Proper Classes: Understanding Ontology Through Paradox.

Author: Neil Alexander Barton
Institution: University College London
Degree: MPhilStud Philosophical Studies

I, Neil Alexander Barton, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Contents

List of Figures

Abstract

Engaging in the philosophical debate surrounding proper classes is a tough task. The waters are both muddy and treacherous. In particular, there is no full and rigorous analysis of the logical geography of the area; what one should accept if one makes certain assumptions about ontology. The water would be substantially cleaned up (if not made any less treacherous) by a thorough and comprehensive treatment of just this.

This thesis can be seen as one part of a comprehensive study. The methodology is to precisely state some ontological assumptions, and then to examine how one should characterise proper classes *given these assumptions.* In Chapter 1 I outline the paradoxes I will be considering and the assumptions I am making. I suggest that a better philosophical understanding of ontology would allow us to motivate a solution to the paradoxes. In Chapter 2 I consider (and reject) just such an ontological thesis; the conception that proper classes are ontologically 'heavyweight'. In Chapter 3 I examine some views that attempt to characterise proper classes using modal resources. I then give reasons as to why I find these standpoints unsatisfactory. Chapter 4 provides an analysis and defence of the stance that proper classes are artefacts of plural reference.

The thesis of this work is that if one holds a certain kind of realism about mathematical ontology, then (under the Iterative Conception of set) proper classes do not exist. However, reference to proper classes can be understood as plural reference to some sets.

Chapter 1

Introduction

1.1 Classes and Paradox.

Firstly, what do I mean by class[1]? We shall see that the notion is more complex than one might immediately expect. Here, however, is a pre-theoretically plausible definition:

> **Definition.** A is a *class* iff A is a collection of 0 or more[2] objects such that for any object x it is definite whether or not x is a member of A[3][4].

The notion of 'definite' given above is one that will recur throughout this thesis. It will therefore serve to be precise about what I mean by the term.

> **Definition.** A proposition P is *definite* iff there is a fact of the matter whether or not P.

So, if for any object x it is definite whether or not x is a member of a class Y, then there is a fact of the matter whether or not x is a member of Y.

[1] I use the term *class* rather than *set* as I shall reserve 'set' for those classes that form sets under the Iterative Conception of set.

[2] I use the locution '0 or more' to respect the fact that I consider the empty class to be a class.

[3] This definition is largely inspired by Cantor; see [Cantor, 1883].

[4] A note on terminology; I use the upper-case letters from the start of the English alphabet ('A', 'B', 'C' etc.) to represent classes, and the lower-case letters from the end of the English alphabet ('x', 'y', 'z' etc.) to denote objects (including sets). The difference between sets and classes will be explained later. The use of the upper-case letters 'X', 'Y', 'Z' etc. is reserved for second-order logic, with the exception of 'P' which denotes a proposition.

Given the metamathematical use to which we can put reasoning about classes[5], we should have cause for concern if our class-theoretical thinking turned out to be inconsistent. As is well known, 'naïve' formulations of class theory have just this fatal flaw.

1.2 The Paradoxes.

I will examine the following paradoxes:

- 1.2.1 Russell's Paradox.
- 1.2.2 Cantor's Paradox.
- 1.2.3 The Burali-Forti Paradox.

1.2.1 Russell's Paradox.

Russell's Paradox has received a great deal of attention[6]. This is partly due to its historical significance, the paradox was what infamously brought down Frege's original system. However, it is also because the derivation of the contradiction is quite simple and elegant.

It will serve first to introduce some notation:

Notation. By '$\{x : \phi(x)\}$' I mean the 'class' of all x such that $\phi(x)$.

The paradox stems from considering the predicate 'x is not a member of itself' (in modern notation '$x \notin x$'). If such a predicate succeeds in defining a class of all objects that are not members of themselves (call it 'R') we would have a class such that $(\forall x)(x \in R \leftrightarrow x \notin x)$. Assuming that R is an object, we may substitute 'R' for 'x' in '$(\forall x)(x \in R \leftrightarrow x \notin x)$'. Then we get;

[5]For excellent reviews of a fraction of its uses see [Giaquinto, 2002] and [Fraenkel et al., 1973].

[6]Including from Russell himself. The following quotation speaks volumes about the difficulties he experienced while trying to solve his own paradox (among others); "I was trying hard to solve the contradictions mentioned above. Every morning I would sit down before a blank sheet of paper. Throughout the day, with a brief interval for lunch, I would stare at the blank sheet. Often when evening came it was still empty.....It was clear to me that I could not get on without solving the contradictions, and I was determined that no difficulty should turn me aside from the completion of *Principia Mathematica*, but it seemed quite likely that the whole of the rest of my life might be consumed in looking at that blank sheet of paper." ([Russell, 1967], p151).

$$R \in R \leftrightarrow R \notin R$$

which is clearly a contradiction.

It is possible to restate the paradox informally. Assuming that R exists, we may ask whether or not it is a member of itself, i.e. is it the case that $R \in R$? If it is a member of itself, then it is in the class of all non-self-membered objects (i.e. R), and hence it is not a member of itself. If, on the other hand, it is not a member of itself, it will be a member of the class of all non-self-membered objects i.e. itself, and thus R will be self-membered. Hence R is a member of itself if and only if it is not a member of itself.

1.2.2 Cantor's Paradox.

Cantor's paradox is, like Russell's paradox, a very clear and straightforward piece of reasoning. Central to the paradox is the notion of *cardinality*:

Definition. Two classes A and B have the same *cardinality* (number of members) iff there is a *bijection* between them.

Definition. A *bijection* between class A and class B is a function with the following three properties:

- i) Total-the function maps every member of A to a member of B.
- ii) Injective-the function maps no two members of A to the same member of B.
- iii) Surjective-every member of B is in the range of the function.

Definition. Class A has a *greater* cardinality than class B iff:

- i) there is no bijection between A and B
- ii) there is a bijection between B and a proper subclass of A.

To generate this paradox, one must consider the class of all classes. If one examines the cardinality of this class (the 'Universal class') and compares it to the cardinality of the class of all subclasses of the Universal class, then one is lead to contradiction.

Let the cardinality of the Universal class be denoted by '$|U|$'. Now consider the power class (i.e. class of all subclasses) of U denoted by '$\mathcal{P}(U)$'. Further, Cantor's Theorem states that for any class A that is a class of all subclasses of some class B, A must have greater cardinality than B. Therefore $|\mathcal{P}(U)|$ is a greater cardinality than $|U|$. However, $\mathcal{P}(U)$ is a class containing only subclasses of the class of all classes. As every subclass is also a class, $\mathcal{P}(U)$ only has classes as members. Therefore, everything in $\mathcal{P}(U)$ is also in U. Thus $\mathcal{P}(U)$ cannot be any more numerous than U. Now we have a contradiction; $|\mathcal{P}(U)|$ both is and is not a greater cardinality that $|U|$.

1.2.3 The Burali-Forti Paradox.

The Burali-Forti paradox is another paradox that has had great historical significance. Moreover, the concept used to generate the paradox (that of *ordinal*) is precise and very mathematically important.

Before going through the paradoxical reasoning, it will be useful to introduce some additional terminology:

Definition. A class A is said to be *well-ordered* by relation R iff

1. For any two elements a and b of class A, the following holds:
   (a) Trichotomy-Exclusively either i) $a\,R\,b$, ii) $b\,R\,a$, or iii) $b = a$.
   (b) Transitivity-If $a\,R\,b$ and $b\,R\,c$, then $a\,R\,c$.
2. Any non-empty subclass of A has an R-least member.

Definition. An *ordinal* is the *order-type* of a class well-ordered under a relation R.

Definition. A function $f(x)$ is an *isomorphism* from class A under relation Q, to class B under relation R iff it is a bijection from A to B such that for any $a^1, a^2 \in A$, $a^1\,Q\,a^2$ implies that $f(a^1)\,R\,f(a^2)$.

Definition. A *section* of a class A well-ordered by relation R is a class S such that for some member x of A, S is the class of all R-predecessors of x in A (i.e. S is a *section* of A iff for some $x \in A$, $S = \{y : y \in A \wedge y\,R\,x\}$).

Now, consider the class of all ordinals (denoted by 'Ω'). It can be shown that Ω is well-ordered by the following relation. For any order-types α and β, where α is the order-type of a class X well-ordered by relation R^1, and β is the order-type of class Y well-ordered by relation R^2, let R be a relation such that;

> $\alpha \, R \, \beta \leftrightarrow_{df}$ [there is an isomorphism from X under R^1 to a proper initial segment of Y under R^2]

It is easy to prove that such a relation would well-order Ω (assuming such a class exists). Thus Ω has its own order-type, to be denoted by '$ord(\Omega)$'.

However, it is a theorem that every section of a well-ordered class is well-ordered. Therefore, the section of ordinals less than $ord(\Omega)$ has a certain ordinal; let it be the denotation of '$ord(S)$'. It is also a theorem that every ordinal α is the ordinal of the section of all ordinals less than α. Therefore, $ord(S) = ord(\Omega)$, and hence $\neg ord(S) \, R \, ord(\Omega)$ (by the exclusivity of 1. (a) in the definition of well-ordering).

It is also a theorem that for any section C of a class A well ordered by R, it is the case that $ord(C) \, R \, ord(A)$. Therefore, $ord(S) \, R \, ord(\Omega)$. Thus we have a contradiction; it is both the case that $\neg ord(S) \, R \, ord(\Omega)$ and $ord(S) \, R \, ord(\Omega)$[7].

There are many variants and other kinds of class paradoxes, but the three just outlined are fairly representative of what sort of reasoning is involved in generating a class paradox and are those that feature most widely in the literature. I will, therefore, restrict myself to only considering these three.

1.2.4 Attacking the Paradoxes.

In order to see what a solution must achieve, it will be instructive to understand why the class paradoxes are so serious, and in what sense they are 'paradoxes'.

The sense in which the class paradoxes are 'paradoxes' is simple enough; they all proceed from (pre-theoretically) plausible assumptions, via seemingly legitimate reasoning to patently false conclusions (contradictions). The reason they present such a threat to our class-theoretic reasoning is that initial investigation reveals no clear fallacy; there is no assumption or inferential step that is obviously faulty.

[7] The derivation of a contradiction on the assumption that there is a class of all ordinals can also be produced using the von Neumann representation of ordinals used in modern Set Theory.

I will now consider what we can expect from a satisfactory solution, and the methodological routes available for blocking the paradoxical reasoning.

1.2.4.1 What Constitutes a Satisfactory Solution?

It is not enough to merely ban the paradoxical reasoning from our class theory on pain of paradox. Such reasoning would be, as Michael Dummett puts it, to merely "wield the big stick" ([Dummett, 1994], p26). Pointing to the problem as evidence for one's solution is not to explain the problem.

In order to be satisfactory, a solution must do more. In order to truly *understand* the paradoxes, we must be sensitive to the features of the mathematical structure about which we are reasoning, and the sense in which we have failed to accurately describe that structure. Therefore, I see four obvious constraints on a satisfactory solution:

1. *Precision*-A solution must be *precise* in that it must identify in which respect our thinking is faulty.

2. *Motivation*-Not only must a solution identify which part of our thinking is faulty, it must also *motivate* this choice of error. Independent reasons must be given regarding the piece of reasoning selected as defective. As such a solution cannot merely be an *ad hoc* ban on the paradoxical reasoning.

3. *Diagnosis*-A solution should also *diagnose* why we have fallen into error in the first place. In this way a solution should reveal a pathological element; it should explain why the faulty reasoning initially seemed so appealing.

4. *No Overkill*-A solution must also avoid *overkill* ([Kirkham, 1995], p273)[8]. It must be minimal in the sense that it should not prohibit accepted and valid forms of reasoning.

With these constraints[9] on possible solutions in place, what are the basic methodological options for locating the error in the paradoxes?

[8] While Kirkham is considering paradoxes of truth rather than class paradoxes, this nonetheless seems like a reasonable constraint to put on a solution to any kind of paradox, including the class paradoxes.

[9] I am by no means claiming that this list is exhaustive of good-making features of responses to paradox. Nor do I claim that they are sufficient for a solution to be satisfactory. Indeed we shall

1.2.4.2 Locating the Error.

I see two main ways of solving any paradox:

1. Deny that the conclusion is actually untrue.

2. Attack a principle used in the derivation of the paradoxes.

(1.) is a route that some have taken (e.g. [Priest, 2002]). It is, however, a very tough bullet to bite. There is little that could be more puzzling or mysterious than the hypothesis that a contradiction is true. Almost any other revisionary claim seems to be more intelligible than the claim that there are true contradictions. In this way, such solutions seem themselves to be paradoxical; the medicine is just as bad as the disease.

(2.) seems a more plausible avenue of inquiry. It is noticeable that all the class paradoxes listed require a principle that if a predicate is precise, there exists a class of all objects that satisfy the predicate. Not only this, but these classes must be *objects* about which one may ask questions of membership, and to which one may apply operations (for example, when one takes all subclasses in Cantor's Paradox). Let us now make a distinction; let a *set* be a definite collection of objects that is itself an object distinct from its elements. Let a *class* be a definite collection of objects that may or may not be an object. It is clear that the derivation of the class paradoxes require us to be able to reason with classes as objects (i.e. sets) and so ask whether they are members *et cetera*. In the case of the Russell paradox, one requires a principle that implies the existence of a *set* of all objects that satisfy the condition $\phi(x)$ such that $\phi(x) \leftrightarrow_{df} x \notin x$. Similar considerations apply in the case of the Burali-Forti paradox ($\phi(x) \leftrightarrow_{df} x = Ord(A)$ for some A), and Cantor's paradox ($\phi(x) \leftrightarrow_{df} x$ is a set).

Such a principle is known as the *Comprehension Principle* and can be formalised as follows:

[COMP] $(\exists C)(\forall x)(x \in C \leftrightarrow \phi(x))$

see that there are some solutions to the class paradoxes that fare quite well with respect to these constraints, but are unsatisfactory for other reasons. This list of *desiderata* is useful, however, for seeing what is good about various responses, and why. For this reason, I will apply them when giving a critical analysis of each proposal I consider.

This states that for any condition $\phi(x)$ there is a set C such that $x \in C \leftrightarrow \phi(x)$. This principle is not explicit in the above presentation of the paradoxes, but is tacitly assumed. At the very least this tacit assumption of [COMP] requires justification. Reasons must be provided for holding that given any precise predicate there is a set of all objects that satisfy it. If one were to deny this principle, one would not be able to justifiably assume the existence of the Russell set, set of all ordinals, or universal set in deriving the contradictions.

We might then seek to restore the consistency of our class theory by denying [COMP]. However, we should keep in mind the above constraints; it would be easy to block the paradoxical reasoning with no justification. Further, [COMP] responds to a strong intuition about sets, specifically that we define sets by using *conditions.* Any class theory that denies [COMP] will want to provide a way of preserving part of this intuition. Indeed, one may modify [COMP] in several ways to restore consistency. How do we choose between the options available? A methodological route one might take in order to avoid falling foul of the above *desiderata* is to try and motivate a solution from *ontology.* If we can understand the nature of sets, then it will be easier to see which of the alternatives available we should endorse. This is the methodology that I will adopt throughout the rest of this thesis.

A comprehensive study of how ontology affects the solutions available is desirable. Unfortunately, space does not permit a full examination of every ontological standpoint. I will, therefore, restrict myself to a critique based on two assumptions about ontology.

1.3 Assumptions of the Thesis.

1.3.1 The Iterative Conception of Set.

The first assumption concerns what sets exist within the universe of Set Theory. The modern conception of set is the *Iterative Conception* of set. In this structure we begin from a base of atomic objects (*urelemente*), and proceed in a series of stages by taking repeated applications of the 'set of' operation.

The conception is succinctly expressed by Shoenfield thus:

"Sets are formed in *stages*. For each stage S there are certain stages which are before S. At each stage S, each collection consisting of sets formed at stages before S is formed into a set. There are no sets other than the sets which are formed at the stages." ([Shoenfield, 1977], p323)

We can see, therefore, that the Iterative Conception of set can be represented as a *hierarchy*. Starting with a definite collection of non-sets as *urelemente*, the sets are built up in stages by constructing all the sets one can from objects available at earlier stages. Thus a set x is in a stage S if and only if each member of x is in a stage before S. The stages are *cumulative* in that every stage contains all the objects from previous stages in addition to the sets formed at that stage.

However, we may disagree on what *urelemente* there are. This is not a problem[10] for the investigation of Set Theory; there are some sets that will be in the hierarchy no matter what *urelemente* are taken. These sets are the *pure* sets:

Definition. A set is *pure* iff all its elements are pure.

Now, as the Iterative Conception begins with a definite collection of *urelemente*, we stipulate that this collection must be a set. It is reasonable, if we have a set, to apply [COMP] restricted to that set. Such a principle is Zermelo's Principle of Separation (or *Aussonderung*):

[SEP] $(\forall x)(\exists y)(\forall z)[z \in y \leftrightarrow (z \in x \wedge \phi(x))]$

Let ϕ be some condition. This axiom schema states that for any set x there is a set y of just the elements within x that satisfy ϕ. We are then guaranteed the existence of the empty set by taking the set of all non-self-identical *urelemente*. The empty set is vacuously pure (as it has no members). Thus the pure sets are those sets which are constructed from the empty set. Since the paradoxes can be constructed with attention restricted to the pure sets, I will (for reasons of simplicity) be primarily concerned with the pure sets. From this point on I will refer to the universe of pure

[10] Though there are interesting differences (both philosophical and technical) between theories that admit *urelemente* and those that do not. Considerations of space prevent an account of such issues here.

sets as the *Cumulative Hierarchy*[11]. We can pictorially represent the Cumulative Hierarchy in the following way:

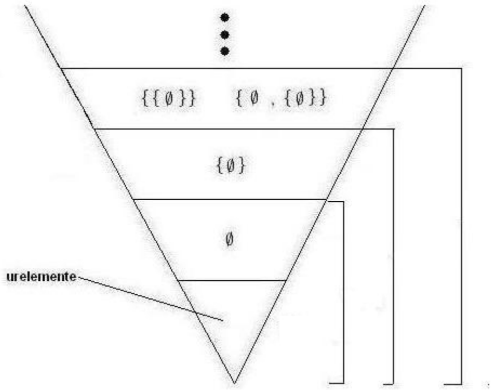

Figure 1.1: An initial segment of the Cumulative Hierarchy.

Assuming that we accept that we may use the sequence of ordinals in our reasoning[12], we can then index the stages as folows:

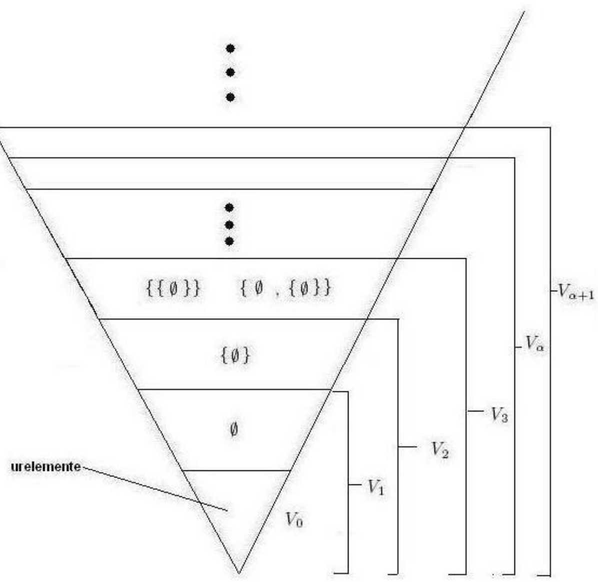

Figure 1.2: Indexing the stages.

