear that mathematics can be done. These are the positive set theory GPK^+_∞ of Esser and reasonably strong extensions of NFU (practical considerations suggest that at least the axioms of Choice and Counting would be desirable).

The negative result (which on reflection might not be too negative) is that no fundamentally different picture of the mathematical universe emerges from the alternative theories. Of course certain advocates of some of these theories might dispute this. Our exact claim is that all these theories are mutually interpretable in well-understood ways (except where one theory is much stronger than another: we can interpret the alternative set theory in Mac Lane set theory but of course not *vice versa*). There is no mathematics which can be done in one theory but not in another (except where very low consistency strength is a barrier, as in the case of pocket set theory or the alternative set theory). We do *not* agree with the claim made by some that a theory such as NFU can only be understood to be consistent via its interpretation in ZFC (and so that it is not an independent approach to mathematics); we have shown elsewhere that a self-contained motivation of and development of NFU up to the consistency proof can be made entirely within the simple theory of types, and motivated in a way which has nothing to do with ZFC.

The one possible exception is NF itself, simply because we do not know what a model of NF looks like: there might possibly be some fundamental idea quite alien to the ZF-iste view of things found in a construction of a model of NF (though we do not expect this). The situation in NFU + Infinity + Choice is well-understood: the advocate of this theory will discover in the theory of isomorphism classes of well-founded extensional relations a structure precisely analogous to an initial segment of the cumulative hierarchy of the usual set theory. In GPK^+_∞, the ranks of the cumulative hierarchy indexed by isolated ordinals are there to be examined, but there is more stuff not in the hierarchy.

There is no justification for a revolution here: the current hegemony of ZFC as the set theory in which mathematics is done in practice does not hamper the progress of mathematics. The most that can be said is that it might benefit a mathematician interested in foundations to know that things *could* be done differently, and study of one or more of these systems would serve that purpose.

BIBLIOGRAPHY

[Abramsky, 1988] S. Abramsky. "A Cook's Tour of the Finitary Non-Well-Founded Sets", in *We Will Show Them: Essays in honour of Dov Gabbay*, S. Artemov and H. Barringer and A. d'Avila Garcez and L. C. Lamb and J. Woods, eds., vol. 1, pp. 1–18, College Publications, 2005.

[Aczel, 1980] P. Aczel. "On the notion of proposition, truth and set", in *The Kleene symposium*, North-Holland, Amsterdam, 1980.

[Aczel, 1988] P. Aczel. *Non-well-founded sets*, CSLI Lecture Notes no. 14, Stanford, 1988.

[Apostoli and Kanda, to appear] P. Apostoli and A. Kanda. "Parts of the continuum: towards a modern ontology of science", to appear in *The Poznan Studies in the Philosophy of Science and the Humanities*, ed. L. Nowak.

[Apostoli and Kanda, 2000] P. Apostoli and A. Kanda. "Approximation spaces of type-free sets", in *Lecture Notes in Artificial Intelligence*, J. G. Carbonell and J. Siekmann, eds., vol. 2005, pp. 98–105, Springer-Verlag, 2000.

[Barwise and Moss, 1996] J. Barwise and L. Moss. *Vicious Circles*, CSLI Lecture Notes, no. 60, Stanford, 1996.

[Boffa, 1972] M. Boffa. "Forcing et Negation de l'axiome de Fondement", *Mem. Acad. Sci. Belg.* vol. 40, fasc. 7, 1972.

[Boffa, 1988] M. Boffa. "ZFJ and the consistency problem for NF", *Jahrbuch der Kurt Gödel Gesellschaft* pp. 102-106, Wien, 1988.

[Boolos, 1971] G. Boolos, "The iterative conception of set", *Journal of Philosophy* **68**, pp. 215–31, 1971.

[Bowler and Forster, 2009] N. Bowler and T. Forster. "Normal subgroups of infinite symmetric groups with an application to stratified set theory" *Journal of Symbolic Logic* **74**, pp 17–26, 2009.

