Executing Gödel's programme in set theory

Executing Gödel's Programme in Set Theory

Neil Barton
PhD Thesis
Birkbeck, University of London

Submitted: January 2016
Revised: September 2016

I hereby declare that all the work presented in this thesis is my own unless indicated otherwise:

____

Neil Barton

Acknowledgements

During my PhD, I have had an enormous amount of help from some wonderful people. Any such list is almost certainly going to be incomplete, and I apologise to those who I have missed. However, I particularly remember receiving assistance in the following areas:

Producing a thesis requires a large number of institutions and individuals contributing to financial support. For this I thank the Arts and Humanities Research Council for funding me during the production of the thesis. I received very kind offers of accommodation from Adam Bradley, Benedict Eastaugh, Ethan Galebach, Richard Lawrence, Tom Pritchard, and Claudio Ternullo. Special mention must be made of Wendy and Kevin Gardner who were fantastically kind in putting me up during an extended research visit in Helsinki.

Philosophically, I have had fruitful conversations with far too many people to list. Some were especially generous with their time, however, and for this I thank Vincenzo Dimonte, Karl Egerton, Monroe Eskew, Dan Isaacson, Joel Hamkins, Jen Hornsby, Leon Horsten, Keith Hossack, Peter Koellner, Nils Kürbis, Øystein Linnebo, Penelope Maddy, John Steel, Florian Steinberger, Claudio Ternullo, Giorgio Venturi, Philip Welch, Hugh Woodin, and Jack Woods. I received very useful help with technical matters from the communities on `https://math.stackexchange.com/` and `https://mathoverflow.net/`, especially Andrés Caicedo, GME, and Asaf Karagila. I am indebted to the Birkbeck Philosophy Department and the variety of reading groups I have either organised or attended in London, Bristol, and Oxford for providing a stimulating and challenging environment for research. In organising these groups, I thank Simmi Pahwah for her hard work and patience.

On various occasions, friends and scholars went above and beyond in providing me with responses at conferences or detailed written comments on material related to this thesis. For this I thank Carolin Antos, Salvatore Florio, Victoria Gitman, Simon Hewitt, Luca Incurvati, Ethan Jerzak, Alex Kocurek, Toby Meadows, Michael Potter, Sam Roberts, Sean Walsh, and John Wigglesworth. The final version of the thesis benefited greatly from close reading, generous comments, and a rigorous examination given by Bob Hale and Alex Paseau.

There are, however, four people who deserve special mention in shaping the thesis, both in spirit and content. My MPhil Stud supervisor Marcus Giaquinto has always been a bastion of support and rigour; if the thesis comes close to the standards he sets for himself, I will be very satisfied. My good friend Chris Scambler gave me exceptionally useful and creative feedback on much of the material throughout its development. Without frequent and patient discussion with Sy Friedman, it would not have been possible to execute the technical part of the project. Finally, my supervisor Ian Rumfitt has been enormously helpful; providing insightful and encouraging help on every aspect of the thesis.

Though my academic debts are large, my personal debts are also great. Throughout the ups and downs of the thesis I have had support from a stalwart group of friends and family. For this, I am especially grateful to Kathrine Cuccuru, Rory Dyson, Karl Egerton, Olly Finegold, Sylvain Garcon, Amy Gardner, Nikki Herwanger, Charlotte Knowles, Tobie Koster, Victoria Lain, Michael Marsh, Charlotte Moultrie, Stefania Polledri, Tom Pritchard, Geoff Reading, Josephine Salverda, Chris Scambler, Hannah Sauer, Jack Woods, and John Wigglesworth. I thank the University of London Boat Club and Richmond Canoe Club for good times and great company during graduate school. My dog Patch was a fantastic companion during the long writing hours.

Special mention goes to my parents. My father Paul was an unfailing source of encouragement and resolve. I thank my mother Jeanne (to whom I dedicate this thesis) for her support, love of education, and convincing a disillusioned teenager that mathematics was beautiful.

For Jeanne:

Lights fade,
Time grows short,
The road gently calls.
But the doors stay open still...

Abstract

The study of set theory (a mathematical theory of infinite collections) has garnered a great deal of philosophical interest since its development. There are several reasons for this, not least because it has a deep *foundational* role in mathematics; any mathematical statement (with the possible exception of a few controversial examples) can be rendered in set-theoretic terms. However, the fruitfulness of set theory has been tempered by two difficult yet intriguing philosophical problems: (1.) the susceptibility of naive formulations of sets to contradiction, and (2.) the inability of widely accepted set-theoretic axiomatisations to settle many natural questions. Both difficulties have lead scholars to question whether there is a *single*, *maximal* Universe of sets in which all set-theoretic statements are determinately true or false (often denoted by 'V'). This thesis illuminates this discussion by showing just what is possible on the 'one Universe' view. In particular, we show that there are deep relationships between responses to (1.) and the possible tools that can be used in resolving (2.). We argue that an interpretation of *extensions* of V is desirable for addressing (2.) in a fruitful manner. We then provide critical appraisal of extant philosophical views concerning (1.) and (2.), before motivating a strong mathematical system (known as'Morse-Kelley' class theory or '**MK**'). Finally we use **MK** to provide a coding of discourse involving extensions of V, and argue that it is philosophically virtuous. In more detail, our strategy is as follows:

Chapter I ('Introduction') outlines some reasons to be interested in set theory from both a philosophical and mathematical perspective. In particular, we describe the current widely accepted conception of set (the 'Iterative Conception') on which sets are formed successively in stages, and remark that set-theoretic questions can be resolved on the basis of two dimensions: (i) how 'high' V is (i.e. how far we go in forming stages), and (ii) how 'wide' V is (i.e. what sets are formed at successor stages). We also provide a very coarse-grained characterisation of the set-theoretic paradoxes and remark that extensions of universes in both height and width are relevant for our understanding of (1.) and (2.). We then present the different motivations for holding either a 'one Universe' or 'many universes' view of the subject matter of set theory, and argue that there is a stalemate in the dialectic. Instead we advocate filling out each view in its own terms, and adopt the 'one Universe' view for the thesis.

Chapter II ('Gödel's Programme') then explains the Universist project for formulating and justifying new axioms concerning V. We argue that extensions of V are relevant to both aspects of Gödel's Programme for resolving independence. We also identify a 'Hilbertian Challenge' to explain how we should interpret extensions of V, given that we wish to use discourse that makes apparent reference to such non-existent objects.

Chapter III ('Problematic Principles') then lends some mathematical precision to the coarse-grained outline of Chapter I, examining mathematical discourse that seems to require talk of extensions of V.

Chapter IV ('Climbing above V?'), examines some possible interpretations of height extensions of V. We argue that several such accounts are philosophically problematic. However, we point out that these difficulties highlight two constraints on resolution of the Hilbertian Challenge: (i) a Foundational Constraint that we do not appeal to entities not representable using sets from V, and (ii) an Ontological Constraint to interpret extensions of V in such a way that they are clearly different from ordinary sets.

Chapter V ('Broadening V's Horizons?'), considers interpretations of width extensions. Again, we argue that many of the extant methods for interpreting this kind of extension face difficulties. Again, however, we point out that a constraint is highlighted; a Methodological Constraint to interpret extensions of V in a manner that makes sense of our naive thinking concerning extensions, and links this thought to truth in V. We also note that there is an apparent tension between the three constraints.

Chapter VI ('A Theory of Classes') changes tack, and provides a positive characterisation of apparently problematic 'proper classes' through the use of plural quantification. It is argued that such a characterisation of proper class discourse performs well with respect to the three constraints, and motivates the use of a relatively strong class theory (namely $\mathbf{MK}$).

Chapter VII ('V-logic and Resolution') then puts $\mathbf{MK}$ to work in interpreting extensions of V. We first expand our logical resources to a system called V-logic, and show how discourse concerning extensions can be thereby represented. We then show how to code the required amount of V-logic using $\mathbf{MK}$. Finally, we argue that such an interpretation performs well with respect to the three constraints.

Chapter VIII ('Conclusions') reviews the thesis and makes some points regarding the exact dialectical situation. We argue that there are many different philosophical lessons that one might take from the thesis, and are clear that we do not commit ourselves to any one such conclusion. We finally provide some open questions and indicate directions for future research, remarking that the thesis opens the way for new and exciting philosophical and mathematical discussion.

Contents

Chapter I

Introduction

I.1 The rise of set theory

The 20^{th} century saw a huge amount of change within mathematics and its philosophy, partly precipitated by the powerful combinatorial methods offered by the development of set theory. On the one hand, the new set-theoretic toolkit facilitated several mathematical discoveries, including Cantor's realisations that the study of sets gave rise to new and various cardinalities and that there is, for any natural number n and space of the form $\mathbb{R}^n$, a bijection between the unit interval and all the points of $\mathbb{R}^n$.[1] Further, using the vast array of mathematical objects postulated, it is possible to reinterpret almost[2] any claim concerning mathematical objects as one about sets. While the foundational significance of such a reduction is contested[3] set theory nonetheless remains philosophically interesting for a number of reasons. First, set theory represents our best mathematical theory of infinity. Second, the combinatorial methods offered provide the current standard for confirming theories as non-vacuous; a natural way to answer consistency concerns in mathematics is

[1] This led Cantor to remark (in correspondence with Dedekind) "Je le vois, mais je ne le crois pas!", Translated: "I see it, but I don't believe it!" ([Ewald, 1996a], p860). It is an interesting question what the Cantor of the turn of the century would make of contemporary set-theoretic methods (in their full transfinite resplendency).

[2] There are possible exceptions to this claim. For example, the study of *category theory* is a discipline that seems, prima facie, to go beyond the methods of set theory. There is a vast literature surrounding this issue, and it is not at all clear that we should be persuaded that this appearance is philosophically committing. For discussion and exposition of some of the technical matters, see [Mac Lane, 1971], [Muller, 2001], and [Linnebo and Pettigrew, 2011].

[3] See, for example, [Quine, 1960] who advocates a total set-theoretic reduction, the seminal [Benacerraf, 1965] arguing that such a reduction cannot tell us what the objects of mathematics are, and more recently [Paseau, 2009] for a defence of reductionism.

to show that a particular theory has a model in the sets. Third, by beginning from set-theoretic principles, we are able to set the standard for rigour in mathematics, and ensure that it is not compromised in moving from one area of mathematics to another.[4]

Despite these clear theoretical virtues, set theory has, almost since its discovery, posed several infamous philosophical and mathematical problems. Initial attempts to axiomatise a theory of extensions (such as [Frege, 1893]) were flat out inconsistent, and a vast number of natural set-theoretic questions have turned out to be irresolvable using the standard axioms.

The philosophical landscape that has emerged as a result of these problems is largely fractured, with many different tenable attitudes to set theory and its role within philosophy. In this thesis, we work within realist philosophies of mathematics (broadly construed), in the sense that we assume that mathematical objects are mind-independent entities, and mathematical truth is determined as a relationship between syntax, interpretation, and this ontology. As we shall see in this chapter, a major question that has emerged within such philosophies is the following: "How many universes of sets are there?". The answer to this question, as shall become apparent, often depends on intuitions concerning the practice of mathematics. We begin, therefore, by briefly characterising some of the foundational issues to be brought into sharper focus in the remainder of the thesis.

I.2 Proper classes

The first issue concerns the problems of paradox that have been a part of set theory practically since its inception as a foundational framework. An appealing principle, when one starts to think of mathematical collections is a principle of *comprehension* for sets. If ϕ is any condition in the language of set theory $\mathscr{L}_\in$, we might think that the following obtains:

[Naive Comprehension] $\exists y \forall x [x \in y \leftrightarrow \phi(x)]$

[4] For a detailed exposition of this position, see the excellent [Burgess, 2015] (esp. Ch 2).

This states that for any particular condition ϕ, there is a set of those things satisfying ϕ. The principle is initially attractive; sets are often thought of as collections over and above their elements with a definite membership relation. Given some precise ϕ then, we might believe that Naive Comprehension should hold; since it seems definite what objects fall under ϕ, we might expect to be able to collect them into a set. As is well known, however, it is possible (in combination with other widely held principles of set theory) to quickly derive contradictions.

We briefly survey the paradoxes to which Naive Comprehension gives rise. As we shall see (in Chapter VI) the exact interpretation to be given to the ranges of particular paradoxical conditions is philosophically significant. In order to clarify the rest of the thesis, we first mention the background set theory we shall use (**ZFC**), and then recast the paradoxes as theorems of **ZFC**.

I.2.1 ZFC

The axiomatic system we shall consider is *Zermelo-Fraenkel Set Theory with Choice* or **ZFC**. It is the most widely[5] used set theory within foundations.[6] **ZFC** has classical first-order logic with identity as the background theory, and the only non-logical predicate is set membership, denoted by '$\in$'. We shall refer to this language as $\mathscr{L}_{\in}$. **ZFC** comprises the following axioms:

Axiom 1. *Axiom of Extensionality.* $\forall x \forall y[\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y]$.

Intuitive characterisation. Any two sets with the same members are identical.

Axiom 2. *Axiom of Pairing.* $\forall x \forall y \exists p \forall z[z \in p \leftrightarrow (z = x \vee z = y)]$.

Intuitive characterisation. For any two sets x and y there is a set containing just x and y.

Axiom 3. *Axiom of Union.* $\forall x \exists y \forall z \forall w[(w \in z \wedge z \in x) \rightarrow w \in y]$.

[5] Though there are alternative systems, see for example the kinds of set theory arising from the *New Foundations* programme, given by [Quine, 1937] and developed by [Forster, 1995] and [Holmes, 1998]. There are also various constructive foundations, see [Feferman, 2009] and [Rathjen, 2012] for some examples.

[6] See [Fraenkel et al., 1973] for a discussion of the genesis of **ZFC** and comparison with other theories.

Intuitive characterisation. For any set x, there is a set of all elements of members of x.

Axiom 4. *Axiom of Choice.* If $\mathcal{F}$ is a set of pairwise-disjoint non-empty sets then:

$\exists c \forall x \in \mathcal{F}\, \exists y (c \cap x = \{y\})$.

Intuitive characterisation. For any non-empty set of pairwise-disjoint non-empty sets, there is a set that picks one member from each.[7]

Axiom 5. *Axiom of Infinity.* $\exists x [\exists y\, y \in x \wedge (\forall z\, z \in x \rightarrow z \cup \{z\} \in x)$.

Intuitive characterisation. There is a non-empty set such that if it contains a set z, it also contains z unioned with its singleton. The axiom thus guarantees the existence of an infinite set.[8]

Axiom 6. *Power Set Axiom.* $\forall x \exists y \forall z (z \in y \leftrightarrow z \subseteq x)$.

Intuitive characterisation. For any set x, there is a set of all subsets of x.

Axiom of Foundation. $\forall x (x \neq \emptyset \rightarrow \exists y \in x\, y \cap x = \emptyset)$.

Intuitive characterisation. Every set contains an element that is disjoint from it. The axiom both rules out self-membered sets and also the existence of infinite descending membership chains.

Axiom 7. *Axiom Scheme of Separation.* If ϕ is a formula in $\mathscr{L}_{\in}$ with y not free then:

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

Intuitive characterisation. Given a set x, one can 'separate' out the ϕs from x into a new set y.

Remark 8. In first-order **ZFC**, this is actually a *scheme*, since there is one axiom for every formula ϕ of the correct form.

---

[7] There are a large number of equivalents of the Axiom of Choice, both within set theory and from other areas of mathematics. We choose this formulation because it most naturally meshes with motivation from the Iterative Conception of Set (see below).

[8] This claim is made slightly more complex by the fact that within set theory there are different definitions of the notion of *infinite set*. This is usually made precise by defining a set to be *finite* iff it is bijective with (a von Neumann representative of) a natural number (and infinite otherwise) or, alternatively, being *Dedekind-infinite* iff it is bijective with a proper subset of itself. The issue is somewhat subtle, as the two notions can come apart in the absence of the Axiom of Choice. See [Jech, 2002], p34 for details.

Separation is derivable from the following axiom (*modulo* the other axioms of **ZFC**):

Axiom 9. *Axiom Scheme of Replacement.* Let $\phi(p, q)$ define a function, in the sense that if $\phi(x, y)$ and $\phi(x, z)$ both hold then $y = z$. Then:

$\forall x \exists y \forall z[z \in y \leftrightarrow \exists p \in x\phi(p, z)]$.

Intuitive characterisation. Again, the above axiom represents an axiom scheme, and states that if ϕ defines a function, then the image of any particular set under ϕ is also a set.

If one wishes to remove the schematic nature of the above axiom, one might instead opt for a *second-order* statement of Replacement (and therefore Separation). For any *second-order* entity[9] representing a function, we have:

Axiom 10. *Second-Order Replacement.* $\forall F \forall x \exists y \forall z[z \in y \leftrightarrow \exists p \in xF(p, z)]$

We shall indicate the amount of higher-order resources allowed with subscripts (so '$\mathbf{ZFC_2}$' denotes *second-order* **ZFC**). In all other cases, where a difference is not obvious from context, '**ZFC**' denotes first-order **ZFC**.

The status of justification concerning axioms of **ZFC** is a thorny philosophical issue in itself. There is some controversy surrounding each of the axioms, with the possible exception of Extensionality.[10] Certainly the justification of the different Replacement Axioms has been questioned (see for example [Potter, 2004]) as has the Power Set Axiom (see [Feferman et al., 2000]). Since we are concerned here with settling questions in **ZFC** and its extensions, we shall largely leave the justification of **ZFC** unexamined, except where it helps to elucidate particular concepts or other justifications.

As it turns out, full Naive Comprehension is inconsistent with **ZFC**. In order to see from where contemporary set theory has arisen, and set up some material that will be essential for arguments later, we provide a brief reconstruction of the

[9]We use the term 'entity' because it is not yet clear how to interpret the second-order variables. We shall see some analysis of possible interpretations in Chapter VI. However, it should be noted that there are other interpretations not canvassed in this thesis (see, for example, [Linnebo, 2006] and [Hale, 2013]).

[10]See, for example, Boolos' comment that if any axiom deserves to be called analytically true, then the Axiom of Extensionality is probably a good candidate ([Boolos, 1971], p230).

paradoxes as negative theorems of **ZFC**, each a counterexample to Naive Comprehension (rather than in their original presentation as actual antinomies).

I.2.2 Russell's Paradox

Russell's paradox cuts to the very heart of set theory, in particular because the reasoning turns upon a simple use of the membership relation.

> **Theorem 11.** [Russell, 1902] There is no set of all non-self-membered sets.
>
> *Proof.* Suppose there is such a set (let it be denoted by 'R'). Since R is a set, we may consider whether or not $R \in R$. Suppose $R \in R$. Then R is self-membered, and so $R \notin R$, contradicting the assumption that $R \in R$, $\perp$. We can conclude, therefore, that $R \notin R$. But then R is non-self-membered, and hence $R \in R$, $\perp$. ■

I.2.3 Cantor's Paradox

Cantor's Paradox turns on the mathematically fruitful notion of *cardinality*, where two sets have the same cardinal number iff there is a bijection between them.[11]

> **Theorem 12.** [Cantor, 1899][12] There is no universal set (set of all sets).
>
> *Proof.* Suppose that there is such a set, let it be denoted by 'U'. We can then apply the Power Set Axiom to U to yield $\mathcal{P}(U)$. By Cantor's Theorem, $\mathcal{P}(U)$ contains strictly more members than U. But as U was meant to contain all sets, every member of $\mathcal{P}(U)$ is a set and hence a member of U, $\perp$. Thus, there is no such set.[13] ■

[11] This method for measuring the size of (possibly infinite) sets, originating with Cantor, is by far the most widely used in modern set theory. However, for an alternative way of measuring set size, see [Benci et al., 2006].

[12] We call this theorem *Cantor's* Paradox because it uses Cantor's Theorem concerning the relative cardinalities of sets and their power sets. There is a question as to whether Cantor himself would have approved of the proof given here due to worries about the range of application of the power set axiom. For textual evidence, see his correspondence with Hilbert in [Cantor et al., 1899].

[13] Often the proof is significantly shortened by simply producing the Russell reasoning and arguing that a Russell set is obtainable by applying Separation (a quick consequence of Replacement) to U with the condition $x \notin x$. The presentation given in terms of Cantor's Theorem is more instructive for understanding the nature of the paradoxes; it shows how notions of cardinality interact with the Power Set Axiom and the existence of a Universal Set.

I.2.4 The Burali-Forti Paradox

The Burali-Forti Paradox (noticed by [Cantor, 1899] and by [Russell and Whitehead, 1910] on the basis of a theorem in [Burali-Forti, 1897]), is also interesting because of the mathematical fruitfulness of the concept in question. The notion of *ordinal* and *order-type* are, informally understood, ways a collection can be ordered in a linear and well-founded manner. These concepts, as we shall see, are central to contemporary set theory. A difficulty when speaking of these notions is that there are a variety of philosophical and mathematical distinctions at play. In order to keep the required level of precision, we will first clarify our notation by making some remarks about the foundational role of set theory.

We remarked in §1 that it was possible to *represent* most (if not all) mathematical objects with sets. There are, therefore, two ways in which we may talk about mathematical entities:

(1) We may speak about the mathematical entities *in themselves*.

(2) We might instead talk about the set-theoretic *representatives* of the entities in question.

It is an open philosophical question whether or not mathematical entities simply *are* their set-theoretic representatives, and one we lack the space to address here.[14] It will be important throughout this thesis to be very precise about what we have in mind when addressing the topic of "ways a collection can be ordered in a well-founded manner".

We first set up some notation that will be useful for distinguishing philosophical and mathematical concepts throughout the thesis:

Notation 13. Concerning ordinals and order-types, we will follow the following conventions:

(i) By *order-type* we mean the mathematical entity corresponding to a

[14] See [Benacerraf, 1965] for an argument that set-theoretic representatives (or any other objects for that matter) cannot simply be what we talk about in mathematics, and [Paseau, 2009] for a defence of a set-theoretic reduction.

well-ordering (whatever that may be). We will denote these by underlined lower-case Greek letters, e.g. '$\underline{\alpha}$', '$\underline{\beta}$', '$\underline{\gamma}$', and so on.

(ii) By *ordinal* we mean a *von Neumann representative*[15] of an order-type $\underline{\alpha}$, and will denote these by the lower-case Greek letters 'α', 'β', 'γ', etc.

(iii) The *class* of all ordinals of a model $\mathfrak{M}$, whatever its final interpretation may be, will be denoted by '$On^{\mathfrak{M}}$'.

(iv) Any *set* of all the ordinals in a model $\mathfrak{M}$ will be denoted by '$Ord(\mathfrak{M})$'.

(v) The *order-type* exemplified by $On^{\mathfrak{M}}$ will be denoted by '$\Omega^{\mathfrak{M}}$'.

Theorem 14. [Burali-Forti, 1897], [Cantor, 1899], and [Russell and Whitehead, 1910] There is no set of all ordinals (i.e. $Ord(V)$ does not exist).

Proof. Clearly $Ord(V)$ is transitive; it contains all and only the transitive pure sets well-ordered by $\in$, and so for any ordinal $\alpha \in Ord(V)$, $Ord(V)$ contains all $\gamma < \alpha$ (since, by definition, every $\gamma < \alpha$ is also a transitive set well-ordered by $\in$), and hence all $\gamma \in \alpha$. However, $Ord(V)$ is also well-ordered by $\in$. To see this, take any $\emptyset \neq X \subseteq Ord(V)$. Suppose X does not have a unique least element. Then, either (i) X contains an infinite descending membership chain, contradicting the Axiom of Foundation, or (ii) X contains two or more least elements, let them be denoted by 'δ' and 'ζ'. Since δ and ζ are both ordinals, they are transitive sets well-ordered by $\in$. Hence $\delta \cup \zeta \cup \{\delta\} \cup \{\zeta\}$ is a transitive set well-ordered by $\in$. However, this immediately yields that either $\delta < \zeta$ or $\zeta < \delta$ (given that $\delta \neq \zeta$). In either case, this contradicts the claim that both δ and ζ are $\in$-least in X. Thus $Ord(V)$ is itself a transitive set well-ordered by $\in$, and hence $Ord(V) \in Ord(V)$, contradicting the Axiom of Foundation, $\perp$. ■

[15] The von Neumann ordinal representatives are defined recursively as follows. The representation of the order-type of zero is the empty set. Next, if α is the representation of $\underline{\alpha}$, $\alpha + 1 = \alpha \bigcup \{\alpha\}$. Finally for limit $\underline{\lambda}$, $\lambda = \bigcup_{\beta < \underline{\lambda}} \beta$. Equivalently, an ordinal is a transitive pure set well-ordered by $\in$, where a set x is *transitive* iff whenever $y \in x$ and $z \in y$ then $z \in x$ (i.e. x contains all elements of its members).

I.2.5 Naive Comprehension examined

Let us take stock. We have seen that Naive Comprehension is inconsistent with the axioms of **ZFC**. However, something of a puzzle remains. For, given any particular structure[16] $\mathfrak{M} \models \mathbf{ZFC}$ and paradoxical condition ϕ, there are some sets in $\mathfrak{M}$ such that each of them satisfies ϕ, but there is no set of all the ϕ in $\mathfrak{M}$. Some have taken this datum to be philosophically problematic. Boolos, for example, writes:

> "Wait a minute! I thought that set theory was supposed to be a theory about all, 'absolutely' all, the collections that there were and that 'set' was synonymous with 'collection'." ([Boolos, 1998], p35)

As Boolos points out, we would like set theory to be our theory of all collections whatsoever, yet it appears that (in a particular $\mathfrak{M}$) there are collections (namely the collection of all ϕ) that are not in $\mathfrak{M}$. One might think that this threatens **ZFC** set theory as our best theory of infinite collections; given a certain domain satisfying **ZFC**, there are infinite collections not in that domain.[17]

Thus far, we have only *stated* the standard axiomatic system (i.e. **ZFC**) that proves the existence of enough sets (and more) to represent all mathematical objects. From a philosophical perspective, however, there is a further salient question: can such an axiomatic system be justified on good philosophical grounds?

I.3 The Iterative Conception of Set

The standard justification given has often been the Iterative Conception of Set. Under such a conception, sets are formed in a well-ordered sequence of *stages*. Shoenfield expresses the conception as follows:

> "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)

[16] A note on notation: Throughout this thesis, uppercase fraktur letters denote arbitrary models of **ZFC** (whether set-sized or proper-class-sized).

[17] We shall examine these issues in more detail in Chapter VI. As we shall see, the interpretation to be given to such discourse over a particular structure is philosophically significant.

Thus, under such a conception, we begin with the empty set[18], and iterate the power set operation through the ordinals, collecting stages together at limits. More formally, we define the *Cumulative Hierarchy of Pure Sets* (or 'V') as follows:

Definition 15. *The Cumulative Hierarchy of Pure Sets:*

$V_0 = \emptyset$

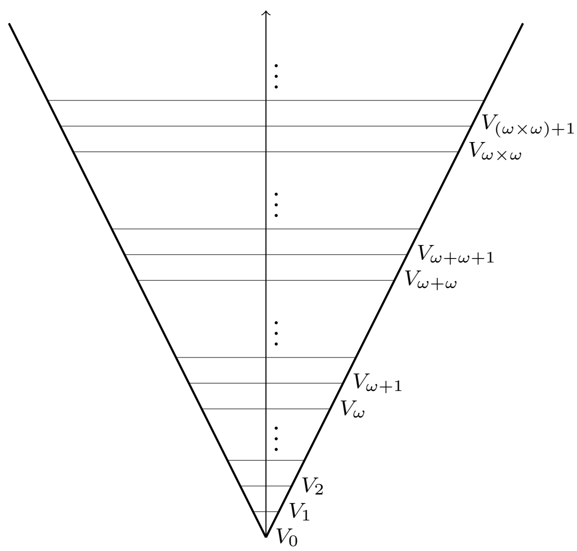
$V_{\alpha+1} = \mathcal{P}(V_\alpha)$

For a limit ordinal λ:

$V_\lambda = \bigcup_{\beta<\lambda} V_\beta$

$V = \bigcup V_\alpha$ for $\alpha \in On$ (where On denotes the class of all ordinals[19]).

One can visualise the Cumulative Hierarchy in the following manner:



The conception is pleasing for a number of reasons. First, the picture provided by the Iterative Conception is theoretically simple and elegant. For example Boolos writes:

[18]We could also have begun with a set of urelemente if we wished to examine impure set theory. Here, we simply remark that for most mathematical purposes the addition of urelemente seems an unnecessary complication (there are usually plenty enough sets to represent some given set of urelemente), and so we refrain from providing discussion. This said, such systems have independent *meta*mathematical interest; see, for example [McGee, 1997].

[19]Of course, we have not said yet how we should understand reference to the class of all ordinals (an especially salient point given the set-theoretic paradoxes). For now we simply leave this unexamined and assume that the term can be given *some* philosophically acceptable interpretation. The issue will, however, be examined in detail in Chapter VI.

"**ZF** alone (together with its extensions and subsystems) is not only a consistent (apparently) but also an independently motivated theory of sets: there is, so to speak, a "thought behind it" about the nature of sets which might have been put forth even if, impossibly, naive set theory had been consistent. The thought, moreover, can be described in a rough, but informative way without first stating the theory the thought is behind." ([Boolos, 1971], p219)

Now, we need not agree with Boolos that the Iterative Conception is such that it might (in some appropriate sense of modality) have been thought of without the discovery of the paradoxes. Such a claim, at the very least, seems controversial. However, we can agree that the picture provided is conceptually simple, and helps the set theorist understand the range of standard[20] models with which she works.[21] The philosophical significance of this conceptual simplicity is hotly contested[22], but nonetheless it is part of what makes the Iterative Conception philosophically engaging.

Second, the Iterative Conception easily blocks the formation of problematic sets, and provides an explanation of why Naive Comprehension is false. Key here is that for each of the paradoxical conditions (given appropriate set-theoretic interpretation) there are satisfiers of the condition unbounded in a particular[23] iterative hierarchy. To see this more clearly, we briefly explain the cases of the conditions considered thus far.

The case of the universal set is clear; new sets appear unboundedly throughout the V_α. Similarly, it is easy to see that *every* set on the Iterative Conception is non-self-membered, and hence the non-self-membered sets also appear unboundedly. The issue with the Burali-Forti Paradox is a little more subtle, as it is unclear what constitutes an order-type in set theory. However, it is standard practice to represent

[20]It is a simple consequence of the Compactness Theorem for first-order theories that **ZFC** has non-standard non-well-founded models. These are clearly *not* founded on the Iterative Conception (however, see [Field, 2003], [Hamkins, 2012], and [Barton, Fa] for discussion of the extent to which we can say such models are unintended).

[21]For more on precisely what is supplied by the Iterative Conception, especially contrasted to category theory, the reader is directed to [Linnebo and Pettigrew, 2011].

[22]See, for example, [Boolos, 1971], [Potter, 2004], [Paseau, 2007], [Maddy, 1988a] and [Maddy, 1988b].

[23]We say "particular" here, because this is true on any iterative hierarchy of any height.

an order-type by a corresponding set-theoretic representative (commonly a von Neumann ordinal[24]). It is then easy to show again that such representatives do not form a set, and hence (at least insofar as set-theoretic representations are concerned), a contradiction will not arise in from considering the representations of these objects.[25] The situation is markedly more complex if we admit order-types as *urelemente* into our theory of sets, as we could then have a set of all order-types appearing in V_1. As our interest is in pure set theory, we put aside these worries, but note that it represents a deep an interesting question for analysing the interaction between ordering properties and set theory.[26]

A final reason that some philosophers have been attracted to the Iterative Conception is that is that many have felt that it is compatible with justifications for the standard axioms of **ZFC**. Extensionality, we note, is justified on independent grounds; it is something like a (possibly partial) definition of what it is to be a set. Pairing follows from the fact that given any two sets x and y existing in some V_α, the pair $\{x, y\}$ is formed at latest at $V_{\alpha+1}$. Union is justified by the fact that for any set x first formed at V_α, all elements of x exist at stages previous to V_α, and hence all members of elements exist at stages prior to V_α, and so the set of all members of elements of x exists (at latest) at stage V_α. The Power Set Axiom is justified by appealing to the fact that at each additional stage we collect together *all* subsets of V_α. Thus for some set x first formed at V_α, all its elements exist prior to V_α, hence all sets of elements of x exist at V_α, and hence $\mathcal{P}(x)$ exists at $V_{\alpha+1}$. The case of Foundation is a little more complex. Clearly there cannot be circular membership chains; if x is first formed at V_α then in order to contain itself it would have to be first formed before V_α, an immediate contradiction. If instead there were an infinite descending membership chain, we would have an infinite descending sequence of stages. Since the stages are indexed by the ordinal number sequence, we would then obtain an infinite descending sequence of ordinals, contradicting the understanding of ordinals as well-ordered.