[11]It should be noted that it is possible to define a non-well-founded (and hence non-identical to the Cumulative Hierarchy) structure that is defined in an iterative manner. See [Forster, 2008] for such an example. For the sake of simplicity, however, I will assume that the Cumulative Hierarchy is the intended structure defined by the Iterative Conception, and will not consider 'deviant' iterative structures.

[12]I shall use the lower-case Greek letters ('α', 'β', 'γ' etc.) to denote ordinals.

Here we can see that for any stage V_α there is a further stage $V_{\alpha+1}$ such that $V_{\alpha+1} = V\alpha \cup \mathcal{P}(V\alpha)$. Next, assume that there is a set of all natural numbers, denoted by 'ω'. It is clear that ω is well-ordered by the less-than relation on the natural numbers. Thus ω has an ordinal. Therefore, there is a stage V_ω such that:

$$V_\omega = \cup V_{n,n\in\mathbb{N}}$$

Then once more we can proceed as follows:

$$V_{\omega+1} = V_\omega \cup \mathcal{P}(V_\omega)$$

At stages indexed by a limit ordinal (an ordinal that is neither zero nor the successor of any ordinal) take the union of all previous stages thus:

$$V_\lambda = \cup V_{\beta,\beta<\lambda}$$

A fuller pictorial representation of the Cumulative Hierarchy is, therefore, the following[13]:

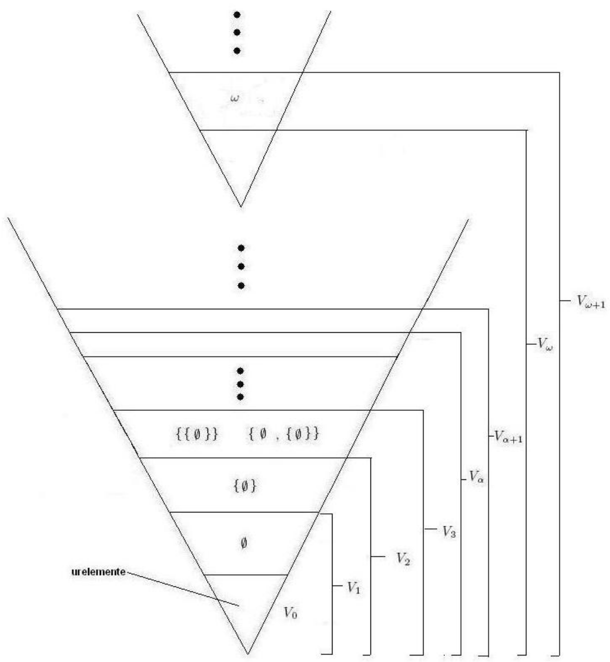

Figure 1.3: The Cumulative Hierarchy

[13]Obviously, there will be many more sets formed at $V_{\omega+1}$ than I can write down. The representation is useful for seeing the rough structure of the Cumulative Hierarchy.

The Iterative Conception is appealing both technically and philosophically. From a technical perspective, any theory that describes the sets as given to us by the Iterative Conception will not allow the paradoxical reasoning. For the conditions used in deriving all the paradoxes I have considered share the following property:

> **Definition.** A condition ϕ is *extendible* iff for any set x of ϕs in the Cumulative Hierarchy, there is a y in the Cumulative Hierarchy such that $\phi(y)$ and y is not in x.

One can see then that the paradoxical classes do not appear as sets at any point in the Cumulative Hierarchy. For if we take any set of ϕs present at some V_α of the Cumulative Hierarchy, then we can use the above extendibility property of the condition ϕ involved to show that there is some ϕ not in that collection. Therefore, the collection of *all* ϕs has not been formed at that stage. If we think that the Iterative Conception describes all sets, then the derivation of the paradoxes is blocked.

Let us see how the contradictions are prevented with respect to each collection. In the case of Cantor's paradox, for any V_α the set of all sets that exist at V_α will not include some sets from other $V_{\beta>\alpha}$. Thus the Universal class will never appear as a set.

The Russell class also never appears as a set. One can see that any set on the Iterative Conception will satisfy the predicate '$x \notin x$'. To see this, assume that there is an x such that $x \in x$, first formed at V_α. However, for any y in the Cumulative Hierarchy $y \in x$ implies that y was formed at some $V_{\beta<\alpha}$. Given then that $x \in x$, it must be the case that x was first formed at some $V_{\beta<\alpha}$. But this contradicts our assumption that x was first formed at V_α. Thus the following is the case on the Iterative Conception:

$$(\forall x) x \notin x$$

As for any V_α there will always be non-self-membered sets at other $V_{\beta>\alpha}$, the Russell class can never appear as a set.

The Burali-Forti paradox is slightly more complicated. What, after all, is an ordinal? It is normal in Set Theory to represent ordinals by sets (known as *ordinal* *numbers*). As long as it holds that for any well-ordered set x there is an ordinal number that is order-isomorphic to x, we can mimic our reasoning about ordinals with our reasoning about the ordinal numbers. Clearly, there will be different well-ordered sets that one might choose to do this job. I shall use the popular von Neumann representation of ordinals and let '$rep(\alpha)$' abbreviate 'the von Neumann representation of α'[14]:

Definition $rep(0) = \emptyset$.

Definition. $rep(\alpha + 1) = rep(\alpha) \cup \{rep(\alpha)\}$.

Definition. For any limit ordinal λ, $rep(\lambda) = \cup_{\beta<\lambda} rep(\beta)$.

Definition. The well-ordering relation on the ordinals is represented by $\in$ on the ordinal numbers.

From this point on, by 'the ordinal number of α' I mean 'the von Neumann representation of α' (i.e. '$rep(\alpha)$'). One can see that for any ordinal α, $rep(\alpha)$ will occur at $V_{\alpha+1}$. This can be shown by the process of transfinite induction. If we show the following three things:

(i) The ordinal number of 0 appears first at V_1.

(ii) If α is an ordinal and $\alpha + 1$ its successor, then if the ordinal number of α appears first at $V_{\alpha+1}$ then the ordinal number of $\alpha + 1$ appears at $V_{\alpha+2}$ (i.e. $V_{(\alpha+1)+1}$).

(iii) For any limit ordinal λ, if the ordinal numbers of all $\beta < \lambda$ are first formed at stage $V_{\beta+1}$, then the ordinal number of λ appears first at $V_{\lambda+1}$.

Then as all ordinals are either 0, a successor, or the limit of a sequence of successors, then showing these three things is sufficient to show that for any ordinal α, its ordinal number is formed at $V_{\alpha+1}$.

(i) The fact that the ordinal number of 0 appears first at V_1 is immediate.

[14] My choice of the von Neumann ordinal numbers reflects ease of use and the fact that it is the canonical representation.

(ii) Assume that the ordinal number of α is first formed at $V_{\alpha+1}$. Then for every member x of $rep(\alpha)$, x is is first formed at some $V_{\beta<(\alpha+1)}$. Therefore, every member x of $rep(\alpha)$ exists at $V_{\alpha+1}$. Given that every member of $rep(\alpha)$ exists at $V_{\alpha+1}$, and that $rep(\alpha)$ is first formed at $V_{\alpha+1}$, and further that $rep(\alpha+1) =_{df} rep(\alpha) \cup \{rep(\alpha)\}$, then $rep(\alpha+1)$ is first formed at $V_{\alpha+2}$.

(iii) Let λ be the ordinal number of any limit ordinal. Assume that for all $\beta < \lambda$, $rep(\beta)$ is first formed at $V_{\beta+1}$, but that $rep(\lambda)$ is not first formed at $V_{\lambda+1}$. Then either:

a) $rep(\lambda)$ is first formed after $V_{\lambda+1}$.

or b) $rep(\lambda)$ is first formed before $V_{\lambda+1}$.

a) *ex hypothesi* the ordinal number of each $\beta < \lambda$ is first formed at $V_{\beta+1}$. As λ is a limit ordinal, for all $\beta < \lambda$, $\beta + 1 < \lambda$. Since i) we have just shown that for all $\beta < \lambda$, $\beta + 1 < \lambda$, ii) the ordinal numbers of all $\beta < \lambda$ are formed at $V_{\beta+1}$, and we already know that iii) $V_\lambda =_{df} \cup_{\zeta<\lambda} V_\zeta$, then V_λ contains all $rep(\beta)$, $\beta < \lambda$. But $rep(\lambda)$ is defined as $\cup_{\beta<\lambda} rep(\beta)$ and so (if $rep(\lambda)$ has not already been formed) $rep(\lambda)$ will be formed at $V_{\lambda+1}$. So $rep(\lambda)$ must exist at some $V_{\alpha\leq(\lambda+1)}$.

b) assume then that $rep(\lambda)$ is formed at some $V_{\beta<(\lambda+1)}$. Clearly $rep(\lambda)$ cannot be formed at any $V_{\zeta<\lambda}$ as then it would be the case that i) *ex hypothesi* $rep(\zeta)$ is formed at $V_{\zeta+1}$, ii) *ex hypothesi* $\zeta < \lambda$ and so $rep(\zeta) \in rep(\lambda)$, and iii) $rep(\zeta)$ is first formed after $rep(\lambda)$ is first formed and so $rep(\zeta) \notin rep(\lambda)$. Therefore; $(*)$ $rep(\lambda)$ is formed at V_λ. However, as λ is a limit ordinal and any stage indexed by a limit ordinal is defined as $V_\lambda =_{df} \cup_{\zeta<\lambda} V_\zeta$, no new sets are formed at V_λ (the ones that already exist are merely collected into a single stage) contradicting $(*)$.

Thus, for any ordinal α, the ordinal number of α is formed at $V_{\alpha+1}$.

□

Given this fact, it is clear that the class of all ordinal numbers never occurs as a set in the Cumulative Hierarchy. This is because for any set x of ordinal numbers in some

V_α, there will be ordinal numbers not in x first formed at all $V_{\beta>\alpha}$. Therefore, the assumption that there is a set of all ordinals contravenes the Iterative Conception of set; there is no faithful representation of the ordinals that would allow the collection of all representatives to appear in the Cumulative Hierarchy.

Further, the Iterative Conception has philosophical merit. The collection forming practice represented by the structure is a natural one. For example, often mathematical inquiry begins by considering a domain of objects (for simplicity let us consider the natural numbers). We then might notice that certain natural numbers stand in certain relations to one another. 16 is the square of 4 would be one such relation between two numbers. We then want to consider all objects that stand in this relation to one another. We can represent such relations as sets of ordered pairs. Thus we may examine the function $f(x) = x^2$ as a single object, represented by a collection of ordered pairs[15]. However, then we might want to look at collections of these functions that share a common property, such as the class of all exponentiation functions on the natural numbers. So we collect these collections of ordered pairs into different collections. Now, we may want to consider functions on functions, such as the function $f(x)$ that given a representation of a function $g(x)$ will output the representation of the function $gg(x)$. Again we may want to collect these together into functions that share common properties, or define functions on these objects. Thus we see how given a starting collection of objects, we examine collections of these objects, and then collections of these collections and so on. Such a process of mathematical development and collection forming is exactly to what the Iterative Conception responds.

Perhaps then, by respecting our collection forming practice in describing the universe of sets, we motivate the rejection of the status of the paradoxical collections as sets. It seems that it would be desirable if the Iterative Conception was an accurate characterisation of the universe of Set Theory. I will, therefore, make the following ontological assumption:

> **Assumption.** The Cumulative Hierarchy, as described by the Iterative Conception of set, contains all the (pure) sets that exist.

[15] If one is unhappy with the notion of ordered pair as a primitive, it is normal in modern Set Theory to use the Kuratowski representation of ordered pair. We may represent the ordered pair $< x, y >$ by the set $\{\{x\}, \{x, y\}\}$.

This tells us what sets exist. However, there is a second substantial ontological question left unanswered; what is the nature of the existence of the Cumulative Hierarchy and the sets it contains?

1.3.2 Monist Realism.

I will assume the following view that provides an answer to the previous question:

> **Assumption.** (Monist Realism)-The Cumulative Hierarchy has the following properties:
>
> 1. The sets are objects.
> 2. The Cumulative Hierarchy is a 'complete' abstract structure; for every ordinal α, V_α exists. No new V_α are being created.
> 3. For any object x, it is definite whether or not x occurs in the Cumulative Hierarchy.
> 4. There is only one Cumulative Hierarchy of pure sets (although isomorphic copies of it may exist within universes of sets with different *urelemente*).
> 5. There is just one interpretation of the Cumulative Hierarchy. Our quantifiers are not ambiguous; it is possible to quantify over *all* sets.

What then is the sense in which the Cumulative Hierarchy of pure sets is a 'complete' structure? Under Monist Realism every stage V_α of the Cumulative Hierarchy exists as part of an abstract, unchanging, mathematical structure. The sets which form the stages are in turn eternal mathematical objects. *No sets ever 'come into being'.* While the Cumulative Hierarchy is explained using the metaphor of formation, it is not undergoing constant construction, it is rather (to put it metaphorically) 'finished' (and always was). Moreover, there is no way in which the universe could be extended by adding more (pure) sets.

This conception of the Cumulative Hierarchy has some nice features. It is conceptually and intuitively simple. When we reason about sets, we are reasoning about objects that are part of a structure based on our normal collection forming practice.

This conceptual simplicity makes the semantics for our theories much easier. A constructivist, for example, will have to explain why we believe our set theoretic statements to be true, even if the domain of objects over which she is quantifying is constantly changing. If the Cumulative Hierarchy was undergoing constant genesis then reference to sets would be different on different occasions. One might think that this might have semantic implications, for example the statement $(\exists x)|x| \geq |\omega|$ could have different truth conditions depending on whether or not we have constructed V_ω. The statement will be false (on a normal understanding) before the construction of V_ω, and true after. Similar problems apply for someone who thinks that the range of set-theoretic quantifiers is ambiguous. There is no such problem with Monist Realism.

The nominalist about sets also has a great deal of work to do in their semantics, as they have no objects with which they can substantiate their set theoretic claims. Monist Realism allows us to easily ground the truth of statements we make about objects within the Cumulative Hierarchy.

Thus Monist Realism seems to be a desirable view about the ontology of Set Theory. However, it seems that a puzzle remains for a Monist Realist who subscribes to the Iterative Conception of set.

1.4 A Puzzle for the Two Assumptions.

Let us pause for a moment. We have seen that the Iterative Conception is a plausible theory about the nature of the set-theoretic universe; one that seems to provide a basis for rejecting the paradoxes. Monist Realism, it would seem, is a desirable view about ontology. A natural question to ask at this point is the following; 'How compatible are the two assumptions about ontology?'. I suggest that there is an important problem left unsolved by a combination of these standpoints.

Whence then the puzzle? The argument is a simple one. If the Cumulative Hierarchy exists in the manner just described, then for any precise predicate it is definite whether or not it is satisfied by any particular set. Therefore there seems to be a definite collection of just the objects satisfying that predicate.

For example, take the following perfectly precise set-theoretical predicate; 'x is an ordinal number'. The predicate is perfectly precise in the sense that for any object presented, there is a fact of the matter whether or not it satisfies that predicate. Further, it is definite what the Cumulative Hierarchy contains. Therefore, 'the ordinal numbers' picks out a definite range of sets within the Cumulative Hierarchy. Thus 'the ordinal numbers' has definite membership and hence is a class. But the class of all ordinal numbers does not appear at any V_α. Indeed, if the class of all ordinal numbers did appear at some V_α, the Iterative Conception would be inconsistent. I will call classes of this sort *proper classes*.

> **Definition.** A class A is a *proper class* iff A is a class with definite membership that does not appear as a set in the Cumulative Hierarchy.

The problem then presented is the following. There are classes that do not appear in the Cumulative Hierarchy. We lack a philosophical explanation of the ontological nature of these classes. The issue is pressing; under Monist Realism the cogency of the Iterative Conception depends on a satisfactory account. So then, given the two assumptions, how should a Monist Realist characterise proper classes?

Chapter 2

Heavyweight Proper Classes

In this Chapter I analyse the view that proper classes are ontologically 'heavyweight' objects. On this conception, proper classes are collections that are also abstract objects over and above their elements.

- In 2.1 I explain the view to be discussed and present some of its positive features. I suggest that, as it stands, a substantial philosophical question is not answered; 'Why are proper classes not sets?'.
- In 2.2 I examine what form a satisfactory explanation of why proper classes are not sets should take. I suggest that an explanation should provide us with the resources for blocking the paradoxes. I note one way to do this, by justifying the principle that proper classes cannot be members of other class-like entities. I then consider three explanations for why proper classes are not sets that could be used to motivate the principle that proper classes are not members:

  [NMH]-The Non-Member Hypothesis: Proper classes fail to form sets because they 'cannot be members'.

  [LSH]-The Limitation of Size Hypothesis: Proper classes fail to form sets because they are 'too big'.

  [OHH]-The Occurrence in the Hierarchy Hypothesis: Proper classes fail to form sets because they do not occur at any stage of the Cumulative Hierarchy.

- I give a critical evaluation of these explanations. I present some good features of each, but argue that [NMH] and [LSH] are unsatisfactory explanations. I then argue that [OHH] is a satisfactory explanation, but cannot be appealed to by the heavyweight theorist.
- In 2.3 I conclude that the Heavyweight Theorist has no satisfactory answer as to why proper classes fail to form sets.

I shall show that the Heavyweight View, while initially appealing, is untenable for a Monist Realist.

2.1 The Heavyweight View.

2.1.1 What is the Heavyweight View?

It was seen in the last Chapter that if we wish to hold Monist Realism true, we need to give a characterisation of proper classes. A natural starting point is to note some salient features of proper classes.

As argued earlier, the Monist Realist must assert that proper classes have definite membership. Further, it would seem that we are able to reason meaningfully about proper classes and also use talk of proper classes in our discourse about sets. Here are two such examples:

(Example 1) The use of proper classes within Set Theory.

Much work into the theory of large cardinals makes use of proper classes. For example, a measurable cardinal can be defined as the critical point of a non-trivial elementary embedding from the universe into a transitive class M[1]. Here, both M and V are proper classes, and any ordered pairs that could represent this mapping would also be a proper class.

(Example 2) The meaningfulness of proper class talk.

[1] There are other equivalent definitions of a measurable cardinal available (see [Kanamori, 2003], p26 onwards). However, often definitions which have apparent reference to proper classes are used, and it is not obvious that this use is eliminable (see [Uzquiano, 2003] for an argument to this effect). A theory that can account for this use in a simple way would, therefore, be preferable to one that cannot.

I can make assertions about proper classes that seem to be truth evaluable. For example, I might say "The class of all sets is the same class as the class of all non-self-membered sets". It would seem that this statement is false just in case there is a self-membered set, and true otherwise (hence it is true under the Iterative Conception). Here I seem to have talked about a proper class and said something meaningful.

These examples might[2] lead one to believe that we have another notion of collection besides that of set. Maybe the sets are not the only collections that exist. Maybe proper classes are another kind of object-like collection, about which we may construct a mathematical theory. If one has such inclinations one might hold the Heavyweight View:

> [Heavyweight View] Proper classes are collections with definite membership that are objects distinct from their elements, but that are not sets.

Often, proponents of the Heavyweight View will talk of the proper classes being 'above' the Cumulative Hierarchy. This should be regarded as loose talk; a heavyweight theorist is not necessarily committed to the view that proper classes reside in a 'domain' or 'stage' beyond all the stages of the Cumulative Hierarchy.

Thus one admits into the ontology of set theory two different kinds of objects; proper classes and sets. Theoretically this is cashed out in different ways; Gödel, for example, makes do with only variables for 'class', but has a sethood predicate [Gödel, 1940]. Bernays on the other hand has two different kinds of variable; set variables and class variables [Bernays, 1958]. Such considerations are clearly irrelevant to ontology; howsoever one chooses to describe the situation the relevant ontology is the same.