[Church, 1974] A. Church. "Set theory with a universal set", in *Proceedings of the Tarski symposium*, L. Henkin ed., *Proceedings of symposia in Pure Mathematics*, vol. XXV, American Mathematical Society, pp. 297–308, 1974.

[Corazza, 2000] P. Corazza. "The Wholeness Axiom and Laver Sequences", *Annals of Pure and Applied Logic*, pp. 157-260, October 2000.

[Crabbé, 1982] M. Crabbé. "On the consistency of an impredicative subsystem of Quine's NF", *Journal of Symbolic Logic*, 47, pp. 131-36, 1982.

[Esser, 1997] O. Esser. "Interprétations mutuelles entre une théorie positive des ensembles et une extension de la théorie de Kelley-Morse", Ph. D. Thesis, Université Libre de Bruxelles, 1997.

[Esser, 1999] O. Esser. "On the consistency of a positive theory", *Mathematical Logic Quarterly*, 45, pp. 105–116, 1999.

[Esser, 2004] O. Esser. *Une théorie positive des ensembles*, Cahiers du centre de logique, no. 13, Centre national de recherches en logique, Academia-Bruylant, Louvain-La-Neuve, 2004.

[Fitch, 1948] F. B. Fitch. *Journal of Symbolic Logic*, 13, pp. 95–106, 1948.

[Flagg and Myhill, 1987a] R. Flagg and J. Myhill. "Implication and analysis in classical Frege structures", *Ann. Pure Appl. Logic*, 34, pp. 33–85, 1987.

[Flagg and Myhill, 1987b] R. Flagg and J. Myhill. "κ-continuous lattices and comprehension principles for Frege structures" *Ann. Pure Appl. Logic*, 36, pp. 1–16, 1987.

[Flagg and Myhill, 1989] R. Flagg and J. Myhill. "A type-free system extending ZFC", *Ann. Pure Appl. Logic*, 43, pp. 79–97, 1989.

[Fraenkel, 1922] A. Fraenkel. "Über den Begriff 'definit' und die Unabhängigkeit des Auswahlaxioms, Sitzungsberichte der Preussischen Akademie derWissenschaften, Physikalisch-mathematische Klasse, 1922 pp. 253–257, translated in [van Heijenoort, 1967], pp. 284–289.

[Forster, 1976] T. E. Forster. *N.F.*, Ph.D. thesis, University of Cambridge, 1976.

[Forster, 2006] T. E. Forster. Permutations and wellfoundedness: the true meaning of the bizarre arithmetic of Quine's NF. *Journal of Symbolic Logic* **71**, pp 227-240, 2006.

[Forti and Hinnion, 1989] M. Forti and R. Hinnion. *The consistency problem for positive comprehension principles*, *Journal of Symbolic Logic*, 54, pp. 1401–1418, 1989.

[Forti and Honsell, 1996a] Forti and F. Honsell. "Choice principles in Hyperuniverses", *Ann. Pure Appl. Logic*, 77, pp. 35-52, 1996.

[Forti and Honsell, 1996b] M. Forti and F. Honsell. "A general construction of hyperuniverses", *Theoretical Computer Science*, 156, pp. 203–215, 1996.

[Forti and Honsell, 1983] Forti, M., and F. Honsell, "Set theory with free construction principles," Annali Scuola normale superiore, Pisa, classe di scienze, s. 4, vol. 10, pp. 493–522, 1983.

[Gilmore, 1974] P. Gilmore. "The consistency of partial set theory without extensionality", in *Axiomatic Set Theory*, Proceedings of Symposia in Pure Mathematics, vol. 13, Amer. Math. Soc., pp. 147–153, Providence, 1974.

[Gödel, 1940] K. Gödel. *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory*, Princeton University Press, Princeton, 1940.

[Grishin, 1969] V. N. Grishin. "Consistency of a fragment of Quine's NF system", *Sov. Math. Dokl.*, 10, pp. 1387-90, 1969.

[Hailperin, 1944] T. Hailperin. "A set of axioms for logic", *Journal of Symbolic Logic*, 9, pp. 1-19, 1944.