[24]The Burali-Forti reasoning is also blocked when representing order-types with other kinds of sets, such as Zermelo's representation (modified appropriately to allow generalisation to infinite sets), or representing ordinals by (restricted) equivalence classes of sets (in the case where we have no restriction to least rank on the equivalence classes, then not even a single ordinal is a set in a given iterative hierarchy).

[25]A fuller exposition of these issues is given in [Barton, 2012].

[26]See [Menzel, 1986] and [Menzel, 2014] for discussion and attempts at providing a framework theory.

The Axioms of Infinity, Choice, and Replacement are more vexed, however, but nonetheless arguments have been offered for each.

Infinity is perhaps a good candidate for an axiom that receives little justification from the Iterative Conception alone, but nonetheless is independently justified. We can think of Infinity as asserting the existence of a set of all von Neumann natural numbers (*modulo* extracting a minimal inductive[27] set using Replacement). The fact that we are able to work coherently with the set of natural numbers, sets of natural numbers, and functions on reals, is witnessed by the independent conceptions offered by number theory and analysis. However, viewed from an iterative standpoint, we do not seem to get outside the finite; iterating on finite levels will only yield more finite levels until we take the union at V_ω. There are thus two main justifications to which one can appeal on the Iterative Conception. One is to argue that it is part of the Iterative Conception that we iterate 'as far as possible', and the apparent consistency of operating under the Axiom of Infinity indicates that it is possible to iterate to infinite stages. On this justification, however, there is the deep problem of explaining how we should understand the notion of 'as far as possible'. A second justification is to note that we seem to work with infinite sets in non-set-theoretic mathematics anyway, and so the Axiom of Infinity functions as something like a prior mathematical assumption upon which we layer set theory.[28] Certainly, however, the Iterative Conception is at least *compatible* with having infinite stages, and amenable to heuristic explanations of why it meshes well with the Axiom of Infinity.

Turning to Choice, several authors have thought that AC follows from the notion of arbitrary subset. We can put the argument informally as follows. Let x be first formed at V_α. We know then that all elements of x exist prior to V_α. Thus, all members of elements of x also exist before V_α. Since we form *all* sets at successor stages, a set C containing exactly one member of every element of x is formed by V_α. Hence C witnesses the truth of AC.

The argument is patently unconvincing to a disbeliever in AC; by saying that the subsets formed at successor stages facilitate the existence of a choice set for x, we

[27]A set y is *inductive* iff y is non-empty and $x \in y$ implies that $x \cup \{x\} \in y$. The existence of an inductive set is precisely what the Axiom of Infinity delivers.
[28]For additional discussion of the Axiom of Infinity see [Maddy, 1988a].

implicitly assume a notion of arbitrary subset that includes some form of choice principle. However, such a justification nonetheless provides explanation for a believer in Choice as to *why* the principle holds in the Cumulative Hierarchy.[29]

The justification for Replacement is at least as bad as Choice. Some have argued that it is part of the iterative conception that we iterate the stages maximally far (for example, [Drake, 1974]). Given this, they argue, our theory of sets should satisfy *first-order* reflection: the claim that for any first-order ϕ in $\mathscr{L}_\in$, if ϕ holds then for every α there is a $\beta > \alpha$ such that $V_\beta \models \phi$. Such a principle is (modulo $\mathbf{ZC}-$Infinity) equivalent to the Axioms of Infinity and Replacement.[30] The extent to which such a principle follows from iterativity *alone* is hotly contested[31] and exactly what such justifications guarantee is a subtle question[32].

Instead, some authors have considered *adding* additional content to our concept of set. Boolos, for example, says the following:

> "Perhaps one may conclude that there are at least two thoughts "behind" set theory." ([Boolos, 1989], p19)

The paper in question provides a justification of several axioms (including Choice and Replacement) based on a modification of Naïve Set Theory denoted by **FN** (for Frege-von-Neumann). Central to Boolos' arguments is a restricted form of Frege's inconsistent Basic Law V (called 'New V') which incorporates a notion of *limitation of size*. It would take us too far afield to study abstraction principles such as New V in detail; however, limitation of size will be essential for parts of the thesis.

The notion of limitation of size encapsulates the idea that some objects form a set just in case they are not 'too big'. This seems to contrast with the Iterative Conception which embodies a 'limitation of rank' principle; some objects form a set just in case they are not unbounded in the V_α. Though the two conceptions are often regarded as competing, it is an interesting question the extent to which the two can be fused.[33]

[29] For a more in-depth review of these sorts of argument, see [Maddy, 1988a]. For a full presentation of the second-order argument for AC based on the Iterative Conception, see [Potter, 2004], §14.8. For a detailed exposition of some of the technical issues surrounding AC, see the excellent [Jech, 1973].

[30] See [Levy, 1960] and [Montague, 1961].

[31] See [Boolos, 1971], [Boolos, 1989] and [Paseau, 2009].

[32] For discussion see [Koellner, 2009], [Welch, 2014], and [Barton, Fb].

[33] See Chapter VI and [Hallett, 1984] for discussion.

Of course what counts as 'too big' is up for debate. Limitation of size comes in two main forms[34]:

Weak Limitation of Size Principle. If some objects can be put in a one-to-one correspondence with a set, then there is a set of those objects.

Strong Limitation of Size Principle. Some objects form a set iff they are not bijective with the universe.

The latter entails the former, but not vice versa (Weak Limitation of Size does not guarantee that if some objects are not bijective with all sets then there is a set containing all of them). We shall see additional discussion of limitation of size in Chapter VI. For now, we note that under both we easily obtain Replacement. By Choice, for every set x there is an ordinal α such that there is a bijection between x and α. Further, if we have a set x and a function f on x, the range of f cannot be *bigger* than x. Hence $y = ran(f)$ is bijective with an ordinal of length at most α, and hence y is both not bijective with the ordinals and bijective with a set (namely some $\gamma \leq \alpha$). Unfortunately, it is far from clear whether either principle follows from *iterativity*.

The literature here is extensive, and it is outside the scope of the current work to analyse the justification of **ZFC** in detail, unless it helps elucidate justificatory procedures.[35] In particular, limitation of size will reappear in Chapter VI, and will form an important part of our final conclusions. However the preceding discussion helps to highlight the following facts: (1.) the Iterative Conception has been seen as the dominant pre-theoretically appealing notion of set, which in turn (2.) is capable of blocking the paradoxes, and (3.) is amenable to at least heuristic explanations of why the main axioms of set theory should hold. For these reasons and the purposes of this thesis, we shall be concerned with philosophy of set theory as concerned with *iterative* set theory in this sense, with **ZFC** taken as the starting point.

Despite this controversy regarding the status of justification of axioms of **ZFC**, it is true that **ZFC** represents an elegant foundational theory (in the sense that it can

[34]This presentation is available in [Boolos, 1989]. For additional versions and historical remarks, see [Hallett, 1984].

[35][Boolos, 1971], [Boolos, 1989], [Maddy, 1988a] and [Potter, 2004] are good first references on the matter.

provide the necessary objects to formalise claims about mathematics). In this way, it provides an arena for facilitating the interrelation of mathematical structures and comparing proofs. The sheer scale of transfinite objects postulated is also philosophically fascinating, and an examination of the vast ontology posited by **ZFC** will help to elucidate how we, as philosophers, conceive of infinity. The Iterative Conception, though it may not provide full and clear explanation of why *every* axiom of **ZFC** should be true, nonetheless indicates a natural conception of a standard model of **ZFC**, and explains why the paradoxes are blocked in such structures. We thus will consider *primarily* **ZFC** set theory and its extensions under the Iterative Conception.

I.4 Extending models

Once we have a picture of sets as given by **ZFC** and underpinned by the Iterative Conception, we note that the truth value of any statement of set theory ultimately boils down to answering the following two questions:

(1) How far do the stages extend upwards?

(2) What sets are formed at successor stages?

The first question we may characterise as issues of *height*; we are interested in what ordinals exist to index the stages. The second we may refer to as questions concerning *width*; we care about what subsets of a given V_α are members of $V_{\alpha+1}$. The interesting issue for present purposes is that it seems that we can extend many models in both of the above two senses.

We present the case of height extension first. Given a particular set-sized transitive model of **ZFC**, we note that we can easily extend it in height. To do this, note that the ordinals of the relevant set model $\mathfrak{M} \models$ **ZFC** have a least upper bound $Ord(\mathfrak{M})$. Clearly, $Ord(\mathfrak{M}) \notin \mathfrak{M}$ by the Burali-Forti reasoning. We can then have $Ord(\mathfrak{M})$ appearing as an ordinal in an extended model $\mathfrak{N}$, and in which $\mathfrak{M}$ then appears as a set. Here, $\mathfrak{N}$ might not satisfy **ZFC**, to remedy this one can close under **ZFC** to form another set-sized model $\mathfrak{N}' \models$ **ZFC** such that $\mathfrak{M} \in \mathfrak{N}'$.[36]

[36]The details here are somewhat subtle, as it is consistent with **ZFC** that there is a single transitive

Important here is that (with some additional assumptions[37]) any classical set-sized structure $\mathfrak{M} \models$ **ZFC** can be lengthened[38] to form another set-sized structure $\mathfrak{N}' \models$ **ZFC** with $\mathfrak{M} \in \mathfrak{N}'$. In this way, a natural view of the Iterative Conception holds that any particular universe can be extended in height.[39]

One can also often extend the *width* of a particular hierarchy. The method of *forcing* shall be discussed in more detail in future chapters. We provide a brief characterisation of the method here. If $\mathfrak{M}$ is a model satisfying **ZFC** that is countable in V, and $\mathbb{P} = \langle P, \leq_{\mathbb{P}} \rangle$ is an atomless[40] partial order in[41] $\mathfrak{M}$, then we can define a new set G (a filter on $\mathbb{P}$ intersecting every subset of $\mathfrak{M}$ that is dense[42] in $\mathbb{P}$). Any such G cannot then be in $\mathfrak{M}$ (since the complement of G would be the dense set that is missed when $G \in \mathfrak{M}$).

A key fact concerning forcing is that it provides a way of adding subsets to certain models of **ZFC** in such a way that **ZFC** is preserved whilst altering the truth value of many set-theoretic statements. For example, if $\mathfrak{M}$ is a model satisfying **ZFC**$+CH$, then there is a simple forcing (Cohen forcing) that changes the value of the continuum to any value not contradicting König's Theorem.[43]. This holds not just for CH, but also a wide variety of sentences. Hamkins puts the point as follows:

> "A large part of set theory over the past half-century has been about constructing as many different models of set theory as possible, often to ex-

set-sized model of **ZFC**: the so called Shepherdson-Cohen minimal model (see [Shepherdson, 1951], [Shepherdson, 1952], [Shepherdson, 1953], and [Cohen, 1963], for details of the construction). Nonetheless the hypothesis that there are unboundedly many transitive models of **ZFC** is relatively weak; indeed it is consistent relative to the existence of a proper class of inaccessible cardinals (which in turn is consistent with $V = L$). The point is conceptual rather than technical; there are relatively weak assumptions that yield unboundedly many set models of **ZFC**, between which we may move with fluidity.

[37]See previous footnote.

[38]In doing so, we might add subsets to the relevant model as well, so such a lengthening need not be *only* in height. For example, Shepherdson-Cohen model (let it be denoted by '$\mathfrak{L}$') is of exactly this form; there is no transitive model $\mathfrak{M} \models$ **ZFC** such that it $\mathfrak{L} = (V_\alpha^{\mathfrak{M}}, \in\restriction V_\alpha^{\mathfrak{M}})$. To see this, note that in a model L_β of $V = L$, first-order ϕ is true iff for some n, ϕ is Σ_n and there exists a satisfaction predicate for Σ_n formulas which says that ϕ is true. These partial satisfaction predicates range over $L_{\beta+1}$ (i.e. are L_β-definable) and thus this yields a truth definition of L_β which is first-order definable over $L_{\beta+1}$ (and therefore belongs to $L_{\beta+2}$). Since $\mathfrak{L}$ is countable and satisfies $V = L$, this truth definition is coded as a real, and so any addition of height to another model of **ZFC** will necessarily add reals.

[39]This thought finds currency in both the historical and contemporary philosophical literature, see [Zermelo, 1930], and more recently [Hellman, 1989] and [Isaacson, 2011].

[40]A partial order is *atomless* iff any element of $\mathbb{P}$ has $\leq_{\mathbb{P}}$-incompatible extensions.

[41]We shall see later that for certain forcing constructions the case where $\mathbb{P}$ is a subset (but not a set of) $\mathfrak{M}$ is also pertinent.

[42]A subset X of $\mathfrak{M}$ is *dense* in $\mathbb{P}$ if for all $p \in \mathbb{P}$, there is a $q \in X$ such that $q \leq_{\mathbb{P}} p$.

[43]König's Theorem states that for $\kappa \geq 2$ and any infinite cardinal λ, $cf(\kappa^\lambda) > \lambda$. It implies that the continuum cannot have cofinality ω. See [Kunen, 2013] for details.

hibit precise features or to have specific relationships with other models. Would you like to live in a universe where CH holds, but $\Diamond$ fails? Or where $2^{\aleph_n} = \aleph_{n+2}$ for every natural number n? Would you like to have rigid Suslin trees? Would you like every Aronszajn tree to be special? Do you want a weakly compact cardinal κ for which $\Diamond_\kappa(REG)$ fails? Set theorists build models to order." ([Hamkins, 2012], p417)

Many of the above kinds of model are obtained by forcing. The method provides a fine-grained way of manipulating the subsets present in set-theoretic structures, allowing us to produce a huge variety of models. Further the models produced by forcing look *standard* as long as we start with an appropriate ground model; if G is $\mathbb{P}$-generic over $\mathfrak{M}$ then $\mathfrak{M}[G]$ is transitive and well-founded iff $\mathfrak{M}$ is, and $\mathfrak{M}[G]$ has the same ordinal rank as $\mathfrak{M}$.

Throughout this thesis, we will discuss width and height extensions separately. However, it is important to note that the two dimensions are intimately interrelated; often an extension in height will necessitate an extension in width. To see this, we first make the following definitions, important in set theory:

Definition 16. Let $\mathfrak{M} = (M, \in)$ be a set-sized model. A set x is *definable over* $\mathfrak{M}$ iff there exists a formula $\phi \in \mathscr{L}_\in$ and $a_1, ..., a_n \in M$ such that $x = \{y \in M | \mathfrak{M} \models \phi(y, a_1, ..., a_n)\}$.

Definition 17. $def(\mathfrak{M}) = \{x \subseteq M | x$ is definable over $\mathfrak{M}\}$

Definition 18. A set x is *constructible* iff it is a member of the following hierarchy:

(i) $L_0 = \emptyset$,

(ii) $L_{\alpha+1} = def(L_\alpha)$, for successor $\alpha + 1$,

(iii) $L_\lambda = \bigcup_{\beta<\lambda} L_\beta$ for limit λ,

(iv) $L = \bigcup_{\alpha \in On} L_\alpha$

L is thus the hierarchy formed by iterating definability through the ordinals (using, along the way, parameters from previous stages). [Gödel, 1940] showed that

L satisfies **ZFC** and is the *minimal* inner model[44] of **ZFC**. Now, by work of Shepherdson[45] and Cohen[46] we know that if there is a transitive set model of **ZFC**, then there is a *minimal* transitive model of **ZFC**; a countable transitive model of the form $L_\alpha \models$ **ZFC** (where α is the least β such that $L_\beta \models$ **ZFC**) and such that for any transitive model $\mathfrak{M} \models$ **ZFC**, $L_\alpha \subseteq \mathfrak{M}$. Any height extension of L_α to a taller transitive model of **ZFC** necessarily adds subsets to L_α. For, adding two more L-levels to yield $L_{\alpha+2}$ generates a new $x \subseteq \mathcal{P}(\omega)$.[47] Thus $V_{\omega+1}^{L_\alpha} \subset V_{\omega+1}^{L_{\alpha+2}}$. In this way, adding height can force a universe to also extend in width, even low down in the model. Thus, when considering model-theoretic extensions of universes, though we have $\mathfrak{M} \in \mathfrak{N}$ whenever $\mathfrak{N}$ extends $\mathfrak{M}$ in height, it is not necessarily the case that $\mathfrak{M}$ is an *initial segment* of $\mathfrak{N}$ (i.e. $\mathfrak{N}$ can disagree with $\mathfrak{M}$ on the identity of the V_α for $\alpha < Ord(\mathfrak{M})$).[48]

I.5 How many universes?

The upshot of the extensions in height and width show that **ZFC** alone dramatically fails to pin down an intended model. For many models of **ZFC**, even if they are transitive and well-founded, there are extensions of the models in both height and width also satisfying **ZFC**. A natural point of departure is thus to study the subject matter and objects we talk about when we do set theory. In other words, given that first-order **ZFC** does not come close to determining the height or width of its standard transitive models, we might be interested in the philosophical significance of this lack of determination.

One answer to this problem is given by the following view concerning set theory:

> **Definite Universism.** There is a maximal and unique universe of set-theoretic discourse in which every sentence has a definite truth value (i.e. is either true or false), denoted by 'V'.

Many scholars feel that Definite Universism facilitates the role of set theory as

[44] An *inner model of* **ZFC** is a structure satisfying **ZFC** and containing all ordinals. Of course a meta-mathematical challenge here is how to understand reference to a 'structure' if it cannot be a set. We shall see discussion of this issue later in the thesis.
[45] See [Shepherdson, 1951], [Shepherdson, 1952], and [Shepherdson, 1953].
[46] See [Cohen, 1963].
[47] For a sketch of the argument, see footnote 38.
[48] Many thanks to Sy Friedman for emphasising to me the importance of these subtleties.

a foundation for mathematics. Earlier, we noted that set theory had a deep foundational role in that it is possible to construe claims about mathematical objects in set-theoretic terms, and further that set theory provides standards for coherency in mathematics; the way to show a mathematical theory $\mathbf{T}$ to be non-vacuous is to find a model for $\mathbf{T}$ in the sets. Definite Universism is easily able to account for this role, it provides "one arena" to act as the "final court of appeal" for questions concerning mathematical proof and existence.[49]

Furthermore, as was noted earlier, recent scholars have proposed understanding the notion of *mathematical rigour* in set-theoretic terms.[50] If this is a project with which one wishes to engage, Definite Universism ensures that rigour is preserved when moving from one branch of mathematics to another; since there is just one maximal set-theoretic universe in which we are representing the claims of different mathematical discourses, the widest possible context cannot shift when considering different disciplines.[51]

In addition, such a view seems (at least prima facie) natural when presented with the Iterative Conception of Set. If we iterate the power set operation through all the ordinals, we generate a particular interpretation of set-theoretic discourse. On the (plausible) assumption that there are facts of the matter concerning whether or not particular ordinals exist and what subsets get formed at successor stages, *prima facie* this yields an interpretation that is maximal, unique, and settles every sentence of set theory.

It is precisely the definiteness in the notions of ordinal height and powerset where we might disagree with this, however. In the previous section, it was noted that there are often ways to extend a particular model in height and width. For height, we assume that the ordinals have a supremum and form a set (along with the relevant universe) in an extended model. For width, we have cogent methods (such as forcing) for defining new subsets.

We might, therefore, opt for one of the following views:

[49]See [Maddy, 1997], p26.

[50]A good example being [Burgess, 2015].

[51]Of course, one might think that Definite Universism is not *needed* for a characterisation of Burgessian rigour. Given a picture on which there is some indefiniteness one might still be able to come up with an account. However, it is a pleasing theoretical feature of Universism that it is clearly *sufficient* for providing such a characterisation.

> **Indefinite Universism.** There is just one maximal interpretation of set-theoretic discourse, but it is indefinite in some way, say because the powerset operation does not yield a definite characterisation[52] or the ordinal height of V is not definite[53].

or

> **Multiversism.** There is not one maximal universe of set-theoretic discourse, but rather many equally legitimate universes.[54]

Each of the above positions represents a cluster of views. Indefinite Universists, for example, can disagree on whether it is the ordinal numbers or the subsets formed at successor stages that are indefinite, or at what point it is that definiteness breaks down (say at the level of $\mathbb{N}$ or $\mathcal{P}(\omega)$ or $\mathcal{P}(\mathcal{P}(\omega))$). Multiversists, on the other hand, disagree as to exactly what the structure of the various universes should be like. We can hold that every universe should satisfy $\mathbf{ZFC_2}$ with the full semantics ([Zermelo, 1930], [Isaacson, 2011]), and hence the universes agree on width but not on height. We could also say that universes are all of the same height, but can be extended in width ([Steel, 2014][55]). Alternatively, we might hold that any universe can be extended in both width and height ([Arrigoni and Friedman, 2013]). Finally, we might be so radical as to say that the universes need only satisfy first-order $\mathbf{ZFC}$, and hence might even contain non-standard natural numbers ([Hamkins, 2012]).

The challenge then for the Definite Universist is to explain what it is that convinces them that relevant ontology and concepts are sufficiently sharp to yield such a definite characterisation. Unfortunately, however, we reach a standoff here. For, as we shall see, the arguments on each side are much too quick.

[52]See, for example, [Feferman et al., 2000].

[53]See, for example, [Linnebo, 2010].

[54]A subtle point here is that what one takes to be a 'legitimate' universe need not merely be a model of $\mathbf{ZFC}$. [Balaguer, 1998] and [Hamkins, 2012] have a view on which *any* such model will do, but we might take certain models to be ontologically privileged. For example, under one interpretation, [Arrigoni and Friedman, 2013] advocates a view on which well-founded models of $\mathbf{ZFC}$ satisfying certain maximality criteria are accorded a distinguished ontological status.

[55][Steel, 2014] is somewhat different from the other proposals in that in addition Steel advocates a theoretical *shift* to a multiverse language.

I.5.1 Arguments for Universism

Again, a full analysis of the arguments for and against the above views would take us too far afield. However, for the purpose of seeing the methodology of the thesis, some brief comments on each are in order. Arguments for Definite Universism that do not simply insist on determinacy in the concept of ordinal, powerset, or quantification over all sets (notions that are clearly question begging against Multiversists and Indefinite Universists), usually depend on *categoricity* arguments. Here, we use either an informal or formal argument to try and establish that all models of set theory satisfying certain constraints must be isomorphic, and hence there is a privileged maximal structure.[56]

There have been several such attempts. [McGee, 1997] provides a proof using urelemente to show that the pure sets of any two models must be isomorphic. However, the proof depends both upon unrestricted first-order quantification, and that the urelemente form a set. Both are likely to be controversial assumptions, and as such dialectically unconvincing. [Martin, 2001] provides an argument that depends on a pairwise comparison of models. However, such a comparison again depends on assuming that there is a unified arena in which we can compare the models, and also that such a comparison process will eventually result in a maximal model (rather than simply being a species of indefinite extendibility).

Koellner diagnoses the problems with such arguments as follows:

> "But this [i.e. categoricity arguments] doesn't get any traction with the advocate of the multiverse since it presupposes absolute conceptions of powerset and infinity and it presupposes that there is a single, univocal conception of set. The advocate of the multiverse will argue that the above argument is circular. "True if one presupposes that there is a univocal conception of set, one which has absolute notions of powerset and infinity, then one can run the categoricity argument. But that just presupposes in the meta-language what one set out to establish. One gets out

[56] A complication here is that such arguments only establish that there is one universe *up to isomorphism*. There might, for all we have said, be many universes all instantiating this maximal structure. Since we plan on dismissing this argument anyway, we set this issue aside.

> what one puts in."" ([Koellner, 2013], p11)

Such arguments, while possibly providing comfort to the Universist, thus do nothing to convince her opponents; we have to assume determinacy in the metalanguage to get determinacy with respect to the interpretation of the object language.

It is important here to be precise concerning the dialectic. The above is not meant to be definitive concerning the debate. Indeed, a full study of the arguments for and against Definite Universism would require a substantive literature in itself, and the debate is ongoing.[57] Nor is the possibility of a convincing argument on behalf of Universism thereby vitiated. For our purposes, we wish to note that extant direct arguments for the position are likely to be dialectically unconvincing when posed to Multiversists. Indeed this is precisely how the issue transpires in the literature. Hamkins, for example, writes:

> "The multiversist objects to Martin's presumption that we are able to compare the two set concepts in a coherent way. Which set concept are we using when undertaking the comparison? Martin's argument employs a background concept of 'property', which amounts to a common set-theoretic context in which we may simultaneously refer to both set concepts when performing the inductive comparison. Perhaps one would want to use either of the set concepts as the background context for the comparison, but it seems unwarranted to presume that either of the set concepts is able to refer to the other internally, and the ability to make external set (or property) concepts internal is the key to the success of the induction. If we make explicit the role of the background set-theoretic context, then the argument appears to reduce to the claim that within any fixed set-theoretic background concept, any set concept that has all the sets agrees with that background concept; and hence any two of them agree with each other. But such a claim seems far from categoricity, should one entertain the idea that there can be different incompatible set-theoretic backgrounds." ([Hamkins, 2012], p427)

[57] In addition to [Koellner, 2013], see [Meadows, 2013] for a discussion of the philosophical significance of categoricity results. Perhaps the most comprehensive recent argument for Definite Universism is available in [Maddy, 2011].

This brings Koellner's point into sharper focus. It is precisely the suppositions required to support the philosophical import of the categoricity theorems that renders them ineffective against the Definite Universist's opponents. The situation is characteristic of the current state of the debate; as things stand the Universist simply begs the question.

I.5.2 Arguments for indefiniteness

Let us turn then to the arguments for Multiversism and Indefinite Universism. As we shall see, such arguments face problems complementary to those suffered by the Definite Universist.

One option, analogous to the Definite Universist's insistence on determinacy in the concepts of ordinal and powerset, is simply to insist that such concepts are not so determinate. Feferman, for example, argues as follows:

> "My own view—as is widely known—is that the Continuum Hypothesis is what I have called an "inherently vague" statement, and that the continuum itself, or equivalently the power set of the natural numbers, is not a definite mathematical object. Rather, it's a conception we have of the totality of "arbitrary" subsets of the set of natural numbers, a conception that is clear enough for us to ascribe many evident properties to that supposed object (such as the impredicative comprehension axiom scheme) but which cannot be sharpened in any way to determine or fix that object itself." ([Feferman et al., 2000], p405)

Of course, such an argument is just as question begging as the Definite Universist's insistence that the ordinal number sequence and power set operation is definite. If the Universist cannot insist that the relevant ontology and concepts are sharp on the basis of (tutored) intuitions, than the Indefinite Universist cannot in good faith insist that they are *not* sharp on intuitive grounds.

One might instead point to the diversity of models arising from model-theoretic constructions, and then argue that holding one particular model privileged is *ad hoc*. Hamkins, for example, writes:

> "This abundance of set-theoretic possibilities poses a serious difficulty for the universe view, for if one holds that there is a single absolute background concept of set, then one must explain or explain away as imaginary all of the alternative universes that set theorists seem to have constructed. This seems a difficult task, for we have a robust experience in those worlds, and they appear fully set theoretic to us. The multiverse view, in contrast, explains this experience by embracing them as real, filling out the vision hinted at in our mathematical experience, that there is an abundance of set-theoretic worlds into which our mathematical tools have allowed us to glimpse." ([Hamkins, 2012], p418)

A vast amount of work in set theory since the discovery of the independence results has been focussed on the study of diverse models and the properties that hold between them in their own right. It seems then, that set theory investigates a vast array of different set concepts rather than a single such.

Again though this is simply to beg the question against the Universist. First, she has an easy interpretation of model theory; she can view it as concerned with substructures of V. Indeed, on the Universist's picture Hamkins' 'Multiverse Axioms' (a series of rules designed to provide intuitive content[58] to his Multiverse picture) are satisfied in the collection of all countable computably saturated models of **ZFC**.[59] Thus, the Universist can perfectly easily interpret Multiversist-inspired mathematics as concerned with such substructures. Second, she feels that she already has good reason to accept her view on the basis of previous considerations (such as the Iterative Conception of Set). Showing that there are, for the Universist, unintended interpretations of **ZFC** where 'V' does not denote V does nothing to shake her from her position. Definite Universists and their opponents both believe they have good reasons to assert their view; they simply disagree on the intuitive force of the relevant considerations and interpretation of the mathematical data.

Again, none of the arguments concerning the above issues are meant to be defini-

[58] There is an issue in that, for technical reasons to do with Hamkins' view, it is hard to see how there could be a formal system axiomatising his claims. This is discussed in more detail in [Hamkins, 2012] (p436) and [Barton, Fa].

[59] See [Gitman and Hamkins, 2010] for details of computably saturated models, references, and the relevant construction.

tive, and the above discussion represents a very coarse grained and simplistic characterisation of what is an ongoing debate of philosophical interest. For the moment, however, it suffices to note that the debate has hit something of a stalemate; neither party finds the arguments dialectically persuasive, and we find ourselves in an apparently intractable philosophical situation.

Partly for this reason[60] we propose a different tack here. Instead of trying to adjudicate the dialectic between the Definite Universist and her opponents, we shall rather concentrate on filling out some of the details of a particular strategy Definite Universists might follow. We shall, therefore, take Definite Universism (which, now the dialectic with the Indefinite Universist has been discussed, we refer to with the term 'Universism') as an assumption of the thesis, and see what is highlighted by working out the view in more detail. As such, the thesis forms part of a wider foundational programme to study elucidations of each view on its own terms, rather than trying to be dialectically convincing against opponents. It is hoped that in this manner we will be able to see what foundational fruits are borne by each philosophical position. This is not to say that our arguments cannot be used in informing the debate between the Definite Universist and her opponents—in fact we shall mention some possibilities in Chapter VIII ('Conclusions'). However, such a project will not lie in the focus of our attention here.

I.6 Structure of the Thesis

Given the independence phenomenon, there is a substantial challenge for the Universist: explain her faith in the existence of a single, maximal, interpretation in the face of the vast zoo of different epistemological possibilities. There are three possible options here:

(1) Accept that there are absolutely undecidable sentences (in the sense that there could never be a well-justified axiom resolving independence), and explain why this does not threaten the claim that there is a maximal, unique, definite universe of set theory.

[60] The other main reason being the intrinsic interest of such a change in philosophical focus.

(2) Justify new axioms that resolve[61] previously independent questions, providing confidence that every set-theoretic sentence has a definite truth value.[62]

(3) Neither accept absolutely undecidable propositions nor that we need to engage in a project of justification for new axioms.[63]

This thesis focusses on an exposition of issues surrounding (2), in particular the kinds of principles and techniques that can be examined in formulating and justifying new axioms. It will be argued that the Universist can provide a philosophically satisfying way of interpreting more than previously thought.

The thesis is structured as follows. After this initial characterisation of the debate, Chapter II ('Gödel's Programme') explains the Universist project for formulating and justifying new axioms concerning V. We argue that extensions of V are relevant to both aspects of Gödel's Programme for resolving independence. We also identify a 'Hilbertian Challenge' to explain how we should interpret extensions of V, given that we wish to use discourse that makes apparent reference to such non-existent objects.

Chapter III ('Problematic Principles') then lends some mathematical precision to the coarse-grained outline of Chapter II, examining mathematical discourse that seems to require talk of extensions of V. Whilst we acknowledge that a Universist is not *forced* to interpret discourse concerning extensions of V, we point out that if she *can* use such talk coherently, new mathematical avenues are open to her.