However, different heavyweight theories postulate the existence of different classes. As proper classes are different kinds of things from sets there is no obstacle to holding [COMP] as a principle about classes (here, for the sake of convenience and clarity, the class membership relation is given by 'η' in order to distinguish it from the set membership relation represented by '$\in$'):

[2] I say 'might', because it is not established that the examples entail the Heavyweight View. Indeed it is my opinion that they do not. Examples such are these, however, allow one to see the initial motivation for the Heavyweight View.

[COMP$^\eta$] $\exists A \forall x(x \eta A \leftrightarrow \phi(x))$

In the system $VNBG$ (originally proposed by von Neumann[3] and subsequently refined by Bernays[4] and Gödel[5]) only predicative conditions are allowed in [COMP$^\eta$]. An impredicative condition is one that has a quantifier whose range includes the class being defined in the defining condition. An example of impredicative definition is that of greatest lower bound (or infimum) of a set:

$$inf(S) =_{df} (\iota x)(\forall y \in S)(x \leq y) \wedge (\forall z)[(\forall y \in S)(z \leq y) \rightarrow z \leq x]$$

Here we see that $inf(S)$ itself must fall within the range of the quantifier '$(\forall z)$'. The point can be put informally as follows. The above sentence states that the greatest lower bound of a set S is a lower bound x of S such that for any lower bound z of S z is less than or equal to x. In the previous sentence the phrase "for any lower bound z of S" quantifies over all lower bounds of S of which x (the object being defined) is one. Hence the definition is impredicative.

Other theories allow for impredicative conditions in [COMP$^\eta$]. Morse-Kelley set theory (MK) is just such a theory. Because $VNBG$ restricts [COMP$^\eta$] to predicative conditions where MK does not, MK will posit the existence of more classes than $VNBG$.

While this is indeed a substantial theoretical difference, it is not important for my discussion. The fact that MK posits the existence of more proper classes than $VNBG$ does not change the fact that both theories posit the existence of proper classes. In both theories the Russell class, Universal class, and class of all ordinal numbers exist. This is also the case with other heavyweight theories (such as Ackermann's system[6]). These are the collections for which we are seeking an explanation.

We now have a precise statement of the Heavyweight View, and have identified an extraneous ontological question that should not distract us from the key tenets of the proposal:

[3] A noticeable theoretical difference between von Neumann's original system and the developments made by Bernays and Gödel is the use of the notion of 'function' as primitive rather than those of 'set' and/or 'class'. This fact has no ontological import; there is a total one-to-one correspondence between functions and classes, and the distinction between classes which are sets and proper classes is exactly matched by a distinction among functions.

[4] In a 1931 letter to Gödel and later in his [Bernays, 1958].

[5] See [Gödel, 1940].

[6] Ackermann Set Theory, however, has other interesting features that will be discussed later.

1. Proper classes have definite membership.

2. Proper classes are objects distinct from their elements.

3. Proper classes are a different kind of object from sets.

I will now examine some positive features of the Heavyweight View. Despite these good aspects of the view I will suggest that the heavyweight theorist must give an explanation of why proper classes are not sets.

2.1.2 Positive Remarks about the Heavyweight View.

Though I plan to argue against the Heavyweight View, it could be argued that it fares quite well with respect to the criteria outlined in Chapter 1.

It is possible for the heavyweight theorist to provide a *diagnosis*. To begin with, we think that we only have one notion of 'collection'. It turns out that we have at least two, one of 'set' and one of 'proper class'. Proper classes are similar to sets; they depend for their identity on their members, they have definite membership, and they are abstract objects. It is understandable then, that we might mistake one for another in our naïve reasoning about classes.

The *overkill* constraint is also respected. We may continue to talk about proper classes in the way we intuitively think that we do; by referring to a particular kind of collection. In this way it nicely meshes with Monist Realism; the puzzle of Chapter 1 seems to indicate that there are precise collections which are not sets. For the heavyweight theorist these are proper classes which exist and about which we may talk in a singular manner.

The Heavyweight View seems to provide an ontological basis for much of our reasoning about large cardinals. For example, embeddings from the universe into other proper classes can be represented as proper classes of ordered pairs. In order to be philosophically satisfied (if these grounds were not present) one would have to explain why our (very precise) talk about such things as embeddings was mere *façon de parler*. This is a problem the heavyweight theorist does not encounter.

The Heavyweight View is also able to identify [COMP] as illegitimate when assumed to only apply to sets. Thus it is *precise*. Recall [COMP]:

[COMP] $(\exists C)(\forall x)(x \in C \leftrightarrow \phi(x))$

One might think that '$(\exists C)$' only ranges over sets. If so, this would imply that there is a set of all sets. This is false for the heavyweight theorist; the class of all sets is a proper class and hence a completely different kind of object from any set. There are proper classes, and so the existential quantifier may assert the existence of a proper class rather than a set. In this way the Heavyweight View is *precise* in that it identifies [COMP] as the guilty principle when used as a principle only about sets. While no *motivation* for rejecting [COMP] as a principle only about sets has yet been provided, the Heavyweight View allows for such motivating reasons. This will now be given fuller consideration.

2.1.3 A Problem for The Heavyweight View; Why are Proper Classes not Sets?

A weak objection to the Heavyweight View would be the following. The Heavyweight View does not present a particularly ontologically parsimonious view of the set theoretic universe. In addition to (the already very extensive) universe of sets, we are postulating the existence of more objects. The objection is weak for the following reason. Set Theory is not particularly concerned with ontological parsimony of this kind. Indeed it seems to run counter to the whole investigative process of the large cardinal discussion; the existence of new sets is postulated, and then the consequences of this drawn out. Generally speaking, the existence of additional entities is not of particular theoretical concern within such a vast ontology.

However, one might press a similar thought. Under the Heavyweight View we postulate the existence of an additional ontological *kind* of object. This is significant in the way that postulating the existence of more sets is not. Postulating a distinct kind of entity adds another layer of complexity to one's theory. If this complexity turns out to be redundant, the Heavyweight View would be unappealing.

It is, at this stage of the dialectic, perfectly open to the heavyweight theorist to point out that the existence of proper classes as heavyweight objects is not a redundant hypothesis. They would point to the motivation for holding their view;

the facts that there are precise conditions that seem to define non-set collections within the hierarchy, and that we require an explanation for our reasoning about large cardinal hypotheses.

However, if one is postulating the existence of objects of a different kind, one should have an account of why this kind of object is different from objects of other kinds. Let us pause briefly to compare the ontological nature of proper classes on the Heavyweight View with that of sets. It seems that proper classes and sets are very similar. Both proper classes and sets are objects. Further, both kinds of object have definite membership.

The resemblance between proper classes and sets is uncomfortable. The heavyweight theorist should not be satisfied with an arbitrary distinction between objects of the same kind. An adequate explanation of the difference between proper classes and sets is, therefore, particularly pressing.

I see three main ways a heavyweight theorist might explain why proper classes are not sets:

> [NMH]-Proper classes are not sets because they cannot be members.
>
> [LSH]-Proper classes are not sets because they are 'too big'.
>
> [OHH]-Proper classes are not sets because they do not occur at any stage of the Cumulative Hierarchy.

I shall argue that any of the three explanations provides the heavyweight theorist with the resources to block the paradoxes. However, I will show that none of these is an acceptable explanation for the heavyweight theorist. [NMH] and [LSH] are unsatisfactory explanations in themselves. [OHH], on the other hand, is a good explanation. However, it cannot be appealed to by the heavyweight theorist.

2.2 Explaining why Proper Classes are not Sets.

2.2.1 What a Satisfactory Explanation of why Proper Classes are not Sets Should Achieve.

Before I examine the above hypotheses, we should examine what an explanation should achieve. For there is a difference between an *explanation* and a *true principle*. To see the difference, suppose that I am teaching an elementary school student how to differentiate equations. I require an explanation for why some solutions for the value of the differential are correct, and why some are faulty. Now, unfortunately my student repeatedly forgets that the differential of a constant with respect to x is zero. I try to explain why she is wrong. There are a number of ways I could do this. One is to point out that substitution of values gives her inaccurate results for the gradients of tangents. This (in a sense) provides her with a reason why her method is wrong; it gets the wrong answers. However, there is a deeper sense in which she is wrong, the sense in which her reasoning does not respect the operation being performed when finding the differential of an equation. To explain this sense would require an explanation of how modern calculus uses the notion of the value of a function as it approaches a limit in order to find the derivative of the function.

What can be seen from this example? In the first case we have a *true principle* in the sense that the principle in play will not get a prediction wrong. If one gets erroneous values when applying the equation one has obtained from differentiation then the solution is incorrect (unless we are being truly obtuse and cannot compute values correctly). In the other case, however, we are giving an *explanation* for why differentiating a constant with respect to x should always be zero. It tells us *why* the equation we have found is incorrect; the reasons given are descriptive of the mathematical structure under consideration.

To make an analogy with the class paradoxes, it may very well be the case that will we find true principles that hold of proper classes but not of sets (or *vice versa*). These principles may be very useful in our reasoning; they will tell us for a particular given collection whether or not it is a set or a proper class. However, in order to be philosophically satisfied, we require *explanation* for why proper classes are not sets.

Such criteria will also need to respect the constraints outlined in Chapter 1.

Further, we require explanations that allow us to block the paradoxes. Merely given the statement that proper classes are a different kind of object from sets does not make it clear how the paradoxes should be prevented. One principle that would allow us to block the paradoxes would be the Non-member Principle ([NMP]):

> [NMP] Proper classes are not members of class-like entities.

Let us now examine how each paradox is blocked using [NMP].

Russell's paradox depends on asking whether or not R is a member of itself. By [NMP], proper classes are not the sort of things that can be members, thus we can say that R is not a member of itself. There is no class of all classes that are not members of themselves; such an object would depend on having proper classes as members. However, there is a (proper) class of all *sets* that are not members of themselves.

The Burali-Forti paradox is solved using slightly different terminology between different authors, but nonetheless still admits of a solution. The following is Gödel's solution from [Gödel, 1940]. For Gödel an *ordinal* is a transitive class well-ordered by the membership relation. He makes a distinction, however, between *ordinals* and *ordinal numbers*; the latter are also sets. Thus there is a class of all ordinal numbers, to be denoted by 'On'. On is indeed well-ordered, and hence is an ordinal. It is *not*, however, an ordinal number. There is no class of all ordinals.

Why is there no class of all ordinals? To do so would presuppose that On could be a member of something (as it is an ordinal). Hence a class of all ordinals would violate [NMP].

Cantor's paradox is also answered. The operation of taking all subclasses of a class is clearly an inappropriate operation for proper classes. This is because every class is a subclass of itself (and hence a member of the class of all its subclasses). To allow such an operation for proper classes would thus allow a proper class A to be a member of another class (namely the class of all subclasses of A). Thus [NMP] prohibits the taking of all subclasses of a proper class. In this way the paradox cannot get off the ground; there is no class of all subclasses of U.

Thus we see (on the assumption of [NMP]) that we must be careful which conditions we allow into [COMP$^\eta$]. For some conditions (such as 'x a non-self-membered class', 'x is an ordinal', or 'x is a class') are not satisfactory conditions for proving the existence of classes (even proper classes). If we could justify [NMP] we would also provide motivation for rejecting [COMP] as a principle only about sets. If it were a principle only about sets, then proper classes would be sets and hence members of sets, thus violating [NMP].

Let us take stock. We have seen that the heavyweight theorist requires an explanation for why proper classes are not sets. Further, an explanation that motivated the acceptance of [NMP] would provide the heavyweight theorist with the resources she needs to block the paradoxes. I shall examine three putative explanations that might be put to this task. All, I shall argue, are of no use to the heavyweight theorist.

2.2.2 The Non-member Hypothesis.

Given the use to which we can put [NMP], maybe the heavyweight theorist should appeal to this as explanation for the difference between proper classes and sets. I will call this view the Non-Member Hypothesis ([NMH]):

> [NMH] Proper classes are not sets because proper classes (unlike sets) cannot be members.

This explanation of why proper classes are not sets would clearly allow us to assert that proper classes are not members. [NMP] would be immediate; inferred from the very explanatory feature of the difference between proper classes and sets.

It should be noted that the phrase 'cannot be members' is ambiguous. Does it mean that proper classes cannot be members of sets, or classes in general? The method by which the class paradoxes are blocked ([NMP]) involves the fact that proper classes cannot be members of other proper classes. There (2.2.1) it was noted that one must be careful to not use conditions in [COMP] that would make proper classes members, even of other proper classes. For example, 'x is a non-self-membered class' was just such a condition. It is thus clear that to be effective [NMH] must state that proper classes are not members of classes of all kinds. With this in mind, let us see if [NMH] is a satisfactory explanation.

2.2.2.1 Positive Remarks about [NMH].

It would be highly desirable for [NMH] to be a satisfactory explanation for why proper classes are not sets for the heavyweight theorist. If it were a good explanation, then the method for preventing the paradoxes and explanation for why proper classes are not sets would perfectly mesh.

Further, [NMH] is *precise*; it tells us clearly which part of our reasoning was faulty. As noted earlier in 2.2.1, we might mistakenly allow conditions that allow proper classes to be members into [COMP$^{\eta}$]. This can be identified as the faulty step in the derivation of the paradoxes.

A *diagnosis* can also be given by the [NMH] theorist. Pre-theoretically, it is unclear why we should expect proper classes not to be members of sets. After all they are precise mathematical objects. Given this, why should they not be members?

2.2.2.2 Why [NMH] is Unsatisfactory.

This, however, points the way to an objection to [NMH]. What is the structural feature of proper classes that prevents them being members (aside from paradox)? Proper classes are precise objects with definite membership. Why then should they not be members? We would like our class theory to be the most comprehensive theory of collections possible. Given this we require reasons why there are objects that cannot be taken as members of classes. [NMH] seems to not be a basic explanation, but rather demands justification. It seems *ad hoc*, a putative explanation designed merely to allow us to block the paradoxes. If [NMP] is true, then we want to know *why* proper classes cannot be members.

The situation is made worse by the fact that having proper classes be members is not in itself the root of the contradiction. For there are systems that allow proper classes to be members of other classes but are consistent relative to ZFC (and hence also $VNBG$[7]). Ackermann's A, for example, is a system that allows a proper class to be a member of another proper class, and is a conservative extension of ZFC ([Fraenkel et al., 1973], p153). Given that it is possible to develop such systems, it

[7]This is because $VNBG$ is a conservative extension of ZFC, i.e. all theorems of $VNBG$ in the language of ZFC are theorems of ZFC.

seems unlikely that proper classes being members is in itself the cause of the paradoxes.

For these reasons, [NMH] is poorly *motivated*. There seems to be no good reason to accept why proper classes cannot be members. In fact, there is pressure to accept the converse. We want our class theory to be the most comprehensive account of collections possible. Given that proper classes are (on the Heavyweight View) precise objects, we should be able to talk about collections thereof. Further to this, we can construct consistent systems where proper classes are members. [NMH] thus fails as an explanation for why proper classes are not sets.

2.2.3 The Limitation of Size Hypothesis.

It seems then we not only require an explanation of why proper classes are not sets, but also principled reasons for rejecting the idea that proper classes can be members of other classes. The Limitation of Size Hypothesis attempts to provide the ground for [NMP].

One may state the Limitation of Size Hypothesis as follows:

[LSH] Proper classes are not sets because proper classes are 'too big'.

If [LSH] were true, one could motivate [NMP]. If the reason that proper classes are not sets is that they are too big then it is plausible that a class could not contain a proper class. For, if a proper class is too big to be a set then one might think that it is also too big to be contained within a class.

What is meant by 'too big'? This, as we shall see, is a hard question to answer. However, one might think that it is somehow part of the Iterative Conception that it only allows sets of a certain size. All the (pure) sets that exist are constructed out $\emptyset$, the power set operation, and the process of taking unions at limits. We might expect then that something larger than anything that occurs at any V_α cannot be a set[8].

This is, as it stands, still quite inexact. However, it can be made more precise. First, let '$Set(A)$' stand for 'A is a set' and '$A \approx B$' stand for 'A and B may be correlated one-to-one'. One may then formulate [LSH] in the following way:

[8] I put aside for the moment questions about whether inaccessible cardinals (that cannot be reached in this manner) are sets. However, this fact about defining stronger axioms of infinity will become important later in evaluating [LSH].

[LSH^F] $(\forall A)(Set(A) \rightarrow \neg A \approx On)$

This states that if A is a set it is not possible to correlate A one-to-one with the ordinal numbers. Hence, if one can correlate A one-to-one with the ordinal numbers, then A is not a set (and thus A is a proper class).

We now have a precise formulation of [LSH]. One must now ask whether or not it is satisfactory as an explanation for why proper classes are not sets.

2.2.3.1 Positive Remarks about [LSH].

It should be noted that [LSH] provides a perfectly precise criterion for when an object is a proper class. [LSH^F] is *true*; a class forms a set unless one can map the class onto the ordinal numbers.

Therefore, theories that use [LSH] are *precise*. Firstly, as noted earlier (2.2.2), [COMP] violates [NMP]. As [NMP] is supported by [LSH], it seems then that [LSH] implies the falsity of [COMP]. However, [LSH] is also precise in the sense that it identifies the use of [COMP] as a principle only about sets as erroneous; a class defined by a condition $\phi(x)$ can be of any size. Furthermore, [LSH^F] provides us with a *true principle* about proper classes and sets.

There is also a *diagnosis* available. It is not obvious why considerations of size would be relevant to whether or not a given class is a set. One notion of size, the concept of cardinality, can be quite unusual (especially in the infinite case). Consider the proof that the set of rational numbers is equinumerous with the set of natural numbers. Such a result is, pre-theoretically, quite unexpected. After all, there are infinitely many rational numbers between the natural numbers 0 and 1! Now, one may not ascribe cardinality to proper classes. Rather, proper classes 'transcend' all cardinality. What is meant by this? It can be elucidated by means of an analogy with the real line. Consider the question 'What is the length of the real line?'. It seems this question has no answer. One should say instead that the real line 'transcends' all length; it is so indefinitely long that one cannot ascribe to it a length. This parallels proper classes and cardinality; they are so indefinitely large that they cannot be given a cardinality. The notion of a one-to-one mapping is central, however, both to cardinality and to [LSH^F]. This makes it plausible that in a similar way, both cardinality results (such as the fact that $\mathbb{N} \approx \mathbb{Q}$) and results involving [LSH^F] (for example; the class of all non-self-membered sets is not a set) might be surprising and unexpected.

Despite these good features of [LSH], it is unsatisfactory. It fails as an *explanation* because it is poorly *motivated*. Moreover, there is a worry that it violates *overkill* that (when scrutinised closely) reveals a circularity within [LSH].

2.2.3.2 Why [LSH] is Unsatisfactory.

I will examine the charge of *overkill* first. If we think that the explanation for why proper classes are not sets is [LSH], we should have an account of why each of the axioms that describes the universe of sets does not produce sets that are 'too large'. One of these axioms will be the Axiom of Power Set.

[POWER] $\forall x \exists y \forall z (z \in y \leftrightarrow z \subseteq x)$

As we can see [POWER] asserts that there is a set of all subsets of a given set. There are reasons to want [POWER] as part of our theory, of which I will now survey some[9].

Firstly, it merely seems intuitive to say that we can take the set of subsets of a given set. If we have the set in question, then we have the elements that make up the set, and hence one might think we have the set of all possible combinations of such elements.