[Henson, 1973] C. W. Henson. "Type-raising operations in NF", *Journal of Symbolic Logic*, 38, pp. 59-68, 1973.

[Hinnion and Libert, 2003] R. Hinnion and Th. Libert. "Positive abstraction and extensionality", *Journal of Symbolic Logic*, 68, pp. 828–836, 2003.

[Holmes, 2001] M. R. Holmes. "Strong axioms of infinity in NFU", *Journal of Symbolic Logic*, vol. 66, no. 1, pp. 87-116, 2001. (brief notice of errata with corrections appeared later in 2001).

[Holmes, 1998] M. R. Holmes. *Elementary Set Theory with a Universal Set*, volume 10 of the Cahiers du Centre de logique, Academia, Louvain-la-Neuve (Belgium), 1998. A revised and corrected version is available online.

[Holmes, 2004] M. R. Holmes. "Paradoxes in double extension set theories", *Studia Logica*, vol. 77, pp. 41-57, 2004.

[Holmes, 2005] M. R. Holmes. "The structure of the ordinals and the interpretation of ZF in double extension set theory", *Studia Logica*, vol. 79, pp. 357-372, 2005.

[Jensen, 1969] R. B. Jensen. "On the consistency of a slight (?) modification of Quine's 'New Foundations'," *Synthese*, 19, pp. 250-63, 1969.

[Kisielewicz, 1998] A. Kisielewicz. "A very strong set theory?", *Studia Logica* 61, pp. 171-178, 1998.

[Kelley, 1976] J. Kelley. *General Topology*, appendix, Springer-Verlag 1976.

[Kemeny, 1950] J. G. Kemeny. *Type theory vs. set theory* (abstract). Journal of Symbolic Logic 15, p. 78, 1950.

[Barwise, 1975] J. Barwise. *Admissible Sets and Structures*, Springer-Verlag, 1975.

[Levy, 1959] A. Levy. "On Ackermann's set theory". *J. Symb. Logic*, 24, pp. 154-166, 1959.

[Libert, 2004] T. Libert. "Semantics for Naïve Set Theory in Many-valued Logics: Technique and Historical Account", in *The Age of Alternative Logics*, J. van Benthem and G, Heinzmann eds, Kluwer Academic, 2004.

[Libert, 2008a] T. Libert. "Positive abstraction and extensionality revisited", *Log. Anal.*, 51, pp. 149–165, 2008.

[Libert, 2008b] T. Libert. "Positive Frege and its Scott-style semantics", *Math. Log. Quarterly*, 54, pp. 410–434, 2008.

[Libert and Esser, 2005] T. Libert and O. Esser. "On topological set theory", *Mathematical Logic Quarterly*, 51, pp. 263–273, 2005.

[Mac Lane, 1986] S. Mac Lane. *Mathematics, Form and Function*, Springer-Verlag 1986.

[Malitz, 1976] R. J. Malitz. "Set Theory in which the Axiom of Foundation Fails", Ph. D. thesis, University of California, Los Angeles, 1976.

[Manakos, 1984] Manakos, J., "On Skala's set theory", *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 30, pp. 541-546, 1984.

[Mathias, 2001a] A. R. D. Mathias. "The strength of Mac Lane set theory", *Annals of Pure and Applied Logic*, Volume 110, Number 1, pp. 107-234, June 2001.

[Mathias, 2001b] A. R. D. Mathias. "Slim models of Zermelo set theory", *Journal of Symbolic Logic*, 66, pp. 487–496, 2001.

[Nelson, 1977] E. Nelson. "Internal set theory, a new approach to nonstandard analysis", *Bulletin of the American Mathematical Society*, 83, pp. 1165-1198, 1977.

[Potter, 2004] M. Potter. *Set Theory and its Philosophy*, Oxford University Press, 2004.

[Quine, 1937] W. V. O. Quine. "New Foundations for Mathematical Logic", *American Mathematical Monthly*, 44, pp. 70-80, 1937.