Chapter IV ('Climbing above V?'), examines some possible interpretations of height extensions of V. We argue that several such accounts are philosophically problematic. However, we point out that these difficulties highlight two constraints on resolution of the Hilbertian Challenge: (i) a Foundational Constraint that we do not appeal to entities not representable using sets from V, and (ii) an Ontological

[61]By Gödel's Second Incompleteness Theorem, no consistent formal theory able to represent arithmetic will ever *prove* every true sentence.

[62]Of course, advocating philosophical research into (2) does not preclude an analysis of the status of set theory in the absence of a satisfactory resolution of independence. One may argue that one's favourite axiom is well-justified, whilst maintaining that even if it is not so justified, Universism remains true for other reasons.

[63]This might seem like a rather odd suggestion—how could an independent sentence ever be *resolved* without justification? However, recently [Williamson, F] has proposed a view on which there are *no* absolutely undecideable sentences in virtue of the possibility of the existence of beings who find such statements primitively compelling. Since our interest is in (2), we put aside this issue.

Constraint to interpret extensions of V in such a way that they are clearly different from ordinary sets.

Chapter V ('Broadening V's Horizons?'), considers interpretations of width extensions. Again, we argue that many of the extant methods for interpreting this kind of extension face difficulties. Again, however, we point out that a constraint is highlighted; a Methodological Constraint to interpret extensions of V in a manner that makes sense of our naive thinking concerning extensions, and links this thought to truth in V. We also note that there is an apparent tension between the three constraints.

Chapter VI ('A Theory of Classes') changes tack, and provides a positive characterisation of apparently problematic 'proper classes' through the use of plural quantification. It is argued that such a characterisation of proper class discourse performs well with respect to the three constraints, and motivates the use of a relatively strong class theory (namely **MK**).

Chapter VII ('V-logic and Resolution') then puts **MK** to work in interpreting extensions of V. We first expand our logical resources to a system called V-logic, and show how discourse concerning extensions can be thereby represented. We then show how to code the required amount of V-logic using **MK**. Finally, we argue that such an interpretation performs well with respect to the three constraints.

Chapter VIII ('Conclusions') reviews the thesis and makes some points regarding the exact dialectical situation. We argue that there are many different philosophical lessons that one might take from the thesis, and are clear that we do not commit ourselves to any one such conclusion. We finally provide some open questions and indicate directions for future research, remarking that the thesis opens the way for new and exciting philosophical and mathematical discussion.

Chapter II

Gödel's Programme

We saw in the last chapter that there are difficult questions raised for Universism stemming from discourse involving extensions of V. For current purposes, one particularly salient challenge is for the Universist to provide a defence of her position that there is one Universe of sets in the face of a large diversity of natural set-theoretic possibilities. It was then noted that one response to the problem (among many) is to provide a means to settle sentences independent of **ZFC** in a well-justified manner.

II.1 Gödel's Programme explained

In his seminal expository article on the Continuum Hypothesis, Gödel said the following concerning the lack of resolution of CH:

> "For if the meanings of the primitive terms of set theory...are accepted as sound, it follows that the set-theoretical concepts and theorems describe some well-determined reality, in which Cantor's conjecture must be either true or false. Hence its undecidability from the axioms being assumed today can only mean that these axioms do not contain a complete description of that reality. Such a belief is by no means chimerical, since it is possible to point out ways in which the decision of a question, which is undecidable from the usual axioms, might nevertheless be obtained." ([Gödel, 1964], p260)

Here, Gödel claims that if one holds Universism, and if our concepts and theorems relate appropriately to objects in V, then CH is either true or false, despite its independence from the axioms of **ZFC**. We might then attempt to propose and justify new axioms to settle CH. In particular, Gödel believed that there were certain concrete methodologies we could use in providing such justifications. I will refer to this process of proposing and justifying new axioms for set theory as *Gödel's Programme*.

Central to the above quotation is the possibility of *deciding* questions through providing a description of set-theoretic reality that allows us to *prove* previously independent sentences. Exactly how such a description might be given is a tricky problem and one of the central philosophical challenges in articulating Gödel's Programme. More precision is thus required in order to provide a philosophically cogent position. We begin by remarking that Gödel's Programme is executed in a two step process:

(1) Propose a new set-theoretic axiom.

(2) Justify said axiom.

As we shall see in Chapter III, both are relevant for the current discussion. In executing Gödel's Programme, the philosophical interpretation of theorems and content of axioms varies according to what resources we allow for talking about V. To give a hint of what is to come, recall from the previous chapter that there were various methods that could be employed for extending models of **ZFC**, either by lengthening the ordinals or by adding more subsets. Now, assuming Universism (and that **ZFC** is true in V) we may ask whether or not it is possible to talk about V using extensions of V (either to prove new results about sets in V or formulate new axioms).

The answer to the question for the Universist is, in a certain sense, going to be obviously "**NO!**". Simply put, for the Universist, V is meant to be all the sets there are, and thus extensions of V do not exist.

However, the contention of this thesis is that there is another sense in which the Universist *can* make use of some resources of this kind. Of course, the reference here will not be to actual extensions of V, naively understood. Rather, how we understand the reference of the term 'extension of V' will receive an interpretation amenable to the Universist's philosophical position.

Immediately, the reader might recoil from such a suggestion. Talk of forcing extensions of V and lengthening the ordinals is normally regarded as an unforgivable sin for the Universist. However, as we shall see in Chapter III, there is a large class of statements with which the Universist might be interested, and which in turn she might wish to be interpret.

Moreover, a simple example indicates that such a position is not as heretical as it might seem. It is a fact (a quick corollary of the Compactness Theorem for first-order theories) that if there is a model of **ZFC**, then there is a non-well-founded model of **ZFC**. Now it is entirely possible that by studying such non-well-founded structures we might come to learn about the sets in V under $\in$. Indeed, in a rather trivial sense, truths about a non-well-founded model $\mathfrak{M}$ just *are* truths about sets in V under $\in$ for the Universist, it is simply that $\in$ is not the same relation as $E_{\mathfrak{M}}$ (i.e. $\mathfrak{M}$'s interpretation of the membership relation). However, a Universist does not accept the existence of *actually* non-well-founded sets in V! She may even reject the existence of non-well-founded sets altogether (not just within V)! Can she then make use of non-well-founded models of **ZFC** (and for that matter, other theories[1] that *directly* postulate the existence of non-well-founded sets)? Of course she can, but in doing so she should be mindful of what such models and theories are *about*, and in particular note that the models are unintended.

So it is with *extensions* of V. What the Universist must attend to is that the discourse involving such constructions *cannot*, on her view, *actually* refer to sets outside V. Rather she has to provide an appropriate codification of the relevant statements that allow her to derive the desirable consequences of the discourse, but do not commit her to 'extra-V' sets.

Before we embark on the journey of unearthing the uses of such extra-V entities, let us first be more precise about what is required for each of the two steps. As we shall see, the details are somewhat more complex than one might anticipate, and it is to these issues that I now turn.

[1] Good examples here are **AFA** or **NFU** (for both of which we can find models inside **ZFC** models, see Ch. 7 of [Devlin, 1979] for **AFA** and [Jensen, 1968] for **NFU**).

II.2 Formulating the axioms

If the Universist is to propose a particular axiom to try and capture some aspect of V, then prima facie she can only use resources which are coherent. As we shall see in Chapter III, the use of 'ideal' objects outside V in formulating axioms leads to triviality or apparent falsehood, even when we are trying to make a claim about objects within V. For the moment, we provide a more general statement of the problem and leave examination of concrete examples until then. However, a toy example is helpful to understand how consideration of extensions might be useful for making a claim about V. Let Φ and Ψ be conditions on universes. A Universist might try to state something about V using the following principle:

> (Principle-Ψ^V) If there is an extension of V such that Φ, then Ψ holds of V.

The problem with Principle-Ψ^V is that it will always come out as true, but fail to capture the intended aspect of V (namely that Ψ holds of V). For the antecedent (on its natural reading) is trivially false, and so the conditional is true. But this provides us with no reason to think that Ψ is actually true of V which, presumably, was the intended consequence of asserting the putative axiom in the first place. We shall see a more fine-grained analysis of this phenomenon in Chapter III, however for now we note an interesting comparison. Surprisingly, given that Universism is often regarded as paradigmatically *realist*, it is instructive to consider Universism as a species of *nominalism*. We can be nominalists about a various kinds of entity: medium-sized dry goods, properties, or even mathematical objects. Our Universist is a nominalist about sets outside V. For each variety of nominalism, we have a body of discourse that we want to put to work in talking about the objects we *do* countenance. Viewing Universism in this nominalistic spirit allows us to draw parallels with other kinds of nominalism and see if there are any insights to be gained. The following example from Stephen Yablo clarifies the situation:

> "Imagine we have a strange, kabbalistic reading of Genesis. Go forth and multiply, God commanded. The "multiply" means, we believe, that

the animals should proliferate at a constant rate; each year's population was to be n times larger than the year before's. The value of n revealed itself when "forth" turned out to be a mistranscription of "fourth." The command was issued on day 5, and we believe on other grounds that the number of animals at that time was three. According to us, then,

(NA) The number of animals on the nth day $= 3 \times 4^{(n-5)}$

Unfortunately for this way of putting it, our reading of Genesis *also* tells us that God never got around to creating numbers. So we can't in consistency regard our hypothesis as true." ([Yablo, 2014], p80)

The predicament of the Kabbalistic Nominalist in Yablo's colourful example is roughly analogous to our Universist's. The Kabbalistic Nominalist wants to make a statement concerning *the animals*, but appears to appeal to the false assumption (NA) in doing so. Our Universist might want to make a substantive claim about *the sets* using Principle-Ψ^V, but fails to do say anything of significance in virtue of her nominalism concerning sets outside V.

We shall see in Chapter III that there are axioms that make mention of such entities, and that the difficulties in making claims about the sets are often more complex than the problem faced by the Kabbalistic Nominalist. For now, we merely note that an interpretation of objects outside V would quite possibly allow us to make new claims about the structure of V that we would otherwise not be able to appropriately state. A Universist is not likely to be interested in any of the 'ideal' entities 'outside' V (such entities do not exist). However, a principle that has consequences *within* V might be of interest, but only if it can be interpreted so that its truth or falsity tells us more about the structure of V.

II.3 Justifying the axioms

There is a second way we might motivate the consideration of resources beyond V. In order to understand *why* such resources might be useful, we need to examine the notion of *justification* in set theory.

What is justification? The question is philosophically problematic in a general context, and as such we will leave the notion informal. However, it is possible to provide some clarifying remarks, and in particular attend to a distinction that will help to make the details of Gödel's Programme more precise. Justification is, at least, an epistemic notion that links our understanding and beliefs with truth in a sufficiently accurate manner[2]. As is widely accepted in contemporary philosophy, [Gettier, 1963] showed that justification need not entail knowledge (even if, as in a Gettier case, a particular justified statement is true). However, we nonetheless expect justification to provide good reason to hold a proposition true (though said proposition may turn out to be false).

Since the issues that surround justification are so thorny (even when restricted to the case of mathematics), we shall analyse the more tractable question of what *mathematical methods* a Universist may employ in justifying a particular statement (as opposed to how statements receive justification or whether or not particular statements are justified). Despite this, and in order to understand what is at stake for the Universist, a well-rehearsed distinction is important. The relevant contrast is between different kinds of justification, namely *intrinsic* and *extrinsic*. The distinction plausibly goes back as far as [Russell, 1907], but is most famously stated in [Gödel, 1947] and [Gödel, 1964]. Characterising intrinsic justification, Gödel writes:

> "First of all the axioms of set theory by no means form a system closed in itself, but, quite on the contrary, the very concept of set on which they are based suggests their extension by new axioms which assert the existence of still further iterations of the operation "set of"...These axioms show clearly, not only that the axiomatic system of set theory as used today is incomplete, but also that it can be supplemented without arbitrariness by new axioms which only unfold the content of the concept of set explained above." ([Gödel, 1964], pp260-261)[3]

[2]Of course, what is meant by 'sufficiently accurate manner' is going to depend on one's theory of justification.

[3]Here, I use [Gödel, 1964] rather than [Gödel, 1947] (p181) for a couple of reasons. First [Gödel, 1964] represents Gödel's more mature philosophical views (and indeed he was more satisfied with his command of English during this period; see [Moore, 1990]). Second, his wording leaves the kinds of justification he has in mind more open; in [Gödel, 1947] he is clearly more concerned with weak reflection principles, referring to "axioms which are only the natural continuation of the series of those set up so

and

> "also there may exist, besides the ordinary axioms,...other (hitherto unknown) axioms of set theory which a more profound understanding of the concepts underlying logic and mathematics would enable us to recognize as implied by those concepts" ([Gödel, 1964], p261)[4]

These quotations require some unpacking before we have a full characterisation. Key to the above passages is the thought that intrinsic justifications are concerned with unfolding and explaining particular concepts. We determine, *via* conceptual analysis, what principles are implied by the mathematical conception with which we are working. Such a characterisation, as it stands, is somewhat unclear; what constitutes a satisfactory 'unfolding' of a concept is itself in need of explaining.[5] Koellner makes the issues a little more precise:

> "One can also gain a sharper understanding of the notion of intrinsic justification by pointing to some of its properties. First, an intrinsically justified statement need not be self-evident, in part because the justification may be quite involved..., in part because it is possible that the underlying conception is problematic (as, for example, was the case with the Fregean conception of extension). On the other hand, the notion of intrinsic justification is intended to be more secure than mere "intrinsic plausibility"... in that whereas the latter merely adds credence, the former is intended to be definitive (modulo the tenability of the conception)." ([Koellner, 2009], p207)

Koellner's picture of intrinsic justification is thus one on which we proceed via (possibly quite involved) conceptual analysis to provide definitive justification of certain principles. A major problem here is that it is not clear that such an 'unfolding of a concept' will be widely agreed upon. Indeed, the large numbers of constructivists and believers in indefiniteness are good evidence that there is unlikely to be

far", rather than those that "unfold the concept of set".

[4]See p182 of [Gödel, 1947] for the relevant passage, essentially similar in content.

[5]For a heroic attempt, see [Hauser, 2006].

any such accord. However, we can view the project of intrinsic justification as providing a conceptual analysis that renders particular statements *more evident*. It is useful to sharpen this understanding somewhat by examining an example of a particular intrinsic justification. Turning again to Koellner, he remarks:

> "Over the base theory RCA_0 (Recursive Comprehension Axiom), the Hilbert Basis Theorem[6] is equivalent to the statement that ω^ω is well-ordered...
>
> ...the Hilbert Basis Theorem is far from immediate. It is the sort of thing that we set out to prove from things that are more evident. In contrast, the statement that ω^ω is well-founded is something that becomes clear by reflecting on the concepts involved. It is not the sort of thing that we set out to prove from something more evident. This is what lies behind the fact that the Hilbert Basis Theorem is called a theorem and not an axiom." [Koellner, 2011]

We can thus analyse the project of intrinsic justification as settling particular questions through proving them on the basis of more evident statements. An intrinsic justification, on this conception, renders a principle more evident than we previously thought. In the context of the above example, we note that ω^ω is an ordinal, so the fact that it is well-orderable is more evident on the basis of the *kind of thing it is* compared to a non-trivial theorem concerning Noetherian rings.

It is unclear whether or not Gödel himself took intrinsic justifications to be "definitive" (certainly many of the terms he used, such as "intrinsic necessity", indicate that this may well have been the case). There is also the separate question of whether or not *we* should take intrinsic justification to be definitive or simply a matter of degree or adding credence. I do not wish to become entangled in these tricky issues here. Since our interest is in what resources a Universist can use in formulating and justifying new axioms, it suffices to note that intrinsic justification of an axiom would consist in examining the axiom relative to a particular mathematical concep-

[6]The Hilbert basis theorem states that any polynomial ring over a Noetherian ring (i.e. a ring such that for any chain of ideals $I_0 \subseteq I_1 \subseteq ... \subseteq I_k \subseteq I_{k+1}, ..$ there is a n and ideal I_n in the chain such that $I_n = I_{n+1} = I_{n+2}...$) is also Noetherian.

tion, without considering any particular proof-theoretic consequences of the axiom.

The notion of intrinsic justification contrasts with that of *extrinsic* justification, and it is with respect to this latter kind that we find the most relevance for extensions of V. Rather than providing reasons to think that a principle results from conceptual analysis of a particular mathematical conception, extrinsic justification concerns its theoretical consequences. Explicating the notion, Gödel writes:

> "Secondly, however, even disregarding the intrinsic necessity of some new axiom, and even in case it had no intrinsic necessity at all, a probable decision about its truth is possible also in another way, namely inductively by studying its "success". Success here means, fruitfulness in consequences, in particular in "verifiable" consequences, i.e., consequences demonstrable without the new axiom, whose proofs with the help of the new axiom however, are considerably simpler and easier to discover, and make it possible to contract into one proof many different proofs." ([Gödel, 1964] p261)[7]

and

> "A much higher degree of verification than that, however, is conceivable. There might exist axioms so abundant in their verifiable consequences, shedding so much light upon a whole field, and yielding such powerful methods for solving given problems (and even solving them constructively, as far as that is possible) that no matter whether or not they are intrinsically necessary, they would have to be accepted at least in the same sense as any well-established physical theory." ([Gödel, 1964], p261)[8]

Again, the philosophical and exegetical issues here are both difficult and subtle. For our purposes, the key fact to note is that extrinsic justification is concerned with the consequences of a putative axiom rather than whether or not the principle is appropriately related to an underlying mathematical conception.

[7]Here p182 of [Gödel, 1947] does not differ substantially from [Gödel, 1964].

[8]Again here, [Gödel, 1947] (pp182-183) is not significantly different from the [Gödel, 1964] revision.

There is a detailed literature[9] on the topic, and a full examination would require a separate thesis. For the purposes at hand, we wish to examine one particular aspect of Gödel's distinction. Essential to the notion of *extrinsic* justification is the notion of a *verifiable* consequence. Now exactly in what sense the consequences of a hypothesis are *verifiable* is, it would seem, somewhat unclear (as is the epistemic link between the consequences of a principle and its truth).[10,11] For our purposes, however, it suffices to note that in order to find whether or not a principle ϕ has *consequences* that are *verifiable*, we must first settle on (i) the kinds of methods we accept for deriving consequences from ϕ, and (ii) a bank of truths concerning V that we take to have been *verified*.

A (very) simple example here is useful to see the sense in which we might have a principle supported by verifiable consequences (as in (ii)). We begin by noting the following celebrated theorem:

> **Theorem 19.** [Gödel, 1931] *Gödel's First Incompleteness Theorem.* Let $\mathbf{T}$ be any ω-consistent[12] theory capable of representing Primitive Recursive Arithmetic (henceforth '$\mathbf{PRA}$'). Then there is a Π_1-sentence $G_{\mathbf{T}}$ such that $\mathbf{T} \nvdash G_{\mathbf{T}}$ and $\mathbf{T} \nvdash \neg G_{\mathbf{T}}$.

In the particular case of *Peano Arithmetic* (henceforth '$\mathbf{PA}$'), the immediate corollary (as is well known) is that there is a Gödel sentence $G_{\mathbf{PA}}$ in the language of $\mathbf{PA}$ that is formally undecidable from $\mathbf{PA}$ alone. There is (of course) a close relationship between the First Incompleteness Theorem and the Second Incompleteness Theo-

[9] See, for example, [Maddy, 1988a], [Maddy, 1988b], [Maddy, 1997], [Maddy, 2011], [Koellner, 2009], [Koellner, 2010], among many others.

[10] Gödel is actually relatively clear on this point, stating that they must be "demonstrable without the new axiom". However, we might take there to be other kinds of "verifiable consequence", say if the axiom proves some intrinsically plausible statement that has not yet been *demonstrated*.

[11] Two additional complications here are that there are clearly false statements (e.g. $0 = 1$) that have a large number of verifiable consequences, and that there might be contradictory axioms with the same (or a similar amount of) verifiable consequences. The former problem is dealt with by noting that such principles normally also have *falsifiable* consequences. The latter is more subtle, but is not so problematic if we do not hold that justification need be an all or nothing matter. We might have two axioms or theories, both of which seem equally well extrinsically justified, but cannot both be true. We simply say that we cannot yet form a judgement on which is correct. A good example here comes from the philosophy of physics; both Relativity Theory and Quantum Mechanics are well extrinsically justified (in that they are both exceptionally good at accounting for verifiable phenomena), but nonetheless are in apparent tension. A natural methodology is to try then to *resolve* this tension, rather than declaring this to be a problem with the project of justifying physical theories in terms of accounting for verifiable data.

[12] A theory $\mathbf{T}$ is *ω-inconsistent* iff $\mathbf{T} \vdash \exists x \neg \phi(x)$, but also $\mathbf{T} \vdash \phi(n)$ for all $n \in \mathbb{N}$. $\mathbf{T}$ is *ω-consistent* iff it is not ω-inconsistent.

rem. The details will be familiar to specialists, but for the sake of clarity we provide a brief exposition. It is possible, *via* an adequate coding of one's choosing, to represent syntax in a countable language by natural numbers. For a formula ϕ, we denote the natural number representing ϕ (or 'Gödel code of ϕ') by '$\ulcorner\phi\urcorner$'. We are similarly able to code proofs as sequences of formulae, where each position in the sequence is either an axiom or follows from the previous steps in the proof *via* a legitimate rule of inference. We let $Prf_{\mathbf{T}}(x, \ulcorner\phi\urcorner)$ be a predicate that holds between the natural numbers x and $\ulcorner\phi\urcorner$ just in case x codes a proof of ϕ in $\mathbf{T}$. We then define the *consistency statement for* $\mathbf{T}$ as follows:

Definition 20. $Con_{\mathbf{T}} =_{df} \neg\exists x Prf_{\mathbf{T}}(x, \ulcorner 0 = 1\urcorner)$

The sentence states that it is not the case that there is a number coding the proof of a contradiction in $\mathbf{T}$. As $\mathbf{PA}$ contains $\mathbf{PRA}$, it has a consistency sentence $Con_{\mathbf{PA}}$. Assuming $\mathbf{PA}$ is consistent, can we prove $Con_{\mathbf{PA}}$ in $\mathbf{PA}$? As is well known, the answer is negative:

Theorem 21. [Gödel, 1931] *Gödel's Second Incompleteness Theorem.* Let $\mathbf{T}$ be an ω-consistent theory capable of representing $\mathbf{PRA}$. Then $\mathbf{T} \nvdash Con_{\mathbf{T}}$.

We have, as an immediate corollary, that $\mathbf{PA} \nvdash Con_{\mathbf{PA}}$. We might then ask, "Is $Con_{\mathbf{PA}}$ justified?".

It is clear that, assuming that $\mathbf{PA}$ has good intrinsic justification from an underlying mathematical conception, then so does $Con_{\mathbf{PA}}$. Presumably, any coherent underlying conception that provides the resources to justify the claim that the axioms of $\mathbf{PA}$ are *true*, thereby justifies the claim that they are *consistent*. Our interest here, however, is in *extrinsic* justification, and so we shall give a putative justification of $Con_{\mathbf{PA}}$ on the basis of its *consequences*. The argument is certainly not meant to be watertight, but serves as a toy example to see the sense in which we might use verifiable consequences in justifying a principle. Less well known than the celebrated First and Second Incompleteness Theorems is a theorem schema for *finite* proofs. There are, in addition to full consistency statements for theories, partial statements of finite consistency:

Definition 22. $Con^n_{\mathbf{PA}} =_{df} \neg\exists x[Prf_{\mathbf{PA}}(x, \ulcorner 0 = 1 \urcorner) \wedge$ "x has fewer than n-many symbols"]

This sentence states that there is no proof of a contradiction in $\mathbf{PA}$ *in under n-many symbols*. By the standard Gödel reasoning[13], for each $n \in \mathbb{N}$, there is no proof in under n-many symbols of $Con^n_{\mathbf{PA}}$. However, there *is* a proof of $Con^n_{\mathbf{PA}}$ in $\mathbf{PA}$; one simply enumerates all the proofs containing n-many symbols and checks them for consistency. We can then view each instance of particular $Con^n_{\mathbf{PA}}$ for $n \in \omega$ as a particular verifiable consequence to be accounted for.

Now, though each $Con^n_{\mathbf{PA}}$ is provable in $\mathbf{PA}$, the proofs are long, cumbersome, and somewhat *ad hoc* (we have to enumerate the relevant formulae each time). However, the following *is* a theorem:

Theorem 23. Take any $n \in \mathbb{N}$. Then $\mathbf{PA} + Con_{\mathbf{PA}} \vdash Con^n_{\mathbf{PA}}$ in a 'small' number of steps.

Proof. By $Con_{\mathbf{PA}}$ there is no contradiction provable in $\mathbf{PA}$, and hence no contradiction provable in fewer than n-many symbols. ■

The above theorem is not fully precise; the notion of 'smallness' is left informal. The point is simply that $\mathbf{PA} + Con_{\mathbf{PA}}$ condenses the form of proof of the relevant verifiable consequences (i.e. each $Con^n_{\mathbf{PA}}$) into one simple and easy proof, thereby conferring extrinsic justification on $Con_{\mathbf{PA}}$.[14]

To see how (ii) would fit into the above example, note that we had to prove (in the previous theorem), that $Con_{\mathbf{PA}}$ has the consequence of providing an easy proof of each $Con^n_{\mathbf{PA}}$. Any such proof depends on a background of resources. In this case, they are benign; we simply use first-order logic over $\mathbb{N}$. However, I could plausibly have taken a (wholly unnecessary) detour through the transfinite in finding consequences of $Con_{\mathbf{PA}}$. The question then might have been raised as to whether or not the use of such resources was justified in determining the proof of every $Con^n_{\mathbf{PA}}$ from $Con_{\mathbf{PA}}$.

Where do resources beyond V come into the picture? Well, it is entirely possible (and indeed, as we shall see in Chapter III *true*) that we might use resources outside

[13] See [Buss, 1994] for book-keeping and details.
[14] For a fuller exposition of these speed-up proofs, see [Buss, 1994].

V to derive consequences about V. In the above example, we had to know (i) that each of the $Con^n_{\mathbf{PA}}$ were true before we could note that $Con_{\mathbf{PA}}$ helps to settle these consequences correctly in a unified and drastically simpler manner, and (ii) needed a background theory to determine the consequences of $Con_{\mathbf{PA}}$. If, for example, we did not accept the use of $\mathbf{PRA}$, we would have problems formalising the relevant metalogical notions required to talk about $Con(\mathbf{PA})$. The case for axioms and V is similar; we want to know what is true in V before we attempt to provide extrinsic arguments concerning verifiable consequences, and how we can assess what consequences a principle might have. It is entirely plausible that extensions of V might have some part to play in determining these facts, and hence have an impact upon the project of extrinsic justification for the Universist.

II.4 The Hilbertian Challenge

Let us take stock. I have argued thus far in this chapter that it is at least open that reasoning using extensions of V might be informative for Gödel's Programme, both from the perspective of formulating new set-theoretic principles and (extrinsically) justifying these proposed axioms. Looking forward to future chapters, we shall see that set-theoretic resources that appear to refer to sets beyond V can be useful both in formulating putative axioms for V and deriving consequences about V. From a Universist perspective, these techniques must be understood as merely useful for talking about V; to admit the existence of such sets is to concede that her position is false. What the Universist requires, if she wishes to use such resources, is an explanation of why they are *reliable*, in that they will not produce false conclusions despite their incoherence on a naive reading of their content.

A historical parallel is emerging. On the Universist picture, we have the potential for executing Gödel's Programme whilst utilising talk about extensions of V. Such extensions would have to be 'ideal'; we do not literally talk about such objects, but we would like to have an assurance that such discourse will not lead us astray concerning truth within V. The similarities between this position and Hilbert's Formalism on the basis of his Finitism are striking. Under one interpretation of Hilbert, there was a period of time (usually understood as after [Hilbert, 1922]) when he held that all legitimate mathematical objects in ontological good standing were *finite*. However, Hilbert nonetheless wanted to use *transfinite* resources in proving theorems about the *finite* objects, much as the Universist wishes to use *extra-V* resources in proving theorems (or formulating axioms) about *intra-V* objects. Hilbert thus tried to prove (via finitary means) that such resources would not lead to a contradiction.[15] Of course, by Gödel's Second Incompleteness Theorem, such a project is doomed from the outset.[16] This does not preclude, however, a Hilbertian from providing *philosophical* reasons to accept that such resources will not lead us astray. This suggests the following challenge to be answered by the Universist:

> **Hilbertian Challenge.**[17] Provide philosophical reasons to legitimise the use of extra-V resources for formulating axioms and analysing intra-V consequences.

Prima facie, it is hard to see how a Universist could provide such reasons. However, we can begin to form an answer by (oddly enough) viewing Universism as a species of *nominalism*, specifically nominalism about the existence of extra-V sets. Just as Hilbert was a nominalist about transfinite entities, so the Universist is a nominalist about extra-V entities. Recall our Kabbalistic Nominalist from earlier. Yablo points out the following:

> "How much should this [the false assumption about numbers] bother us? ... it's enough for us if *The number of animals on the n^{th} day is* $3 \times 4^{(n-5)}$ is true about the animals, or more generally the physical world." ([Yablo, 2014], p81)

Suppose then that the kabbalistic reading is true; animals proliferate at a constant rate, and numbers do not exist. The key point for Yablo is that despite the *appar-*

[15]This receives resounding expression in the catechism of the Hilbertian position that "No one shall be able to drive us from the paradise Cantor created for us." ([Hilbert, 1925], p376).

[16]Though salient here is a heroic defence of Hilbert's aims in the work of Michael Detlefsen (see [Detlefsen, 1986], [Detlefsen, 1990], and [Detlefsen, 2001]).

[17]There is a substantial and interesting question as to *how* Hilbertian this challenge really is. We certainly do not claim that it accords with all of Hilbert's writing. Indeed, one might take Hilbert as requiring *conservativity* rather than merely a lack of contradictions. All we wish to identify here is that there are certain parallels with Hilbertian Finitism and Universism. Thanks to Giorgio Venturi for emphasising this point.

ent reference to (non-existent) mathematical objects, we can construe our statements about *the animals*. Now the Universist is in a similar position; she countenances the existence of *sets* but not extra-V sets. One way of responding to the Hilbertian Challenge then is to provide a detailed analysis of how we can interpret discourse involving extensions of V in such a way that we talk about objects countenanced by the Universist. By doing so, we would be able to determine what objects and structures we actually study when using talk involving extensions of V, thus allowing us to determine the exact content of axiom formulations and relevant mathematical consequences.

An answer to this question would provide a new dimension to Gödel's Programme, opening up the consideration of new axioms and techniques for establishing consequences in V. In much the same way, had (*per impossibile*) Hilbert's Programme been successful, it would have allowed the Finitist to consider resources mentioning infinite structures for determining truth about finite objects.

We recall one final example to illustrate how such a solution might operate. We noted earlier that although the Universist will not assent to the existence of non-well-founded sets in V, she may assent to the legitimacy of talk involving non-well-founded sets as concerned with pathological set-theoretic structures. Use of non-well-founded set theory is thus acceptable for deriving consequences in V, as long as it is kept clear which structures are being talked about in consideration of non-well-founded models of **ZFC** or models of non-well-founded set theories (such as **AFA** and **NFU**). We would like to do the same for extensions of V by providing an answer to the question of why such talk is acceptable, and what entities it concerns. As things stand, however, when we use an extension of V in proving facts about V, it is unclear what we are talking about and whether or not the discourse is philosophically acceptable.[18] Our project here will be to show that we can reliably use talk of extensions, even when we allow 'V' to denote V.

[18]There are existing interpretations in the literature. We examine these in Chapters IV and V.

Chapter II: Conclusions

Let us take stock. Gödel's Programme for justifying new axioms depends on a two step process: (1.) Formulate and suggest new axioms, and (2.) Provide a justification of these axioms. The use of resources outside V could be plausibly relevant to both. First, we may require talk of extra-V entities to even formulate an axiom. Second, even when a formulation of a particular axiom is given, we may wish to use extra-V resources, both for finding implications of the axiom and ascertaining the data that count as having been verified. In both cases we would like an explanation of what this discourse can be taken to be *about*. As we shall see, it is possible to talk about extensions of V for a far broader class of techniques than usually thought.