Mere intuition, however, is not enough. Nonetheless, it seems that the notion of Power Set is somehow 'written into' the Cumulative Hierarchy. Recall how the Cumulative Hierarchy was defined in Chapter 1; for any V_α there is a stage $V_{\alpha+1}$ such that $V_{\alpha+1} = V_\alpha \cup \mathcal{P}(V_\alpha)$. In order to know about the sets at any $V_{\alpha+1}$ we must effectively consider all subcollections of V_α. The Power Set operation is thus integral to our understanding of the Cumulative Hierarchy.

Further, there is a sizeable amount of mathematical pressure to want [POWER]. As Hallett, notes, the power set of any set x represents the set of all extensions of

[9] There are many more reasons that have been discussed in the literature (see [Hallett, 1984] for discussion). As my project is not to justify the Power Set Axiom, I have not given it a thorough treatment. It is necessary to see, however, that it is a highly desirable principle to have in one's theory, and one that responds to the Iterative Conception.

properties on x ([Hallett, 1984], p206). [POWER] is also very useful in other branches of mathematics; in probability theory, for example, the power set of the sample space is the set of all possible events.

Given that it is desirable to adopt [POWER], how can the [LSH] theorist argue that we should indeed use it as an axiom?

It is clear that power sets will be 'small' in the finite case. I suggest, however, that it is very unclear how someone who holds [LSH] can argue that power sets are small in the infinite case, even for the smallest infinite set.

This set is ω; the set of all natural numbers. Let '$\mathcal{P}(\omega)$' denote its power set. Let us take one measure of set size; cardinality. What can we say about the cardinality of such a class? Is it 'small'?

The fact of the matter is, we have little idea of the cardinality of $\mathcal{P}(\omega)$. One thing we do know (by Cantor's theorem) is that it must be strictly greater than $|\omega|$.

That is, however, pretty much all we know about it. Thanks to the work of Gödel and Cohen ([Gödel, 1940], [Cohen, 1963]), we know that the cardinality of $\mathcal{P}(\omega)$ is independent of the axioms of ZFC[10]. It is impossible to prove, from axioms that arise naturally out of the Iterative Conception, exactly what the cardinality of $\mathcal{P}(\omega)$ is.

The [LSH] theorist may respond that $\mathcal{P}(\omega)$ is in fact 'small', citing the fact that its cardinality is strictly smaller than $|\mathcal{P}(\mathcal{P}(\omega))|$. This is, however, no response; it is precisely the nature of the set $\mathcal{P}(\omega)$ that under scrutiny. This argument requires $\mathcal{P}(\omega)$ to be a set (and hence the sort of thing to which we may apply the power set operation) in order to be effective.

A different response available to the [LSH] theorist is that we cannot map a subclass of $\mathcal{P}(\omega)$ onto the ordinals and thus it is small. This is an acceptable response to the problem of the cardinality of $\mathcal{P}(\omega)$ only if one thinks that not being able to map a class one-to-one with the ordinal numbers is an acceptable measure of set 'smallness'. I think that there is good evidence to suggest that the [LSH] theorist's appeal to the ordinal numbers as the 'measuring stick' of class size is not satisfactory. Indeed it seems to point to a deep philosophical circularity.

[10]Due to the fact that $VNBG$ is a conservative extension of ZFC, this will mean that the cardinality of $\mathcal{P}(\omega)$ is also independent of the axioms of $VNBG$.

Using [LSH^F], how does one show that a class fails to form a set? One must show that one can map a subset of the class onto the ordinals. One may then conclude that the class is at least as big as the ordinals and (as the ordinals are a proper class) the class in question cannot be a set. So far so good. But what reason can the [LSH] theorist give for why the ordinal numbers do not form a set? They might simply say that the ordinal numbers are 'too big'. But how do we know that they are 'too big'? One could note that one can map the ordinal numbers one-to-one with the ordinal numbers. But this is obviously circular! The only other option seems to be to point to the fact that paradox would ensue if we did not assume them to be 'too big'. But this is clearly no response if [LSH] is to provide an *explanation.* [LSH] theory operates by taking a class that one assumes to be too big, and then showing that other classes are too big in relation to this class. But to do so, as Linnebo notes, is merely to "move in a tiny circle" ([Linnebo, 2010], p154).

Given that it is (at the very least) dubious whether or not [LSH] can avoid *overkill*, let us examine its *motivation.*

It seems to me that [LSH] fundamentally mistakes the nature of sets. Why should we think that considerations of size have bearing on whether or not a class is a set?

Sets are abstract objects, some of which are extremely large. Why should they reach a certain point and then 'overflow'? For any set in the Cumulative Hierarchy there is always another set of greater cardinality. Paradoxes arise because we naïvely hold two or more contradictory principles about a notion. It must be that when we assume that a proper class is a set we have inadvertently accepted a proposition that entails the negation of one we already hold. It is hard to see, once we accept the varying cardinalities and sheer enormity of the transfinite sets, how it should be the *size* of proper classes that is the root of the contradiction.

This point is backed up by current work in Set Theory. It is a fact of set theoretical practice that where set theorists can they have defined larger and larger sets. This is done by extending the standard axioms of ZFC by so called 'large cardinal axioms' that assert the existence of *very* large sets. Such practice can be done (we think consistently), against the backdrop of the Cumulative Hierarchy. Large cardinal axioms effectively operate like the Axiom of Infinity that asserts the existence of the first infinite set, ω. They simply posit that other (stronger) kinds of infinite set exist, and the Cumulative Hierarchy can then proceed from these objects. If size considerations were somehow 'written into' the Iterative Conception, one would think that this would not be possible. Given that we can define sets of larger and larger size, it seems arbitrary to insist that size is the key difference between sets and proper classes.

[LSH] then, while associated with a *true principle*, is not *explanatory.* It seems unable to account for [POWER], looks unavoidably circular, and fails to chime with our basic notion of set. The heavyweight theorist should, therefore, look elsewhere in drawing the distinction between proper classes and sets.

2.2.4 The Occurrence in the Hierarchy Hypothesis.

However, examination of [LSH] reveals a different explanation one might give for why proper classes do not form sets.

While [LSH] was seen to be unsatisfactory, one can learn a great deal from the way it was expressed by its original proponents.

Bar-Hillel, Fraenkel, and Levy express the view as follows:

> "...we do not admit very comprehensive sets in order to avoid the antimonies" ([Fraenkel et al., 1973], p135)

Fraenkel himself (in [Fraenkel, 1927]) states that (if we use his axioms):

> "...the scope of the new set is never boundless" (p116)

and also

> "...in this way, the possibility that a set could be constituted by the completely limitless assignment of elements is avoided from the beginning" (p118)

When expressing [LSH], its advocates used words such as 'bound', 'limit', and 'extent'. There is a sense in which the fact that proper classes not being sets is bound up with such notions. The reason proper classes are not sets is that (owing to the extendible nature of the conditions by which they are defined) they do not appear at any V_α. This was noted to be a structural feature of the Iterative Conception.

Perhaps then the heavyweight theorist should make us of this fact and appeal to the following explanation for why proper classes are not sets:

> [OHH] Proper classes are not sets because sets occur at some V_α of the Cumulative Hierarchy whereas proper classes do not appear at any V_α.

We might motivate [NMP] from [OHH] as follows. In order to be a member of either a set or a proper class, an object must appear in the Cumulative Hierarchy. Proper classes themselves, however, appear at no V_α. Thus they cannot be members of either sets or proper classes.

I shall argue that [OHH] is indeed a good explanation. It is, however, one to which the heavyweight theorist cannot appeal.

2.2.4.1 Positive Remarks about [OHH].

[OHH] is *precise*. Assuming the adequacy of [OHH] as an explanation, because [OHH] supports [NMP] it will imply that using conditions that would allow proper classes to be members in [COMP] is illegitimate. However, we can also see that [COMP] is not a correct principle to tell us what *sets* exist. Any sets that exist occur at some V_α and so are not defined by extendible conditions. [COMP], however, admits extendible conditions and so (if assumed to be a principle only about sets) will prove the existence of such sets. Thus [OHH] shows that using [COMP] as a principle that applies only to sets, combined with the Iterative Conception, is equivalent to accepting that certain sets both do and do not exist.

Moreover, [OHH] is clearly well *motivated*. For it responds exactly to a structural feature of the Cumulative Hierarchy in relation to the extendibility property of certain conditions that was noted in Chapter 1. In this way, it is obtained directly from our notion of set given by the Iterative Conception.

[OHH] further provides a *diagnosis*. It is not obvious that for certain conditions there will be no set of all its satisfiers in the Cumulative Hierarchy. This only becomes apparent when we closely examine the nature of the Cumulative Hierarchy, and the fact that certain conditions have an extendibility property that will prevent a set of all their satisfiers occurring at some V_α.

Moreover [OHH] does not violate *overkill.* It is a controversial issue what large cardinal sets exist. Such an issue in itself deserves a good deal of philosophical consideration. Unfortunately I lack the space to examine the question closely here. However, if a particular large cardinal set exists it can be introduced into the Cumulative Hierarchy and the resulting structure researched without problem[11]. The process is not much different from the way in which the Axiom of Infinity asserts the existence of the first infinite set, ω. Disagreements on what large cardinal sets exist will be reflected in a disagreement as to the extent of the Cumulative Hierarchy. If one thinks that a certain set exists, then there is a stage of the Cumulative Hierarchy that contains it. Thus the Cumulative Hierarchy does not *prohibit* reasoning about such sets, provided justification can be given for their existence.

2.2.4.2 Why the Heavyweight Theorist Cannot Appeal to [OHH].

We have seen then, that [OHH] is a satisfactory explanation for the difference between proper classes and sets. This is not the end of the story, however; one must also show that one's theory is able to use [OHH]. It is my contention that the heavyweight theorist cannot use [OHH].

Recall the failure of [NMH]. It was seen there (2.2.2.2) that, given the fact that proper classes are precise objects, they should be members of other classes.

Thus proper classes should be members of non-set classes[12]. Following Fraenkel, Bar-Hillel, and Levy I will call such things *hyper-classes*[13] ([Fraenkel et al., 1973], p142). Every class is a hyper-class. Further, hyper-classes have definite membership. Given that the heavyweight theorist accepts that proper classes are objects, it seems that they should also accept that hyper-classes are objects. This is because there is no reason for hyper-classes to not be objects that would not also apply to proper classes.

[11] Providing, of course, that the relevant large cardinal axiom is consistent.

[12] As the focus of this thesis is philosophical, I will put aside the technical problem of how to block the paradoxes with proper classes as members. There are systems (such as Ackermann's theory) that block the paradoxes but allow proper classes as members.

[13] One might, instead of defining a new kind of class, simply say that proper classes are members of other proper classes. The difference is purely one of nomenclature; I use the term 'hyper-class' simply to mark the fact that we have moved on from the picture where proper classes could not be members (and hence could not have proper classes as members).

The claim that they should not be objects as they have proper classes as members is clearly not going to work. The above arguments show that proper classes (under the Heavyweight View) are objects in a similar manner to sets (except that they do not appear at any V_α of the Cumulative Hierarchy). Given the similarity between proper classes and hyper-classes, the heavyweight theorist should accept that hyper-classes are also a kind of object. Therefore, any hyper-classes can be a member of some other hyper-class[14].

Let hyper-classes be denoted by the letters '$\mathfrak{A}$', '$\mathfrak{B}$', '$\mathfrak{C}$' and so on. Let the hyper-class membership relation be represented by '$\Subset$'. Immediately we may want to formulate some plausible principles about hyper-classes. A natural principle to adopt might be the following:

[H-COMP] $(\exists\mathfrak{A})(\forall\mathfrak{B})[\mathfrak{B} \Subset \mathfrak{A} \leftrightarrow \phi(\mathfrak{B})]$

Of course, [H-COMP] is going to be inconsistent. If we let $\phi \leftrightarrow_{df} \mathfrak{B} \not\Subset \mathfrak{B}$ we will get exactly parallel reasoning to the original Russell paradox about sets:

We have:

$(\forall\mathfrak{B})[\mathfrak{B} \Subset \{\mathfrak{A} : \mathfrak{A} \not\Subset \mathfrak{A}\} \leftrightarrow \mathfrak{B} \not\Subset \mathfrak{B}]$ (by [H-COMP].)

Therefore:

$\{\mathfrak{A} : \mathfrak{A} \not\Subset \mathfrak{A}\} \Subset \{\mathfrak{A} : \mathfrak{A} \not\Subset \mathfrak{A}\} \leftrightarrow \{\mathfrak{A} : \mathfrak{A} \not\Subset \mathfrak{A}\} \not\Subset \{\mathfrak{A} : \mathfrak{A} \not\Subset \mathfrak{A}\}]$ (by substituting $\{\mathfrak{A} : \mathfrak{A} \not\Subset \mathfrak{A}\}$ for $\mathfrak{B}$ bound by universal quantifier)

$\square$

Thus we have a problem; our naïve hyper-class theory is inconsistent. However, hyper-classes are just another object-like collection with definite membership. Not only this, proper classes seem to be 'formed' in a similar manner to sets. To see this, consider any V_α of the Cumulative Hierarchy. What is the relation of sets formed at $V_{\alpha+1}$ to V_α? The sets formed at $V_{\alpha+1}$ are subcollections of V_α that are not members of V_α. Now let the Cumulative Hierarchy be denoted by 'V'. Proper

[14] At this point, I depart from Bar-Hillel, Fraenkel, and Levy's exposition of the hyper-class structure. This is because Fraenkel *et al* do not draw out the philosophical consequences (i.e. that hyper-classes are objects that can be members) of holding both that proper classes are objects with definite membership, and that hyper-classes have definite membership.

classes are subcollections of V that are not members of V. Thus proper classes bear the same relation to V as sets formed at $V_{\alpha+1}$ bear to V_{α}. One might argue then that hyper-classes should (like sets) obey our principles of collection forming. In order to maintain cohesion with our basic notion of collection, the hyper-classes should be part of a hierarchical structure extending the original Cumulative Hierarchy. Thus, while the Cumulative Hierarchy contains stages, the stages of the hyper-class structure are distinct from these stages . I shall, therefore, use a different symbol to represent the stages the hyper-class structure. Let the stages of the hyper-class structure be indexed by ordinals and denoted by '$\mathcal{V}_{\alpha}$' (as opposed to the V_{α} of the Cumulative Hierarchy). We may then define the hyper-class structure as follows:

> Let $\mathcal{V}_0$ contain all hyper-classes that are either a pure set or a proper class containing only pure sets[15].
>
> For any ordinal number α let $\mathcal{V}_{\alpha+1} = \mathcal{V}_{\alpha} \cup \mathcal{P}(\mathcal{V}_{\alpha})$
>
> For limit ordinal λ there is a $\mathcal{V}_{\lambda}$ such that $\mathcal{V}_{\lambda} = \cup \mathcal{V}_{\beta, \beta<\lambda}$.

This structure I shall refer to as '$\mathcal{V}$' (a pictorial representation of $\mathcal{V}$ is provided in Figure 2.1 at the end of this Chapter). The heavyweight theorist might thus claim that she has a satisfactory way of allowing proper classes to be members of other class-like objects while maintaining cohesion with our collection forming practice. She can hold on to [OHH] as a principle about the Cumulative Hierarchy, while allowing that proper classes appear in stages of $\mathcal{V}$.

If she were to claim this, however, she would be wrong. I see two objections to the theory. One of the criticisms is weak, the other fatal.

Firstly, one might quite simply find the picture intuitively distasteful. It seems that we are replicating a cumulative structure on top of our first Cumulative Hierarchy. The hyper-classes that occur as part of $\mathcal{V}$ look very similar to the sets of V. Are we not just replicating the same structure again?

The obvious response is to deny that V and $\mathcal{V}$ are the same structure. No hyper-class within the Cumulative Hierarchy contains a hyper-class one may map one-to-one

[15]Once again, the restriction to pure sets and proper classes containing pure sets is merely a matter of simplicity. One could quite easily extend the structure to include hyper-classes that contain *urelemente*.

with the ordinals. $\mathcal{V}$ does contain such a hyper-class. At no point in V do we get classes that contain a proper class and a set. In $\mathcal{V}$ we have many such things, for example $\{\emptyset, On\}$ will be a hyper-class.

However, there is a very strong objection to the structure $\mathcal{V}$. Any theory of proper classes should be able to solve the paradoxes and the puzzle for Monist Realism set out in Chapter 1. Anyone who holds Monist Realism and uses the structure $\mathcal{V}$, should hold the following:

(Monist Realism)$^{\mathcal{V}}$ The structure $\mathcal{V}$ has the following properties:

1. The hyper-classes are objects.
2. $\mathcal{V}$ is a 'complete' abstract structure; for every ordinal α, $\mathcal{V}_\alpha$ exists. No new $\mathcal{V}_\alpha$ are being created.
3. For any object x, it is definite whether or not x occurs in $\mathcal{V}$.
4. There is only one $\mathcal{V}$ (although isomorphic copies of it may exist within universes of sets with different *urelemente*).
5. There is just one interpretation of $\mathcal{V}$. Our quantifiers are not ambiguous; it is possible to quantify over *all* hyper-classes.

Now, consider the following definitions used to define the concept *hyper-ordinal*. I shall use the capital Greek letters 'Γ', 'Δ', 'Λ' etc. to represent *hyper-ordinals*; a certain kind of hyper-class:

Definition. (1) $\{0n\}$ is a *hyper-ordinal*.

Definition. (2) If Γ is a *hyper-ordinal* then so is $\Gamma \cup \{\Gamma\}$.

Definition. (3) Λ is a *hyper-ordinal* if for some limit ordinal λ:

$\Lambda = \cup\{\Gamma : \exists\beta < \lambda, \Gamma \text{ is a hyper-ordinal in } \mathcal{V}_\beta\}$.

Definition. (4) Nothing but the hyper-classes satisfying one of the above clauses count as *hyper-ordinals*[16].

[16] The term 'hyper-ordinal' was chosen as the construction of the hyper-ordinals in $\mathcal{V}$ clearly mimics the construction of the von Neumann ordinal numbers in V. There are some disanalogies, however. For example, the hyper-ordinals are not transitive hyper-classes; there are members of a member of $\{0n\}$ that are not members of $\{0n\}$. To see this, observe that all the von Neumann ordinal numbers are members of On, but not members of $\{0n\}$.

Now we may prove a result about hyper-ordinals:

Theorem. The hyper-ordinals do not appear as a hyper-class at any $\mathcal{V}_\alpha$ of $\mathcal{V}$.

Proof. To prove this, it will be sufficient to prove that for any ordinal α indexing $\mathcal{V}_\alpha$ there are hyper-ordinals outside $\mathcal{V}_\alpha$. This will ensure that all the hyper-ordinals are never all present at some particular $\mathcal{V}_\alpha$, and thus there is no hyper-class of all hyper-ordinals formed at any $\mathcal{V}_{\alpha+1}$.

We use transfinite induction on the ordinals.

(i) We first must show that there are hyper-ordinals outside $\mathcal{V}_0$. It is clear that there are hyper-ordinals outside $\mathcal{V}_0$ (e.g. $\{On\}$).

(ii) We must then show that if α indexes $\mathcal{V}_\alpha$ and there are hyper-ordinals outside $\mathcal{V}_\alpha$, then there are hyper-ordinals outside $\mathcal{V}_{\alpha+1}$. Assume then that α indexes $\mathcal{V}_\alpha$ and that there are hyper-ordinals outside $\mathcal{V}_\alpha$. Then there are are hyper-ordinals in some $\mathcal{V}_{\beta>\alpha}$ that are not in $\mathcal{V}_\alpha$. If this is the case, then either:

a) There are hyper-ordinals in $\mathcal{V}_{\alpha+1}$ that are not in $\mathcal{V}_\alpha$.

or b) There are hyper-ordinals in $\mathcal{V}_{\gamma>(\alpha+1)}$ that are not in $\mathcal{V}_\alpha$.