[Quine, 1945] W. V. O. Quine. "On ordered pairs". *Journal of Symbolic Logic*, 10, pp. 95-96, 1945.

[Ramsey, 1926] F. Ramsey. "The foundations of mathematics", *Proc. Lond. Math. Soc. 2*, vol. 25, p. 338-384, 1926.

[Reinhardt, 1970] W. Reinhardt. "Ackermann's set theory equals ZF". *Ann. of Math. Logic*, 2, pp. 189-249, 1970.

[Kunen, 1971] K. Kunen. "Elementary embeddings and infinitary combinatorics", *Journal of Symbolic Logic* vol. 36, pp. 407-413, 1971.

[Robinson, 1966] A. Robinson. *Nonstandard Analysis*, North-Holland, London, 1966.

[Rosser, 1953] J. B. Rosser. *Logic for Mathematicians*, second edition, Chelsea, New York, 1973.

[Russell, 1903] B. Russell, *The Principles of Mathematics*, appendix, W.W. Norton Company, New York, 1996.

[Scott, 1955] D. Scott. "Definitions by abstraction in axiomatic set theory", *Bull. Amer. Math. Soc.*, vol. 61, p. 442, 1955.

[Scott, 1975] Scott, D., "Combinators and classes", in *λ-Calculus and computer science*, Lecture Notes in Computer Science, vol. 37, C. Böhm, ed., pp. 1–26, Springer, Berlin, 1975.

[Seldin, 2009] J. Seldin. "Logic from Church to Curry", in *Logic from Russell to Church*, Dov M. Gabbay and John Woods, eds., pp. 819-874. Volume 5 of the *Handbook of the History of Logic*. Amsterdam: North-Holland, 2009.

[Skolem, 1960] T. Skolem. "Investigations on a comprehension axiom without negation in the defining propositional functions", *Notre Dame Journal of Formal Logic*, 1, pp. 13–22. 1960.

[Solovay, preprint] R. Solovay. *The consistency strength of NFUB*, preprint, available through logic e-prints on the WWW.

[Skala, 1974] H. Skala. "An alternative way of avoiding the set-theoretical paradoxes", *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 20, pp. 233–237, 1974.

[Specker, 1953] Specker, E. P., "The axiom of choice in Quine's 'New Foundations for Mathematical Logic' ", *Proceedings of the National Academy of Sciences of the U. S. A.*, 39, pp. 972-5, 1953.

[Tarski and Givant, 1987] A. Tarski and S. Givant, *A Formalization of Set Theory without Variables*, American Mathematical Society, Providence, 1987.

[van Heijenoort, 1967] J. van Heijenoort, (ed). *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, Jean van Heijenoort, ed., Harvard University Press, 1967.

[von Neumann, 1925] von Neumann, John, "An Axiomatization of Set Theory" (appeared 1925): English translation in [van Heijenoort, 1967].

[Vopěnka, 1979] P. Vopěnka. *Mathematics in the alternative set theory*, Teubner-Verlag, Leipzig, 1979.

[Wang, 1970] H. Wang. *Logic, Computers, and Sets*, p. 406, Chelsea, 1970.

[Weydert, 1989] E. Weydert. "How to Approximate the Naïve Comprehension Scheme Inside of [*sic*] Classical Logic", Ph. D. thesis, Bonner Mathematische Schriften Nr.194, Bonn 1989.

[Whitehead and Russell, 1910-13] A. N. Whitehead and B. Russell, *Principia Mathematica*, 3 vols, Cambridge University Press, 1910, 1912, and 1913. The abridgment is generally easier to find: Whitehead, Alfred North, and Russell, Bertrand, *Principia Mathematica (to *56)*, Cambridge University Press, 1967.

[Wiener, 1914] Wiener, Norbert. "A simplification of the logic of relations", *Proceedings of the Cambridge Philosophical Society*, 17, pp. 387-390, 1914.

[Zermelo, 1914] E. Zermelo. "Untersuchen über die Grundlagen der Mengenlehre I", *Mathematische Annalen*, 65, pp. 261-281, 1908.

INDEX