Chapter III

Problematic principles

We saw in Chapters I and II that (i) there is good reason to analyse how Gödel's Programme might be filled out on a Universist picture, and (ii) in doing so we might want to use resources beyond V. The time has come to lend some mathematical precision to these issues. In this chapter we explain specific examples of mathematical axioms and tools that appear to make reference to objects outside V, and what the implications are for a Universist unwilling to accept interpretation of such talk.

The structure of the chapter is as follows. After these introductory remarks, we examine principles of both width and height. We first (§1) make more precise exactly what resources we will consider when going beyond V. In particular we provide explanation of the notions of *direct height extensions*, *definable long well-orders*, *set forcings*, *class forcings*, and *sharps*. §2 analyses how some interesting set-theoretic axioms are formulated using resources beyond V. §3 then argues that there are methods in the set theorist's toolkit that the Universist may want to use in proving facts about V, but the reliability of which seems to require resources beyond the Universist's ontology. Especially of note here are the perspectives provided by outer models, understanding relatively 'small' sets with generic embeddings, and the use of long well-orders in fine structure theory. It is concluded that there are several directions in which we may want to pursue Gödel's Programme using extra-V resources, each of which is without clear interpretation for the Universist.

III.1 Extending V

We first provide more mathematical detail to the phenomenon sketched in Chapter I concerning extensions of V. As was noted there, we have two main ways of extending models of set theory: either by adding subsets (or, as we shall see, subclasses) to the model or lengthening the ordinals (or both). I will refer to the former dimension as 'extension in width' and the latter as 'extension in height'. In this section, our main aim is to provide a brief exposition of some of the relevant techniques in order to give the reader a feel for the constructions and be precise about the sense in which they go beyond V. First, however, a remark concerning the methodology of this chapter is in order.

Throughout the chapter, we will be talking about extending models, and in particular extensions of V. Since we are, at this stage, mainly concerned with exposition and how extensions could be *useful*, we shall speak naively about extensions and assume that they are always available, pausing occasionally to outline a difficulty when V is involved. The reader who suffers from any metamathematical queasiness is invited to return to *terra firma* and interpret everything said in their favourite countable transitive model, where, on the assumption that V is uncountable and satisfies **ZFC**, extensions are readily available.[1]

III.1.1 Direct extensions

As it is perhaps (to begin with) the less technical of the two dimensions, we shall start with lengthenings of the ordinals of a model, and discuss width extensions later.[2]

This first method has already been discussed in the Introduction. We might just talk about the ordinals of some model $\mathfrak{M}$ having a least upper bound in some larger universe $\mathfrak{M}'$. Hellman, for example, formulates this claim as follows:

> "Every [**ZFC**] structure...has a proper extension, both in the sense of inclusion and in the sense that it, or some copy, occurs as a "member" of

[1]See Chapter V for further discussion of this issue.

[2]We should keep in mind that, as noted in Chapter I with the case of the Shepherdson-Cohen minimal model, sometimes extending a universe in height forces it to extend in width too. While we consider the two dimensions in isolation here, this does not mean that they do not sometimes change in tandem.

its proper extensions (i.e. in the domain of the relevant membership relation)." ([Hellman, 1989],p59)

On this picture, when given a model $\mathfrak{M} \models \mathbf{ZFC}$, we simply assume that $\mathfrak{M}$ forms a set in some larger model $\mathfrak{M}'$. The ordinals of $\mathfrak{M}$ also form a set in $\mathfrak{M}'$, and have an associated von Neumann representative $Ord(\mathfrak{M}) \in \mathfrak{M}'$. It need not be the case that $\mathfrak{M}$ is an initial segment of $\mathfrak{M}'$,[3] nor that $\mathfrak{M}' \models \mathbf{ZFC}$ (in fact the first few levels extending $\mathfrak{M}$ will not). However, if desired one can simply close under the operations of $\mathbf{ZFC}$ to yield some particular $\mathfrak{N} \models \mathbf{ZFC}$ such that $\mathfrak{M} \in \mathfrak{N}$.

Of course, for the Universist, V is just one structure[4] satisfying $\mathbf{ZFC}$ among many (albeit a highly privileged one). We might then try and apply the above method to V, assuming that it forms a completely legitimate set in some extension V'. The ordinals of V are lengthened, and On^V has an associated von Neumann representation $Ord(V)$.

Clearly though, the idea of V being a member of a proper length extension V' is anathema to the Universist's position, at least insofar as said extension contains only sets.[5] For the Universist, V is meant to be all sets there are, and so Universism does not permit the existence of any such V' containing more sets.[6]

III.1.2 Definable long well-orders

A second way that we might consider 'lengthenings' of the ordinals of a model $\mathfrak{M}$, *without* explicitly adding sets, is through defining well-orders longer than $\Omega^{\mathfrak{M}}$ from *within* $\mathfrak{M}$. The following is an example:

$\alpha \prec_{\Omega+1} \beta$ iff either:

(i) $\alpha \geq 1 \wedge \beta \geq 1 \wedge \alpha < \beta$, or

[3]It should be noted that [Hellman, 1989] *does* have in mind extensions where the ground model is an initial segment of the extension.

[4]Of course, there is a difficult question of how to formalise this, given that V is not a set and the model theory of sets is often coded in set theory itself (we will see some palliative suggestions later). For the moment we will talk naively about proper-class-sized structures and models satisfying various theories and sentences, and leave an answer to the tricky problem of formalisation for Chapters V–VII.

[5]In Chapter IV, we shall see that dropping the requirement that these extensions are *sets* provides an internally consistent though philosophically unsatisfactory response to this difficulty.

[6]Indeed, such a problem arises for *any* proper-class-sized model and not just V. We will see further discussion of these issues in Chapters IV and VI.

(ii) $\alpha \geq 1 \wedge \beta = 0$.

Such an ordering (expressible in $\mathscr{L}_{\in}$) effectively puts $\emptyset$ past the end of all the ordinals of $\mathfrak{M}$, defining a well-order of length $\Omega^{\mathfrak{M}} + 1$. The example can be pushed further:

$\alpha \prec_{\Omega.2} \beta$ iff either:

(i) α is a successor and β is a limit, or

(ii) α and β are both limits and $\alpha < \beta$, or

(iii) α and β are both successors and $\alpha < \beta$.

Such a definition (expressible in $\mathscr{L}_{\in}$) *prima facie* defines an ordering of length $\Omega^{\mathfrak{M}}.2$. We can provide definitions of still longer well-orderings. The following defines an ordering on ordered pairs of ordinals that is (*prima facie*) $\Omega^{\mathfrak{M}}$ times as long as $\Omega^{\mathfrak{M}}$:

$\langle \alpha, \beta \rangle \prec_{\Omega.\Omega} \langle \gamma, \delta \rangle$ iff

(i) $\alpha < \gamma$, or

(ii) $\alpha = \gamma$ and $\beta < \delta$.[7]

Intuitively speaking, such an ordering defines an $\Omega^{\mathfrak{M}}$-length sequence of ordered pairs for every ordinal in $\mathfrak{M}$. Clearly it is possible (by moving to ordered triples, quadruples, etc.) to iterate the definition to ordinally multiplying $\Omega^{\mathfrak{M}}$ by itself any finite number of times. Again, if we let $\mathfrak{M} = V$ a puzzle emerges; we *appear* to be defining orders longer than Ω^{V}, but there is no set-theoretic ordinal corresponding to these order-types.

The issue here is, from a philosophical standpoint, importantly asymmetric from the case of simply assuming that there *exist* extensions in which V features as an element. There, we noted that if V was taken to be all sets, it was impossible to expand to a larger V' containing more sets. On the assumption that V is all the sets, it is *incoherent* to add more sets to V. Here, however, there is no incoherence in

[7]The orderings $\prec_{\Omega+1}$, $\prec_{\Omega.2}$, and $\prec_{\Omega.\Omega}$ are taken from [Shapiro and Wright, 2006].

saying that there are formulae defining well-orders longer than Ω^V. It is simply that these well-orders cannot have ordinal representatives in V (on pain of the Burali-Forti contradiction). This then raises a *conceptual* problem: if we are able (prima facie) to coherently compare these long well-orders, then why are there no ordinal-like representatives among the pure sets? If these order-types have mathematical use, what secures their reliability in the face of a lack such a representation?[8]

We will analyse these issues in greater detail later in this chapter, and indeed throughout the thesis. For the moment, we turn our attention to width extensions. As we shall see later, both height and width extensions are useful for formulating axioms and proving theorems about V.

III.1.3 Set forcing

We have seen that there are ways of talking about V that appear to involve extending the ordinals. We shall now provide an exposition of techniques for extending the *width* of models.

The first construction for expanding width is *set forcing*. Here, we begin with a partial order with domain P, ordering $\leq_{\mathbb{P}}$, and maximal element $\mathbb{1}_{\mathbb{P}}$, denoted by '$\mathbb{P} = \langle P, \leq_{\mathbb{P}}, \mathbb{1}_{\mathbb{P}} \rangle$', and have $\mathbb{P} \in \mathfrak{M}$ for some **ZFC** model[9] $\mathfrak{M}$. The relevant $p \in P$ are known as *conditions* and effectively operate by providing partial information about membership of the new object to be defined. We then, *via* a careful choice of names (known as '$\mathbb{P}$-*names*')[10] and evaluation procedure[11] add G to $\mathfrak{M}$ and close under the operations definable in $\mathfrak{M}$. The end result is a model $\mathfrak{M}[G]$ that (i) satisfies **ZFC**, (ii)

[8] As noted earlier, we might get such a representation through the use of impure set theory and use of the devices in [Menzel, 1986] and [Menzel, 2014]. For the purposes of this thesis, we confine our attention to *pure* sets.

[9] A couple of points of clarification are in order here. First, we do not always force over models of full **ZFC**, and forcing over models of weaker theories is well studied. Second, often extra constraints are put on $\mathfrak{M}$ in order to deal with any awkward metamathematics; for example, many set theorists assume that the models they work with are countable to ensure the existence of generics. A clear presentation of some of these issues is available in [Kunen, 2013]. For the moment, we simply note that we are, in this thesis, taking **ZFC** for granted as a base theory, and that the metamathematics of forcing will be considered in detail in Chapter V.

[10] A $\mathbb{P}$-name is a relation τ such that $\forall \langle \sigma, p \rangle \in \tau[$"$\sigma$ is a $\mathbb{P}$-name" $\wedge p \in \mathbb{P}]$. In other words, τ is a collection of ordered pairs, where the first element of each pair is a $\mathbb{P}$-name and the second is some condition in $\mathbb{P}$ (the definition is not vacuous in virtue of the empty set trivially being a $\mathbb{P}$-name).

[11] We evaluate $\mathbb{P}$-names by letting the value of τ under G (written '$val(\tau, G)$' or 'τ_G') be $\{val(\sigma, G) | \exists p \in G(\langle \sigma, p \rangle \in \tau)\}$. The valuation operates stepwise by analysing the valuation of all the names in τ and then either adding them to τ_G (if there is a $p \in G$ and $\langle \sigma, p \rangle \in \tau$) or discarding them (if there is no such $p \in G$).

has exactly the same ordinal height as $\mathfrak{M}$, and (iii) is strictly larger than $\mathfrak{M}$ (in the sense that $\mathfrak{M} \subset \mathfrak{M}[G]$)[12]. We provide a couple of examples to illustrate the practice which, as we shall see later in the thesis, are relevant for analysing interpretations of forcing:

Example 24. *Cohen forcing.* Let κ be a cardinal in a model $\mathfrak{M} \models \mathbf{ZFC}+CH$ such that $cf(\kappa) > \omega$. Let $\mathbb{P} = Fn(\kappa \times \omega, 2)$ be the poset of all finite partial functions from $\kappa \times \omega$ to 2, ordered under extensions of functions (i.e. $p \leq_{\mathbb{P}} q$ just in case p extends q as a finite partial function). Then in the generic extension $2^{\aleph_0} = \kappa$.[13]

We can think of the above forcing as adding κ-many reals (conceived of as functions from ω into 2) to $\mathfrak{M}$. Because the relevant partial order also satisfies the *countable chain condition* (i.e. all antichains are at most countable), the forcing preserves cardinalities and cofinalities, and the value of the continuum is moved to κ.

Whilst the above forcing largely preserves the *cardinal structure* of $\mathfrak{M}$ (in that all the cardinals and cofinalities remain the same), it is important to note just how many of the cardinality properties we can change using forcing. The above technique moves the continuum to κ, leaving the rest of the ground model relatively untouched. We can, however, seriously modify the structure of the ground model, even to the point of destroying cardinals:

Example 25. *The Lévy Collapse.* Let κ be a regular cardinal and λ be an inaccessible cardinal greater than κ. We define The Lévy Collapse (denoted by '$Col(\kappa, < \lambda)$') as follows. The relevant partial order $\langle \mathbb{P}, <_{\mathbb{P}}, \mathbb{1} \rangle$ consists of functions (each denoted by 'p') on subsets of $\lambda \times \kappa$ such that (i) $|dom(p)| < \kappa$, and (ii) $p(\alpha, \zeta) < \alpha$ for each $\langle \alpha, \sigma \rangle \in dom(p)$. In the extension every α such that $\kappa \leq \alpha < \lambda$ has cardinality κ, and every cardinal $\leq \kappa$ and $\geq \lambda$ remains a cardinal.[14]

[12]It should be noted that in order for the forcing to be non-trivial, $\mathbb{P}$ has to be *non-atomic* (i.e. every $p \in P$ has incompatible extensions in $\mathbb{P}$).

[13]For details, see [Kunen, 2013].

[14]One can also consider collapses where we use an arbitrary regular cardinal $\lambda > \kappa$ (rather than the more strict condition that λ be inaccessible). We will also denote these forcings by '$Col(\kappa, < \lambda)$', with context determining meaning when clear.

The forcing shows how we can manipulate the cardinal structure of a model of set theory in a very fine-grained manner, choosing which cardinals we want to collapse and to where, whilst leaving the rest of the structure of the ground model intact.[15]

Forcing is an especially interesting philosophical construction for a number of reasons. First, it is historically significant in that it has been used to settle many open questions concerning (the most famous examples being the independence of CH and AC). Second, it is of central importance in virtue of its ubiquity across modern set-theoretic mathematics; much of set theory concerns constructing one model from another using forcing arguments. However, especially *philosophically* interesting is that it keeps models *standard*[16]. Assuming that the ground model $\mathfrak{M}$ is transitive, well-founded, and satisfies **ZFC**, the forcing extension $\mathfrak{M}[G]$ (i) has the same ordinals as $\mathfrak{M}$, (ii) satisfies **ZFC**, (iii) is transitive, and (iv) is well-founded. Later in the chapter we shall see some further concrete examples of forcing in action. For now though, it suffices to note that the fact that forcing keeps the models *standard* is significant; generic extensions of a standard model of **ZFC** are also **ZFC**-satisfying cumulative hierarchies.

The issue concerning forcing and V is, of course, that if we wish to perform a non-trivial forcing where 'V' denotes the Universist's V as the ground model, the relevant generic G must lie outside V. But V was meant to be all the sets there are, and so such a G does not exist. Hence, without an appropriate intra-V codification establishing the reliability of such a method, using talk of set forcing extensions of V is forbidden.

III.1.4 Class forcing

Class forcing is very similar to set forcing, except we drop the requirement that the set be a member of $\mathfrak{M}$ and instead permit *proper-class-sized* partial orders. The development of class forcing goes through largely the same as set forcing, with a few

[15]There are some limitations in this respect. For example, it is not known how to arrange $(\aleph_{\omega+1})^{\mathfrak{M}} = (\aleph_2)^{\mathfrak{M}[G]}$, and often whether a construction can be performed depends upon properties of the ground model. Thanks to Sy Friedman for pointing this out to me.

[16]A model $\mathfrak{M}$ is normally called *standard* iff it has the real $\in$-relation. See [Kunen, 2013], §IV.2 for verification of the basic properties of forcing.

additional intricacies and features.[17]

When performing a class forcing, we generally force over models of the form $L(A) = \bigcup\{L(A \cap V_\alpha)|\alpha \in On\}$. Any model (M, A) of **ZF** (where we include Replacement for formulas mentioning A) can be changed to a model of this form by expanding it to a model (M, A^*) where $A^* = \{\langle 0, x\rangle | x \in A\} \cup \{\langle 1, V_\alpha^M\rangle | \alpha \in On^M\}$.

We then, as before, have a partial order with a maximal element $\langle \mathbb{P}, \leq_\mathbb{P}, \mathbb{1}_\mathbb{P}\rangle$, and add a generic to our ground model $\mathfrak{M}$. The difference here being, of course, that G is now a class rather than a set.

Class forcing has some interesting properties not enjoyed by standard set forcing. A good example is that if we allow class forcing over L, then there are reals that we can add using class forcing that cannot be added by set forcings.[18] Further, using class forcing we can produce models that violate **ZFC**. For example, the partially ordered class $Col(\omega, On)$ (i.e. functions p from finite subsets of ω into On ordered by reverse inclusion) is (without further constraints) perfectly legitimate, and, since $\mathfrak{M}[G]$ can see an ω-length cofinal sequence in $On^{\mathfrak{M}[G]}$, Replacement fails (as long as we can mention G as a predicate in the Replacement Scheme). This is unlike the case of set forcing where $\mathfrak{M}[G]$ is guaranteed to satisfy **ZFC** if $\mathfrak{M}$ does. If **ZFC** preservation is desired[19] some care[20] is required in defining the relevant $\mathbb{P} \subseteq \mathfrak{M}$ to be used in forcing.

The situation with V and class forcing is a little more complex. Whilst the majority of class forcings also add sets, there are (non-**ZFC**-preserving) class forcings that add only a class. For example, if we restrict $Col(\omega, On)$ to those $p \in Col(\omega, On)$ whose domains are initial segments of ω and force with the resulting poset, we obtain a model $\mathfrak{M}[G]$ that satisfies **ZFC** as long as G is *not* allowed as a class predicate, as the first-order domains of $\mathfrak{M}$ and $\mathfrak{M}[G]$ are identical.[21] However, if we admit G as

[17] Details of the presentation given here are available in [Friedman, 2000], Chapter 2.

[18] See [Friedman, 2010], p559 for details.

[19] We shall argue later that for the purposes of talking about forcings over V, there is no reason why **ZFC** preservation is especially desirable.

[20] The relevant conditions are *pretameness* and *tameness* of the partial order, corresponding respectively to preservation of Replacement and Power Set. See [Friedman, 2000] for details.

[21] To see this, note that for any $\mathbb{P}$-name σ for this poset and for each condition p in the intersection of the transitive closure of σ with $\mathbb{P}$, $ran(p) \subseteq rank(\sigma)$. We then define the dense set $D = \{p \in \mathbb{P}|rank(\sigma) \in ran(p)\}$. D is then both dense and definable over $\mathfrak{M}$. Letting $\sigma^p = \{\tau^p|\exists q \in \mathbb{P}[\tau, q \in \sigma \wedge p \leq_\mathbb{P} q]\}$ We then have $\sigma^p = \sigma^G \in M$ whenever G is $\mathbb{P}$-generic over $\mathfrak{M}$ and $p \in D \cap G$, because p either extends or is incompatible with any condition in the transitive closure of σ. Hence, whenever G is $\mathbb{P}$-generic over $\mathfrak{M}$, they contain exactly the same first-order objects.

a predicate into the language, Replacement fails. This is because G codes a cofinal sequence from ω to $On^{\mathfrak{M}[G]}$ and there is no set in $\mathfrak{M}[G]$ corresponding to $On^{\mathfrak{M}[G]}$.[22]

The problems for interpreting class forcing over V are thus twofold. First, there is the standard problem (as in set forcing) of interpreting talk of class forcings that add sets. Secondly, however, the addition of G also adds a *class*, and there can be (non-**ZFC**-preserving) class forcings can add classes whilst leaving the sets untouched. Assuming then that the Universist is already happy to try interpreting class forcings, there are thus two dimensions to be accounted for: (1.) interpret the sets added, and (2.) interpret how classes get added. (1.) we assume will go through exactly as in the set forcing case, however (2.) presents a new and interesting challenge, especially when there are no sets added to underpin a change in classes. The Universist has a couple of ways she could deal with this: she could either (a) reject talk of such class forcings as illegitimate, and provide a philosophical explanation of why they are different from other instances of class forcing, or (b) provide an explanation of how to interpret forcing talk that adds classes but leaves the sets untouched.

III.1.5 Sharps and mice

The final variety of width extension we shall explicitly consider involves the taking of sharps. The notion of a sharp is important within the highly technical field of *inner model theory* and the consideration of objects known as *mice*. As we shall see in the rest of the chapter, issues in inner model theory are relevant not just for both the formulation of new axioms and the proving of theorems about V, but also provide an interesting case where height and width extensions interact.

The situation is perhaps most easily seen through starting with a consideration of L. We first need to set up some familiar definitions from the model theory of sets. A mathematical notion that will appear throughout this thesis is that of an *elementary embedding*:

Definition 26. Let $\mathfrak{N}$ and $\mathfrak{M}$ be (set-theoretic) structures. Then an *elementary embedding* is a mapping $j : \mathfrak{N} \longrightarrow \mathfrak{M}$ such that:

[22] For details, see [Holy et al., F].

(i) j is one-to-one.

(ii) j preserves first-order truth: i.e. $\mathfrak{N} \models \phi(x_0, ..., x_n)$ iff $\mathfrak{M} \models \phi(j(x_0), ..., j(x_n))$.

We furthermore say that an embedding is *non-trivial* iff it is not the identity map.

Definition 27. The least ordinal moved by j (if there is one) is called the *critical point* of j and is denoted by '$crit(j)$'.[23]

We can now begin to explain the theory of sharps, and how this might relate to extensions of V. To illustrate the rough structure of the construction, we begin by considering the case of L. Suppose that there is a cardinal κ and an ultrafilter U on $\mathcal{P}^L(\kappa)$ with the following properties:

(i) $Ult(L_\kappa, U)$ is well-founded and its collapse produces an embedding $j_0 : L \longrightarrow L$.[24]

(ii) $crit(j_0) = \kappa$.

(iii) $j_0(\kappa) = \kappa_1$.

Suppose further that it is then possible to take the ultrapower of L_{κ_1} by an ultrafilter U_1, producing an embedding j_1 with the same features except $crit(j_1) = \kappa_1$ and $j_1(\kappa_1) = \kappa_2$. Assume also that we can continue this process, taking the direct limit of the ultrapowers, continuing through the ordinals whilst keeping the ultrapower well-founded. The process will produce (by the elementarity of the relevant j_α) a class of indiscernibles for L and code an elementary embedding $j : L \longrightarrow L$. More formally, we define the following object:

Definition 28. [25] A *mouse* is a structure[26] $J^U_\alpha = (J_\alpha, U)$ such that:

[23] In all the elementary embeddings we shall consider, the non-triviality of j implies the existence of a critical point.

[24] The *ultrapower construction* is a technique whereby, using an ultrafilter on $\mathcal{P}(\kappa)$, we can define an equivalence relation on ${}^\kappa\mathfrak{M}$, producing equivalence classes for every $f \in {}^\kappa\mathfrak{M}$. This gives us an embedding $i_0 : \mathfrak{M} \longrightarrow Ult(\mathfrak{M}/U)$, where '$Ult(\mathfrak{M}/U)$' denotes the ultrapower of $\mathfrak{M}$ by U. Assuming that the ultrapower relation E_U (also defined using U) is well-founded, extensional, and set-like, we can then use the Mostowski Collapse Lemma to obtain another embedding $i_1 : Ult(\mathfrak{M}/U) \longrightarrow \mathfrak{N}$ to produce a transitive set-theoretic structure $\mathfrak{N}$. The relevant j are then the composition of these two embeddings. For details of the ultrapower construction, (including the Łoś Theorem) see [Jech, 2002].

[25] See [Jensen, 1995] and [Schimmerling, 2001] for excellent aerial surveys of sharps and mice, and [Jech, 2002], Ch 35 for the details (but beware of typographical errors).

[26] The J hierarchy is an alternative presentation of L (in the sense that $J[A]$ is always the same structure as $L[A]$) in terms of rudimentary functions, but has some additional pleasing fine structural features (such as the closure of the J_α under the formation of ordered pairs).

(i) U is a normal, iterable[27], κ-complete ultrafilter on some $\kappa < \alpha$.

(ii) All iterated ultrapowers of J^U_α by U are well-founded.

(iii) For κ and finite $p \subseteq \alpha$, $J^U_\alpha = Hull_n^{J^U_\alpha}(\kappa \cup p)$. i.e. J^U_α is the Σ_n Skolem hull (for some $n \in \omega$) of κ unioned with a finite $p \subseteq \alpha$.[28]

Mice are used in the highly technical context of *fine structure theory*: the study and construction of models resembling L yet containing non-constructible sets. We shall see several issues concerning mice and extensions of V. First, however, we define an ordering on mice. Let J^U_α be a mouse at κ, λ be a regular cardinal greater than κ^+, and C_λ be the club filter on λ (i.e. C_λ contains all and only the closed and unbounded subsets of λ). We then know that the λ^{th} iterate of J^U_α is such that:[29]

$Ult^\lambda_U(J^U_\alpha) = J^{C_\lambda}_\beta$, for some β.

We can now provide the following definition:

Definition 29. Let $J^U_\alpha = Hull_n^{J^U_\alpha}(\gamma \cup p)$ and $J^{U'}_{\alpha'} = Hull_m^{J^{U'}_{\alpha'}}(\gamma' \cup p')$ be mice and λ be any sufficiently large regular cardinal. Let $i_{0,\lambda} : J^U_\alpha \longrightarrow J^{\mathcal{C}_\lambda}_\beta$ and $i'_{0,\lambda} : J^{U'}_{\alpha'} \longrightarrow J^{\mathcal{C}_\lambda}_{\beta'}$ be the respective iterated ultrapowers witnessing their mice-hood. Then $J^U_\alpha = Hull_n^{J^U_\alpha}(\gamma \cup p) <_M J^{U'}_{\alpha'} = Hull_m^{J^{U'}_{\alpha'}}(\gamma' \cup p')$ iff:

(i) $\beta < \beta'$, or

(ii) $\beta = \beta'$ and $\gamma < \gamma'$, or

(iii) $\beta = \beta'$, $\gamma = \gamma'$, and $q < q'$ in the descending lexicographic ordering.

This ordering well-orders the class of all mice. We can now define $0^\sharp$ as follows:[30]

Definition 30. $0^\sharp$ is the minimal mouse (if it exists).

[27] An ultrafilter U is *iterable* iff whenever there is a $< \kappa$-length sequence of members of U in a model $\mathfrak{M}$, then the set of those members is also in $\mathfrak{M}$. This condition facilitates the iteration of the ultrapower construction. For details, see [Jech, 2002], p354.

[28] A Skolem hull of a set X is the closure of a set of Skolem functions witnessing the true formulas of some $\mathscr{L}$ on a structure $\mathfrak{M}$. A Σ_n Skolem hull is simply a Skolem hull where we only worry about Σ_n formulas.

[29] See, [Jech, 2002], pp661-662.

[30] $0^\sharp$ admits of many alternative definitions. See [Jech, 2002] and [Kanamori, 2009] for discussions of several available options.

Interestingly, if we assume that other L-like models also have ultrafilters on cardinals producing well-founded ultrapowers, we can also have sharps for these models. For example, given a non-constructible real x, we can have an embedding $j_x : L[x] \longrightarrow L[x]$, with corresponding indiscernibles and a sharp for $L[x]$ (denoted by '$x^\sharp$'). In fact, there is a sharp hierarchy, and we can proceed through the ordinals defining further sharps.

The study of the fine structure of L and its extensions is both abstract and difficult. We shall, therefore, avoid delving into the technical details too far in order to keep philosophical upshots clear. For now, we need only note that mice represent natural objects of set-theoretic study. Importantly though, mice cannot be within their models; the relevant indiscernibles provide a truth definition of the structure in question (and hence cannot be in the structure by Tarski's Theorem on the undefinability of truth).

However, more important still is that there is no known forcing-like technique for constructing $0^\sharp$ from within L (and similarly for other mice over their relevant inner models). Thus $L[0^\sharp]$ is a *non*-forcing extension of L; $0^\sharp$ cannot be reached with known forcing constructions.

Further, once we have all mice, it is possible to continue. Mice are essential in the definition of the core model (often denoted by 'K'). First we need to expand the definition of L relative to a predicate A. We introduce a predicate $A(x)$ into the language of **ZFC**, such that $A(x)$ holds iff $x \in A$. We then can have the notion of being *definable over* $\mathfrak{M}$ *relative to* A:

Definition 31. $def_A(\mathfrak{M}) = \{X \subseteq M | X$ is definable over $(M, \in, A \cap M)\}$

We then can define the following relativised constructible universe:

Definition 32.

(i) $L_0[A] = \emptyset$,

(ii) $L_{\alpha+1}[A] = def_A(L_\alpha)$, for successor $\alpha + 1$,

(iii) $L_\lambda[A] = \bigcup_{\beta<\lambda} L_\beta$ for limit λ,

(iv) $L[A] = \bigcup_{\alpha \in On} L_\alpha$

Effectively, we allow ourselves the expressive resources provided by a predicate for an additional (possibly set-sized) class for defining sets in the usual way in L. Combining this idea with the notion of a mouse, we are led to the notion of the *Core Model*:

Definition 33. The *Core Model* is the following structure:

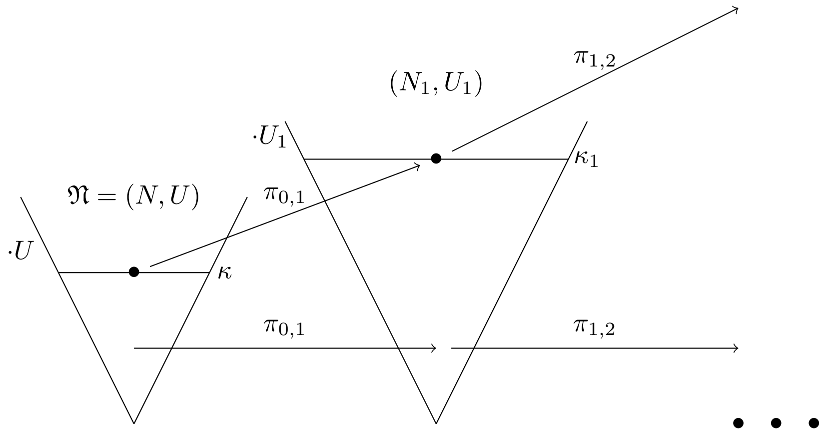
$K = L[\{J^U_\alpha | \text{“}J^U_\alpha \text{ is a mouse” }\}]$

Essentially K is an L-like model containing all mice (and indeed $K \models \mathbf{ZFC}$). However, under the assumption of the existence of an L-like inner model containing a measurable cardinal κ (denoted by '$L[U]$', where U witnesses κ's $L[U]$-κ-completeness), there is a non-trivial elementary embedding $j : K \longrightarrow K$ with attendant indiscernibles. Further, we can continue, having an embedding $j : L[U] \longrightarrow L[U]$, with more indiscernibles and ultrapower maps, producing new L-like inner models.

The details become exceptionally complex very quickly. Currently, the production of inner model theory has stalled at the level of many Woodin cardinals, or (with additional assumptions on the kind of iteration available) at supercompact cardinals.[31] Despite this, work continues, and the set-theoretic picture and philosophical upshots of the foregoing discussion remain; the consideration of mice and sharps is a very general enterprise, holding for a wide variety of models, and the models in question are not obtainable by known forcing constructions. Further, we can provide a more general definition for sharps in a wider context (a visual representation is provided in Figure III.1):

[31] For discussion see [Sargsyan, 2013] and [Steel, 2014].

Figure III.1: A visual representation of the initial ultrapowers corresponding to a sharp (N, U)



Definition 34. A structure $\mathfrak{N} = (N, U)$ is called a *sharp with critical point* κ, a *sharp*, or just a $\sharp$, iff[32]:

(i) $\mathfrak{N}$ is a model of $\mathbf{ZFC}^-$ (i.e. $\mathbf{ZFC}$ with the power set axiom removed) in which κ is the largest cardinal and is strongly inaccessible.