(ii)b) If there are hyper-ordinals at some $\mathcal{V}_{\gamma>(\alpha+1)}$ there are hyper-ordinals outside $\mathcal{V}_{\alpha+1}$.

(ii)a) Therefore, assume that there are hyper-ordinals at $\mathcal{V}_{\alpha+1}$. Let Γ be one. But then (by the definition of hyper-ordinals) $\Gamma \cup \{\Gamma\}$ is also a hyper-ordinal. Since (by assumption) Γ is not in $\mathcal{V}_\alpha$, $\Gamma \cup \{\Gamma\}$ is first formed at $\mathcal{V}_{\alpha+2}$. Hence there are hyper-ordinals outside $\mathcal{V}_{\alpha+1}$.

(iii) Finally, let λ be a limit ordinal indexing some $\mathcal{V}_\lambda$ of $\mathcal{V}$. Assume that for all $\beta < \lambda$ there is a hyper-ordinal outside $\mathcal{V}_\beta$. Now (by Definition 3) the hyper-class $\Lambda = \cup\{\Gamma : \exists\beta < \lambda, \Gamma \text{ is a hyper-ordinal in } \mathcal{V}_\beta\}$, is a hyper-ordinal. Further, for any hyper-ordinal Γ formed at a $\mathcal{V}_\beta$, where $\beta < \lambda$, there is a hyper-ordinal outside $\mathcal{V}_\beta$ at $\mathcal{V}_{\beta+1}$, namely $\Gamma \cup \{\Gamma\}$[17]. Moreover, as

[17] Though we already have the fact that there is a hyper-ordinal outside $\mathcal{V}_\beta$ (by assumption) it is useful to see that at every $\mathcal{V}_\beta$ there is a new hyper-ordinal formed at each $\mathcal{V}_{\beta+1}$.

λ is a limit ordinal (and hence not a successor), for any $\beta < \lambda$, $(\beta+1) < \lambda$. Therefore, as

$\Lambda = \cup\{\Gamma : \exists\beta < \lambda, \Gamma \text{ is a hyper-ordinal in } \mathcal{V}_\beta\}$

Λ cannot appear at any $\mathcal{V}_{\beta<\lambda}$ for the reason that there will be a hyper-ordinal Δ in the hyper-ordinal Λ such that Δ is outside $\mathcal{V}_\beta$. In addition, $\mathcal{V}_\lambda$ is (by definition) $\cup\mathcal{V}_{\beta,\beta<\lambda}$. Thus, no new hyper-classes are formed at $\mathcal{V}_\lambda$. Hence, Λ must be formed at some $\mathcal{V}_\alpha$, $\alpha > \lambda$. Therefore, there are hyper-ordinals outside $\mathcal{V}_\lambda$.

Thus we have shown that (i) There are hyper-ordinals outside $\mathcal{V}_0$, (ii) If α is an ordinal indexing $\mathcal{V}_\alpha$ and there are hyper-ordinals outside $\mathcal{V}_\alpha$ then there are hyper-ordinals outside $\mathcal{V}_{\alpha+1}$, and (iii) If λ is a limit ordinal indexing $\mathcal{V}_\lambda$, and for every $\beta < \lambda$ there are hyper-ordinals outside $\mathcal{V}_\beta$, then there are hyper-ordinals outside $\mathcal{V}_\lambda$. By transfinite induction then, there are hyper-ordinals outside $\mathcal{V}_\alpha$ for all α. Hence there is no hyper-class of all hyper-ordinals, as all the hyper-ordinals do not appear at any $\mathcal{V}_\alpha$ to then appear as a hyper-class at $\mathcal{V}_{\alpha+1}$.

$\square$

But now we have an analogous problem to the one that faced the heavyweight theorist about the original Cumulative Hierarchy. The predicate '$\mathfrak{A}$ is a hyper-ordinal' is definite; for any hyper-class $\mathfrak{A}$ there is a fact of the matter whether or not it is a hyper-ordinal. Further, it is definite what hyper-classes occur in $\mathcal{V}$. Thus we have a condition with a definite range of satisfiers (so in some sense a 'collection'), that is not a hyper-class. What then is the ontological nature of this collection of hyper-classes?

Given that when a virtually identical puzzle was presented about sets in Chapter 1, and the heavyweight theorist responded to it by postulating the existence of heavyweight proper classes, we can expect a similar response. Similarly, we can then expect an analogous principle to [OHH] to be appealed to to explain the difference between these 'collections' and hyper-classes. Let us pause briefly in order to clear up nomenclature with respect to hyper-classes:

Definition. Let a 0-*hyper-class* be any hyper-class that occurs first in the Cumulative Hierarchy V (i.e. the sets).

Definition. Let a 1-*hyper-class* be any hyper-class that occurs first in $\mathcal{V}$.

We can then restate [OHH] as follows:

[OHH]^1 The difference between 0-hyper-classes and 1-hyper-classes is that 0-hyper-classes appear in the 0-hyper-class structure where 1-hyper-classes do not.

To deal with the non-hyper-class forming collections, the heavyweight theorist is forced to introduce the following objects:

Definition. Let a 2-*hyper-class* be a collection of hyper-classes that does not appear in $\mathcal{V}$.

Once again, the heavyweight theorist will want to develop plausible principles about this kind of object. Again, they will be driven to defining a new, '2-hyper-class' structure that extends $\mathcal{V}$. And once again, they will need to give an ontological characterisation of these 2-hyper-classes as follows:

[OHH]^2 The difference between 1-hyper-classes and 2-hyper-classes is that 1-hyper-classes appear in the 1-hyper-class structure where 2-hyper-classes do not.

Yet again, I will be able to define some extendible condition for this kind of object that will raise a similar puzzle to the one outlined in Chapter 1. The heavyweight theorist will then be forced to move to 3-hyper-classes, with a corresponding [OHH]^3 for the 3-hyper-class structure. Should we continue and move to 4-hyper-classes? n-hyper-classes for any natural number n? Even given this, one will always be able to define some extendible condition that is definite, but does not occur in any structure. Should we move to α-hyper-classes for any ordinal α?

One can do this, all the while pushing the puzzle back further. I think it is time to review, however, the situation in which the heavyweight theorist now finds herself. There are three problems with her position.

First, the game being played is no longer convincing. Proper classes were motivated as a way of dealing with collections that (if assumed to be sets) produce paradoxes. It is now obvious that the heavyweight theorist's principles have forced her to merely move the puzzle of Chapter 1 further and further back. There is no solution to the puzzle here, just a reallocation to a different order of class.

Second, we should carefully scrutinise what our theory of classes is *for*. It seemed to be a formalisation of our notion of collection. It might be plausible that there is more than one notion of 'collection' we use in our reasoning[18]. It is not plausible that we have infinitely many notions of collection packed into our more general concept of collection.

The heavyweight theorist might retreat and point out that any collection I point to will at least appear in *some* α-hyper-class structure, even if it is an $(\alpha+1)$-hyper-class structure of α-hyper-classes. This, however, points to the third problem for the heavyweight theorist. There are extendible conditions such that the relevant collection of *all* satisfiers never appears in any α-hyper-class hierarchy. The condition '$x = x$', when applied to any α-hyper-class will have a definite range of satisfiers that is not an α-hyper-class. But consider *all* α-hyper-class hierarchies. Within these hierarchies there is still a definite range of objects that satisfy '$x = x$' (i.e. everything). This definite range of objects does not form an α-hyper-class of any kind. In the end, the Heavyweight theorist will always be unable to give an account of certain definite collections of objects.

2.2.4.3 Conclusion about [OHH].

We have seen that [OHH] is a good explanation for why proper classes are not sets. However, we have seen that the heavyweight theorist cannot appeal to [OHH]. Any attempt to do so would (using other principles she holds), result in a proliferation of hierarchies. This merely shifts the puzzle higher and higher. Furthermore, it seems to splinter our seemingly simple notion of collection into infinitely many different notions. Finally, in the end, the Heavyweight View is still unable to account for certain definite collections.

[18] I shall argue later that this is indeed the case.

2.3 Conclusions.

Let us take stock. We have seen that the Heavyweight View, while it is appealing at first, does not in itself provide an explanation for why proper classes are not sets. Further, the putative explanations considered are not satisfactory for the heavyweight theorist. [NMH] and [LSH] are deeply flawed as explanations. [OHH], on the other hand, could not be appealed to by the heavyweight theorist. However, during our discussion we did identify [OHH] as an adequate explanation for why proper classes are not sets; proper classes (unlike sets) do not appear at any V_α of the Cumulative Hierarchy. It would be sensible, therefore, to examine proposals that make use of this fact.

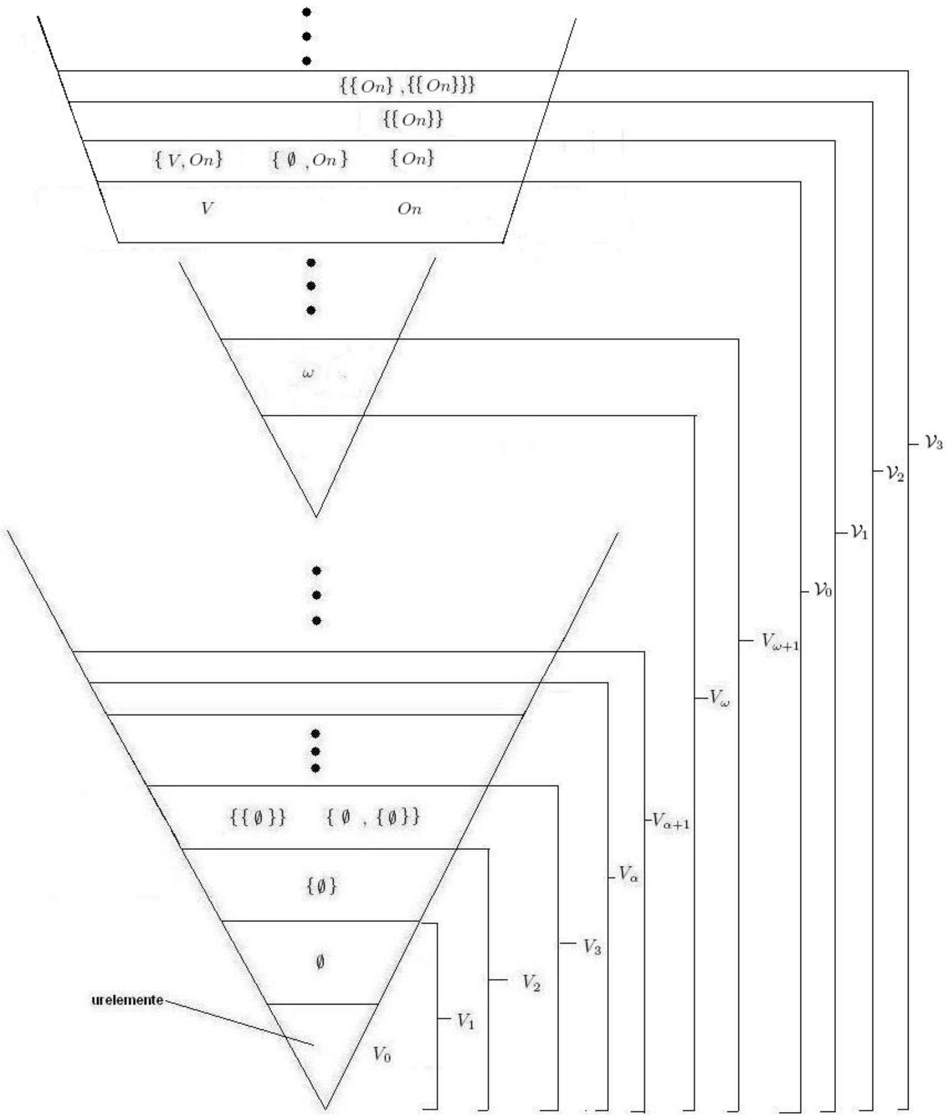

Figure 2.1: The structure $\mathcal{V}$.

Chapter 3

Modal Views of Proper Classes.

We saw in the last Chapter that there were significant problems with taking proper classes to be heavyweight objects. An obvious route to take then is to deny the objecthood of proper classes. However, merely stating this attitude does not yet explain why proper classes can figure meaningfully in our discourse.

In this Chapter, I examine two accounts that deny the objecthood of proper classes that may be extracted from the work of Charles Parsons[1]. My strategy is as follows:

- In 3.1 I examine some excerpts from Parsons' (and related authors') writings. I draw out the features of proper classes presented by these texts.
- In 3.2 I examine and reject one view of proper classes that may be extracted from the above writings; 'The Modal Function View'. I argue that a basic notion of the view is problematic.
- In 3.3 I analyse a different account about proper classes that one might hold based on the passages of section 3.1; 'The Projection View'. I reject it for the reason that it cannot explain one of the notions central to its expression.

[1] His view is largely developed in several papers in his *Mathematics in Philosophy: Selected Essays*, Cornell University Press, 1983.

I will show that none of the accounts I consider are satisfactory from the perspective of a Monist Realist.

3.1 What Parsons Said.

A substantial problem when trying to give an exegesis of Parsons' view of proper classes is that the exposition of his ideas is not clear. This is partly due to the fact that his view is partially articulated across several papers[2]. However, it is also because the main aim of the paper in which his most detailed discussion of proper classes occurs ('What is the Iterative Conception of Set') does not have as its main aim a characterisation of proper classes. Rather, Parsons sought to give an explanation of the Iterative Conception of Set, without using the metaphor of construction in stages.

Thus his remarks on proper classes are something of a side note. The views he expresses there are by no means developed. Nonetheless, Parsons' account may provide the Monist Realist with the resources she needs to characterise proper classes.

For this reason, I will not give a detailed exegesis of Parsons. Rather I will select some passages from his papers, and state what features of proper classes we may extract from these. I will then draw out and evaluate two characterisations of proper classes one might give based on some of these features.

We saw in the last Chapter that we require an account of proper classes on which proper classes are not heavyweight objects. Parsons' view provides just such a characterisation. He says the following:

> (A) "...we should think of predicates whose 'extensions' are proper classes as really not having *fixed* extensions." ([Parsons, 1977], p291)

I take an extension to be a heavyweight object[3]. Here we see Parsons arguing that if the objects that satisfy a predicate do not form a set, then the predicate does not actually have an extension (in the normal sense of 'extension').

Parsons' account provides the resources for a *modal* characterisation of proper classes. The point is put in his article 'What Is the Iterative Conception of Set?' in

[2] See [Parsons, 1974b], [Parsons, 1974a], [Parsons, 1977] and [Parsons, 1983b].

[3] If an extension is not a heavyweight object, then the term 'extension' is at least misleading and at most incoherent.

the following way:

> (B) "Indeed Reinhardt has suggested that proper classes differ from sets in that under counterfactual conditions they might have had different elements" ([Parsons, 1977], p286)

The Reinhardt paper indicated is 'Remarks on Reflection Principles, Large Cardinals, and Elementary Embeddings' and contains the following on proper classes:

> (C) "A proper class P may...be distinguished from a set x in the following way...if there were more ordinals...x would have the same members, whereas P would necessarily have new elements. We could say that the extension of x is fixed but that of P depends on what sets exist. Roughly, x is its extension, whereas P has more to it than that" ([Reinhardt, 1974], p196).

What can we glean from these two passages? (B) states that the difference between proper classes and sets is that if A is a proper classes then in some non-actual circumstance A would have members it does not actually have or A would lack members it actually has. Sets on the other hand have the same members in every possible circumstance, both actual and non-actual.

The distinguishing feature of proper classes is put in a slightly different way in (C). This claims that if A is a proper class, then if there were more ordinals, A would have member(s) it does not actually have. However, if A is a set, then if there were more ordinals, A would have just the members it actually has.

Parsons thus argues that *set* membership is rigid where proper class membership is not rigid[4]. (B) and (C) put this point slightly differently. (B) suggests that proper

[4] We might formalise this notion of rigidity as follows:

(R $\in$) $x \in y \rightarrow \Box(x \in y)$
(R $\notin$) $x \notin y \rightarrow \Box(x \notin y)$

However, as Parsons points out ([Parsons, 1983b], p298-301; [Parsons, 1977], p286-287) the status of the *existence* of proper classes is at issue. He therefore proposes an *existence dependent* treatment of rigidity which he formalises by the following three principles:

(E_0 $\in$) $x \in y \rightarrow E(x) \wedge E(y)$
(R_0 $\in$) $x \in y \rightarrow \Box(E(y) \rightarrow x \in y)$
(R_0 $\notin$) $x \notin y \wedge E(y) \rightarrow \Box(x \notin y)$

classes in some possible world *might* have different members. (C), on the other hand, states that if there were more ordinals a proper class *would* have more members.

Central to (B) and (C) is a notion of modality. This will be very important for the discussion of Parsons' view. A modal notion one might use is that of *intensions*. Parsons says the following:

> (D) "Reinhardt himself suggests...a class x is an intension.." ([Parsons, 1977], p287)

Parsons (as he endorses Reinhardt's characterisation of proper classes[5]) is committed to proper classes being *intensional*. Exactly what is meant by 'intensional' is not entirely clear. Carnap, when introducing the notion of intension said the following:

> **4-14.** The *extension of a predicator* (of degree one) is the corresponding class.
>
> **4-15.** The *intension of a predicator* (of degree one) is the corresponding property. ([Carnap, 1947], p19).

So, the statement that proper classes are *intensional* could be interpreted as saying that proper classes are properties. However, it seems that we want to say that proper classes are extensional, in the sense that if they have the same members then they are the same class. In this way we are able to say that the Russell Class and Universal Class are the same class[6]. It is also clear that the *property* of being non-self-membered is quite a different property from the property of being self-identical. Thus it does not seem right to say that proper classes are properties, it seems to imply that two proper class that we take to be the same are in fact not the same.

I do not think that Parsons provides a fully clear account of how to interpret these principles. I will, therefore, leave 'rigidity of membership' as an unformalised notion. The core idea is simple enough; sets have the same members in every possible world, proper classes could have different members in some possible world.

[5]See, for example, the following passage from [Parsons, 1977] (p286): "I am endorsing this suggestion [i.e. Reinhardt's Proposal] as an explication of the intuitions about 'inconsistent multiplicities' [i.e. proper classes]".

[6]Indeed Parsons refers to proper classes as 'attributes' at certain points (see [Parsons, 1983b], p304). He also explored the idea that proper classes do not obey [EXT] but rather the following principle he calls 'intensionality':

$$\Box(\forall z)(z \in x \leftrightarrow z \in y) \to x = y$$

As I think it fairly clear that proper classes are not properties I will not consider this further here.

We therefore require a different way of spelling out what is meant by saying that proper classes are intensional. Further, any characterisation should make sense of the definite membership of proper classes. I will now examine two ways one might make this claim more precise; the Modal Function View and the Projection View.

3.2 Proper Classes as Modal Functions.

3.2.1 The Modal Function View Explained.

We require a way of making the claim that proper classes are intensional more precise. In possible world semantics, intensions are represented as functions from possible worlds w to subclasses of the domains of the w. Maybe then we can transform this semantic claim into a metaphysical one.

First, however, we must make sense of what is meant by 'possible world'. We might think of possible worlds as way the Cumulative Hierarchy might have been. Each of these universes has different contents[7]. If we examine (B) we see that Parsons suggests that a proper class (in some universe) might have had different elements. Let '$\mathbb{V}$' represent an arbitrary universe of Set Theory. One could hold the following view about proper classes:

> [Proper Classes as Modal Functions]
>
> (1) For any condition ϕ there is a function f_ϕ from all universes such that for any universe $\mathbb{V}$, $f_\phi(\mathbb{V})$ = the class of all ϕs in $\mathbb{V}$.
>
> (2) The class of ϕs is a proper class iff there are distinct universes $\mathbb{V}'$ and $\mathbb{V}''$ such that $f_\phi(\mathbb{V}') \neq f_\phi(\mathbb{V}'')$ (i.e. there are two universes where ϕ has different elements.)
>
> (3) If the class of ϕs is a proper class, then the class of ϕs is the function f_ϕ.