(ii) (N, U) is amenable (i.e. $x \cap U \in N$ for any $x \in N$).

(iii) U is a normal measure on κ in (N, U).

(iv) $\mathfrak{N}$ is iterable in the sense that all successive ultrapowers starting with (N, U) are well-founded, providing a sequence of structures (N_i, U_i) and corresponding Σ_1-elementary iteration maps $\pi_{i,j} : N_i \longrightarrow N_j$ where $(N, U) = (N_0, U_0)$.

The above definition provides the resources to talk about extensions of *arbitrary* models with the relevant ultrafilter needed to construct embeddings. Such sharps cannot be within their respective models; as with mice[33] the indiscernibles they generate provide a truth definition for the relevant structure.[34] The question then is the following: could the Universist codify talk of a sharp for V? This is especially difficult, as such an object could not be defined by standard forcing techniques (assuming

[32]This way of defining sharps is due to [Friedman, S] and [Friedman and Honzik, 2016].
[33]Mice are just a *particular kind* of sharp.
[34]See [Friedman and Honzik, 2016] for details.

of course that the Universist has an interpretation of both set and class forcing). This fact, as we shall see in Chapter V, will turn out to be significant. For now, however, we shall explore how we might put these techniques and objects to work in the study of V.

III.2 Formulating axioms that go beyond V

We first examine axioms, relevant for the analysis of Gödel's programme, that can be formulated using extensions of V. Of especial interest will be how consideration of extensions lets us formulate certain kinds of reflection principle, postulate the existence of inner models, and define new kinds of embedding.

III.2.1 Inner model density

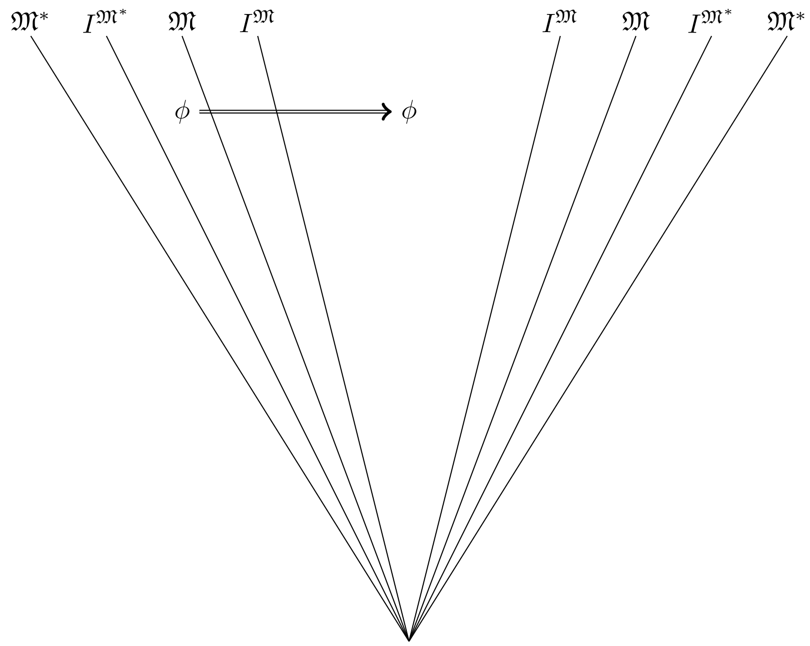
Extensions of V are useful for postulating the existence of many inner models. The *Inner Model Hypothesis* does exactly this, using extensions of a model $\mathfrak{M}$ in order to make claims about the inner models of $\mathfrak{M}$ (see Figure III.2 for a visual representation of an application of the IMH):

> **Definition 35.** [Friedman, 2006] Let ϕ be a parameter-free first order sentence. $\mathfrak{M}$ satisfies the *Inner Model Hypothesis* (henceforth 'IMH') iff whenever ϕ holds in an inner model $I^{\mathfrak{M}^*}$ of an outer model $\mathfrak{M}^*$ of $\mathfrak{M}$, there is an inner model $I^{\mathfrak{M}}$ of $\mathfrak{M}$ that also satisfies ϕ.

The IMH thus states that $\mathfrak{M}$ has a high density of inner models, in the sense that any sentence ϕ true in an inner model of an outer model of $\mathfrak{M}$ is already true in an inner model of $\mathfrak{M}$. In this way, $\mathfrak{M}$ has been maximised with respect to *internal consistency*.

There are a number of reasons to find the IMH interesting, not least because it maximises the satisfaction of consistent sentences within structures internal to $\mathfrak{M}$. The IMH is thus (if true) foundationally significant: it gives us an inner model for any sentence model-theoretically compatible with the initial structure of V, and thus serves to ensure the existence of well-founded, proper-class-sized structures in

Figure III.2: A visual representation of an application of the IMH



which we can do mathematics. However, it is also interesting in that it has various *anti*-large cardinal properties, indeed certain versions of the IMH prove that there are no inaccessibles in $\mathfrak{M}$.[35] The principle (and each of its variants) is thus worth scrutiny; the IMH provides the possibility of motivating an axiom that substantially reduces the 'cap' on the height of the ordinals.[36]

Whence the problem then for the Universist? If the Universist wishes to use the IMH as a new axiom about V,[37] she has to examine issues concerning extensions of V. If they ascribe *no* meaning to claims concerning extensions, then the IMH is utterly trivial. Under this analysis, everything true in an inner model of an outer model of V is also true in an inner model of V, as either (i) the outer model is proper,

[35]See [Friedman, 2006], p597 for details.

[36]Talk of a 'cap' on the ordinals is somewhat difficult, as usually the term is taken to talk about properties of *cardinals* that cannot exist. Thus, the term 'cap' denotes a relationship between height and width, rather than only height. For example, one can have countable models with a highly impoverished conception of the power set operation that believe they contain supercompact cardinals. For this reason, even assuming a definite power set operation (and hence fixing of this aspect of the cardinal properties of V), what one takes to be the cap will depend on other properties of V. If $V = L$, the cap appears as early as $0^\sharp$. Assuming AC, there cannot be a Reinhardt cardinal (i.e. there is no non-trivial elementary $j : V \longrightarrow V$). The point here is that the IMH pulls this cap all the way down to one of the *smallest* kinds of large cardinal.

[37]Here, and throughout, we will use the term 'IMH' both to refer to the property that a model can satisfy, and also V's satisfaction of said property. We let context determine meaning.

does not exist, and hence nothing is true in an inner model of that proper outer model of V, or (ii) the outer model is V itself, and obviously anything true in an inner model of V is true in an inner model of V.

However, even supposing that the Universist allows *some* interpretation of extension talk, the content that the IMH has is going to vary according to the resources one allows. If for example, one permits the interpretation of set forcing and no more, the IMH only goes slightly stronger than **ZFC**. It does, however, imply $V \neq L$:

Theorem 36. If the IMH is true restricted to set forcings, then $V \neq L$.

Proof. Assume $V = L$ and that the IMH is true restricted to set forcings. Then there there is an inner model of an outer model in which $V = L$ is false (the addition of a single Cohen real x over L to $L[x]$ will suffice, with the relevant inner model simply being the forcing extension $L[x]$). By the IMH there is an inner model of L in which $V \neq L$. But L is the smallest inner model, and so $V = L$ and $V \neq L$, $\perp$. ■

However, though it is sufficient to get us a certain density of inner models (enough to break $V = L$) we get more if we restrict to class forcings. This is brought out in the following:

Theorem 37. There is a model satisfying the IMH for set forcing that does not satisfy the IMH for class forcing.

Proof. Let $\mathfrak{M}$ be a model of $V = L$ containing a reflecting cardinal[38] κ. Next, perform the Lévy Collapse to move κ to ω_1. The extension $L[G]$ satisfies the IMH for set forcing. However, by [Jensen, 1972] one can define class-generic reals that are not set-generic. Let H be a generic yielding such a real in an extension $L[H]$. We then note that $L[G]$ does not have an inner model with a real that is not set-generic over L, and so the IMH for class forcing fails. ■

[38]In the present context, κ is a *reflecting cardinal* iff (i) κ is regular, and (ii) whenever ϕ is a sentence with parameters from V_κ, if ϕ holds in some V_α (in V), then ϕ holds in some V_β with $\beta < \kappa$. The definition is equivalent to saying that κ is regular and V_κ is Σ_2 elementary in V. In terms of consistency strength, it is stronger than inaccessible but weaker than Mahlo. Many thanks to Sy Friedman for communicating to me this proof technique.

Thus, we see how the IMH varies in its content and implications dependent on what we allow as extensions. Supposing that we do wish to assert that the IMH is true of V. Then we need to give meaning (in whatever appropriate codification) to the claim that V has various kinds of extension. The intra-V consequences provable from the IMH will then vary depending on the kinds of width extensions we can interpret.

We have seen that it is possible to formulate a principle with substantial anti-large cardinal features using extensions. A natural question then is: Does the consideration of extensions feed into formulation of standard large cardinal axioms? Two of the most used ways of formulating large cardinal axioms are (i) the consideration of reflection properties over V, and (ii) postulating the existence of elementary embeddings. As we shall see, extensions are relevant for both these ways of formulating new axioms.

III.2.2 Direct reflection

We start with reflection properties. When put *very* informally, a *reflection principle* is of the following general form:

> Any property held by V is held by some initial segment of V.

Of course, what one takes to be a property of V is going to be difficult to cash out, especially on a Universist picture. Furthermore, care is needed as reflection with unrestricted third-order parameters is inconsistent![39] However, key for our discussion will be the possible use of extensions in formulating axioms that postulate the witnessing of reflection by initial segments. Most discussion of reflection centres on properties held by V using a parameter over V. For a (higher-order) parameter A over V a reflection principle states:

$$(V, \in, A) \models \phi \rightarrow \exists \alpha (V_\alpha, \in, A \cap V_\alpha) \models \phi^{V_\alpha}$$

In informal terms, if V satisfies ϕ relative to some parameter A over V, there is a V_α that also satisfies ϕ when A, quantifiers in ϕ, and parameters in ϕ are restricted

[39] For the result, see the suggestive remarks in [Reinhardt, 1974] and more precise exposition in [Tait, 2005].

to V_α. There is a detailed literature on the small large cardinals that one can derive from such principles[40], and also problems of justification and consistency[41] but we focus here on the possible use of extensions in formulating reflection.

The use of extensions is particularly relevant in providing characterisations of axioms that postulate the witnessing of reflection principles by individual V_α. Friedman and Ternullo, for example, have formulated the following version of reflection:

> **Definition 38.** [Friedman and Ternullo, S] $\mathfrak{M}$ satisfies the *extended reflection axiom*[42] (henceforth 'ERA') iff $\mathfrak{M}$ has a lengthening $\mathfrak{M}' \models$ **ZFC** (i.e. a model satisfying **ZFC** containing $\mathfrak{M}$ as an element) such that for all first-order formulas ϕ and subclasses $A \subseteq \mathfrak{M}$ belonging to $\mathfrak{M}'$, if $\phi(A)$ holds in $\mathfrak{M}'$ then $\phi(A \cap V_\alpha^{\mathfrak{M}})$ holds in $V_\beta^{\mathfrak{M}}$ for some pair of ordinals $\alpha < \beta$ in $\mathfrak{M}$.

So, for a universe $\mathfrak{M}$ to satisfy the ERA, it must have a **ZFC**-satisfying lengthening $\mathfrak{M}'$ such that if $\mathfrak{M}'$ satisfies ϕ relative to the parameter A, then $\mathfrak{M}$ already contains a pair of ordinals α and β, with $\alpha < \beta$, such that V_β can see a level (namely V_α) that reflects ϕ. Effectively, $\mathfrak{M}$ can already see pairs of ordinals witnessing various reflection axioms. The challenge for the Universist is that if she wishes to assert that the ERA holds of V, we have to be able to refer to extensions of V. We have to state that there is a *lengthening* V' of V, such that V already has witnesses for any reflection occurring in V' in its own V_α. Of course this is hard to interpret for the Universist, since there are *no* height extensions of V. Thus, without further interpretation, the ERA will always come out as trivially false.

III.2.3 $\sharp$-generation

We have discussed how we might use extensions to directly formulate notions of reflection. It is interesting to note that it is possible to encapsulate the large cardinal consequences of reflection properties through the use of sharps. Earlier we said that

[40] See, for example, [Levy, 1960] and [Bernays, 1961].

[41] For discussion, see [Koellner, 2009] and [Barton, Fb].

[42] Friedman and Ternullo in fact use the term 'ordinal maximality of $\mathfrak{M}$' instead of '$\mathfrak{M}$ satisfying the extended reflection axiom' largely because [Friedman and Ternullo, S] is concerned with *maximality criteria* on universes. As we are interested in *axiom formulation* for a Universist, we opt for the term 'extended reflection axiom'.

a sharp is a structure $\mathfrak{M} = (M, U)$ with various iterability and amenability properties. Recall that part and parcel of a sharp (N, U) is the existence of a sequence of structures (N_i, U_i) and corresponding Σ_1-elementary iteration maps $\pi_{i,j} : N_i \longrightarrow N_j$. We can then provide the following definition:

Definition 39. [Friedman, S] A model $\mathfrak{M} = (M, \in)$ is *sharp-generated* (or just $\sharp$-*generated*) iff there is a sharp (N, U) and an iteration $N_0 \longrightarrow N_1 \longrightarrow N_2 ...$ such that $M = \bigcup_{\alpha \in On^{\mathfrak{M}}} V_{\kappa_\alpha}^{N_\alpha}$.

In other words, a model is sharp-generated iff it arises through collecting together the $V_{\kappa_i}^{N_i}$ (i.e. each level indexed by the critical point of the i^{th} iteration map, according to the model indexed by i) resulting from the iteration of a sharp through the ordinal height of $\mathfrak{M}$.

A model's being sharp-generated engenders some pleasant features. In particular, it implies that any satisfaction obtainable in an extension of $\mathfrak{M}$ (possibly with parameters) is already reflected to an initial segment of $\mathfrak{M}$.[43] In this way, we are able to coalesce many reflection principles into a single property of a model. A natural axiom then would be:

Axiom 40. *Axiom*$^\sharp$. V is sharp-generated.

which would allow us to assert in one fell swoop that V satisfies many reflection axioms (rather than having to assert them in a piecemeal fashion). However, such an axiom is also clearly problematic; claiming that V is sharp-generated depends upon the existence of a sharp for V, which cannot be in V by design. Thus, the claim that V is sharp-generated comes out as trivially false; there simply is no such sharp.

III.2.4 Generic embeddings

We have seen that we can use extensions to formulate large cardinal principles based on ideas concerning *reflection*. We now wish to see if we can formulate new set-theoretic axioms via the use of embeddings in combination with talk of extensions.

In the case of reflection principles, we formulate new and stronger large cardinal principles by postulating higher-orders of reflection and incorporating steadily more

[43] See [Friedman, S] and [Friedman and Honzik, 2016] for discussion.

parameters. A different way of asserting the existence of large cardinals (often quite strong), is through the use of elementary embeddings. Recall from III.1.5 that we could characterise $0^\sharp$ through a non-trivial $j : L \longrightarrow L$. Such an embedding pushes us (unlike standard reflection principles[44]) outside $V = L$; the embedding cannot be within L by design. We can generalise further, to consider embeddings from arbitrary models $j : \mathfrak{N} \longrightarrow \mathfrak{M}$, producing a further hierarchy of large cardinals. For example, measurable cardinals are a natural object of study:

> **Definition 41.** A cardinal κ is *measurable* iff κ is the critical point of a non-trivial elementary embedding $j : V \longrightarrow \mathfrak{M}$.

These cardinals push us well outside $V = L$, and imply the existence of all small[45] large cardinals consistent with L. When defining a large cardinal through an embedding $j : \mathfrak{N} \longrightarrow \mathfrak{M}$, the strength of the embedding depends mainly on two parameters:

(i) The size of $\mathfrak{N}$ and $\mathfrak{M}$.

(ii) Where j sends the ordinals.

We have already seen the minimal case for an embedding between proper class models; namely the existence of a non-trivial $j : L \longrightarrow L$. If we assume that $dom(j) = V$, we strengthen to the level of a measurable cardinal, and break $V = L$. We know that the existence of a non-trivial $j : V \longrightarrow V$ is inconsistent (*modulo* **ZF**) with AC.[46] Despite this we can study intermediate cardinals by modifying the properties of j and $\mathfrak{M}$. For example, we can use the following pair of definitions to strengthen the notion of measurable along the dimensions of (i) and (ii):

> **Definition 42.** A cardinal κ is λ-*supercompact* iff it is the critical point of a

[44] See [Koellner, 2009] for a discussion of why standard reflection is not able (yet) to push us outside $V = L$. There is another class of axioms (also often referred to by the term 'reflection principles') that combine reflection with embeddings, see [Reinhardt, 1974], [Welch, 2014], and [Barton, Fb] for discussion.

[45] A note on terminology is needed here. When distinguishing between cardinals, there are two especially pertinent points in the hierarchy of large cardinal strength; (i) when a cardinal breaks $V = L$, and (ii) when a cardinal breaks AC. Cardinals consistent with $V = L$ I shall call *small* large cardinals, those that are known to be inconsistent with $V = L$ but not known to be inconsistent with AC I shall call *middling* large cardinals, and those known to be inconsistent with AC but not known to be inconsistent with ZF I shall call *very* large cardinals.

[46] See [Kunen, 1971] for the result. We shall see further discussion of this phenomenon in Chapter VI.

non-trivial elementary embeddings $j : V \longrightarrow \mathfrak{M}$, such that $j(\kappa) > \lambda$ and ${}^{\lambda}\mathfrak{M} \subseteq \mathfrak{M}$ (i.e. $\mathfrak{M}$ is closed under λ-sequences).

Definition 43. A cardinal κ is *supercompact* iff it is λ-supercompact for all $\lambda \in On$.

The definition of supercompact uses the dimensions of (i) and (ii) to increase the strength of the embedding. We postulate a higher degree of similarity between V and $\mathfrak{M}$ (in terms of closure under λ-sequences for the relevant λ), and stipulate that j sends κ above λ.

Standard discussions of middling large cardinals proceed from this template. However, we can generalise the construction to *generic* embeddings. Given a forcing construction adding a generic G over a model $\mathfrak{N}$, a generic embedding is of the form $j : \mathfrak{N} \longrightarrow \mathfrak{M} \subseteq \mathfrak{N}[G]$. In other words, we begin to study embeddings from structures to inner models *of their forcing extensions*.

Recently, there has been an increased focus on such embeddings. Indeed, the study of generic embeddings has become widespread, as Foreman (in a handbook article on generic embeddings) illustrates:[47]

> "The main aim of the chapter is to illustrate that there is a coherent theory here, that there are unifying fundamental ideas that occur frequently in many different contexts. These include master condition ideals, natural and induced ideals, disjointing, self-genericity, the role of diagonal unions for representing Boolean sums, good elementary substructures—the list is long." ([Foreman, 2010], p890)

Generic embeddings are thus useful for studying certain natural mathematical properties. Furthermore, the involvement of extensions in the consideration of embeddings provides an additional dimension in which we may vary the nature of the construction. Not only does the embedding depend upon the size of the domain and range of the embedding and where the ordinals are sent, but also on a third parameter:

[47] See also, [Foreman, 1986] for several key results, and [Foreman, 1998] for a more informal overview.

(iii) The nature of the forcing required to define j.[48]

The fact that we have an extra dimension in which we can vary the structure of these kinds of embeddings makes them an intriguing subject matter. However, even more interesting is that the critical points of these axioms can be rather small. For example, we have:

> **Theorem 44.** If I is an ω_2-saturated ideal on ω_1 and U is generic for the poset of I-positive sets, then in $V[U]$ the ultrapower $Ult(V, U)$ is well-founded and we get a map $j : V \longrightarrow \mathfrak{M} \subseteq V[U]$ with $crit(j) = \omega_1$ and $j(\omega_1) = \omega_2$.

Such an embedding from V to another (necessarily fatter) model $\mathfrak{M}$ has ω_1 as its critical point, far below the size of a measurable cardinal.[49] Despite the smallness of the critical points, however, these embeddings have a significant amount of large cardinal strength.[50] Thus, these embeddings provide significant combinatorial power whilst facilitating proof concerning *small* uncountable sets. Foreman, for example, writes:

> "The advantage of allowing the embeddings to be generic is that the critical points of the embeddings can be quite small, even as small as ω_1. For this reason they have many consequences for accessible cardinals, settling many classical questions of set theory." ([Foreman, 2010], p887)

So, generic embeddings provide a new way of combining large cardinal strength with a seemingly very direct way of talking about the small uncountable sets. Of course the same problem arises as with standard forcing constructions; the generic used to engender the embedding cannot be in V, and j itself is coded as a class of ordered pairs in $V[G]$. Thus any theorem proceeding from the supposition of the existence of a generic G facilitating an embedding in an extension (as above) is true,

[48]For further exposition of this line of thinking, see [Foreman, 1998] and [Foreman, 2010].

[49]We know, for instance, that $|\omega_1| \leq |\mathcal{P}(\omega)|$, making it accessible. To get an idea of the scale of the difference, if κ is measurable then it has to be an inaccessible limit of inaccessible cardinals.

[50]For example, the existence of both a saturated ideal on ω_1 (and associated generic embedding) and a measurable cardinal implies the existence of an inner model with a Woodin cardinal, whereas the consistency strength of a measurable cardinal is far below that of a single Woodin. See [Steel, 1996] for details.

but tells us nothing about the structure of V; the conditional is made true vacuously by the non-existence of such a G.

III.2.5 Remarkable behaviour in extensions

It is not just through postulating the existence of embeddings *directly* between V and models within forcing extensions that allow us to formulate new axioms with significant large cardinal strength. Through analysing properties of ordinals in extensions, we can come to characterisations of new varieties of cardinal.

An example of Schindler is especially pertinent here:

> **Definition 45.** [Schindler, 2000] A cardinal κ is *remarkable* iff in the $Col(\omega, <\kappa)$ forcing extension $V[G]$, for every regular $\lambda > \kappa$ there is a cardinal $\lambda_0 < \kappa$, λ_0 regular in V, and $j : H^V_{\lambda_0} \longrightarrow H^V_\lambda$ such that $crit(j) = \gamma$ and $j(\gamma) = \kappa$.

We are able to characterise the notion of a cardinal being remarkable if, when we collapse all cardinals less than κ to ω through forcing, in this $Col(\omega, <\kappa)$ extension $V[G]$, for every regular $\lambda > \kappa$ in $V[G]$ there is a V-regular cardinal $\lambda_0 < \kappa$ such that the heriditarily λ_0-sized sets elementarily embed into the hereditarily λ-sized sets. Thus, by studying how sets are embeddable *in the extension*, we are able to ascribe large cardinal properties to ordinals *in* V. The consistency strength of a remarkable cardinal, for instance, does not break $V = L$ but is substantially stronger than a weakly compact cardinal[51]. The problem here is that we are predicating a large cardinal property of κ, an object *in* V, but using resources from extensions to define what it is to have said large cardinal property.

III.3 Using tools that go beyond V

Thus far, we have examined some axioms that are formulated through the use of extensions of V. However, recall from the previous chapter that it was not just for formulation of axioms that we wished to provide interpretation of extensions of V,

[51] Weakly compact cardinals are so named in virtue of being characterisable through compactness properties on infinitary languages. They admit of a diverse number of equivalent characterisations. For details, see [Kanamori, 2009].

but also for ascertaining *verifiable consequences*. We now review several ways in which we can use extensions of V to prove facts about V.

III.3.1 Perspective of extensions

The first point uses well-known facts about the model theory of sets. We know that any Π_1-sentence of set theory ϕ is *downward absolute*, in that if it holds in some model $\mathfrak{M}$ and $\mathfrak{N} \subseteq \mathfrak{M}$ then ϕ also holds in $\mathfrak{N}$. We then know that if we are able to interpret talk of extensions V, and prove Π_1-facts in said extensions, then such facts will also be true in V. Thus, talk of extensions can facilitate proof of intra-V facts.

Indeed, it turns out that certain extensions are useful for proving theorems about ground models. A good example here is the following:[52]

> **Theorem 46.** [Baumgartner and Hajnal, 1973] $\omega_1 \longrightarrow (\alpha)^2_n$ for all finite n and countable α (i.e. For all finite n and countable α, every partition of the two-element subsets of ω_1 into a finite number of pieces has a homogeneous[53] set of order-type α).

The proof proceeds by finding a homogeneous set in a forcing extension where MA holds. This establishes that a certain tree from the ground model is non-well-founded in the extension. We then know, by the absoluteness of well-foundedness, that the tree is also non-well-founded in the ground model, establishing the theorem. If we permit the use of forcing constructions over V, we are thereby able to establish the results of [Baumgartner and Hajnal, 1973] as true about the *real* ω_1. It is thus desirable, in line with the Hilbertian Challenge, that we have an account of why moving to an extension of V where MA holds is acceptable for proving a theorem about V.

III.3.2 Large cardinals: redux

In the previous section, we discussed how the use of extensions aided in the formulation of axioms of significant consistency strength. However, it is not just in the

[52] I am grateful to Andrés Caicedo for pointing out this example.

[53] Here, a homogeneous set is a subset X of ω_1 such that every 2-element subset of X is in the same member of the partition.

formulation of the axioms that extensions are useful. They are also helpful for determining consequences of an axiom V given an *alternative* formulation that avoids mention of extensions.

An interesting feature of both remarkable cardinals and generic embeddings is that both admit of an equivalent formulation *internal* to V. We begin with the case of remarkable cardinals. The following is a theorem:

> **Theorem 47.** [Schindler, 2001] A cardinal κ is *remarkable* iff for every regular $\lambda > \kappa$, there are countable transitive models $\mathfrak{M}$ and $\mathfrak{N}$ with embeddings:
>
> (i) $\pi : \mathfrak{M} \longrightarrow H_\lambda$ with $\pi(\kappa_0) = \kappa$,
>
> (ii) $\sigma : \mathfrak{M} \to \mathfrak{N}$ such that
>
> (a) $crit(\sigma) = \kappa_0$,
>
> (b) $Ord(\mathfrak{M})$ is a regular cardinal in $\mathfrak{N}$, and $\mathfrak{M} = H^{\mathfrak{N}}_{Ord(\mathfrak{M})}$
>
> (c) $\sigma(\kappa_0) > Ord(\mathfrak{M})$.

Thus, the notion of remarkable cardinal admits of a formulation using only objects within V. Instead of talking about properties in the $Col(\omega, {<}\kappa)$ forcing extension, we instead talk about embeddings between countable transitive models $\mathfrak{M}$ and $\mathfrak{N}$ in V, and, for regular cardinals $\lambda > \kappa$, embeddings between $\mathfrak{M}$, $\mathfrak{N}$, and the hereditarily-${<}\lambda$-sized sets. The characterisation in terms of behaviour in a forcing extension is perhaps more natural, but this nonetheless shows that remarkable cardinals can be understood within V, rather than having to move to the $Col(\kappa, \omega)$ extension.

Is the same true for generic embeddings? Often, we can use metamathematical results in order to find equivalent first-order statements within V. Foreman discusses the technique:

> "The language of ideals, together with the mechanics of forcing provide the same kind of vehicle for stating generalized large cardinal axioms in the language of set theory. Assuming the existence of a proper class

> of Woodin cardinals, Burke's Proposition[54]...shows that every countably complete ideal is pre-precipitous. More directly: the existence of an elementary embedding $j : V \longrightarrow \mathfrak{M} \subseteq V[G]$ where $G \subseteq \mathbb{P}$ is generic and $j"\lambda \in \mathfrak{M}$ is easily seen to be equivalent to the existence of a $\mathbb{P}$-term for an ultrafilter $\dot{U} \subseteq \mathcal{P}(\mathcal{P}(\lambda))^V$ is normal for regressive functions in V and fine and is such that there is no descending ω-sequence of U-equivalence classes of functions from V. The idea of an *induced ideal* allows us to restate this combinatorially as a normal, fine, precipitous ideal I on $\mathcal{P}(\lambda)$ such that the quotient algebra $\mathcal{P}(\mathcal{P}(\lambda))/I$ inherits some of the properties of the original partial ordering $\mathbb{P}$. Finally, moving along the "F" axis [the nature of forcing required to define j] in the direction of greater strength, the saturation properties of ideals play exactly the same role for generalized large cardinals as ultrafilters do for conventional large cardinals." ([Foreman, 2010], p1128)

The details here are exceptionally technical and research is ongoing. However, the philosophical upshot is the following. For many large cardinals, the existence of an embedding $j : \mathfrak{M} \longrightarrow \mathfrak{N}$ is equivalent to the existence of some first-order combinatorial sets. For instance, the existence of a measurable cardinal is equivalent to there being an uncountable cardinal κ with a κ-complete, non-principal ultrafilter on $\mathcal{P}(\kappa)$. Foreman's point is that the same holds (under the assumption of the existence of the requisite Woodin cardinals) for generic embeddings. The properties of the embedding can be traced through the mechanisms of forcing[55] to properties of the ideal (namely the level of saturation it exhibits). Thus, the existence of a generic embedding usually corresponds to the existence of an ideal in V with a particular saturation property.

So, we see that both remarkable cardinals and generic embeddings admit of a formulation that refers to objects solely in V. We might then ask if these examples are so problematic; one might think that the Universist can just regard the use of

[54]*Burke's Proposition* states that if I is a countably complete ideal on a set Z, and $\delta > |Z|$ is a Woodin cardinal, then I is pre-precipitous. For details, see [Foreman, 2010], p1113.

[55]Some of these mechanisms will be discussed in more detail in Chapter V. In particular, the role of '$\mathbb{P}$-terms' will be made clear.

extensions as a heuristic for talking about the relevant intra-V formulations. Aside from the fact that it is often the most *natural* formulation that uses extensions of V (and therefore it might be regarded as desirable to find a more faithful interpretation of these formulations), we note that if we have *already accepted* the existence of objects given by the intra-V formulations, then there are cases when the easiest proof proceeds using extra-V resources. For this reason, we would like an explanation of why the proof methods that proceed via extra-V resources are guaranteed to result only in truth.

The issues surrounding remarkable cardinals and generic embeddings are thus twofold in character. First, we would like an explanation of why using extra-V resources for formulation will result in truth concerning V. However, we would also like reassurance that, when using an intra-V formulation, any steps of proof that proceed through extensions will not lead us astray.

III.3.3 Long well-orders and fine structure

We noted in §1 that there were several definitions that appeared to define well-orders 'longer' than Ω^V. We then noted that we wanted to talk about objects known as *mice*, and collect them together into the core model K. For the sake of clarity, we repeat the definition of a mouse:

Definition 28. A *mouse* is a structure J^U_α such that:

(i) U is a normal, iterable, κ-complete ultrafilter on some $\kappa < \alpha$.

(ii) All iterated ultrapowers of J^U_α by U are well-founded.

(iii) For κ and finite $p \subseteq \alpha$, $J^U_\alpha = Hull_n^{J^U_\alpha}(\kappa \cup p)$. i.e. J^U_α is the Σ_n Skolem hull (for some $n \in \omega$) of κ unioned with a finite $p \subseteq \alpha$.

As can be seen from how we define K, there are situations in which we want to talk about all mice. Often, the class of all mice is considered under its natural ordering (defined earlier). Again, for the sake of clarity, we recall its definition. Letting J^U_α be a mouse at κ, λ be a regular cardinal greater than κ^+, and C_λ be the club filter on λ, and $Ult_U^\lambda(J_\alpha^U) = J_\beta^{C_\lambda}$, for some β, the canonical well-ordering on the class of all mice was defined as follows:

Definition 29. Let $J_\alpha^U = Hull_n^{J_\alpha^U}(\gamma \cup p)$ and $J_{\alpha'}^{U'} = Hull_m^{J_{\alpha'}^{U'}}(\gamma' \cup p')$ be mice and λ be any sufficiently large regular cardinal. Let $i_{0,\lambda} : J_\alpha^U \longrightarrow J_\beta^{\mathcal{C}_\lambda}$ and $i'_{0,\lambda} : J_{\alpha'}^{U'} \longrightarrow J_{\beta'}^{\mathcal{C}_\lambda}$ be the respective iterated ultrapowers witnessing their mice-hood. Then $J_\alpha^U = Hull_n^{J_\alpha^U}(\gamma \cup p) <_M J_{\alpha'}^{U'} = Hull_m^{J_{\alpha'}^{U'}}(\gamma' \cup p')$ iff:

(i) $\beta < \beta'$, or

(ii) $\beta = \beta'$ and $\gamma < \gamma'$, or

(iii) $\beta = \beta'$, $\gamma = \gamma'$, and $q < q'$ in the descending lexicographic ordering.

As noted earlier, the details are highly technical, and we do not wish to obfuscate philosophical purposes by delving too deep. Important for our current purposes, is the fact that the ordering under which we speak about all mice is a Σ_2-definable ordering that appears to be of length $\Omega.3$.

The use of such an ordering has gained some interest in the philosophical literature.[56] The problem here is exactly the same as with a definable well-order longer than Ω^V; we wish to talk about all mice under such an ordering, but it is unclear what guarantees the reliability of such talk. Shapiro and Wright (in enumerating responses to the problem of long well-orders) express the problem as follows:

> "Allow the unrestricted quantifications and the definitions of the troublesome predicates, but deny that they are associated with ordinals (order-types). Cost: transfinite inductions and recursions of the relevant 'lengths' then come into question (at least on the assumption that transfinite recursions and inductions require an associated order-type) which are part of expert practice and seemingly quite intelligible." ([Shapiro and Wright, 2006], p293)

Exactly what is required for a practice to be 'intelligble' in this context is a somewhat difficult question. For example, if we know that talking about long well-orders

[56]See, for example, [Shapiro and Wright, 2006]. The discussion there is a little odd; Shapiro and Wright appear at points to be conflating issues with a proper class j yielding a mouse (which is undoubtedly a tricky issue in itself), with the length of the ordering under which we talk about the mice.

will always yield truth, but think that the discourse is not to be associated with any *actual* long well-orders, does this make the practice 'intelligible'? Or is something more required? These issues aside, Shapiro and Wright's point is well put: we wish to know *what* it is that licenses such talk as reliable. As it stands, we have a seemingly cogent practice, without adequate explanation of its reliability.

Chapter III: Conclusions.

In Chapter I, we saw that there was a debate between the Universist and her opponents that could be made precise by examining the question of whether there was just one maximal definite interpretation of set-theoretic discourse. In particular we noted that there were two main dimensions to be analysed in settling set-theoretic questions; namely *height* and *width*.

Then, in Chapter II, we argued that there are reasons to analyse the extent to which extensions of V can be interpreted within the Universist framework. In particular, the possibility of using extensions to derive consequences about V (to be used in extrinsic justifications) and formulating new axioms was discussed.

In this chapter, we have seen concrete examples from several areas of set theory of just this phenomenon. Axioms can be formulated using expressive resources that purport to refer to different extensions of V, and vary in their content dependent upon what talk can be given interpretation. Further, extensions are important for the consideration of deriving consequences about V. In particular, we can use extensions to prove facts using absoluteness, analyse the consequences of axioms with both intra-V and extra-V formulations, and use long well-orders in fine structure theory. Examining the extent to which an interpretation meeting the Hilbertian Challenge can be given to talk involving extensions of V is thus deserving of scrutiny. As we shall see, it is possible to provide such an interpretation in a philosophically motivated fashion. However, we must first analyse some extant attempts to meet the Hilbertian Challenge. We shall see that many of these strategies are unsatisfactory, and their failure highlights some additional features we would like an interpretation of extensions of V to satisfy.

Chapter IV

Climbing above V?

Thus far, we have seen that if the Universist wishes to use certain resources in executing Gödel's Programme, she needs to explain why discourse that appears to go 'beyond' V in height will not lead her astray in proving facts about V.

Recall that there are two main techniques that require interpretation:

(1.) *Direct extensions*: Be able to interpret talk concerning models in which V appears as a set.

(2.) *Long well-orders*: Provide reasons to think that discourse involving definable well-orders longer than Ω is acceptable for proving facts about V.

In this chapter, we consider some possibilities for answering the Hilbertian Challenge with respect to these methods. Whilst each view faces seemingly insurmountable philosophical difficulties, we argue that the areas in which they struggle highlight two constraints on satisfying the Hilbertian Challenge. Our strategy is as follows:

§1 explains and discusses the challenges facing a view (Simple Nominalism) on which we baldly deny that we cannot have the symbol 'V' denoting V when considering a height extension. We argue that though the view can interpret the relevant mathematics as *non-vacuous* it cannot secure us satisfaction of the Hilbertian Challenge. §2 then examines the prospects for the polar opposite of Simple Nominalism, on which we accord full ontological reference to height extensions of V. It is noted that while such a position avoids the problems of Simple Nominalism, it goes against the spirit of Universism as providing a unified foundational arena for mathematical discourse. §3 then analyses a slight modification of the strategy of §2, to accord height extensions of V the status of possibilia. Again, however, it is noted that the view faces deep philosophical challenges, and many of the problems of §2 recur. Finally, §4 ties the previous sections together, and notes that several of the problems correspond to the violation of either of two constraints, one ontological and one foundational.

IV.1 Simple Nominalism

The Universist certainly does not want to countenance the existence of sets outside V. A natural position then is to hold nominalism towards the objects purportedly referred to by height extensions. Direct extensions and long well-orders on this view have no ontological reference. Shapiro and Wright express the view concisely concerning long well-orders:

> "Allow the unrestricted quantifications and the definitions of the troublesome predicates, but deny that they are associated with ordinals (order-types)." ([Shapiro and Wright, 2006], p293)

We might thus suggest that we can have discourse concerning long well-orders (or in the case of direct extensions, taller universes), but simply demur from allowing the existence of the relevant extra-V objects. Such a response seems to fly in the face of the Hilbertian Challenge, however. The content of the Hilbertian Challenge is *precisely* to provide an ontology that gives us confidence in the use of such resources for discovering new facts about V. Here, we have merely declared by fiat that we refuse to supply such an interpretation or ontology.

One response to this issue would be to restrict the use of such resources to particular transitive models $\mathfrak{M} \in V$, such that $Ord(\mathfrak{M}) \in On$. We can then perfectly well have height extensions of $\mathfrak{M}$ that provide an ontology for such discourse. It is just that when a troublesome extension is used, we restrict quantification and interpret *locally* over $\mathfrak{M}$ rather than *globally* over V.

Such an interpretation is fine for reassuring us that discourse involving extensions is non-vacuous. For instance, we can compare the order types $\prec_{\Omega+1}$ and $\prec_{\Omega.2}$ through saying that for any transitive model $\mathfrak{M}$ such that $Ord(\mathfrak{M}) = \alpha \in On$, $\prec_{\Omega+1}$ defines an ordering of length $\alpha+1$ over $\mathfrak{M}$ and $\prec_{\Omega.2}$ defines an ordering of length $\alpha.2$ over $\mathfrak{M}$. Similarly, for the ERA, we can talk about a model $\mathfrak{M}$ with $Ord(\mathfrak{M}) \in On$, and how $\mathfrak{M}$ relates to its extensions (in terms of there being some $\mathfrak{M}'$ such that all reflection in $\mathfrak{M}'$ is already realised within levels of $\mathfrak{M}$). In this way, the mathematics involved in the formulation of the axioms is shown to be *non-vacuous*.

The central problem here is that this tactic simply will not do for some of the uses to which we wish to put these extensions. It is all well and good to have the result that the mathematics in question is non-vacuous; however, we want to have an assurance that the relevant kind of talk is rigorous when the domain of discourse is the widest possible (namely when 'V' denotes V). In the case of long well-orders, Shapiro and Wright put the problem as follows:

> "transfinite inductions and recursions of the relevant 'lengths' then come into question (at least on the assumption that transfinite recursions and inductions require an associated order-type) which are part of expert practice and seemingly quite intelligible." ([Shapiro and Wright, 2006], p293)

For the case of long well-orders we can elaborate on Shapiro and Wright's point along the following lines. Recall that in III.1.5 we wanted to talk about the class of all mice when defining the following model:

$K = L[\{J^U_\alpha | \text{"}J^U_\alpha \text{ is a mouse"}\}]$

A natural way to speak about the class predicate used in the definition of K is through talking about the class of all mice. If we want to talk about *all* mice though (rather than just a restricted class of mice over some set-sized model $\mathfrak{M}$) under their natural ordering, then the fact that the definition of the mouse ordering refers to a bona fide ordinal over set-sized $\mathfrak{M}$ is no guarantee of its reliability where the context of discourse is V.

The situation is brought into even sharper relief with the ERA. For, we are interested in seeing whether we could formulate the ERA as true of V. If we interpret the ERA locally, though we are thereby guaranteed the non-vacuity of the discourse, we nonetheless have no assurance that the consequences of the ERA hold of V. We would simply be talking about particular $\mathfrak{M}$ that can see levels witnessing any reflection in an extension of $\mathfrak{M}$. But the existence of such an $\mathfrak{M}$ is no guarantee that V itself satisfies the ERA.

IV.2 A different kind of object?

We saw in the last section that there was a challenge with interpreting height extensions locally, in that we would like our interpretation of the discourse to have its intended consequences for V proper. One way of achieving this is to find an interpretation that actually gives ontological referents to extensions of V, where 'V' really does denote V. We can begin by noting that (in the case of long well-orders) the Burali-Forti contradiction is only problematic if the new order-types have representatives *in* V. Dummett, for example, says the following:

> "What the paradoxes revealed was not the existence of concepts with inconsistent extensions, but of what may be called indefinitely extensible concepts. The concept of an ordinal number is a prototypical example. The Burali-Forti paradox ensures that no definite totality comprises everything intuitively recognisable as an ordinal number, where a definite totality is one quantification over which always yields a statement determinately true or false." ([Dummett, 1991], p316)

Now Dummett has in mind here a distinction between definite and indefinite concepts, and in particular the notion of an indefinitely extensible concept. One might, then, take Dummett's point to simply be a flat rejection of Universism; he does not hold that all statements about 'the ordinals' are determinately true or false in virtue of the indefinite extensibility of the concept of an ordinal number. However, we could also take it that there are, instead, ordinal-like objects that do not define ordinals as understood as sets in V, but rather a new kind of object (call them 'super-ordinals'). Shapiro and Wright, picking up on some of these Dummettian ideas, express the point as follows:

> "Allow the quantification and the predicates, allow the associated order-types, but deny that they are ordinals as originally understood—rather, they are 'higher-order' ordinals, 'proper' ordinals, 'super-ordinals', or whatever." ([Shapiro and Wright, 2006], p293)

This ontology would then sanction the use of a long well-order over V. For any such use there would be a legitimate mathematical object corresponding to the long well-order, namely a super-ordinal of the required length. Aside from the fact that it is unclear how to deal with the extensions of V to be supported by the ERA (we would have to have extensions of V by non-set-like objects that nonetheless obey extensionality and are well-founded), Shapiro and Wright have little truck with the suggestion, describing it as:

> "Hypocrisy. Recall that Ω was supposed to encompass the ordinals in a maximally general sense of ordinal, common to all types of well-orderings. Also, the option is unstable. If we are now saying that Ω does not encompass a maximally general sense of ordinal, and that we need to distinguish (how many?) successive orders of ordinals, then just consider all of these, and the dialectical situation repeats itself" ([Shapiro and Wright, 2006], p293)

There are two separate problems at play here. The first difficulty is that the *exact* structure of the problem replays in the case of the super-ordinals. By simply replicating the Burali-Forti reasoning, there can be no super-ordinal representing the super-order-type of all super-ordinals. Presumably, then this will necessitate some super-duper-ordinal objects, and then super-duper-trooper-ordinals, and so on. Clearly there is no end to this process. We can then ask if there is an 'ordering type' to *all* these higher-order-types, and we have exactly the same situation as with our original ordinals (unless we deny the ability to quantify over all higher order-types).

This raises two subproblems. First, it highlights that there is little conceptual gain in shifting to super-ordinals. Even if the move is made, we still get similar problems for the (non-set) mathematical objects that appear outside V. Second, and more importantly, the repetition of the *exact structure of reasoning* and very *ordinal-like* properties of the super-ordinals might lead us to question the extent to which we are talking about a genuinely new object. When discussing super-ordinals, it looks a lot like we were simply talking about the ordinals over a particular set-sized, transitive $\mathfrak{M}$ with $Ord(\mathfrak{M}) = \kappa$, and have merely established that the definable orderings over $\mathfrak{M}$ have new ordinals above κ. The super-order-types, just like the garden-variety ordinals, are properties corresponding to well-orders. In other words, if it looks like an ordinal, walks like an ordinal, and quacks like an ordinal, then the object we are talking about is probably an ordinal, no matter if we use the term 'super-ordinal' to refer to it.

The second problem, and one that cuts right to the heart of the Universist's motivation for her perspective, is that this method of interpreting long well-orderings utterly vitiates the role of V as the unified arena for analysing foundations. While the explicit contradiction of asserting that there are sets outside V is avoided, the price is high under this solution. If she wishes to maintain V as the arbiter of mathematical truth and rigour, it is thus desirable for the Universist to find a solution that does not admit the existence of *any* mathematical entities that cannot be represented using sets from V.

IV.3 Modal notions

In the previous section, we saw that there was a problem with admitting the existence of objects that could not be represented within V. We might, therefore, deny that such objects *actually* exist, and instead hold that they *could* exist. One might interpret the work of Reinhardt in [Reinhardt, 1974] and [Reinhardt, 1980] in this manner. The two papers are notoriously difficult in both technical content and philosophical interpretation, and so we will not provide a full exposition. We can, however, say a few words about the basic idea, and why it immediately appears deeply problematic.

Reinhardt is explicitly concerned with formulating large cardinal axioms, but wishes to do so through considering height extensions of V. In particular, his interest is in considering elementary embeddings between V and its possible extensions. We will not delve into the details of the embeddings; the ideas are notoriously complex, the underlying conceptions are problematic[1], and we already have many varieties of extension to consider. The key fact is that the 'sets' Reinhardt considers in extensions are not *actual* but rather *possible*. The thought is that while V is a certain size, it *could* have been taller, with longer ordinals. We can then interpret height extensions of V as concerned with these 'possible' sets above V.

There are several immediate objections one can raise for such a modal proposal from a Universist perspective. Firstly, most (if not all) Universists hold that mathematical objects exist out of necessity if at all. Thus, the idea that there 'could' have been more sets than there actually are seems anathema to a normal core tenet held by the Universist.

This said, we might not want to burden the Universist with too many additional philosophical principles. For the moment then, let us assume that she does leave room for non-actual mathematical possibilia. Just as in the previous section, we can point to the role that the Universist sees for V. In I.5, we noted that the Universist saw a substantial benefit of her position as the ability to have a single unified domain in which we interpret mathematical discourse. The modal interpretation precisely gives this up. For now, it is not just sets in V that determine what holds mathematically. Rather, mathematical truth is settled by sets in V combined with *possible* sets. Moreover, it is not simply the modal properties *of sets in* V that is being analysed. In an essential way, the objects of study are *literally outside* V (in some appropriate possible world) and are also objects of mathematical study. In particular, the properties of these possible sets have direct bearing upon what goes on in V. The point is simply this, whilst we avoid talk of extant objects that are unrepresentable in V, the use of modal notions buys us little; we still have mathematical discourse with interpretation that goes beyond V in an essential way.

[1] See [Koellner, 2009] for discussion.

The above arguments are not meant to be conclusive against the modal view. Indeed, the mathematics produced by Reinhardt is fascinating, and the search for a satisfactory underlying conception continues. However, the problems presented for the advocate of modal notions are especially tricky when viewed from the perspective of the Universist. Indeed, it is hard to see how the difficulties of the modal view could be overcome by the Universist without her providing a suitable coding that would remove the need for modal notions anyway. An explanation that avoids use of objects not representable within V is thus preferable for the Universist.

IV.4 The Ontological and Foundational Constraints

It seems then that each of the above suggestions for interpreting height extensions of V falls flat (or at least faces deep problems). However, we can learn from these difficulties. Recall the Hilbertian Challenge:

> **The Hilbertian Challenge.** Provide philosophical reasons to legitimise the use of extra-V resources for formulating axioms and analysing intra-V consequences.

The above failures highlight that not just any way of responding to the Hilbertian Challenge will do. Rather there are additional philosophical constraints on what is to count as a *philosophically satisfactory* solution to the problem presented. In this section, we diagnose some of the above failures and formulate two additional constraints on resolution of the Hilbertian Challenge.

The first constraint makes precise and explicit a feature and motivation of the Universist's position that was outlined in I.5 and used repeatedly in the above arguments. We noted that a substantial motivation for her view was that she was able to provide a single unified arena that set the standards for mathematical rigour and correctness. Thus, in responding to the Hilbertian Challenge, the Universist should aim to adhere to the following constraint:

> **The Foundational Constraint.** In responding to the Hilbertian Challenge, do so in a way that does not necessitate the use of resources that

> cannot be represented by sets within V.

Such a constraint puts significant requirements on our interpretation of extensions of V. A failure of both super-ordinals and the modal approach was that the Foundational Constraint was violated, thereby undermining the Universist's motivation for her position. A coding that only talks about sets in V would thus mesh better with her philosophy.

However, an additional problem was noticed in the discussion of super-ordinals. There the question was raised of why talk of super-ordinals did not just represent covert talk about ordinals definable over some particular set-sized $\mathfrak{M}$. Super-ordinals looked very similar to ordinals (in that they were well-ordering properties) thus raising the question of whether there was a meaningful distinction being made. This raises the following:

> **The Ontological Constraint.** Any interpretation of extra-V resources should make clear the *ontological difference* between the interpretation of extensions and normal sets within V. In other words, any interpretation must make clear in what sense the interpretation does not *literally* refer to extra-V sets.

Satisfaction of the ontological constraint avoids the problems outlined above. In particular, if we can make clear the ontological difference between our interpretation of extensions and sets within V, we will avoid the worry that talk of extensions of V is merely covert talk about sets. Hence, we will avoid the problem that our discourse involving extensions fails to have the content we would like concerning V.

Chapter IV: Conclusions

We conclude by noting a few salient features of the above discussion. Simple Nominalism was given a rather hard time in §1, however it is important to note that it was on the right track in many respects. Especially important to note is the fact that some sort of nominalism concerning extra-V resources deals with the Foundational Constraint immediately and directly. For, if we hold that extension talk does not refer to *anything* outside V, then there is no worry that the referents might be unrepresentable in the sets. We can then retain V as our unified arena for mathematics.

Unfortunately, the problem with Simple Nominalism was that our interpretation of the term 'V' did not bear enough resemblance to V. What we require then is the interpretation of 'V' to bear as much resemblance to V as possible, whilst keeping in line with the above constraints. We shall see in Chapters VI and VII just how this can be done. For now, we turn to questions of width.

Chapter V

Broadening V's horizons?

After setting up the Hilbertian Challenge for the Universist in Chapters I-III, we saw in the last chapter that extant attempts to interpret height extensions of V were philosophically unsatisfactory. We then noted that there were two constraints (one foundational and one ontological) on a resolution of the Hilbertian Challenge. On the ontological side, we wish to interpret extensions in such a way that a conceptual difference between *bona fide* sets and the interpretation of extensions is made clear. To conform to the Foundational Constraint, we need to be able to represent our proposed interpretation using sets within V. In this chapter we shall analyse some possible interpretations of width extensions. Again, we shall find each suggestion unsatisfactory. However, we shall also see that another constraint is highlighted by their shortcomings, one concerning *methodology*. Our strategy is as follows:

§1 considers the use of Boolean-valued models. They are found wanting in that there are significant difficulties in accounting for class forcing, and also certain aspects of the interpretation do not conform to *how* we reason about extensions. §2 provides an exposition of a modification of this approach (the Boolean ultrapower) that in many ways makes the interpretation of extensions more *natural*. Again it is noted that problems of interpreting class forcing remain, but there is an added problem concerning forcing arguments involving relatively small uncountable sets. §3 explains how we might use the *forcing relation* to interpret forcing over V. It is noted that while it provides an interpretation for a wider class of forcing arguments, the interpretation is somewhat unnatural, there is a class of forcing arguments that it cannot interpret, and the treatment of sharps is left untouched. §4 discusses the use of countable transitive models to interpret extensions of V. It is argued that while such a strategy gets frustratingly close to a satisfactory interpretation of extensions of V, certain axioms do not yet have satisfactory interpretation. Finally, §5 discusses the shortcomings of these approaches, and identifies a *methodological* constraint on responses to the Hilbertian Challenge. It is argued that there is an apparent tension in satisfying the three constraints simultaneously.

V.1 Boolean-valued models

We begin with a discussion of Boolean-valued models.[1] Such an interpretation pertains only to *forcing* extensions. Starting with a forcing poset $\mathbb{P}$, we can find a separative[2] partial order $\mathbb{Q}$, equivalent to $\mathbb{P}$ for forcing, and a (unique up to isomorphism) *Boolean completion* of $\mathbb{Q}$ (denoted by '$\mathbb{B}(\mathbb{Q})$').[3,4] We then consider the class of $\mathbb{B}(\mathbb{P})$-names (denoted by '$V^{\mathbb{B}(\mathbb{P})}$'), and assign values from $\mathbb{B}(\mathbb{P})$ to atomic relations between them. More formally, we define:

Definition 48. Let $\mathbb{B}$ be a Boolean algebra. A *Boolean-valued model* $(A, F_{\mathbb{B}})$ consists of a Boolean universe A and assignment of Boolean-values $F_{\mathbb{B}}$ from $\mathbb{B}$ to the relations $=$ and $\in$ obeying the following constraints (for any x, y, z, w, v):

(1) $[\![x = x]\!] = \mathbb{1}_{\mathbb{B}(\mathbb{P})}$

(2) $[\![x = y]\!] = [\![y = x]\!]$

(3) $[\![x = y]\!] \cdot [\![y = z]\!] \leq [\![x = z]\!]$

(4) $[\![x \in y]\!] \cdot [\![v = x]\!] \cdot [\![w = y]\!] \leq [\![v \in w]\!]$

[1]The Boolean-valued approach was developed by Scott and Solovay, with additional contributions by Vopěnka (among others). See [Smullyan and Fitting, 1996], p273 for historical details and references.

[2]A partial order $\mathbb{P} = (P, <_{\mathbb{P}})$ is *seperative* iff for all $p, q \in P$, if $p \not\leq_{\mathbb{P}} q$ then there exists an $r \leq_{\mathbb{P}} p$ that is incompatible with q.

[3]For details of Boolean algebras (from which our presentation is derived) see [Jech, 2002], Chapter 7. A discussion of Boolean completions is available in ibid. Chapter 14.

[4]We will (mildly) abuse notation and use $\mathbb{B}(\mathbb{P})$ to refer to the relevant Boolean completion even when $\mathbb{P}$ is not separative (i.e. the Boolean completion obtained from a separative partial order $\mathbb{Q}$, such that $\mathbb{Q}$ is equivalent to $\mathbb{P}$ for forcing).

We then, for $a_1, ..., a_n \in A$, define $[\![\phi(a_1, ..., a_n)]\!]$ as follows:

(i) The value of atomic formulas is given by (1)–(4).

(ii) $[\![\neg\psi(a_1, ..., a_n)]\!] = -[\![\psi(a_1, ..., a_n)]\!]$

(iii) $[\![\psi(a_1, ..., a_n) \wedge \chi(a_1, ..., a_n)]\!] = [\![\psi(a_1, ..., a_n)]\!] \cdot [\![\chi(a_1, ..., a_n)]\!]$

(iv) $[\![\psi(a_1, ..., a_n) \vee \chi(a_1, ..., a_n)]\!] = [\![\psi(a_1, ..., a_n)]\!] + [\![\chi(a_1, ..., a_n)]\!]$

(v) $[\![\psi(a_1, ..., a_n) \to \chi(a_1, ..., a_n)]\!] = [\![(\neg\psi \vee \chi)(a_1, ..., a_n)]\!]$

(vi) $[\![\psi(a_1, ..., a_n) \leftrightarrow \chi(a_1, ..., a_n)]\!] = [\![((\psi \to \chi) \wedge (\chi \to \psi))(a_1, ..., a_n)]\!]$

(vii) $[\![\exists x\psi(x, a_1, ..., a_n)]\!] =$

$$\sum_{a \in A} [\![\psi(a, a_1, ..., a_n)]\!]$$

(viii) $[\![\forall x\psi(x, a_1, ..., a_n)]\!] =$

$$\prod_{a \in A} [\![\psi(a, a_1, ..., a_n)]\!]$$

We are now in a position to define a Boolean-valued model over V:

Definition 49. Let $\mathbb{P}$ be a forcing poset and $\mathbb{B}(\mathbb{P})$ be its Boolean-completion. Then the *Boolean-valued model* $V^{\mathbb{B}(\mathbb{P})}$ is defined via the following transfinite recursion:

(1) $V_0^{\mathbb{B}(\mathbb{P})} = \emptyset$,

(2) $V_{\alpha+1}^{\mathbb{B}(\mathbb{P})} =$ The set of all functions f with $dom(f) \subseteq V_\alpha^{\mathbb{B}(\mathbb{P})}$ and values in $\mathbb{B}(\mathbb{P})$.

(3) $V_\lambda^{\mathbb{B}(\mathbb{P})} = \bigcup_{\beta<\lambda} V_\beta^{\mathbb{B}(\mathbb{P})}$, for limit λ,

(4) $V^{\mathbb{B}(\mathbb{P})} = \bigcup_{\alpha \in On} V_\alpha^{\mathbb{B}(\mathbb{P})}$

Letting $\rho(x)$ be the least α such that $x \in V_{\alpha+1}^{\mathbb{B}(\mathbb{P})}$, we inductively define Boolean values on pairs $(\rho(x), \rho(y))$ as follows:

(a) $[\![x \in y]\!] = \sum_{t \in dom(y)}([\![x = t]\!] \cdot y(t))$

(b) $[\![x \subseteq y]\!] = \prod_{t \in dom(x)}(-x(t) + [\![t \in y]\!])$

(c) $[\![x = y]\!] = [\![x \subseteq y]\!] \cdot [\![y \subseteq x]\!]$

It is then routine to show that $V^{\mathbb{B}(\mathbb{P})}$ is a Boolean-valued model of **ZFC**: it satisfies clauses (i)–(viii) and every axiom (and hence every theorem) of **ZFC** has Boolean-value $\mathbb{1}_{\mathbb{B}(\mathbb{P})}$ in $V^{\mathbb{B}(\mathbb{P})}$.[5] Moreover, for the purposes of consistency proofs, we know that if we can assign ϕ a Boolean-value greater than $\mathbb{0}_{\mathbb{B}(\mathbb{P})}$, then $\neg\phi$ is not a consequence of **ZFC** (as if $\neg\phi$ is a consequence of **ZFC**, then ϕ receives Boolean value $\mathbb{0}_{\mathbb{B}(\mathbb{P})}$). In fact, an assignment of a Boolean value greater than $\mathbb{0}_{\mathbb{B}(\mathbb{P})}$ to ϕ exactly mimics the satisfaction of ϕ in some $V[G]$, for V-generic G.

Thus, by discussing the Boolean-valued model $V^{\mathbb{B}(\mathbb{P})}$, we are able to capture the intra-V content of talking about set forcing extensions of V. For example, suppose that we wish to show that the satisfaction of ϕ in a set forcing extension by $G \subseteq \mathbb{P} \in V$ has consequence ψ within V. We can then take the Boolean completion $\mathbb{B}(\mathbb{P})$, and show that we can assign ϕ Boolean-value greater than $\mathbb{0}_{\mathbb{B}(\mathbb{P})}$ in $V^{\mathbb{B}(\mathbb{P})}$. By tracing the Boolean-values back to V, we then know that V satisfies ψ.

There are several problems with the use of Boolean-valued models, however. Aside from the fact that it leaves extension by sharps untouched, it has two particular limitations when it comes to forcing. The first is that it is unclear how to interpret class forcing on the present approach. For, in class forcing, the relevant partial order $\mathbb{P}$ is proper-class-sized, and hence unbounded in the V_α. When defining the Boolean completion $\mathbb{B}(\mathbb{P})$ we then encounter a difficulty. The usual method for defining a Boolean completion is to find a separative partial order equivalent to $\mathbb{P}$ for forcing (known as the separative quotient), and embed it into a Boolean algebra[6]. Effectively, we turn a copy of $\mathbb{P}$ 'upside down', and add it 'above' $\mathbb{P}$.[7]

In the present context, however, it is unclear how to do this. As we have already 'reached the top' of V, there is nowhere for the new sets to go above $\mathbb{P}$. This is not to say that it is *impossible* to provide a Boolean completion for class forcings; indeed some interesting results have emerged from explorations in this direction[8],

[5] See [Jech, 2002], Chapter 14.

[6] More formally, for any set-sized partial order $\mathbb{P}$, there is a Boolean algebra $\mathbb{B}(\mathbb{P})$ and an embedding $e : \mathbb{P} \longrightarrow \mathbb{B}(\mathbb{P})^+$ (where $\mathbb{B}(\mathbb{P})^+$ is the set of non-zero elements of $\mathbb{B}(\mathbb{P})$) such that for $p, q \in \mathbb{P}$: (i) if $p \leq_{\mathbb{P}} q$, then $e(p) \leq_{\mathbb{B}(\mathbb{P})} e(q)$, (ii) p and q are compatible iff $e(p) \wedge e(q)$, and (iii) $\{e(p) | p \in \mathbb{P}\}$ is dense in $\mathbb{B}(\mathbb{P})$.

[7] For the full details, see [Jech, 2002], Chapter 14.

[8] See, for example, [Holy et al., F].

and work continues. However, there are difficult mathematical challenges, and the exact landscape is still to be discovered.

A second problem, however, is that this way of coding extensions of V makes our interpretation seriously deform our normal set-theoretic thinking. When reasoning with extensions, set theorists often proceed via thinking combinatorially and classically, i.e. they reason in a two-valued manner about the ways in which sets can be combined, how they can be mapped to one another, and so on. Often their thinking does not have the character of reasoning about Boolean-valued 'probabilistic' sets.[9] This issue is brought out exceptionally clearly when we consider generic embeddings. These provide us with quintessentially combinatorial kinds of reasoning; we want to see what ordinals are moved by j (and where) and what the structure of the remaining sets looks like given the existence of j. An interpretation on which we can account for the phenomenology of this thinking is thus preferable to one on which we cannot.

V.2 Boolean ultrapowers and quotient structures

There are, however, ways of modifying the Boolean-valued models to proper-class-sized two-valued structures. This is done by means of a *Boolean ultrapower* and *quotient structures*. We take it as read that $\mathbb{B}$ is the completion of some partial order $\mathbb{P}$, and henceforth drop the notation $\mathbb{B}(\mathbb{P})$.

Our target will be the following theorem:

> **Theorem 50.** [Hamkins and Seabold, 2012] (*The Naturalist Account of Forcing*). If V is the universe of set theory and $\mathbb{B}$ is a notion of forcing, then there is in V a definable class model of the theory expressing what it means to be a forcing extension of V. Specifically, in the forcing language with $\in$, constant symbols $\check{x}$ for every $x \in V$, a predicate symbol $\check{V}$ to represent V as a ground model, and a constant symbol $\mathring{G}$, the theory asserts:
>
> (1) The full elementary diagram of V, relativised to the predicate $\check{V}$,

[9] One can informally think of a Boolean-valued model $V^{\mathbb{B}}$ as assigning 'probabilities' from $\mathbb{B}$ to membership and equality.

using the constant symbols for elements of V.

(2) The assertion that $\check{V}$ is a transitive proper class in the (new) universe.

(3) The assertion that $\mathring{G}$ is a $\check{V}$-generic ultrafilter on $\check{\mathbb{B}}$.

(4) The assertion that the new universe is $\check{V}[G]$, and **ZFC** holds there.