[7] A natural way to explain this notion further would be through an examination of (C). There Reinhardt argued that if there were more ordinals a proper class would have different members. Under the Iterative Conception there is a stage V_α for every ordinal α. Therefore, if there had been more ordinals, there would have been more V_α and hence the universe would have been larger. As I do want want to commit to these universes being extensions of one another I will not assume that this is the case for this view.

3.2.2 Evaluation of the Modal Function View.

I will argue that the Modal Function View is problematic for the reason that it raises more problems than it answers. However, let us first note some positive features of the view.

3.2.2.1 Initial Positive Remarks.

The Modal Function View has some points to its credit. It is *precise* in that it identifies [COMP] as the guilty principle when only about sets. For [COMP] implies that proper classes are in fact objects. This is not the case, proper classes turn out to be functions from different universes of Set Theory to subclasses of those worlds. Thus (as proper classes are functions from worlds rather than objects in those worlds) they may not be substituted for an object variable in the paradoxical reasoning.

For this reason a *diagnosis* is also given for why we fell into error. It is not apparent, when we examine Set Theory, why there should be different possible universes of sets, or indeed why certain classes are functions from those universes to subclasses of their domains rather than sets.

3.2.2.2 Why the Modal Function View Fails.

The Modal Function View is unsatisfying as a philosophical standpoint. As stated, it does not actually provide us with a characterisation of the things for which we are seeking explanation. We want to know what the ϕs that do not form a set at *this* world are. These are the proper classes in which we are interested; the *actual* proper classes. But it is simply a fact that these are not functions from universes to subclasses of universes.

This worry can be made more precise as follows. Let the actual world be denoted by V. Let ϕ be some condition such that there is no set of all ϕ in V. We want to know what the class of *actual* ϕs is, namely $f_\phi(V)$. It is this that is the proper class, *not* f_ϕ as a whole.

The modal function theorist can, however, modify their view to accommodate this problem by dropping (3) in their view (the proposition that a proper class is the entire function) and modifying (2). Their view would then read as follows:

[Proper Classes as Modal Functions]$'$

(1) For any condition ϕ there is a function f_ϕ from all universes such that for any universe $\mathbb{V}$, $f_\phi(\mathbb{V})$ = the class of all ϕs in $\mathbb{V}$.

(2) The class of ϕs, namely $f_\phi(V)$, is a proper class iff there is some $\mathbb{V}$ distinct from V such that $f_\phi(\mathbb{V}) \neq f_\phi(V)$ (i.e. there is some universe distinct from V where ϕ has different elements.)

This modification would certainly avoid the above complaint. It attempts to provide a characterisation of the objects we are interested in, namely the ϕs in the Cumulative Hierarchy.

However, the view is still unsatisfactory. For it seems that there are insurmountable problems for a view that posits the existence of these functions. Take any f_ϕ used in the explanation of proper classes. It is clear that every $\mathbb{V}$ is a member of the domain of f_ϕ (to be denoted by '$dom(f_\phi)$'). Now we can see that $dom(f_\phi)$ cannot be in any $\mathbb{V}$. To see this, assume that $dom(f_\phi)$ is in some $\mathbb{V}$. Then it is the case that $dom(f_\phi)$ is a member of $\mathbb{V}$ is a member of $dom(f_\phi)$. Thus $\mathbb{V}$ is a non-well-founded universe of sets. But then, as *every* universe of sets is (on the Iterative Conception) well-founded, $\mathbb{V}$ is not a universe of Set Theory at all. Thus the domain of f_ϕ does not exist in any $\mathbb{V}$. But then f_ϕ does not exist in any $\mathbb{V}$; f_ϕ requires its domain to exist in order for it to exist. This result is paradoxical; how could anything (especially something doing significant explanatory work) not exist in any universe[8]?

It seems then, that the Modal Function View is unsatisfactory. An attempt to identify proper classes with functions does not give us a characterisation of the objects for which we required an explanation. While the view can be modified to avoid this problem, it turns out that the functions to which the view appeals are themselves paradoxical[9].

[8]Indeed, there is substantial pressure to think that a function should exist in the *actual* universe. For, in what sense is the statement "For any condition ϕ there is a function f_ϕ from all universes such that for any universe $\mathbb{V}$, $f_\phi(\mathbb{V})$ = the class of all ϕs in $\mathbb{V}$" true if the function in question does not exist at the actual universe?

[9]Parsons was well aware of the fact that such an explanation of proper classes was unsatisfactory. See, for example, the following passage from [Parsons, 1977] "It seems that we cannot consider a proper class as given by an *intension* that is definite in the sense of, say, possible world semantics as a function from possible worlds to extensions." (p290).

3.3 The Projection View.

3.3.1 The Projection View Explained.

So, proper classes cannot be 'intensional' in the sense of functions from possible universes to subclasses of the domains of those worlds. Indeed we do not need to talk about functions in order to understand the central claims of (B) and (C). The point there is that proper classes are able to change members dependent on what ordinals exist. This might lead us to the following view:

> [The Projection View] What makes the class of all ϕs a proper class (rather than a set) is that for some non-actual universe $\mathbb{V}$, the class of ϕs in $\mathbb{V}$ is different from the class of ϕs in V. One class has members the other lacks.

How would this work? An understanding of what is meant can be arrived at from examining the *model theory* of Set Theory. A *model* can be thought of as a domain of objects, the functions that exist on the domain, and the relations that hold on the objects within the model. Let us examine some models and observe a phenomenon about what happens when some objects within a model do not form a set within that model.

For example take the model H_{ω_0}. This model has as its domain the hereditarily finite sets, where a set is hereditarily finite iff it is either the empty set or a finite set $\{a_1, a_2, ..., a_k\}$, where $a_1, ..., a_k$ are all hereditarily finite. The only relation on the model is the standard interpretation of '$\in$'. Now, consider the property of being a von Neumann ordinal number. Present in this structure are the von Neumann representations of all the finite ordinals. Thus, we may identify what objects constitute 'the ordinal numbers' in this model. As the model only includes the hereditarily finite sets, it does not include the set of all finite ordinals. In fact the negation of the axiom of infinity is satisfied on this model[10]. Thus, from the viewpoint of the model, the ordinal number of ω_0 is a proper class. Now consider the model H_{ω_1}. This is the set

[10]It may be objected here that for this reason H_{ω_0} does not satisfy full ZFC, and hence does not satisfy the Iterative Conception. Thus it is not a possible world of Set Theory. While this is true, H_{ω_0} is still a structure that provides a nice analogy with what the Projection theorist argues happens in the case of the actual Cumulative Hierarchy.

of hereditarily countable sets, where a set is hereditarily countable iff it is a countable set of hereditarily countable sets. On this model the ordinal number of ω_0 will be a set. However, the ordinal number of ω_1 will not be a set, but will appear to be a proper class 'from the perspective of the model'. Again, if we consider different models, 'the ordinal numbers' will be different. Indeed, if we consider models that extend one another, at each successive expansion of a model what appeared to be the proper class of all ordinal numbers will be a set at the next model. The condition 'x is a von Neumann ordinal number', however, is consistent across the models.

Maybe then the Cumulative Hierarchy bears a similar relation to the non-actual $\mathbb{V}$s as a model bears to an extension of that model. Thus the objects that satisfy the condition 'x is an ordinal number' can be different depending on what ordinals exist. The actual Cumulative Hierarchy contains all ordinal numbers of actual ordinals. If there had been more ordinals, however, there would have been more ordinal numbers at stages beyond those contained in the actual Cumulative Hierarchy. Thus, in *projected universes* the ordinal numbers would have more members.

Moreover, the objects that constitute the ordinal numbers in the actual Cumulative Hierarchy would form a possible set in an extended hierarchy. Given that every ordinal number is formed at some V_α of the Cumulative Hierarchy, if there are V_β, $\beta > \alpha$ then these objects will constitute a set at some successor stage V_β. There are just such V_β in some of the $\mathbb{V}$s.

We should be careful, here, however. For all I have said, it is not necessarily the case for two distinct universes $\mathbb{V}'$ and $\mathbb{V}''$ that either $\mathbb{V}''$ extends $\mathbb{V}'$ or vice versa[11]. The core idea is that for paradoxical conditions ϕ, the satisfiers of ϕ can be different objects depending on what ordinals there are (and hence what is contained within the hierarchy).

3.3.2 Evaluation of the Projection View.

I think, however, that this view of proper classes is problematic. It is especially clear that the view is wholly untenable for a Monist Realist. Nonetheless, let us first

[11] It should be noted, however, that implicit in both [Parsons, 1974b] and [Reinhardt, 1974] is a view on which possible universes are indeed extensions of one another.

examine some positive features of the account.

3.3.2.1 Initial Positive Remarks.

The Projection View is *precise* in identifying [COMP] as the erroneous principle and attempts to *motivate* this choice of error. Sets and proper classes are very different kinds of thing on the Projection View; the condition which defines a proper class might have had different satisfiers. Hence [COMP] is unsatisfactory; proper classes are very different objects and proper class variables should not be substitutable for set variables in our reasoning.

In this way a *diagnosis* is also provided. For there is a sense in which the ϕs in V might constitute a set. This is because in some of the $\mathbb{V}$ there is a possible set with just those objects as members. It so happens that the Cumulative Hierarchy has the length it actually has, and so these things do not constitute a set. All it took, however, was for there to be enough ordinals, and then the class in question would have been a set. Since we do not know exactly what ordinals exist, it is understandable that we might get the length of iteration wrong, and think that a possible set is an actual set.

The solution also avoids *overkill*. Firstly, although [COMP] is shown to be faulty, one can preserve the intuition of plausibility of [COMP] without inconsistency. We can replace [COMP] with a similar modal principle:

$$[\mathrm{COMP}^{\Diamond}]\ \Diamond(\exists x)(\forall y)[y \in x \leftrightarrow \phi(y)]$$

Such an axiom states that for any condition ϕ evaluated at world $\mathbb{V}$, there is a possible world at which all the ϕs (in $\mathbb{V}$) are collected into a set.

Further, in this way the projection theorist is able to incorporate much talk about large cardinals as talk about *possible* objects. For example, by considering counterfactual projections of the actual universe of sets, the objects that constitute the ordinal numbers of all actual ordinals in a larger universe form a set that is a measurable cardinal ([Parsons, 1977], p288)[12].

[12] This is true given some assumptions about the actual universe. See [Reinhardt, 1974] for full technical details.

3.3.2.2 Why the Projection View is Unsatisfactory.

Despite these considerations the Projection View is highly problematic for a Monist Realist. This is because it is impossible for the Monist Realist to make sense of the modality in play.

In giving an exposition of the modality, it was said that if there were more ordinals proper classes would have more members. But what is the content of the phrase "if there were more ordinals"? What *modality* is in play here?

Clearly, it is implausible to view the modality as either physical or metaphysical; mathematics is invariant over such possibilities. One might try to argue that metaphysical modality does in fact have a part to play; one might say that it is a contingent matter what *urelemente* exist. This is no response. The pure sets will be the same in universes with different *urelemente* (as they are all constructed from the empty set). Thus, allowing what the *urelemente* are to be a contingent matter will have no effect on what pure sets exist. As noted in Chapter 1, the paradoxes that I am considering are ones reproducible in the pure sets. Further, as the projection theorist requires there be more ordinals at different worlds, there must be additional pure sets in those worlds representing those possible ordinals. Therefore, the modality cannot be either physical or metaphysical.

Given that the modality in question is not clearly a familiar notion such as physical or metaphysical modality, the projection theorist is still faced with the task of explaining to what their notion of modality responds. It is not philosophically satisfactory to appeal to a technical notion without some account of what that technical notion formalises.

One might try to elucidate the modality in terms of supposition. We might say that the content of the statement that there could have been more ordinals should be understood as the statement that we can suppose that more ordinals exist (even if they do not). This seems to be a phenomenon that appears in some other areas of mathematics. For example, consider the introduction of i as a solution to the equation $x^2 = -1$. Now, one might think that i nonetheless does not exist. It is not an actual mathematical object, but rather produces a nice formal theory with some interesting uses. Further, expansion can give us plausible results about the 'actual' structure we are considering; the real numbers. As the real numbers are embedded in the complex numbers, universal theorems about the complex numbers restricted to the reals and with no imaginary part are theorems about real numbers. Maybe this is how to understand the modality; we suppose there are more ordinals, and see what conclusions we can draw from this fact. As V is part of all $\mathbb{V}$ that are extensions of V we can learn about V from studying the $\mathbb{V}$ even though they do not exist.

There is, however, a substantial disanalogy between the 'supposition' in the complex numbers and in the possible sets. In the case of the complex numbers, we have an explanation of our reasoning concerning complex numbers through understanding complex arithmetic and analysis as based on a plane rather than a line. The reals then are just one line within this plane. In this way our original supposition is shown to have an intelligible interpretation. This is precisely what has not yet been given for the Projection View. Merely stating that we can suppose that there are more ordinals is not enough, we must understand to what this supposition amounts.

Providing such an interpretation will be an impossible task for the Monist Realist. The universe exists and is 'finished'. It describes all the sets there are, and could ever be (for a given initial starting set of urelemente). It is a principle of Monist Realism that the universe *cannot* be extended in any way. To say 'there could have been more ordinals' is false; if there were more ordinals, there would be more V_α indexed by those ordinals, and so we would have extended the Cumulative Hierarchy.

Maybe then the projection theorist, rather than saying that there could have been more ordinals (and hence extensions of the Cumulative Hierarchy), should say that there could be universes *smaller* than the actual Cumulative Hierarchy. Using this, they could still explain the non-rigidity of proper class membership by arguing that in a universe containing *less*, proper classes would have different members.

I see three ways that we might explicate the idea of 'smaller' universes.

1. A smaller universe is a V_λ for some limit ordinal λ that, together with $\in$ interpreted as membership restricted to V_λ, is a model of Set Theory.

2. A smaller universe is a model of Set Theory with a domain that contains all the ordinals, but which is narrower than the universe, with $\in$ interpreted as

membership restricted to that domain (An example of this sort would be the Constructible Hierarchy 'L'[13]).

3. A smaller universe is a model with both the above two properties (i.e. A model of Set Theory that has as its domain a proper subset of some V_λ for limit λ with $\in$ interpreted as membership restricted to that domain).

However, none of these suggestions is going to work with the Projection View. For, given any of the above three characterisations, objects that are uncontroversially sets will turn out to be proper classes.

Let us first consider the cases of both (1) and (3). Take some V_λ as the domain of the sub-universe in question. Let κ be the successor of λ. Now consider the class of all ordinal numbers of ordinals less than κ. This will contain the ordinal number of λ (as $\lambda < \kappa$). However, the class of ordinal numbers of ordinals less than κ in V_λ does not contain the ordinal number of λ (as $rep(\lambda)$ is not yet formed at V_λ). Hence, if the Projection View is correct and we use either (1) or (3), the class of ordinals less than κ is a proper class. This is false, it is clearly a set.

In the case of (2), let us consider the class of sets of natural numbers. Suppose that (as many believe) that not every set of natural numbers is constructible. Therefore, the class of natural numbers has members in V that it lacks in L. Thus, under the Projection View using (3), the class of sets of natural numbers is a proper class. Again, however, this class is uncontroversially a set.

Thus the Projection View fails. It quite simply cannot give an account of the modality central to its exposition under Monist Realism.

3.4 Conclusions.

There is a thought that we might be able to characterise proper classes using modal resources. However, using functions from possible worlds to subclasses of domains is unsatisfactory; the functions are themselves paradoxical. Instead we might explain

[13] L is a sub-universe of the Cumulative Hierarchy where, instead of including *all* subsets of V_α as sets at $V_{\alpha+1}$, we instead only include sets that are first-order definable by a formula that only contains parameters from the previous stages and has its quantifiers restricted to those stages.

proper classes as defined by conditions that change satisfiers in different possible universes. However, it was seen that there is no satisfactory way to understand the notion of ‘possible world’ used. For this reason, the Modal Views that I have considered are not a satisfactory characterisation of proper classes for a Monist Realist.

Chapter 4

The Plural Account of Proper Classes.

In this chapter I examine the view that proper classes are merely 'some things'. Proper classes do not exist. Instead reference to proper classes is loose talk and should be understood via plural reference to some objects. My strategy is as follows:

- In 4.1 I outline the Plural Account.
- In 4.2 I give an evaluation of the view and defend it against some objections:
  - 4.2.1-I discuss some positive aspects of the account.
  - 4.2.2-I reject an objection to the plural theorist that plural reference cannot capture the whole of class theory.
  - 4.2.3-I discuss and argue against Linnebo's criticism that there is a collapse of pluralities to sets.
  - 4.2.4-I develop an objection to the Plural Account based on superplural quantification. I argue that it is a challenge that can be overcome.
- In 4.3 I conclude that the plural account is a satisfactory account for the Monist Realist.

4.1 What is the Plural Account?

We saw in Chapter 2 that the heavyweight theorist was unable to give an account of why proper classes are not sets. The best explanation for this fact ([OHH]) could not be appealed to for the reason that proper classes, as precise *objects*, were the sort of things that should be members of other kinds of class (hyper-classes). This, we saw, resulted in further problems being developed for these α-hyper-class hierarchies. In the last Chapter, we examined some weakenings of the Heavyweight View. On these views proper classes were not objects, but rather were explained using modal resources. However, these were also seen to be unsatisfactory. It seems then that we require a way of referring to many things at once that does not presuppose the objecthood of a collection distinct from the elements, but also does not make use of the modality of the previous Chapter. Do we have such a notion?

4.1.1 Plural Reference.

Indeed it does seem that we have a way of referring to many things without presupposing the existence of a collection as an object distinct from the things. Such a notion is the notion of *plural reference*.

If one examines natural language, it appears that we are able to refer to more than one object at once using a single referring term. For example if I say "The smarties in this bag are green." I appear to be making reference to precisely the smarties in the bag (of which there are more than one[1]). Such a phenomenon I will call *plural reference*. Plural reference is most easily seen with the use of words like 'some', 'most' and combinations of names (as in 'Russell and Whitehead')[2].

Now, for many examples of plural reference one may give an equivalent sentence in singular terms. For example, the smarties example above one might paraphrase in the following way (with quantification restricted to the contents of the bag):

$$\forall(x)[\text{Smartie}(x) \rightarrow \text{Green}(x)]$$

[1] It should be noted that this sentence does admit of a first-order paraphrase. This is simply an introductory example to show that pre-theoretically one might think that there is a device that is present in natural language that allows us to refer plurally.

[2] There are other examples suggested in the literature, for example music groups. See [Uzquiano, 2011] for a review.

Such a statement simply says that anything that is a smartie is green, and makes no mention of plural reference. One might be tempted to think then that we can deal with all apparent occurrences of plural reference in a singular (and hence first-order) manner. One could simply use singular reference to refer to *urelemente* and sets, and when one needs to refer to many things use the relevant set and membership relation to do the work of plural reference. It seems, however, that plural reference cannot be fully captured by singular reference.

There are a number of reasons why one might think this. For starters, one might think that it is simply intuitive that we are able to refer to more than one thing without referring to anything over and above those things. We often refer to more than one object with a single phrase, with the set of those things seeming to play no role. As Boolos remarked "It is haywire to think that when you have some Cheerios you are eating a *set*" ([Boolos, 1984], p448). Statements made about some objects seem to be about the objects in question, not any set of them.