Immediately, we need to identify a salient point before we proceed to explain the Naturalist Account of Forcing in more detail. Since V cannot have access to its own elementary diagram, really the above is a theorem scheme. In fact, when we examine the exact semantic content of the Naturalist Account of Forcing, it says the following:

Theorem 51. [Hamkins and Seabold, 2012] For any notion of forcing $\mathbb{B}$, a complete Boolean algebra, the set-theoretic universe V has an elementary extension to a structure $(\ddot{V}, \ddot{E})$, a definable class in V, for which there is in V a $\ddot{V}$-generic filter G for $\ddot{\mathbb{B}}$ (the image of $\mathbb{B}$).

$V \precsim \ddot{V} \subseteq \ddot{V}[G]$

In particular, the entire extension $\ddot{V}[G]$ and embedding is a definable class in V.

Important in the proof of this theorem is the relationship between two different structures given an ultrafilter U on $\mathbb{B}$, namely the *Boolean ultrapower* $\check{V}_U$ and the *quotient structure* $V^{\mathbb{B}}/U$. We tackle these in reverse order.

First, we define a class predicate $\check{V}$:

Definition 52. $\check{V} =_{df} \{\langle \check{x}, \mathbb{1}_{\mathbb{B}}\rangle | x \in V\}$

We then take an ultrafilter U on $\mathbb{B}$ and use it to define two dyadic predicates on the $\mathbb{B}$-names:

Definition 53. $\tau \equiv_U \sigma$ iff $[\![\tau = \sigma]\!] \in U$

Definition 54. $\tau E_U \sigma$ iff $[\![\tau \in \sigma]\!] \in U$

and also define a monadic predicate for the ground model:

Definition 55. $\tau E_U \check{V}_U$ iff $[\![\tau \in \check{V}]\!] \in U$

Next, for every $\mathbb{B}$-name τ, we define the *restricted equivalence class* of τ under U:

Definition 56. $[\tau]_U =_{df} \{\sigma | \sigma$ is of minimal rank[10] and $\tau =_U \sigma\}$

We are then able to define the *quotient structure* of $V^{\mathbb{B}}$ as follows:

Definition 57. $V^{\mathbb{B}}/U =_{df} \{[\tau]_U | \tau \in V^{\mathbb{B}}\}$

One can then verify that $V^{\mathbb{B}}/U \models$ **ZFC** and also that if ϕ has Boolean-value greater than $\mathbb{0}_{\mathbb{B}}$ in $V^{\mathbb{B}}$, then $V^{\mathbb{B}}/U \models \phi$. Importantly, there is no need for the ultrafilter to be V-generic[11], and hence U can perfectly well be in V. Especially interesting is the relationship that this quotient structure has to a particular ultrapower map on V. We define:

Definition 58. The *Boolean ultrapower of V by U* is the following class:

$\check{V}_U =_{df} \{[\tau]_U | [\![\tau \in \check{V}]\!] \in U\}$.

and comes with an associated embedding (known as the *Boolean ultrapower embedding*):

$j_U : x \mapsto [\check{x}]_U$

Theorem 51 is proved by letting $\ddot{V} = (\check{V}_U, E_U)$. All of $\check{V}_U$, $V^{\mathbb{B}}/U$, and j_U are definable in V from the parameter U. The key fact for the purposes of interpreting forcing, however, is that $V^{\mathbb{B}}/U$ is precisely the forcing extension of $\check{V}_U$ by the filter $[\mathring{G}]_U$,[12] which is $j_U(\mathbb{B})$-generic over $\check{V}_U$.

Here, we map V to a *subclass* of itself (much as we do with a measurable cardinal embedding). Since $\check{V}_U$ is not the whole of V when U is in V (and hence not V-generic), we have plenty of sets available to be our generic for $\check{V}_U$. The fact that we can interpret a forcing construction as the quotient $V^{\mathbb{B}}/U$ shows that we have a great deal of traction between the two structures through a forcing construction.

[10]This condition is necessary to keep the equivalence classes as sets rather than proper classes and is analogous to the use of Scott's trick in the standard ultrapower construction.

[11]See, [Hamkins and Seabold, 2012] for details. The presentation of the Boolean ultrapower, quotient structures, and Naturalist Account of Forcing is also derived from the same paper.

[12]$\mathring{G} =_{df} \{\langle \check{p}, p\rangle | p \in \mathbb{B}\}$. It is a $\mathbb{B}$-name that always tracks the generic filter, whichever one we choose.

Our interpretation of forcing might then be as follows. We note that $\check{V}_U$, whilst not isomorphic to V, nonetheless looks a *lot* like V; it is a proper-class-sized elementary extension of V. Instead of using the Boolean-valued model $V^{\mathbb{B}}$ (with its attendant difficulties regarding classicality and combinatorial properties), we could examine the structures $\check{V}_U$ and $V^{\mathbb{B}}/U$ which both behave in a combinatorial and two-valued manner.

There are, however, several problems with this approach. We should first note that the use of Boolean ultrapowers and quotient structures does nothing to assuage the worries of class forcing or discussion of sharps. No extra light is shed on proper-class-sized Boolean completions by these structures, and non-forcing extensions remain out of reach. We have, however, moved to an interpretation on which the reasoning is both classical and combinatorial.

However, this can often come at the price of well-foundedness. Observe that if the Boolean ultrapower map is to be well-founded, it must elementarily embed V into an inner model thereof, and hence must have a critical point κ that is measurable. More precisely:

Theorem 59. [Hamkins and Seabold, 2012] If U is an ultrafilter in V on the complete Boolean algebra $\mathbb{B}$, then the following are equivalent:

(1) $\check{V}_U$ is well-founded.

(2) $\check{V}_U$ is an ω-model (i.e. has the standard natural numbers).

(3) U meets all countable maximal antichains of $\mathbb{B}$ in V.

(4) U is countably complete over V (i.e. if $\langle a_n | n < \omega \rangle \in V$ is an ω-sequence of $a_n \in U$, then $\wedge_n a_n \in U$).

(5) U is weakly countably complete over V (i.e. if $\langle a_n | n < \omega \rangle \in V$ is an ω-sequence of $a_n \in U$, then $\wedge_n a_n \neq \emptyset$).

By (4), for the Boolean ultrapower to remain well-founded, we need significant large cardinal properties attaching to the completeness of the ultrafilter U.[13] Recall

[13]The existence of a countably complete non-principal ultrafilter is equivalent to the existence of a measurable cardinal.

that many of the generic embeddings we wished to talk about had very small critical points. We are now in a position to prove the (very quick):

Theorem 60. Let U be an ultrafilter on a complete Boolean algebra $\mathbb{B} \in V$, and assume that $\check{V}_U$ is well-founded with j_U such that $crit(j_U) = \kappa$. Then $\check{V}_U$ cannot be used to interpret forcing constructions that change the structure of sets below V_κ.

Proof. Since κ is the critical point of j_U, we know that j_U preserves V_κ (if it did not, this would imply a different critical point below κ). ∎

The key philosophical consequence of the theorem is that if we are to interpret a forcing construction that involves the structure of sets below a measurable cardinal, then the ultrapower cannot be well-founded. This goes especially for generic embeddings. If we wish to move ω_1 using a generic embedding, then any Boolean ultrapower $\check{V}_U$ interpreting this construction will have differences in relatively 'small' levels' (e.g. $V^{\check{V}_U}_{\omega_1+2}$ and $V^{V^{\mathbb{B}}/U}_{\omega_1+2}$ will have different sets as $\omega_1^{\check{V}_U}$ is countable in $V^{\mathbb{B}}/U$). By the previous theorem, there cannot be any difference in these levels without $\check{V}_U$ becoming non-well-founded. Thus, when interpreting a generic embedding $i : V \longrightarrow \mathfrak{M} \subseteq V[G]$ as an embedding $i' : \check{V}_U \longrightarrow \mathfrak{M} \subseteq V^{\mathbb{B}}/U$, well-foundedness often fails.

Nonetheless, we can still trace results derived from the study of these embeddings back through i' and j_U to V. However, our thinking in the embedded model will be severely limited. Since the non-well-foundedness of the models implies a high degree of non-absoluteness (the satisfaction predicate itself is not even absolute[14]), we cannot simply use many of our normal assumptions regarding the relationship of sets in V to those in $\check{V}_U$ and $V^{\mathbb{B}}/U$.

Moreover, we might question, as we did with plain Boolean-valued models, the extent to which this sort of interpretative strategy respects the phenomenological character of the reasoning of the set theorist. To all intents and purposes, when she works with a generic embedding, she seems to be reasoning about *well-founded* struc-

[14] For details, see [Hamkins and Yang, 2013]. To give an example of just how extreme the phenomenon is, one can have two models that have *the same objects* as natural numbers, but disagree about whether a particular (non-standard) n is odd or even.

tures in a combinatorial and two-valued manner. The use of Boolean ultrapowers and quotient structures gets the latter aspects of the reasoning correct, but in doing so often destroys well-foundedness.[15]

It seems then, that though the use of Boolean ultrapowers and quotient structures provides a way of modifying Boolean-valued structures into a two-valued framework, the price is high, presenting difficulties of both a technical and philosophical character.

V.3 The forcing relation

Instead of pursuing a model-theoretic strategy, we might try to capture width extensions *syntactically* by defining a relation that captures the consequences of extensions without actually committing to the existence of any models. For forcing, this can be done by defining the following relation:

> **Definition 61.** Let $\mathbb{P} = \langle P, \leq_{\mathbb{P}} \rangle$ be a forcing poset. For $p \in P$, ϕ in the forcing language[16] of $\mathbb{P}$, and names τ, θ, σ in $V^{\mathbb{P}}$, we define the relation $p \Vdash^*_{\mathbb{P}}$ recursively as follows:[17]
>
> For atomic ϕ, we define (via double recursion):
>
> (i) $p \Vdash^*_{\mathbb{P}} \tau = \theta$ iff $\forall \sigma \in dom(\tau) \cup dom(\theta) \forall q \leq_{\mathbb{P}} p[q \Vdash^*_{\mathbb{P}} \sigma \in \tau \leftrightarrow q \Vdash^*_{\mathbb{P}} \sigma \in \theta]$.
>
> (ii) $p \Vdash^*_{\mathbb{P}} \pi \in \tau$ iff $\{q \leq_{\mathbb{P}} p | \exists \langle \sigma, r \rangle \in \tau[q \leq_{\mathbb{P}} r \wedge q \Vdash^*_{\mathbb{P}} \pi = \sigma]\}$ is dense below p.
>
> $\Vdash^*_{\mathbb{P}}$ is then defined for composite formulas as follows:
>
> (iii) $p \Vdash^*_{\mathbb{P}} \phi \wedge \psi$ iff $p \Vdash^*_{\mathbb{P}} \phi$ and $p \Vdash^*_{\mathbb{P}} \psi$.
>
> (iv) $p \Vdash^*_{\mathbb{P}} \neg\phi$ iff $\neg \exists q \leq_{\mathbb{P}} p[q \Vdash^*_{\mathbb{P}} \phi]$.
>
> (v) $p \Vdash^*_{\mathbb{P}} \phi \rightarrow \psi$ iff $\neg \exists q \leq_{\mathbb{P}} p[q \Vdash^*_{\mathbb{P}} \phi \wedge q \Vdash^*_{\mathbb{P}} \neg\psi]$.

[15] Interestingly, this opens a new area of enquiry; examine the cases where forcing and large cardinals combine to keep the ultrapower well-founded. We might, for example, consider a generic embedding with a critical point above a measurable cardinal. See [Hamkins and Seabold, 2012] for discussion.

[16] The *forcing language of* $\mathbb{P}$ is the collection of all formulas that can be formed by the usual logical operators from the language $\mathscr{L}_{\in}$ combined with a constant symbol for every name in $V^{\mathbb{P}}$ (the $\mathbb{P}$-names).

[17] We could also have defined $\Vdash^*_{\mathbb{P}}$ in terms of the Boolean-valued models approach, where $p \Vdash^*_{\mathbb{P}} \phi$ iff p is below the Boolean-value of ϕ in the Boolean completion of $\mathbb{P}$.

(vi) $p \Vdash^*_{\mathbb{P}} \phi \vee \psi$ iff $\{q | [q \Vdash^*_{\mathbb{P}} \phi] \vee [q \Vdash^*_{\mathbb{P}} \psi]\}$ is dense below p.

(vii) $p \Vdash^*_{\mathbb{P}} \phi \leftrightarrow \psi$ iff $p \Vdash^*_{\mathbb{P}} \phi \rightarrow \psi$ and $p \Vdash^*_{\mathbb{P}} \psi \rightarrow \phi$.

(viii) $p \Vdash^*_{\mathbb{P}} \forall x \phi(x)$ iff $p \Vdash^*_{\mathbb{P}} \phi(\tau)$ for all $\tau \in V^{\mathbb{P}}$.

(ix) $p \Vdash^*_{\mathbb{P}} \exists x \phi(x)$ iff $\{q \leq_{\mathbb{P}} p | \exists \tau \in V^{\mathbb{P}} [q \Vdash^*_{\mathbb{P}} \phi(\tau)]\}$ is dense below p.

One can then verify:

(1) If $\phi_1, ..., \phi_n \vdash \psi$ and $p \Vdash^*_{\mathbb{P}} \phi_i$ for each i, then $p \Vdash^*_{\mathbb{P}} \psi$.

(2) $p \Vdash^*_{\mathbb{P}} \phi$ for every axiom of **ZFC**.

(3) If $\phi(x_1, ..., x_n)$ is a formula known to be absolute for transitive models, then for every p and all sets $a_1, ..a_n$; $p \Vdash^*_{\mathbb{P}} \phi(\check{a}_1, ..., \check{a}_n)$ iff $\mathbb{1}_{\mathbb{P}} \Vdash^*_{\mathbb{P}} \phi(\check{a}_1, ..., \check{a}_n)$ iff $\phi(a_1, ..., a_n)$ is true in V.

Essentially, $\Vdash^*_{\mathbb{P}}$ lets us talk about what would be satisfied in the extension $V[G]$ by analysing what sentences conditions $p \in P$ force. In particular, if we can show that there is a $p \in P$ such that $p \Vdash^*_{\mathbb{P}} \phi$, we can behave and talk as if such a forcing extension exists. By (3), any theorem proved 'in $V[G]$' will be verified by the check names and hence by specific sets in V. Similarly, if we wish to formulate an axiom about V using a forcing extension, we can do so by finding a p that forces the required sentence about objects in the ideal extension.

Again, the use of the forcing relation is absolutely fine for relative consistency proofs. We know that if we can find a $p \in \mathbb{P}$ such that $p \Vdash^*_{\mathbb{P}} \phi$, then we cannot prove $\neg\phi$ (assuming the consistency of **ZFC**). Indeed, the technique avoids many of the problems suffered by the use of quotient structures and the Boolean ultrapower. Since the check names all determinately end up referring to their respective sets, there is no danger of the ideal extension appearing non-well-founded from the perspective of V (as was the case with $\check{V}_U$ and $V^{\mathbb{B}}/U$). Moreover, the reasoning when using the forcing relation is *classical*—there is no mention of probabilistic sets anywhere.[18]

[18]Of course, one can formulate $\Vdash^*_{\mathbb{P}}$ in the Boolean-valued approach as well. From a mathematical perspective, the two approaches are equivalent; one can easily go back and forth. Conceptually, however, they are very different, and the extent to which each corresponds to the thinking of set theorists is thus a philosophically interesting matter.

We will, however, identify several problems with this approach to interpreting width-extension talk. First, even if the use of the forcing relation worked perfectly for those mathematical constructions it was designed to interpret, there are problems of scope. For example, the use of $\Vdash^*_{\mathbb{P}}$ will not help in trying to interpret the use of non-forcing width extensions such as sharps. More seriously, however, even within forcing there are limitations to the use of the forcing relation.

The difficulty lies in the fact that the forcing relation need not be definable when the forcing poset in question is *proper-class-sized*. For example, consider the following forcing:

Definition 62. Let $\mathfrak{M}$ be a model for **ZFC**. Then the *Friedman* poset (denoted by '$\mathbb{F}^{\mathfrak{M}}$') is a partial order of conditions $p = \langle d_p, e_p, f_p \rangle$ such that:

(i) d_p is a finite subset of ω.

(ii) e_p is a binary acyclic relation relation on d_p.

(iii) f_p is an injective function with $dom(f_p) \in \{\emptyset, d_p\}$ and $ran(f_p) \subseteq \mathfrak{M}$.

(iv) If $dom(f_p) = d_p$ and $i, j \in d_p$, then ie_pj iff $f_p(i) \in f_p(j)$.

(v) The ordering on $\mathbb{F}^{\mathfrak{M}}$ is given by:

$$p \leq_{\mathbb{F}^{\mathfrak{M}}} q \leftrightarrow d_q \subseteq d_p \wedge e_p \cap (d_q \times d_q) = e_q \wedge f_q \subseteq f_p.$$

This defines a proper-class-sized partial order as the individual f_p include every function from some finite subset of ω to a (sub)set of $\mathfrak{M}$, and hence there are proper-class-many such ordered triples (relative to $\mathfrak{M}$). The partial order adds a bijection $F_{\mathbb{F}}$ between ω and M, and a relation $E_{\mathbb{F}} \in \mathfrak{M}[G]$ such that $\langle \omega, E_{\mathbb{F}} \rangle$ and $\langle M, \in \rangle$ are isomorphic. If the forcing relation for $\mathbb{F}$ were definable, $\mathfrak{M}$ would then have access to its own truth definition (contradicting Tarski's Theorem).[19] Thus we have:

Theorem 63. [Holy et al., F] (attributed to Friedman) $\Vdash^*_{\mathbb{F}}$ is not uniformly definable for $\mathbb{F}$.

[19] For the details of the proof, and further discussion of the Truth and Definability lemmas in context of class forcing, see [Holy et al., F].

Hence, there are forcings for which there is no definition of the forcing relation in the ground model. This is true, despite the fact that $\mathbb{F}^{\mathfrak{M}}$ itself is definable over $\mathfrak{M}$. If we wish to use $\mathbb{F}^V$ in proving facts about V then, we cannot do so through consideration of a forcing relation in V. Since tameness implies pretameness[20], which in turn implies that the forcing relation is definable, we know that any extension using $\mathbb{F}$ must violate **ZFC**.[21] One may feel that this provides a response: we should not consider such forcing arguments as legitimate for proving facts about V because the resulting 'model' we are trying to talk about is pathological—it is non-**ZFC**-preserving.

Interestingly, the Universist should have little truck with such a claim. It is true that the resulting extensions are pathological in violating **ZFC**. However, we should note that (from the Universist's perspective) the whole enterprise with which we are engaged (namely considering extensions of V) is pathological. We are trying to code in V the effects of objects that do not, strictly speaking, exist. Thus, there seems to be no objection to considering models where, say, there is a bijection between ω and V (as is the case when forcing using $\mathbb{F}^V$). If consideration of such pathologies has interesting consequences within V then, given that we are already flirting with *prima facie* incoherent concepts, there seems little reason to prohibit their examination.

Moreover, the use of $\Vdash^*_{\mathbb{P}}$ still does not accord with the phenomenological character of the set theorist's reasoning. They wish to reason about sets which can be combinatorially manipulated, embeddings which move ordinals, and so forth. Here, however, the reasoning is fully syntactic; we analyse which formulas particular $p \in \mathbb{P}$ force, and so are not explicitly working with sets in the above manner. In this way, though we retain the well-foundedness and denotation of terms in V (through the

[20]A partially ordered class $\mathbb{P}$ is *pretame* iff whenever $\langle D_i | i \in a \rangle$ is a $\langle \mathfrak{M}, A \rangle$-definable (where $\langle \mathfrak{M}, A \rangle$-*definability* is definability over the ground model augmented with the predicate A, as in Ch I) sequence of dense classes, $a \in M$, and $p \in \mathbb{P}$, then there is $q \leq_{\mathbb{P}} p$ and $\langle d_i | i \in a \rangle \in M$ such that $d_i \subseteq D_i$ and d_i is predense below q for each i (where d_i is *predense* below q iff $r \leq_{\mathbb{P}} q$ implies that r is compatible with an element of d_i). To define *tameness* we first need the definition of a *predense below p partition*: a pair $\langle D_0, D_1 \rangle$ such that (i) $D_0 \cup D_1$ is predense below p and (ii) if $p_0 \in D_0$ and $p_1 \in D_1$ then p_0 and p_1 are incompatible. Two sequences of predense below p partitions $\langle \langle D^i_0, D^i_1 \rangle | i \in a \rangle$ and $\langle \langle E^i_0, E^i_1 \rangle | i \in a \rangle$ are *equivalent below p* iff for each $i \in a$, $\{q | q \text{ meets } D^i_0 \leftrightarrow q \text{ meets } E^i_0\}$ is dense below p. Then $\mathbb{P}$ is *tame* iff (i) $\mathbb{P}$ is pretame and (ii) for each $a \in M$ and $p \in \mathbb{P}$ there is a $q \leq_{\mathbb{P}} p$ and α in the ordinals of $\mathfrak{M}$ such that whenever $\vec{D} = \langle \langle D^i_0, D^i_1 \rangle | i \in a \rangle \in M$ is a sequence of predense below q partitions, $\{r | \vec{D}$ is equivalent below r to some $\vec{E} = \langle \langle E^i_0, E^i_1 \rangle | i \in a \rangle \in V^{\mathfrak{M}}_\alpha\}$ is dense below q. Pretameness implies the preservation of Replacement in a class forcing, and tameness additionally requires that the forcing preserve the Power Set Axiom. For details, see [Friedman, 2000].

[21]See [Holy et al., F].

relationship the check names bear to their objects), we lose the character of the reasoning that the Boolean ultrapower and quotient structure represented well.

V.4 Countable transitive models

A model-theoretic approach is thus desirable. The countable transitive model strategy comes in several forms. Initially, the method was designed to deal with the apparently problematic metamathematics of forcing from *within* **ZFC**. If we do not wish to strengthen **ZFC** substantially, there are two main ways of executing the strategy:[22]

Suppose that we wish to prove that some statement ϕ is independent from **ZFC**. We then suppose that ϕ (or $\neg\phi$) has a proof in **ZFC** (from now on we only consider the case where we wish to show that $\neg\phi$ is unprovable). If $\neg\phi$ were provable, we would have $\mathbf{ZFC} + \phi \vdash 0 = 1$. Since proofs are finite, we then know that this proof would only use a finite set of axioms of **ZFC** (let it be denoted by 'Γ'). We then know (by the forcing method) that given such a Γ, there is a larger finite set of axioms of **ZFC** (let it be denoted by 'Λ') such that **ZFC** proves the conditional:

> "If there is countable transitive model for Λ, then there is a countable transitive model for $\Gamma + \phi$."

However, now we can use the Reflection Theorem, Löwenheim-Skolem Theorem, and Mostowski Collapse Lemma to then obtain a countable transitive model $\mathfrak{N}$ for Λ, and hence have a countable transitive model for $\Gamma + \phi$, contradicting our supposition that $\Gamma + \phi \vdash 0 = 1$.[23]

Alternatively, we could expand $\mathscr{L}_\in$ to $\mathscr{L}_{\in,C,F}$ by two constant symbols C and F. We then add axioms to **ZFC** as follows:[24]

Definition 64. $\mathbf{ZFC}^*$ is a system in $\mathscr{L}_{\in,C,F}$ with the following axioms:

1. **ZFC**

[22]See [Kunen, 2013], IV.5.1 for details.

[23]Observe here that the choice of $0 = 1$ was not essential, the same argument applies to any arithmetical sentence (i.e. If $\mathbf{ZFC} + \chi \vdash \psi$, where χ is independent from **ZFC** and ψ is an arithmetical sentence, then $\mathbf{ZFC} \vdash \psi$).

[24]For this specific approach, see [Shoenfield, 1967] as well as [Kunen, 2013].

2. C is a transitive set.

3. F is a bijection from ω onto C.

4. ϕ^C for every axiom ϕ of **ZFC** (note that, by Gödel's Second Incompleteness Theorem, this is an axiom scheme).

We know (by the Reflection Theorem, Löwenheim-Skolem Theorem, and Mostowski Collapse Lemma) that **ZFC*** is a conservative extension of **ZFC**. We can then treat C as our countable transitive model, and conduct our construction there.

Of course, for any countable transitive model $\mathfrak{M}$, it is entirely possible to take a width or height extension (on the assumption that V is uncountable and satisfies Replacement). Since even the smallest uncountable set will contain subsets of objects not in $\mathfrak{M}$, there is no issue with $\mathfrak{M}$ having a width extension—we move from one particular countable transitive model to another by adding subsets. For height extensions, since any countable transitive model $\mathfrak{M}$ is such that $Ord(\mathfrak{M}) < \omega_1$, a height extension to another countable transitive model is always available (in fact, there are uncountably many such height extensions).

Such methods are fine as far as they go. For the purpose of allowing us to analyse forcing arguments establishing relative consistency proofs from within **ZFC** the technique performs adequately; any inconsistency of ϕ with **ZFC** could be traced to a countable model, and the relevant forcing argument shows that there is no such inconsistency (on the assumption that **ZFC** itself is consistent). We want more than mere consistency results, however. For, while V presumably satisfies **ZFC** for the Universist, there is no guarantee that a given countable transitive model of (a fragment of) **ZFC** represents V with respect to first-order truth *in general*. While we can, for any appropriate given set of assumptions Γ, assume that there is a countable transitive model satisfying Γ, V might satisfy sentences independent from or in conflict with Γ. In order then to mimic the behaviour of V as closely as possible, we would like a countable transitive model $\mathfrak{M}$ that resembles V sufficiently well to allow us to represent the consequences of extensions of V in $\mathfrak{M}$ (and hence in V via the resemblance).

One step in the right direction is to assume the existence of a countable transitive model of **ZFC**. The assumption implies the consistency of **ZFC**, $\mathbf{ZFC} + Con(\mathbf{ZFC})$, $\mathbf{ZFC} + Con(Con(\mathbf{ZFC}))$, and so on, for any finite iteration of the consistency sentence for the previous theory. On the assumption that V satisfies **ZFC**, if ϕ is a first-order consequence shown to hold in a countable transitive model $\mathfrak{M}$ on the basis of $\mathfrak{M}$'s **ZFC** satisfaction by taking an extension, then V must also satisfy ϕ. All that was required was that $\mathfrak{M} \models \mathbf{ZFC}$. Since V also satisfies **ZFC**, we know that ϕ holds there.

However, what we would really like is a countable transitive model $\mathfrak{M}$ that satisfies *exactly the same* parameter-free first-order sentences as V. Cohen constructs an informal argument for the existence of such an $\mathfrak{M}$ that can be adapted to fit this end:

> "The Löwenheim-Skolem theorem allows us to pass to countable sub-models of a given model. Now, the "universe" does not form a set and so we cannot, in **ZF**, prove the existence of a countable sub-model. However, informally we can repeat the proof of the theorem. We recall that the proof merely consisted of choosing successively sets which satisfied certain properties, if such a set existed. In **ZF** we can do this process finitely often. There is no reason to believe that in the real world this process cannot be done countably many times and thus yield a countable standard model for **ZF**." ([Cohen, 1966], p79)

While Cohen is primarily interested in the existence of a countable transitive model for **ZF**, we can import his argument to the case of V as follows. For a finite subset Γ of sentences in $\mathscr{L}_{\in}$ satisfied by V we move to a countable transitive model of Γ by using the Reflection Theorem to find a $V_\alpha \models \Gamma$. We then use AC to find a set of functions F^{V_α} such that for every existential statement $\exists x \phi(x)$ true in V_α, there is an $f_\phi \in F$ that picks a single witness a such that $\phi(a)$ holds. We then form a countable $\mathfrak{M} \models \Gamma$. Since the relation (i.e. $\in$) on V_α is extensional, well-founded, and set-like, so is the relation on $\mathfrak{M}$, and we then use the Mostowski Collapse Lemma to collapse to obtain a countable transitive model $\mathfrak{M}' \models \Gamma$. Turning now to V, we simply note that V is one model of **ZFC** among many. Thus, we can posit the existence of a set of Skolem functions F^V for V (by 'choosing[25]' a witness for the countably

[25] As Cohen notes, this is not possible in $\mathbf{ZF}(\mathbf{C})$ by Gödel's Second Incompleteness Theorem.

many existential statements satisfied by V with a countable sequence of choices). Then, by Skolemising and Collapsing, we obtain a countable transitive model $\mathfrak{V}$ that satisfies exactly the same parameter-free first-order sentences as V.

The main issue here is that, by Tarski's Theorem on the undefinability of truth, that while V can see F^V, it does not know that F^V provides its own set of Skolem-functions. A way around this, very similar to the second countable transitive model approach, was noticed by [Feferman, 1969]:

Definition 65. Let $\mathscr{L}_{\in,\mathfrak{V}}$ be the language $\mathscr{L}_{\in}$ augmented with a single constant symbol $\mathfrak{V}$. $\mathbf{ZFC}^{\mathfrak{V}}$ is then a theory in $\mathscr{L}_{\in,\mathfrak{V}}$ with the following axioms:

(i) **ZFC**

(ii) $\mathfrak{V}$ is countable and transitive.

(iii) For every ϕ in $\mathscr{L}_{\in}$, $\phi \leftrightarrow \phi^{\mathfrak{V}}$ (by Tarski's Theorem, this is an axiom scheme).

Again, we know by the Reflection, Löwenheim-Skolem, and Mostowski Collapse Theorems that $\mathbf{ZFC}^{\mathfrak{V}}$ is a conservative extension of **ZFC**. However, $\mathfrak{V}$ then satisfies *exactly* the same parameter-free first-order sentences of **ZFC** as V. Before we proceed, we make the following:

Remark 66. One can accomplish the same effect with a different trick. If we introduce a truth predicate for V into $\mathscr{L}_{\in}$, add the Tarski T-axioms, and then permit the use of the T-predicate in the Axiom of Replacement, we can perform the usual Reflection, Skolemisation, and Collapse construction, yielding a countable transitive elementary substructure of V.[26]

We can then interpret any extension as concerned with $\mathfrak{V}$. For example, if we wish to formulate an axiom that uses an extension (say a generic embedding), we simply formulate it as concerned with $\mathfrak{V}$ (where extensions are uncontroversially available) and then know that any first-order consequence of the axiom true in $\mathfrak{V}$ is also true of V. Of course, it is a separate question whether or not the objects relevant

[26]Thanks to Sam Roberts for pointing out to me this method.

for construction of the extension exist in V (such as, for example, a saturated ideal I required to facilitate a generic embedding). Nonetheless, if such an object I does exist in V, we will have a corresponding object $I' \in \mathfrak{V}$, and then there *will* be the required embedding $i' : \mathfrak{V} \longrightarrow \mathfrak{M} \subseteq \mathfrak{V}[G]$.

Moreover, such a strategy is not limited to only forcing constructions. Given $\mathfrak{V}$, we can ask whether not it is plausible that $\mathfrak{V}$ is sharp generated. If there are good arguments for this claim, then there can perfectly well be such a sharp for $\mathfrak{V}$ in V. Further, any parameter-free first-order consequences of the sharp generation of $\mathfrak{V}$ in V will be mirrored in the relevant structural features of V.

There are also a number of pleasing philosophical features of the countable transitive model strategy. First, unlike many of the previous methods considered, it allows a very natural interpretation of extensions. Many of our naive ways of thinking about extensions turn out to be represented; $\mathfrak{V}$ is always transitive, well-founded, and the reasoning is combinatorial and classical. The only exception to this is that the reasoning concerns *countable* sets in $\mathfrak{V}$, rather than *uncountable* sets in V. However, we can still move freely between $\mathfrak{V}$ and V, and it is also questionable how much of the set theorist's reasoning depends on the *literal* uncountability of the objects with which they are reasoning. We might think that all that her phenomenology requires is that the objects are uncountable relative to the structure with which she works (namely $\mathfrak{V}$).

In the present context, however, there is a limitation of the countable transitive model strategy. As it stands, $\mathfrak{V}$ is only accurate for *first-order* statements about V. Because of the inherent incompleteness in second-order properties over V, $\mathfrak{V}$ does not perfectly mirror V's second-order properties.