We should not be content with mere intuition, however. Thankfully, there are other cases that are more convincing. For example, if I say "Matthew, Steve, James, and Tim carried the boat to the pontoon." I am not saying anything about the set of Matthew, Steve, James and Tim. It is also true that no individual one of them carried the boat to the pontoon; they carried it together. Further, the claim that "Matthew, Steve, James, and Tim carried the boat to the pontoon." does not admit of an obvious paraphrase using quantification or conjunction. The sentence "Each of Matthew, Steve, James and Tim carried the boat down to the pontoon." should be subject to quantifier elimination and so it should be true that Matthew carried the boat to the pontoon. This, taken literally, is false; he helped to carry the boat to the pontoon, but did not (by himself) carry it to the pontoon. Similarly if we take the conjunction "Matthew carried the boat down to the pontoon and Steve carried the boat down to the pontoon and...." it should be true that each of the conjuncts is true. For exactly the same reason, this is false.

Further, there are some legitimate, grammatical sentences of English that do not admit of a first-order paraphrase. The most famous example is probably the Geach-Kaplan sentence:

(GK) Some critics admire only one another.

Such a sentence was shown by David Kaplan to be non-first-orderisable. While the legitimacy of this statement can disputed[3], it is nonetheless easy to generate other statements that do not admit of a first-order paraphrase. A selection of these were reviewed by Boolos in his seminal paper "To Be is to be the Value of a Variable (or to be Some Values of Some Variables)" [Boolos, 1984]. For example:

(G) There are some gunslingers each of whom has shot the right foot of at least one of the others.

Is most naturally formalised in second order logic as:

(Gi) $(\exists X)((\exists x)X(x) \wedge (\forall x)[X(x) \rightarrow (\exists y)(X(y) \wedge (y \neq x \wedge \text{ ShotRF}(xy)))])$

If we were to substitute '$x = y+1$' for 'ShotRF(xy)' we get a sentence that is true in all nonstandard models of arithmetic, but not in the standard model ([Boolos, 1984], p435)[4].

While (Gi) is formulated in second order logic, this should not detract from the plural nature of the natural language sentence (G). Second order logic was merely used to show its non-first-orderisability. Further, Boolos showed that it is possible to give an interpretation of monadic second-order logic in plural logic. This work has been extended by Hewitt to a plural interpretation of full second-order logic[5].

I take the above examples to show that plural reference is a feature of thought. Further, it is accompanied by a precise formal logic, which makes its use appropriate within Set Theory. While a full defence of plural logic is pertinent, my focus is to see whether or not Set Theory can make use of these plural resources. Therefore, from this point on, I shall assume plural reference and its logic. How then might we use plural reference within Set Theory? How could plural reference contribute to an understanding of 'proper class' talk?

[3]This is a fact that Boolos acknowledged in [Boolos, 1984].

[4]To see why, observe that a nonstandard model permits infinite descending successor chains. The above sentence implies that there is some non-empty X such that if $X(x)$ then $X(x-1)$. This is false on the standard model, as if $X(0)$ then $X(0-1)$ which is false; there is no predecessor of 0. However, in a non-standard model there can be infinite descending successor chains, and so if $X(x)$ holds of a non-standard number with infinitely many predecessors the sentence can be made true.

[5]See [Hewitt, 2012].

4.1.2 Use within Set Theory.

Plural reference comes complete with its own precise logic where we standardly use double lower-case letters (such as 'xx') as variables for plural terms and '$x \prec xx$' to mean 'x is one of the xx'. Often, a variable 'xx' is referred to as 'a plurality'. Such locutions should be regarded as loose talk. The use of a plural reference term does not presuppose the existence of a collection as an object over and above the objects referred to by use of the plural term. Such talk is acceptable only when it is shorthand for a paraphrase in only plural terms (and even then is misleading). I will try and avoid lapses into such 'singularism' where possible, and flag that I am using a singular term in a loose fashion where a plural paraphrase would be particularly clumsy.

The fact that plural reference does not presuppose the existence of a collection over and above the objects to which one refers suggests a way out of the paradoxes. We have seen in previous chapters that part of the problem with viewing proper classes as abstract objects is that it is not clear why such collections would not be sets. If we instead view proper classes not as set-like collections, but artefacts of plural reference to sets, the paradoxes would dissolve.

We then have a system where the inconsistent [COMP] is replaced by a similar principle about plural reference:

[P-COMP] $(\exists x)\phi(x) \rightarrow (\exists xx)(\forall y)[y \prec xx \leftrightarrow \phi(y)]$[6]

This states that for any (nonempty) condition ϕ there are some things that satisfy ϕ. [P-COMP] (at least in some form[7]) certainly has intuitive pull, if we have a condition such that it is definite for any object whether or not it satisfies that condition, then it seems that there are some things that satisfy that condition. Further, [P-COMP] allows us to talk about proper classes without having them be objects. Let us now examine how the paradoxes are avoided.

[6]It is noticeable that this differs from the set-theoretic comprehension axiom in that it has an antecedent stating that there is something that satisfies ϕ. This is because plural logic does not (normally) admit an empty plurality, a fact which will be discussed later. It should also be noted here that I assume a single object does constitute 'some things'. To see this, consider the sentence "The students who take this course will benefit from it.". Here 'the students' is a plural reference term, but the sentence is true if just one student takes the course and benefits from it.

[7]There are those who think that it is not correct in its full generality. See, for example, [Linnebo, 2010].

Reference to the Russell class should be understood as plural reference to all the non-self-membered objects. If I say 'the Russell class exists', what I should be understood as saying is that 'there are some things such that they are not members of themselves'. This is clearly true for a Monist Realist on the Iterative Conception.

The Russell paradox would be avoided, as substitution of R for x in '$(\forall x)(x \in R \leftrightarrow x \notin x)$' would be illegitimate. R is merely a plural term, as this is not a singular referring term it cannot be substituted for a singular variable.

Similarly reference to the Universal class should be understood as plural reference to all the sets. The statement of the existence of the Universal class should be understood as a statement that 'there are some sets that are self-identical'. Again, if Monist Realism is true this seems obvious. The paradox is avoided, as we cannot talk about the power-set of some sets. There is no singular set to which the Power Set operation can be applied[8].

The Burali-Forti paradox is also given an account. Reference to the proper class of all ordinal numbers should be understood merely as reference to the well-ordered von Neumann ordinal numbers. However, assuming that if we have some things we can apply ordering techniques to them, the von Neumann ordinal numbers can be well-ordered by the membership relation. They are collectively transitive and well ordered in the sense that (a) For any transitive well-ordered sets x, y, Trichotomy holds for the relation $\in$, (b) the $\in$ relation is transitive for any sets within the von Neumann ordinal numbers, (c) if there are some transitive well-ordered sets there will be an $\in$-least set, and (d) if x is a transitive well-ordered set that is one of the ordinal numbers then every member of x is a transitive well ordered set. Thus these things do have an order-type; there is a certain structure placed on them by the membership relation.

It seems then that in order to prevent paradox we must deny that there is a representation (i.e. ordinal number) that corresponds to the order-type that the ordinals exemplify collectively. Should this bother us? I do not think it should. I see no good reason to accept that the von Neumann representation of ordinals should provide anything other than a partial representation of order-type.

[8]One can, however, modify this Cantorian argument slightly. This reformulation of the problem will be considered in 4.2.4.

The situation is parallelled with respect to cardinality and proper classes. In 2.2.3 it was noted that proper classes may always be mapped one-to-one. Thus [LSH^F] was seen to be a true principle about proper classes. Thus proper classes have 'cardinality' of a sort. Strictly speaking, however, proper classes do not exist and 'transcend' cardinality. It is true, nonetheless, that if we have some things that do not form a set then they may be mapped one-to-one with some other things that do not form a set. However, it is uncontroversial in this case that there is no set taken as the canonical representation of this 'cardinality'. This is the same with order-type; there are cases where some things may exhibit an order-type, but nonetheless there is no representation of this order-type.

4.2 Evaluation of the Plural Account.

4.2.1 Positive Remarks about the Plural Account.

The Plural Account has a number of good features. Firstly, it is *precise* and provides *diagnosis*. [COMP] is satisfactory when used to prove the existence of some things all of which satisfy a condition, but its use is fallacious if used to prove the existence of a set of those things. It is not clear, before we realise that there are cases where for some things xx there is no set of the xx, why proper classes are not sets. The success of the use of Set Theory to represent other cases of plural reference (such as reference to 'the natural numbers', or 'the real numbers') leads us to think that for any case of plural reference there is a corresponding set. This turns out to be false.

The *motivation* for such a response is also clear. As shown in the previous section, plural reference is a well-established feature of natural language with a precise formal analysis. It is another way to talk about many things collectively other than by referring in a singular manner to the set of those things. We can see this with respect to the puzzle of Chapter 1. There it was noted that we lacked an explanation of the fact that it is definite for any object x within the Cumulative Hierarchy whether or not x satisfies a particular paradoxical condition. Therefore, there seem to be some objects such that each definitely satisfies the condition. This talk of "some objects" is quite naturally understood in plural terms.

The Plural Account also fares very well with respect to *overkill*. Firstly, it preserves the thought that [COMP] has some intuitive plausibility. As noted above, for any condition ϕ there are some objects within the Cumulative Hierarchy that each satisfy ϕ.

Further, the plural response is able to account for our talk of large cardinals. We can ground our talk of embeddings by referring plurally to some ordered pairs, even if there is no set of those pairs[9].

We are also able to make true statements concerning proper classes. If I wish to say that the Russell class and Universal class are the same class, without problematically committing myself to the objecthood of the paradoxical classes, I can do so by stating that anything that is one of the non-self-membered things is also one of the sets.

It seems then that the plural response is in a fairly strong position. It is, however, not without its dissenters. I will now examine some objections one might raise against the Plural Account. All, I shall argue, are answerable.

4.2.2 The Empty Set.

One might argue the following. After discovering the paradoxes, we think that our class-theoretical discourse *extends* our talk of sets. There are some classes that are not sets, but all sets are classes. This is a fact that an account of classes should incorporate. If one thinks that our class talk is characterised by plural reference, then for any class there are some things that are in that class. This is not the case, however, there is one very important set (and hence class) for which there are not some things that are its members; namely $\emptyset$. There is no empty plurality; 'plurality' is merely loose talk to refer to some things, and in the case of the members of $\emptyset$ there are no things to which we may refer. As there is no 'empty plurality' to correspond to the empty set, plural talk is not a satisfactory way of interpreting class talk; it cannot account for all classes.

I regard this as no objection. The reason we originally got into trouble with the class paradoxes was because imprecision in the notion of class resulted in us

[9] A fuller exposition of the relationship between plural quantification over sets and large cardinal hypotheses is available in [Uzquiano, 2003].

equivocating between plural reference to sets and singular reference to a set of sets. The term, 'class' can be understood as talk that may be characterised as singular reference to a set or plural reference to some sets. All that is required for it to be legitimate to use the term 'class' is definite membership. In some cases 'class' can be understood in either plural terms or singular terms (such as when I refer to the class of all natural numbers). Other times reference must be understood plurally (such as when one (misleadingly) refers to the 'class' of ordinal numbers). And at other times, reference to a class must be understood only through singular reference (such as when I refer to the empty class). This account would actually explain rather nicely a pedagogical phenomenon; the bafflement of many students at the notion of the empty set. Such bafflement can be understood as conflation of the notions of 'set' and 'plurality' due to pre-theoretic contact with both notions through reference to many objects.

4.2.3 Collapse.

One might want to contend instead, that pluralities do in fact always have a corresponding heavyweight object. One could do this by arguing that the semantics for plurals nearly always rely on set-theoretic resources. Often, plurals are analysed in our semantics as sets, so why should we not think that they just are sets?

This is a very weak objection. Just because our formal semantics nearly always uses sets does not mean that the objects in question are sets. Benacerraf showed as much for natural numbers in his seminal 'What numbers could not be' [Benacerraf, 1965]. Our formal semantics may simply be inadequate, and only a partial articulation. Alternatively, it might be a notational systematisation, useful for formalising our plural talk but merely a heuristic with no ontological import. Either way, there are many things that we analyse in our semantics as sets, that are not actually sets.

However, there is a stronger argument that some things always have a corresponding heavyweight object. The argument proceeds via an appeal to Extensionality as definitional of set. The Axiom of Extensionality may be stated as follows:

[EXT] $(\forall x)(\forall y)[x = y \leftrightarrow (\forall z)(z \in x \leftrightarrow z \in y)]$

This states that two sets x and y are identical iff they have exactly the same members. It is then argued that it is sufficient for some things to be a set that they have definite membership and satisfy an extensionality axiom. As has been noted, proper classes on the Plural Account have definite membership (it is definite, for any object x whether or not it is one of some things that do not form a set). Now, there is no identity relation for some things (some things, after all, are not an object and hence cannot be identical to anything). However, there is a sameness relation between some things xx and some things yy. Let this relation be denoted by '$\equiv$'. It seems clear that we should take as an axiom the principle that some things xx are the same things as some things yy iff every thing that is one of the xx is one of the yy. This is an extensionality axiom of sorts for our logic of plurals, and may be formalised as follows:

[P-EXT] $(\forall xx)(\forall yy)[xx \equiv yy \leftrightarrow (\forall z)(z \prec xx \leftrightarrow z \prec yy)]$

Thus as the logic of plurals contains an axiom of extensionality and for some things xx it is definite for any object x whether or not $x \prec xx$; every plurality is a set. The argument is nicely summed up by by Linnebo as follows.

> "The semantics of plural logic ensures that a plurality consists of a determinate range of objects. But a set is completely characterized by its elements. A plurality thus provides the resources for a complete and precise characterization of a set. So what could prevent us from collecting the given plurality into a set?" ([Linnebo, 2010], p149)

The thought is the following. Sets are completely characterised by their elements. This is what is meant by Extensionality, the core notion of set. When we refer to some things, it is a definite whether or not an object is one of those things. Therefore, if we simply substitute 'is a member of y' for 'is one of the xx' , we have a complete characterisation of the membership relation of a set. Hence some things must always form a set. The principle that some things always form a set I shall call (following Linnebo) [COLLAPSE]. One can formalise [COLLAPSE] as follows:

[COLLAPSE] $(\forall xx)(\exists y)(\forall x)(x \prec xx \leftrightarrow x \in y)$[10]

It is obvious that [COLLAPSE] combined with [P-COMP] is inconsistent (Just use $x \notin x$ for ϕ in [P-COMP] then use [COLLAPSE] to get the Russell set). Linnebo suggests that we reject [P-COMP][11], and hence that reference to proper classes should be characterised as plural reference to sets[12].

The challenge is a serious one. For [COLLAPSE] appears to be plausible. To begin with, it seems pre-theoretically intuitive; sets are definite collections of objects, and for any things xx it is definite whether or not some object x is one of the xx. Why should all xx not then have a corresponding set? Further, [COLLAPSE] is motivated by one of the core principles of Set Theory; sets depend for their identity on their members. If there is a fact of the matter for any object x whether or not it is one of the xx, it looks like it should be easy to characterise the membership relation for a set.

Further, [COLLAPSE] seems to be a principle at play in our conception of the original definition of the Cumulative Hierarchy. A guiding principle there was that at each $V_{\alpha+1}$ we formed sets of every subclass of V_α. Whenever we have some things at a V_α that are not all members of some set at V_α, then the set of those things will be first formed at $V_{\alpha+1}$. This 'process' responds to the same intuition as [COLLAPSE]; if we have some things at some V_α, their set is formed at $V_{\alpha+1}$.

It seems then that the plural theorist must provide compelling reasons why it is not the case that some things always form a set.

Linnebo argues that anyone who rejects [COLLAPSE] must accept [LSH] as their explanation for why it is not the case that some things always form a set. If true, this would be devastating for the plural theorist; I argued earlier [LSH] cannot explain why for some things there is not always a set of those things.

Linnebo argues from two principles. The first is a plural version of Replacement that states that if there is a function that maps some things xx onto some things yy, and xx form a set, then yy too form a set.

[10]My formulation differs very slightly from Linnebo's (he abbreviates the formal statement of some things forming a set). It is easy to see the two formulations are equivalent.

[11]Though he still preserves a modified version of [P-COMP]. See [Linnebo, 2010].

[12]It would also seem that in order to accept this, Linnebo rejects Monist Realism.

The second is the principle of Cardinal Comparability [CC] which states that for any two pluralities xx and yy either the xx are fewer than the yy or the xx are at least as many as the yy. Linnebo uses these two principles to reach the conclusion[13] that the plural theorist who wishes to deny [COLLAPSE] must accept:

> [P-LSH] Some things form a set iff they are fewer than the ordinals.

Thus the plural theorist is committed to [LSH] and hence their position must be false.

Is this the case? The plural theorist should accept both premises. Further they should accept [P-LSH][14]. This is, however, not a *reductio* of their position.

Earlier, I remarked that [LSH] is *true* insofar as it is true that if there is a one-to-one mapping between a class A and the ordinal numbers then A is not a set. I argued, however, that it was not a satisfactory explanation for why A is not a set. One can have a correct method for determining when someone has got the wrong answer for the differential of a function in that the result fails to predict the correct value for the gradient of the tangent at a point. This is not an explanation that responds to the structure at play, and the notion of taking the limit of a function as it approaches a point. I submit that the plural theorist can accept [P-LSH], but merely as a *true principle* about when some things form a set, rather than an explanation for why they do not form a set.

One can object to Linnebo as follows. We should attend to the following part of Linnebo's argument:

> "A plurality thus provides the resources for a complete and precise characterization of a set. So what could prevent us from collecting the given plurality into a set?" ([Linnebo, 2010], p149)

It is true that a 'plurality' (to use Linnebo's term) provides the resources for a complete and precise characterisation of a set in the following sense. If we have a non-empty set, there will be some things that are just the things within that set. We are able to characterise certain sets by plural reference to their elements.

[13] See the Appendix to [Linnebo, 2010] for the proof.

[14] Indeed, the most obvious justification for the axiom of Replacement comes from an understanding of limitation of size, so it is hardly surprising that Replacement should imply [LSH].

However, it is not the case that for any things xx there is a set of those things. Why not? What the plural theorist needs is a satisfactory *explanation* for the fact that some things do not form sets. As noted in Chapter 2, [OHH] was just such an explanation. Thus the plural theorist may say the following; some things form a set just in case at some V_α they are all present. This was considered in the heavyweight case. It was argued, however, that the heavyweight theorist could not appeal to [OHH] and fully solve the puzzle of Chapter 1. This was due to the fact that proper classes were precise objects that thus should be able to be members of classes.

This is not a problem for the plural theorist. Some things are not an object over and above those things. Those that criticise the Plural Account (such as Linnebo) would agree with this. Where Linnebo would disagree, however, is that he would argue that there is always a set that has just those objects as members. At this point, however, the plural theorist may say that she has very good reason to deny that for any things there is always a corresponding set of those things; owing to the extendible condition used in the case of proper classes it is not the case that for any things there is some V_α containing those things. Thus the plural theorist has good reason to reject [COLLAPSE].

4.2.4 Superplurals.

A different way one could attempt to develop a puzzle for the plural theorist would be to appeal to additional resources. One could try to develop a puzzle through *superplural reference*. If we think that considerations of natural language are relevant to questions of reference it seems legitimate to appeal to languages other than English when trying to determine questions of reference.

Linnebo says the following with respect to Icelandic:

> "In Icelandic, for instance, the number words have plural forms which count, not individual objects, but pluralities of objects that form natural groups. Here is an example:
>
> 'einn skór' means one shoe
>
> 'einir skór' means one pair of shoes

‘tvennir skór’ means two pairs of shoes” [Linnebo, 2012]

Assuming that when I refer to a pair of shoes I am referring to nothing over and above the shoes (one right and one left), it seems that we are able (in Icelandic) not just to count some things, but to count ways in which some things might be organised. We plurally refer to some things organised in a particular way. This I shall call *superplural reference* or *2-plural reference.*

Aside from natural organisation on some things (such as pairs of shoes), Linnebo notes that there appears to be a difference between how some things may be presented. If we let ‘O’ represent a Cheerio for the moment, eight Cheerios may be presented in the following way:

OOOOOOOO

However, the same eight Cheerios could also be presented as follows:

OO OO OO OO

There seems to be a substantial difference here. The first is plural reference to eight Cheerios. The latter is 2-plural reference to eight Cheerios arranged into pairs of Cheerios. There is no commitment to anything above the eight cheerios other than “an additional layer of structure” ([Linnebo, 2003], p87).

There are also examples from English, such as the following:

> “imagine a video game in which any finite number n of teams can play against each other in an n-way competition. Then consider the sentences:
>
> (9a) These people and those people play against each other.
>
> (9b) These people, those people and these other people play against each other.” ([Linnebo and Nicolas, 2008], p193)

In this situation, it seems that the predicate ‘are playing against each other’ is being satisfied by some people organised into teams. The extra articulation of the structure (into teams) is relevant, and is not being captured by simple plural reference.

One might question whether the Icelandic and English examples are actually indicative of superplural reference. Hanoch Ben-Yami, for instance, rejects that the examples show superplural reference. He then gives his own account of how one may account for such cases without making use of superplural reference ([Ben-Yami, U]).

While superplural reference certainly is controversial, I take the above examples to show that it is at least plausible that such a phenomenon exists. Certainly we require an account of why some things may be arranged in different ways. Further, I will argue that even if we grant superplural reference (for the sake of argument) this does not create a problem for the Monist Realist. For these reasons I will assume that superplural reference is a legitimate form of reference.

If we accept 2-plural reference, it seems we can develop the articulation of structure further. Consider the eight Cheerios from earlier:

(i) OOOOOOOO

A case of 2-plural reference would be the following:

(ii) OO OO OO OO

However, we can organise the same eight Cheerios as follows:

(iii) OO OO   OO OO

If we had sixteen Cheerios, they might be structured in the following manner:

(iv) OO OO   OO OO   OO OO   OO OO

What is happening here? (iii) is plural reference to eight Cheerios arranged as a pair of pairs of pairs. We seemed to have pushed reference to another level of structure. Let this be 3-plural reference. (iv) Is plural reference to sixteen Cheerios arranged as a pair of pairs of pairs of pairs. The level of structure has been given at another level, hence this is 4-plural reference.

If one accepts that we may push the level if plural reference higher, there seems to be no barrier to having n-plural reference for all $n \in \mathbb{N}$. If we look at the above examples, we can see that for any case of n-plural reference, we can construct $(n+1)$-plural reference by referring plurally to n-plural reference.

Now, consider the following ω long sequence. For the sake of argument let there be ω many cheerios. Let (n) be shorthand for the contents of the n^{th} row. Let us examine the following column of Cheerios:

1st Row. The first 2 Cheerios arranged as follows: OO

2nd Row. The next 2^2 Cheerios arranged as follows: (1) (1)

3rd Row. The next 2^3 Cheerios arranged as follows: (2) (2)

4th Row. The next 2^4 Cheerios arranged as follows: (3) (3)

.

.

.

n+1st Row. The next $2^{(n+1)}$ Cheerios following the first $2+2^2+2^3+...+2^n$ Cheerios arranged as follows: (n) (n)

.

.

.

This will provide an example of ω-plural reference. Given even more Cheerios, we could then push reference higher as before to yield $(\omega + 1)$-plural reference. Given that we can provide an intuitive picture of ω-plural reference, do we have any reason to suspect that we could not provide a picture of λ-plural reference for any limit ordinal λ? While I have not given an explicit definition of λ-plural reference the above picture suggests that one should be available[15]. Further, substantial technical advancements have been made in extending superplural reference. n-plural quantification is formally equivalent to Simple Type Theory ([Linnebo, 2003], [Hazen, 1997]). This can then be extended into the transfinite and given a plural interpretation (see [Linnebo and Rayo, F], Appendix A for a full exposition).

There are two challenges raised by α-plural reference. I shall argue that both can be answered by the α-plural theorist.

[15]One suggestion that might be given further consideration is as follows. For limit λ, plural reference to $|\lambda|$ things is λ-plural reference iff those things are arranged into parts such that for each $\beta < \lambda$ there is a part so arranged that reference to the things in that part as arranged is β-plural reference. Many thanks to Marcus Giaquinto for this suggestion.

4.2.4.1 Preserving our Notion of Collection.

An argument against the α-plural theorist might run as follows. When analysing the Heavyweight View, it was seen that α-hyper-class structures splintered our notion of 'collection' into infinitely many different notions of collection. In the case of α-plural reference do we have a parallel problem? Each kind of α-plural reference is a different way of referring to many things. We seem to have 'set', '1-plural reference', '2-plural reference' etc. Is there a problem here?

The α-plural theorist does not split our notion of 'collection' into infinitely many notions. The problem with the Heavyweight View was that each kind of $(\alpha + 1)$-hyper-class had to be a different kind of collection-like object from α-hyper-classes. In the plural case, however, the interpretation is not of different kinds of 'collection'. Strictly speaking α-plural reference is not a notion of 'collection'. It is rather a way of referring to many things arranged a certain way; there is no 'collection' to figure as an object. That ways of referring should be infinitely many is far more plausible than the view that we have infinitely many notions of collection.

4.2.4.2 Puzzles for the α-plural Theorist.

A good objection to the α-plural theorist would be to generate puzzles analogous to the puzzles developed in Chapter 1. The α-plural theorist would then find herself in a similar position to the heavyweight theorist; unable to account for certain features within the Cumulative Hierarchy.

It is improbable that it is possible to generate outright contradictions using α-plural theory. As noted earlier, a theory with n-plural quantification is technically equivalent to Simple Type Theory[16]. This in turn can be extended to the transfinite case[17]. The fact that any theory using α-plural reference must respect type considerations will mean that it is unlikely to be outrightly inconsistent as a result of the α-plural reference. Despite this fact, we might wonder if the α-plural theorist faces puzzles in the sense that there are important phenomena for which their theory cannot account.

[16] See [Rayo, 2006] for an argument to this effect.
[17] See [Linnebo and Rayo, F].

Can we construct such puzzles? Essentially the same methodology as the original puzzle in Chapter 1 is not available in the context of α-plural reference. This is because (unlike the Heavyweight View) the α-plural reference to some things is very different from singular reference to a set. There are no direct analogs of the predicates 'x is a member of itself', 'x is an ordinal number', and 'x is a set' in the α-plural case. Such predicates do not even apply to α-plural reference; what is denoted by α-plural reference is not a single object[18].

Indeed it is possible to construct a Cantorian style puzzle for the α-plural theorist. Consider everything, including sets, and forms of α-plural reference etc. There are fewer of these than cases of α-plural reference to at least one of them (by Cantor's Theorem). So there are more cases of α-plural reference to at least one thing than there are things. But each case of α-plural reference to at least one thing is itself a thing. Therefore, there are not more cases of referring α-plurally to at least one thing than there are things. Thus we seem to have contradiction.

There are two ways to respond to the above 'paradox'. The first is to note that this 'paradox' is informal, it cannot be formulated in the system outlined above as it contravenes type restrictions. When I say "there are more cases of α-plural reference", "each case of α-plural reference" I am quantifying over all types of α-plural reference. This is not allowed by a system that incorporates type considerations. Nonetheless it might by argued that this presents an informal puzzle; why should we not be able to quantify over all levels of α-plural reference? I shall argue that the Monist Realist can answer this question.

It should first be noted that the Monist Realist may perfectly well accept absolute generality about sets and objects while denying that we can quantify over all levels of α-plural reference. As I have argued, α-plural reference is not a kind of object. Thus while the Monist Realist must accept that we can quantify over all objects; she need not accept absolute generality with respect to levels of α-plural reference.

Further, it is plausible that we should not be able to quantify over all instances of α-plural reference. α-plural reference is articulation of structure. If we are analysing a statement about α-plural reference, a natural question to ask is the following: "At

[18]Though obviously a single object may be what constitutes the objects involved in a case of α-plural reference.

what level of structure is the statement operating?". To take a simple example, consider four cheerios organised like this:

(a) OOOO

and like this:

(b) OO OO

Now are (a) and (b) 'the same'? This question seems to be ambiguous. This is because from a 1-plural perspective they are the same; they comprise the same objects. However on a 2-plural level they are different; the arrangement of the objects in (b) is different from the one in (a).

Thus we see that for any question about α-plural reference it is legitimate to ask at what level it is directed. The earlier Cantorian argument is directed at no particular level of α-plural quantification. It is thus not a well formed argument.

Thus we see how α-plural reference allows the plural theorist to hold onto Monist Realism, whilst being able to deny a certain kind of absolute generality. Thus they do not have similar puzzles to the one outlined in Chapter 1 to answer, they have principled reasons to argue that it is not possible to quantify over *all* α-plural levels.

One might, however, feel a Gödelian objection that the statement "It is not possible to quantify over *ALL* levels of α-plural reference." is self-undermining. For, in order to make that statement I must violate it. It is indeed impossible to precisely state the position that it is not possible to quantify over all levels of α-plural reference. However, the above example shows that the standpoint is indeed correct, even if not stateable.

However, this worry about how to state the position about quantification over levels of α-plural reference might push us to consider alternative ways to reject the Cantorian argument. Indeed, I think there is another response available.

The Cantorian argument rests on the use of the term 'everything'. This is most naturally understood as employing the condition 'x is a thing' where 'thing' ranges over objects, kinds of α-plural reference etc. We must then consider all 'things'. But we must remember, the 'paradox' will only have force if 'x is a thing' is *definite*. That was why the puzzle of Chapter 1 was so forceful, we had conditions that had definite satisfiers. Is the condition 'x is a thing' definite?

I think there is reason to suppose that it is not. For the meaning of 'thing' cannot be 'object'. α-plural reference is not an object, it is merely some things organised a certain way. As I have argued, there is not always an object that corresponds to a case of α-plural reference. In what sense then is a case of α-plural reference a 'thing'? The term 'thing' seems to have no clear meaning, and thus the claim that 'x is a thing' is a definite condition may be resisted.

4.3 Conclusions.

The Plural Account represents a satisfactory characterisation of proper classes for the Monist Realist. Though it has received a forceful attack from Linnebo using [COLLAPSE], the nature of proper classes on the Plural Account allows the plural theorist to appeal to [OHH] and reject [COLLAPSE].

A consideration of superplural reference motivates the acceptance of α-plural reference for any ordinal α. The interpretation of this phenomenon as ever deeper articulation of structure on some things allows the plural theorist to reject the claims that she either separates our notion of collection into infinitely many notions or that it is possible to generate a puzzle by quantifying over *all* α-plural levels of reference.

Thus the Monist Realist has a satisfactory account proper classes on the Iterative Conception of set; reference to proper classes is to be understood as merely reference to some sets.

Chapter 5

Conclusion

A combination of Iterative Conception of Set and Monist Realism present us with a puzzle. There are conditions ϕ such that it is definite for any object x whether or not $\phi(x)$ but no set of all ϕ appears in the Cumulative Hierarchy.

Studying the view that proper classes are objects distinct from their elements reveals a problem; what is the difference between sets and proper classes? It was then seen that [OHH] is the best explanation for why proper classes are not sets. However, the use of [OHH] alongside the Heavyweight View resulted in a proliferation of different cumulative structures. This, it was argued, is unsatisfactory.

Explaining proper classes in terms of modal resources proved to be fruitless. The first view examined explained proper classes in terms of functions from possible universes of Set Theory to subclasses of the domains of those universes. The functions appealed to by this view were shown to be in themselves paradoxical. A different explanation of proper classes in virtue of the non-rigidity of their defining conditions was also shown to be unsatisfactory. There it was seen that the Monist Realist cannot appeal to any mathematical modality.

A promising solution is to deny the existence of proper classes but to account for our talk of proper classes through using the resources of plural reference. This view was seen to respond well to Linnebo's criticism that there is a collapse of pluralities to sets, for the reason that (as proper classes are not precise objects) the plural theorist is able to appeal to [OHH]. Further examination of the view in relation to superplural reference lead to the plural theorist being pushed to α-plural quantification for any ordinal α. A Cantorian puzzle was then presented for this view. However, it was shown that the puzzle could be resisted, either by denying that it is possible to quantify over all levels of α-plural reference, or by rejecting the condition 'x is a thing' as definite.

I conclude that if one accepts Monist Realism and the Iterative Conception of Set then one should deny the existence of proper classes, but account for our talk of proper classes through plural reference.

Bibliography

[Ben-Yami, U] Ben-Yami, H. (U). Superplurals versus articulated reference. Unpublished.

[Benacerraf, 1965] Benacerraf, P. (1965). What numbers could not be. *Philosophical Review*, 74(1):47–73.

[Bernays, 1958] Bernays, P. (1958). *Axiomatic Set Theory*. North Holland Publishing Co.

[Boolos, 1984] Boolos, G. (1984). To be is to be a value of a variable (or to be some values of some variables). *The Journal of Philosophy*, 81:430–449.

[Cantor, 1883] Cantor, G. (1883). Foundations of a general theory of manifolds; a mathematico-philosophical investigation into the theory of the infinite. In Ewald, W., editor, *From Kant to Hilbert; A Source Book in the Foundations of Mathematics*, volume 2. Oxford University Press.

[Carnap, 1947] Carnap, R. (1947). *Meaning and Necessisty: A Study in Semantics and Modal Logic*. University of Chicago Press.

[Cohen, 1963] Cohen, P. (1963). The independence of the continuum hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 50:1143–1148.

[Dummett, 1991] Dummett, M. (1991). *Frege: Philosophy of Mathematics*. Harvard University Press.

[Dummett, 1994] Dummett, M. (1994). What is mathematics about? In George, A., editor, *Mathematics and Mind*. Oxford University Press.

[Fine, 1981] Fine, K. (1981). First-order modal theories i–sets. *Nous*, 15(2):177–205.

[Forster, 2008] Forster, T. (2008). The iterative conception of set. *The Review of Symbolic Logic*, 1:97–110.

[Fraenkel, 1927] Fraenkel, A. (1927). *Zehn Vorlesungen ber die Grundlegung der Mengenlehre*. Leipzig, Berlin, B.G. Teubner.

[Fraenkel et al., 1973] Fraenkel, A., Bar-Hillel, Y., and Levy, A. (1973). *Foundations of Set Theory*. Elsevier.

[Giaquinto, 2002] Giaquinto, M. (2002). *The Search for Certainty: A Philosophical Account of Foundations of Mathematics*. Oxford University Press.

[Gödel, 1940] Gödel, K. (1940). *The Consistency of The Continuum Hypothesis*. Princeton University Press.

[Hallett, 1984] Hallett, M. (1984). *Cantorian Set Theory and Limitation of Size*. Oxford University Press.

[Hazen, 1993] Hazen, A. P. (1993). Against pluralism. *Australasian Journal of Philosophy*, 71:132–144.

[Hazen, 1997] Hazen, A. P. (1997). Relations in lewis's framework without atoms. *Analysis*, 57:243–248.

[Hewitt, 2012] Hewitt, S. (2012). *A Plural Interpretation of Full Second-Order Logic*. PhD thesis, Birkbeck College.

[Hewitt, F] Hewitt, S. (F). Modalising plurals. Forthcoming in *Journal of Philosophical Logic*.

[Horsten and Welch,] Horsten, L. and Welch, P. Reflecting on the absolutely infinite. Unpublished.

[Kanamori, 2003] Kanamori, A. (2003). *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings*. Springer.

[Kanamori, 2009] Kanamori, A. (2009). Bernays and set theory. *The Bulletin of Symbolic Logic*, 15(1):43–69.

[Kirkham, 1995] Kirkham, R. (1995). *Theories of Truth: A Critical Introduction*. MIT Press.

[Lewis, 1991] Lewis, D. (1991). *Parts of Classes*. Blackwell.

[Linnebo, 2003] Linnebo, O. (2003). Plural quantification exposed. *Nous*, 37:71–92.

[Linnebo, 2007] Linnebo, O. (2007). Burgess on plural logic and set theory. *Philosophia Mathematica*, 15:79–93.

[Linnebo, 2010] Linnebo, O. (2010). Pluralities and sets. *Journal of Philosophy*, 107(3):144–164.

[Linnebo, 2012] Linnebo, O. (2012). Plural quantification. Stanford Encyclopedia of Philosophy. Accessed 5 July 2012.

[Linnebo and Nicolas, 2008] Linnebo, O. and Nicolas, D. (2008). Superplurals in english. *Analysis*, 68:186–197.

[Linnebo and Rayo, F] Linnebo, O. and Rayo, A. (F). Hierarchies ontological and ideological. Forthcoming in *Mind*.

[Parsons, 1967] Parsons, C. (1967). Frege's theory of number. In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Parsons, 1971a] Parsons, C. (1971a). Ontology and mathematics. In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Parsons, 1971b] Parsons, C. (1971b). A plea for substitutional quantification. In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Parsons, 1974a] Parsons, C. (1974a). The liar paradox. In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Parsons, 1974b] Parsons, C. (1974b). Sets and classes. *Nos*, 8(1):1–12.

[Parsons, 1977] Parsons, C. (1977). What is the iterative conception of set? In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Parsons, 1983a] Parsons, C. (1983a). Introduction. In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Parsons, 1983b] Parsons, C. (1983b). Sets and modality. In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Parsons, 1986] Parsons, C. (1986). Quine on the philosophy of mathematics. In *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Parsons, 2008] Parsons, C. (2008). *Mathematical Thought and its Objects*. Cambridge University Press.

[Priest, 2002] Priest, G. (2002). Paraconsistent logic. In *Handbook of Philosophical Logic, Volume 6, edited by D. Gabbay and F. Guenthner*. Kluwer Academic Publishers.

[Rayo, 2006] Rayo, A. (2006). Beyond plurals. In Uzquiano, G. and Rayo, A., editors, *Absolute Generality*. Oxford University Press.

[Reinhardt, 1974] Reinhardt, W. (1974). Remarks on reflection principles, large cardinals, and elementary embeddings. *Proceedings of Symposia in Pure Mathematics*, 13:189–205.

[Resnik, 1988] Resnik, M. (1988). Second-order logic still wild. *Journal of Philosophy*, 85:74–92.

[Rumfitt, 2005] Rumfitt, I. (2005). Plural terms: Another variety of reference? In Bermudez, J. L., editor, *Thought, Reference, and Experience: Themes from the Philosophy of Gareth Evans*. Oxford University Press.

[Russell, 1967] Russell, B. (1967). *The Autobiography of Bertrand Russell*. Unwin.

[Shoenfield, 1977] Shoenfield, J. (1977). Axioms of set theory. In *Handbook of Mathematical Logic*. Elsevier.

[Uzquiano, 2003] Uzquiano, G. (2003). Plural quantification and classes. *Philosophia Mathematica*, 11:67–81.

[Uzquiano, 2011] Uzquiano, G. (2011). Plural quantification and modality. *Proceedings of the Aristotelian Society*, 111:219–250.

[von Neumann, 1923] von Neumann, J. (1923). On the introduction of transfinite numbers. In van Heijenoort, J., editor, *From Frege to Gödel: A Source Book in Mathematical Logic, 1979-1931*. Harvard University Press.

[von Neumann, 1925] von Neumann, J. (1925). An axiomatisation of set theory. In van Heijenoort, J., editor, *From Frege to Gödel: A Source Book in Mathematical Logic, 1979-1931*. Harvard University Press.