This is mildly problematic for the countable transitive model strategy. For some of the axioms that we wish to analyse are slightly greater than first-order over V. The IMH, for example, makes a claim about the density of inner models of V. However, the existence of an inner model is a second-order property; it involves the quantification over a proper class. Thus, there is no guarantee that if $\mathfrak{V}$ satisfies the IMH, then V also has the corresponding inner models (and vice versa).

This is especially interesting given that a large part of set theory comprises the structural relationship between models. Simply because an axiom is not first-order is not a reason (without significant further argument) to establish that it is not of independent interest. What we would therefore like is not only a response to the Hilbertian Challenge that satisfies the Ontological and Foundational Constraints, but one that provides a way of interpreting greater than first-order properties of V expressed through the use of extensions.

V.5 The Methodological Constraint

The countable transitive model strategy did not quite deliver everything we wanted for interpreting extensions of V, but it certainly came frustratingly close. Of particular interest was that it accounted very *naturally* for extensions of V; our naive thinking concerning extensions was easily captured. We were then able to export many results about $\mathfrak{V}$ back up to V. The only reason for rejecting the strategy as not fully satisfactory is an issue of scope: there is no guarantee that $\mathfrak{V}$ mirrors greater than first-order properties of V.

This relative success of the countable transitive model strategy contrasts with some of the failures of the other interpretations. The use of Boolean-valued models did not capture the combinatorial and two-valued nature of the set theorist's reasoning. Quotient structures and Boolean ultrapowers were often non-well-founded where the reasoning appeared to involve well-founded structures. The use of the forcing relation was overly syntactic, and failed to account for the combinatorial nature of thinking. This suggests the following constraint on responses to the Hilbertian Challenge:

> **The Methodological Constraint.** In responding to the Hilbertian Challenge, do so in a way that accounts for our naive thinking about extensions and links them to structural features of V.

The satisfaction of such a constraint would ensure not just that we respond to the Hilbertian Challenge, but that the solution proposed meshes well with the way set theorists go about their daily practice. However, we should pause to reflect on the constraints set up in the previous Chapter:

The Foundational Constraint. In responding to the Hilbertian Challenge, do so in a way that does not necessitate the use of resources that cannot be represented by sets within V.

The Ontological Constraint. Any interpretation of extra-V resources should make clear the *ontological difference* between the interpretation of extensions and normal sets within V. In other words, any interpretation must make clear in what sense the interpretation does not *literally* refer to extra-V sets.

We are immediately faced with an apparent tension between these three constraints. On the one hand, the Foundational Constraint demands that we keep all discourse strictly regimented within V. On the other hand, the Ontological Constraint requires that our interpretation of extra-V resources makes it clear how 'extensions' (suitably interpreted) are different from sets. On a third hand[27], the Methodological Constraint obliges us to take account of our naive reasoning of extensions. How can we possibly provide an interpretation of extensions that (i) makes it clear how our interpretation of extensions differs from garden-variety sets, (ii) in doing so uses only sets from V, whilst (iii) capturing our naive talk and relating it to the structure of V? This seems like an insurmountable task. In the next two chapters we show that it is possible to climb the mountain of conflict and provide such an interpretation.

Chapter V: Conclusions.

In this chapter, we saw that there are various different ways we might interpret width extensions of V. All have problems of scope; Boolean-valued models and Boolean ultrapowers have difficulties with class forcing, the forcing relation cannot interpret non-**ZFC**-preserving class forcings, and both these methods cannot account for non-forcing extensions of V. Even the countable transitive model strategy (which performed far better), could not necessarily mirror greater than first-order properties of V. Moreover, each view had problems accounting for the naive yet

[27]Normally, we do not speak as if we have more than two hands. Since we are currently engaged in the philosophy of set theory, we set aside this ontological excess.

fruitful way the set theorist often reasons with extensions (the only exception here being the countable transitive model strategy). These difficulties highlight an additional methodological constraint on any resolution of the Hilbertian Challenge. However, there is a problem: the three constraints outlined in the thesis appear to be in tension with one another. The reader may feel that in imposing them we are trying to have our cake and eat it. The project of the final two chapters will be to show that we can.

Chapter VI

A Theory of Classes

Thus concludes the negative part of the thesis. We find ourselves in something of a pickle; we want to respond to Hilbert's Challenge, but are unsure of how to do so whilst satisfying the three constraints (and dissolving the apparent tension therein). We shall now embark upon the first of two steps in resolving Hilbert's Challenge; motivating a strong class theory that provides the foundational resources to support the technical results of Chapter VII.

The chapter is structured as follows. After these brief introductory remarks, §1 revisits the issue of proper classes for the Universist and examines a distinction between 'first' and 'second' philosophy. §2 then explains one way of interpreting class talk, namely by viewing discourse involving proper classes as shorthand for the satisfaction of a first-order definable formula. In §3 and §4, we argue that there are reasons to accept the legitimacy of *non*-first-order definable class talk for a Universist. §3 does so using a philosophy first methodology, and §4 provides second philosophical arguments. §5 considers two possible ways of interpreting non-first-order definable class talk: either (a) proper classes are ontologically robust objects distinct from sets, or (b) proper class talk should be understood via the use of plural reference and quantification. We argue that (b) performs far better with respect to the three constraints. §6 then provides a justification of **MK** on the basis of the plural interpretation of proper class discourse. We conclude that **MK** is reasonably well-motivated on her picture.

VI.1 Proper classes, first philosophy, and second philosophy

Let us remind ourselves of some earlier mathematical and philosophical features of Universism. In Chapter I, we noted that there are conditions that do not define sets. For many conditions $\phi(x)$, though there is a fact of the matter for any set x whether or not $\phi(x)$ holds, there is no set of all ϕ. Good examples here were "x is an ordinal", "x is a set", and (equivalently) "x is non-self-membered". It was also noted that the Iterative Conception of Set provides an explanation of why proper classes do not form sets: their elements are not all available for collection in some V_α.

For this reason, a natural position for the Universist to hold is that class talk should be completely expunged from set-theoretic discourse. We will argue against this view, and in fact justify the use of Morse-Kelley class theory in formalising our class-theoretic reasoning.

First, however, we explain the kinds of reasons we shall accept for justifying particular mathematical theories of classes. We have in mind a famous distinction between *first* and *second* philosophy. The distinction, especially with respect to set theory, has its roots in [Quine, 1969a], and has been subsequently refined and developed in the work of Maddy[1].

The term 'first philosophy' in the sense we shall use here, emerged in the work of Descartes.[2]. Under one interpretation, Descartes' main focus was to provide a justification for his physics on the basis of more secure principles. The following is a typical passage at the very start of the First Meditation:

> "Some years ago I was struck by the large number of falsehoods that that I had accepted as true in my childhood, and by the highly doubtful nature of the whole edifice that I had subsequently based on them. I realized that it was necessary, once in the course of my life, to demolish everything completely and start again right from the foundations if I wanted to establish anything at all in the sciences that was stable and

[1] See [Maddy, 1997], [Maddy, 2007], and [Maddy, 2011].
[2] Especially [Descartes, 1637] and [Descartes, 1641].

likely to last." ([Descartes, 1641], p17)

Descartes then doubts his beliefs (via progressively stronger assumptions, culminating in the supposition that an all-powerful deceiving demon exists) in order to build his philosophy (and hence his physics) on a fully secure foundation. Now we might doubt whether or not Descartes is successful in his aims, or whether the methodology of reasoning in the face of such extreme scepticism is well-advised.[3] However, we can characterise the 'philosophy first' approach as one on which we, from a given bank of accepted truths, reason about what should hold without further scientific enquiry. Such a methodology has had scorn heaped upon it by some scholars in the philosophy of mathematics. For example, Quine claims:

> "The old tendency[4] was due to the drive to base science on something firmer and prior in the subject's experience; but we dropped that project." ([Quine, 1969a], p87)

Maddy extends this feeling to the pejorative:

> "Philosophy undertaken in such complete isolation from science and common sense is often called 'First Philosophy"' ([Maddy, 2011], p41)

Certainly, first philosophy understood in Descartes' radical fashion is a difficult methodology to follow. However, we might wonder if legitimate justifications can be provided in a first philosophical spirit. More precisely, can we provide reasons to justify the use of particular mathematical resources in virtue of what we think those resources concern (rather than analysing the mathematical fruits borne by the relevant theories)?

A first philosophical methodology is thus one which examines what holds without especially mathematical ends. When considering a subject matter, we examine what we think should hold given the nature of the subject matter in question. If we find, on philosophical grounds, a certain part of mathematics to be conceptually

[3]See [Broughton, 2002] for analysis.

[4]The 'old tendency' of which Quine speaks is to associate observation sentences with a sensory subject matter. In the context of [Quine, 1969a], the principle forms an integral part of epistemology 'conducted from the armchair'; justifying epistemological claims based on reasoning from the nature of experience, rather than through scientific observation.

bankrupt, then we reject that part of mathematics. Good examples of first philosophical mathematics include Constructivist rejection of certain parts of set-theoretic mathematics based on the view that mathematical objects are mental constructions, or intrinsic arguments for small large cardinals on the basis of reflection principles.

Second philosophy, on the other hand, concerns the *ends* of mathematics:

> "We've seen that sets were posited in the first place in the service of explicit mathematical goals...In broad overview, these goals range from relatively local problem-solving, to providing foundations, to more open-ended pursuit of promising mathematical avenues...Given what set theory is intended to do, relying on considerations of these sorts is a perfectly rational way to proceed: embrace effective means toward desired mathematical ends." ([Maddy, 2011], p52)

Thus second philosophy is concerned with what is required for us to use our mathematical theories rather than the nature of the subject matter as we understand it (considered independently from specific mathematical theorems). While the First Philosopher does not concern himself with the possible mathematical applications of the particular theory we take to be up for justification, the Second Philosopher takes such ends as the main source of justification.

At times, Maddy's Second Philosopher seems to have *some* sympathy for our First Philosopher's methodology:

> "To round out this quick portrait, consider the contrast with philosophy understood as starting either before science begins or after all scientific evidence is in, that is, philosophy as an entirely independent enterprise. Notice that if such a philosophical undertaking intends to correct science, or even to justify it in some way, then it isn't effectively separated from our inquirer's sphere of interest: working without any litmus test for 'science' or 'non-science', she will view it as a potential part of her own project, out to revise or buttress her methods" ([Maddy, 2011], p40)

Important here is the fact that the distinctions between science and philosophy, and first and second philosophy need not be sharp. A second philosopher might show interest in first philosophical methods, insofar as such an enterprise informs scientific and mathematical practice. While there are arguments that are neither clearly first nor second philosophical, Maddy holds that it is more fruitful to work under the more second philosophical end of the spectrum.

Maddy *does* argue, however, that metaphysical enquiry is largely irrelevant for mathematical development. Considering the particular case of impredicative definition[5] she writes:

> "It's often suggested that the answer lies in metaphysics, in the nature of the abstract subject matter of mathematics...[a view] might hold that mathematical entities exist only insofar as they are defined, and thus that impredicative definitions—which define an object in terms of a collection to which that very object belongs—should not be allowed...The Second Philosopher notes that the controversy was eventually resolved in favor of allowing impredicative definitions and that ontological debates over the existence and nature of sets remain unresolved to this day. This strongly suggests that metaphysical agreement did not underlie this methodological outcome." ([Maddy, 2007], pp347-348)

and more generally:

> "After uncovering corresponding methodological argumentation in a range of cases, the Second Philosopher concludes that though metaphysical theories on the nature of mathematical truth and existence undeniably do turn up in such debates, they are not in fact decisive, they are in fact distractions from the underlying purely mathematical considerations at work." ([Maddy, 2007] p349)

Thus, Maddy regards metaphysical engagement on non-mathematical grounds as a largely redundant enterprise. According to Maddy, mathematicians will study the mathematics they want, without taking notice of metaphysical scruples or the

[5] A definition is *impredicative* iff it quantifies over that which is being defined. A natural example is that x is the tallest person in the room iff x is a person in the room and for all y, if y is a person in the room and $x \neq y$ then x is taller than y. The definition is impredicative in that x itself falls under the quantifier "for all y". For exactly the same reason, the definition: x is a least upper bound of Y iff for all z, if z is an upper bound of Y and $z \neq x$ then $x < z$, is also impredicative.

kinds of objects they talk about. While we might dispute[6] Maddy's claims here, our interest is in providing justifications of our own. Now we have a characterisation of the more first philosophical and more second philosophical methods, we need not detain ourselves any further with these tricky exegetical issues. For, taking the current state of set theory for granted, what we shall argue is that there are *both* first philosophical and second philosophical reasons to accept non-definable class talk and **MK** as a class theory.

VI.2 Definable classes

If we are motivated by considering set theory as a discipline, and what philosophical views might underlie it, there is a substantial amount of talk concerning classes to be accounted for. For example, as noted earlier, one can define a measurable cardinal as the critical point of a non-trivial elementary embedding $j : V \longrightarrow \mathfrak{M}$. Here both V and $\mathfrak{M}$ are proper classes, and a natural way of coding j is as a proper class of ordered pairs. Similarly much of set-theoretic practice (as noted in the discussion of fine structure theory) involves comparing proper-class-sized models. If we wish to completely expunge discourse involving proper classes from our discourse, how might we do so?

We take it as read that we should want to provide an interpretation of talk involving proper classes, even if it is just to show that it can be completely paraphrased using first-order talk about sets. To reject certain areas of set theory as totally meaningless is simply not to engage with the debate at hand. We are interested in how a Universist interprets the set theorist's talk, unrestrictedly and without cherry-picking those aspects of the discourse that best suit the philosophical position under consideration.[7]

Instead of countenancing class talk as legitimate in its own right, we might try

[6]A full consideration of the issues is well outside of the scope of the present work. However, it bears mentioning that the extent to which one's opinions on mathematical ontology are symbiotically entangled with one's views on mathematical fruitfulness (and indeed, area of mathematical expertise) is a question deserving of further scrutiny.

[7]One might feel that this conflicts with our desire to provide first philosophical justifications. For the sake of clarity we re-emphasise the following point: We are not providing a first philosophical justification for set theory as it stands, but rather take the current state of set theory for granted, and then justify **MK** from this point of view.

to paraphrase the class talk through the use of the relevant ϕ that define the classes. Hamkins, for example, says the following:

> "One traditional approach to classes in set theory, working purely in **ZFC**, is to understand all talk of classes as a substitute for the first-order definitions that might define them..." ([Hamkins, 2012], p1873)

To take some simple examples, we can paraphrase "$x \in R$", "$x \in V$", and "$x \in On$", as "x is non-self-membered", "x is a set", and "x is an ordinal" respectively. Similarly, if we wish to state that $V = R$, we can do by stating that "$\forall x(x = x \leftrightarrow x \notin x)$".

Interestingly, we can provide first-order definitions for more complicated kinds of class. For example, let $j : V \longrightarrow \mathfrak{M}$ be an embedding witnessing the measurability of an uncountable cardinal κ. We can (using a parameter U for a κ-complete non-principal ultrafilter on κ) define a first-order formula $\phi(x, y, z)$ such that $j(x) = y$ iff $\phi(x, y, U)$ holds in V. Then, one can show:

(1) $\phi(x, y, U)$ relates every x to at most one y (i.e. $\phi(x, y, U)$ is function-like).

(2) $\phi(x, y, U)$ relates no two x to the same y (i.e. $\phi(x, y, U)$ is one-to-one).

(3) $\phi(x, y, U)$ relates every set in V to a set in $\mathfrak{M}$ (i.e. $\phi(x, y, U)$ is total on V).

(4) There is at least one x and y such that $\phi(x, y, U)$ and $x \neq y$ (i.e. $\phi(x, y, U)$ is non-trivial).

(5) For any $x_0, ..., x_n$ and $y_0, ..., y_n$ if $\phi(x_n, y_n, U)$ holds for both sequences then for any parameter-free first-order formula $\psi(z_0, ...z_n)$, $\psi(x_0, ..., x_n) \leftrightarrow \psi(y_0, ..., y_n)$ (i.e. $\phi(x, y, U)$ preserves first-order truth).[8]

(6) There is an ordinal x (namely κ) such that $\phi(x, y, U)$ and y is an ordinal greater than x (i.e. $\phi(x, y, U)$ identifies the critical point of j).

[8] As this holds for any first-order formula ψ, this will be a schema of theorems.

All this can be shown in a first-order fashion[9]. We can thus use the relation $\phi(x, y, U)$ to do the work of the prima facie second-order entity j, whilst only talking about sets. The above formula effectively moves through the hierarchy of sets relating the sets in V and $\mathfrak{M}$, identifying a critical point along the way, without ever talking about actual proper classes.

It might be attractive then to regard all class talk as simply covert first-order definable talk about sets. We will now argue that such a view is misguided. Our strategy will be two-pronged. We begin by providing first philosophical reasons to reject the claim that the Universist should hold that only first-order definable class talk is legitimate, and then move to a second philosophical methodology.

VI.3 Conceptual Interlude: First philosophical reasons to countenance the use of non-definable classes

We should begin by noting some difficulties with providing any first philosophical justification of the use of non-definable classes. Since we wish to argue for a *more powerful* system, it will be tough to argue that there is conceptual incoherence lurking in the position that we should not countenance any non-first-order definable class talk. After all, the view that only first-order definable class talk is legitimate is easily mimicked within the position that there are non-definable classes simply by employing a restricted class quantifier. We can, however, point to several reasons why a ban on non-definable classes meshes poorly with the Universist's wider philosophical commitments.

We should note first that such a prohibition on non-definable classes creates a large disparity between the behaviour of class talk over a particular V_α compared to over V. Given an infinite V_α, the Universist thinks that there are non-first-order definable classes *relative* to V_α. This is witnessed by the existence of non-first-order definable sets in $V_{\alpha+1}$. Now for V, we do not have stages above V to witness the coherence of non-first-order definable class talk. However, we should note that, given that V is *far* richer in the number of sets it contains than *any* V_α, the idea that the

[9]See [Suzuki, 1999] and a very clear exposition in [Hamkins, 2012] for the full technical details.

class-theoretic structure of V is so impoverished fits poorly with the Universist's philosophical view about V and the relationships it bears to the V_α. The immediate response is that the legitimacy of non-first-order definable class talk is incoherent over V, and hence we should expect disparity here. We see little reason to accept the claim that such talk is incoherent. Theories involving non-definable classes are relatively weak, and are highly likely to be consistent.[10] The claim that non-first-order definable class talk is illegitimate thus needs bolstering.[11]

Further, the restriction to first-order definable classes seems ad hoc when compared to other Universist presuppositions. Simply put, her very insistence that there is a maximal universe in which we can understand set-theoretic discourse and on which every sentence of set theory receives a definite truth value, shows that the Universist already countenances the claim that there is more to set theory than can be expressed using mere first-order formalism. In particular, the very statement of her own view requires greater than first-order definable classes (e.g. the claim that V satisfies either ϕ or $\neg\phi$ for any $\phi \in \mathscr{L}_\in$ is not formalisable using a first-order definable class).

This plays out with respect to various mathematical claims the Universist should wish to make, but which a restriction to only first-order definable classes prohibits. For example, the following sentence:

"The ordinals are well-ordered by $\in$."

is clearly *true* for the Universist: among any ordinals whatsoever there is one that is $\in$-least. However, in order to state such a claim, we have to be able to formulate the notion that a class is well-ordered. This is a second-order statement that is not definable by any first-order formula. An insistence on only the use of first-order definable class talk would thus leave the Universist in a rather strange situation; there would be statements that seem *prima facie* true, but nonetheless she is forced to pass over in silence.

From the perspective of the First Philosopher, it is thus preferable to be able to

[10]For example,**MK** itself (to be discussed later in the chapter) postulates the existence of a large number of non-first-order definable classes and is consistent relative to the existence of an inaccessible cardinal.

[11]One response would be to argue that there is not a satisfactory interpretation of the class variables. We shall see in §5 that we can provide such an interpretation.

interpret non-first-order definable class talk. A puzzle remains, however, in that we have not provided an explanation of what such talk amounts to. Before we analyse specific characterisations to which the Universist may subscribe, we provide some additional *second philosophical* reasons to accept the use of non-first-order definable classes.

VI.4 Mathematical Intermission: Second philosophical reasons to have non-definable classes

In this section, we argue that the use of non-definable classes provides a better and more natural interpretation of mathematical discourse from a *second* philosophical perspective. That is, we will argue that there are set-theoretic results that are better interpreted if we countenance non-definable class talk.

Reflection principles are just one such area. A reflection principle is of the following general form:

$$\exists\alpha(\phi \rightarrow \phi^{V_\alpha})$$

In other words, if ϕ is true then ϕ is satisfied by some initial segment V_α (with quantifiers and parameters restricted to V_α). A salient fact is that often Universists consider reflection properties that are given by second-order parameters over V, and use the principles to study small large cardinals. For example, the second-order reflection principle states that, for any second-order parameter A over V:

$$(V, \in, A) \models \phi \rightarrow (V_\alpha, \in, A \cap V_\alpha) \models \phi^{V_\alpha}$$

Such a principle is most naturally understood when A is able to refer to non-first-order definable parameters over V, and produces many orders of large cardinals consistent with $V = L$. Without the use of such non-definable classes, we lose interpretation of the relevant A, and hence lose the consequences we would like within V (such as, in the case of second-order reflection, inaccessibles and Mahlo cardinals). Two problems are attendant here. First, second-order reflection as it is normally understood guarantees the truth of full impredicative comprehension in the class theory. To see this[12], note that if any instance of impredicative compehension fails in the class theory of V, then (by the second-order reflection principle) there must be a V_α for which impredicative comprehension fails. However, this is impossible: since the restricted second-order variables are interpreted as restricted to subsets of V_α (i.e. as ranging over $V_{\alpha+1}$), the truth of impredicative comprehension in the second-order theory of V_α is guaranteed by the strength of the Power Set Axiom. Moreover, any modification of the principle to allow for reflection of definable classes would have to explain why (say) the full inaccessibility of the ordinals was reflected, rather than mere *definable* inaccessibility (a far weaker notion).

Further, the study of large cardinal embeddings is given a far more natural interpretation if we have the use of non-definable classes. As noted earlier, we can characterise a measurable cardinal as the critical point of a non-trivial elementary embedding j from V to some transitive inner model $\mathfrak{M}$. Also, as we noted earlier, we could characterise this embedding using an ultrafilter parameter U and a first-order formula $\phi(x, y, U)$. A natural question is whether or not this method makes good sense of *all* theorems concerning embeddings.

There are reasons to think that the definable formula interpretation does not. We mention two such theorems, one negative and one positive. We deal with the negative first:

> **Theorem 67.** [Kunen, 1971] There is no non-trivial elementary embedding $j : V \longrightarrow V$.

Kunen's Theorem is relatively involved. It was conjectured by Reinhardt that there could be such an embedding, and took roughly a year to solve.[13] Moreover, the theorem built on other results in infinitary combinatorics (such as [Erdős and Hajnal, 1966]). Recent presentations use a result of Solovay that any stationary set S on a regular cardinal κ can be partitioned into κ-many stationary sets, and although they substantially simplify the proof[14] the result remains non-trivial. Contrast this with the

[12] I am grateful to Sam Roberts for this observation.

[13] The timings are somewhat hard to determine in virtue of the fact that [Solovay et al., 1978] was 'about' to be published from at the latest 1970 (Kunen himself mentions Reinhardt and cites the paper in [Kunen, 1971]). The philosophically relevant point still stands; the possibility of a $j : V \longrightarrow V$ was conjectured, relatively well-known, and took some time to refute.

[14] See [Schindler, 2014] for exposition.

result for *first-order definable* elementary embeddings:

> **Theorem 68.** [Suzuki, 1999] There is no non-trivial elementary embedding $j : V \longrightarrow V$ definable from parameters.
>
> *Proof.* This result is far simpler than any proof of Kunen's Theorem. Consider a j with $\kappa = crit(j)$. Let $\phi(x, y)$ define j (we suppress any parameters). We know that since ϕ is first-order, then we can define a first-order formula $\psi(x)$ that holds iff x is the least ordinal moved by j. Since $\psi(\kappa)$, by the elementarity of j we have that $\psi(j(\kappa))$ in the target model. But since $dom(j) = V$ and $ran(j) = V$, we have that $V \models \psi(\kappa)$, $V \models j(\kappa) > \kappa$, and $V \models \psi(j(\kappa))$. Hence κ both is and is not the least ordinal moved by j, $\perp$. ■[15]

The proof does not require any deep analysis of the nature of sets to prove. All we do is follow through the consequences of j being first-order definable and make some elementary observations about the nature of j in terms of its domain and range. Thus there seems to be some discord between the claim that all embeddings are first-order definable and the complexity involved in Kunen's Theorem. On the subject of Kunen's Theorem and the definability of j, Hamkins says the following:

> "Our view is that this way of understanding the Kunen inconsistency does not convey the full power of the theorem. Part of our reason for this view is that if one is concerned only with such definable embeddings **j** in the Kunen inconsistency, then in fact there is a far easier proof of the result, simpler than any of the traditional proofs of it and making no appeal to any infinite combinatorics or indeed even to the axiom of choice." ([Hamkins et al., 2012], p1873)

There are several points to note here. First, it is simply a fact that many set theorists are interested in the possibility of non-definable elementary embeddings, lending the meaningfulness of such talk second philosophical weight. Second, the view that all elementary embeddings are first-order definable substantially *trivialises*

[15]For full thoroughness (including checking that the notion of elementary embedding can be formalised in a first-order theory), see [Suzuki, 1999].

Kunen's Theorem, in that it makes his result relatively easy when it appears to concern deep facts about the combinatorial nature of the sets. Third, definability is unaffected by whether or not the Axiom of Choice holds. Kunen's Theorem (and subsequent modifications) depends essentially on use of AC. Currently, it is regarded as an open question whether or not there could be a non-trivial elementary embedding $j : V \longrightarrow V$ if AC turns out to be false in V (or indeed in any proper-class-sized model of **ZF** where AC fails). Regarding all embeddings as first-order definable would immediately answer this question: since there can be no definable embedding with or without AC, there is no embedding in the particular case where AC is false.[16]

It is not just with respect to *negative* theorems concerning the non-existence of elementary embeddings that we see this problem, however. *Prima facie*, set theorists talk about the existence of embeddings that cannot be first-order definable. The following is a good example:

> **Theorem 69.** [Vickers and Welch, 2001] Suppose $I \subseteq On$ witnesses that the ordinals are Ramsey[17]. Then, definably over $(V, \in, I)$, there is a transitive model $\mathfrak{M} = (M, \in)$, and an elementary embedding $j : (M, \in) \longrightarrow (V, \in)$ with a critical point.

Here, I is a proper class of good indiscernibles for On. If we introduce a predicate '$I(x)$' into the language to talk about those indiscernibles (so $I(x)$ holds iff $x \in I$), we can define (using $I(x)$) a non-trivial elementary embedding from $\mathfrak{M}$ to V. However, we should also be mindful of the following result:

[16] I am grateful to Sam Roberts for emphasising the importance of triviality and the settling of open questions to me, and also for pointing out Kunen's Theorem as a place where these issues arise.

[17] The details of Ramsey properties are somewhat technical and inessential for seeing the philosophical issues, and so we relegate them to a footnote:

To define Ramseyness, we first need the notion of a *good set of indiscernibles*. Let $I \subseteq \mathfrak{A} = L_\kappa[A, \in, \vec{B}, ...]$ be a first-order structure. Then I is a *good set of indiscernibles* for $\mathfrak{A}$ if for any $\gamma \in I$:

(i) $\mathfrak{A} \restriction =_{df} L_\gamma[A \restriction \gamma, \in, \vec{B} \restriction \gamma, ...] \prec \mathfrak{A}$,

(ii) $I \backslash \gamma$ is a set of indiscernibles for $\langle \mathfrak{A}, \langle \zeta \rangle_{\zeta < \gamma} \rangle$

We then say that κ is *Ramsey* iff any first-order structure with $\kappa \subseteq |\mathfrak{A}|$ has a good set of indiscernibles of length κ. To define Ramseyness for the particular case of the proper class On (the previous definitions only apply to set-sized structures). We say that On is *Ramsey* iff there is a class $I \subseteq On$, unbounded, of good indiscernibles for $(V, \in)$. More details and uses of these definitions are available in [Vickers and Welch, 2001].

Theorem 70. [Suzuki, 1999] Let $j : \mathfrak{M} \longrightarrow V$ be a definable elementary embedding such that $\mathfrak{M}$ is transitive and $On \subset \mathfrak{M}$. Then j has no critical point.

By this theorem, the Vickers-Welch embedding cannot be first-order definable over V. However, it seems that we are able to talk about such an embedding in a perfectly rigorous manner. It is not just j that cannot be definable in the above theorem. I cannot be definable as one can define a satisfaction relation for $(V, \in)$ over $(V, \in, I)$.[18]

Insisting that all embeddings be first-order definable in fact vitiates the possibility of an entire area of study. In the Introduction to the paper containing the above result, Vickers and Welch say the following:

> "It is quite natural to study the properties of elementary embeddings $j : V \longrightarrow M$ for M some inner model, since many such embeddings, if they exist, have first order formulations within **ZFC**. The question of reversing the arrow and looking at a non-trivial $j : M \longrightarrow V$ in general does not readily admit of such formulations. So we study in this paper what might be considered the **ZFC** consequences of the second order statement that there are proper classes j, M such that..."
> ([Vickers and Welch, 2001], p1090)

Thus, insisting that all classes be first-order definable prohibits an area of study that may produce fruitful mathematics with consequences for V. Placing a ban on the use of non-definable classes is thus not amenable to the "pursuit of promising mathematical avenues", and hence the Second Philosopher should be open to the use of non-first-order definable classes.

VI.5 Characterisations of classes

Let us take stock. We have seen that there are both first and second philosophical reasons to accept the use of non-first-order definable class talk for a Universist. How-

[18] See [Vickers and Welch, 2001] for details.

ever, a puzzle is now raised; how should we characterise this talk? One advantage of considering definable classes is that such an interpretation performs exceptionally well with respect to both the Foundational and Ontological constraints. Since talk of classes is paraphrased as satisfaction of formulas in V, we interpret our talk as only about sets. Further, the use of definable formulas represents a thoroughgoing class nominalism, and so it is clear how sets are different from our interpretation of classes. However, we now want to interpret a kind of discourse that cannot be interpreted through the use of first-order formulas. How then should we interpret non-first-order definable class talk? One option is to note that, naively understood, proper classes behave a lot like sets; they are extensional entities on which the membership relation is well-founded. However, on pain of contradiction, such objects *cannot* be sets.

Notice here that in interpreting this talk we are providing a response to a new Hilbertian Challenge. The wider context of this thesis is to answer the Hilbertian Challenge concerning extensions of V. For the moment, we have mathematical discourse that appears to not be solely about sets in V, but rather concerns problematic 'proper classes' that we would like to use in formulating axioms and proving theorems about V. We thus have a similar form of philosophical challenge as with extensions, with a corresponding Hilbertian Challenge. We should also, therefore, hold any possible response to the same standards with respect to our three constraints.

I will survey two possible options. Both, I argue, can provide an interpretation of the required discourse. However, as we shall see, under the three constraints one comes out as preferred.

VI.5.1 Ontologically 'heavyweight'

One option is to hold that proper classes are objects in ontological good standing. In other words, proper classes are objects over which we may freely quantify. Given that we do not want proper classes to appear anywhere *within* V, we have to provide a story of exactly how these objects behave and are to be characterised. What we require is an underlying conception that can underwrite our claims about classes.

This is especially so given that *any* conception of classes as heavyweight ontological objects is going to come up against an immediate objection: Why are these objects not sets? We appear to be talking about extensional, well-founded collections that are objects distinct from their elements. What then distinguishes the classes from the sets? Pressing a similar worry, Maddy remarks:

> "The problem is that when proper classes are combinatorially determined just as sets are, it becomes very difficult to say why this layer of proper classes atop V is not just another stage of sets we forgot to include. It looks like just another rank; saying it is not seems arbitrary. The only difference we can point to is that the proper classes are banned from set membership, but so is the κ^{th} rank banned from membership in sets of rank less than κ." ([Maddy, 1983], p122)

One response to this problem has been recently proposed by [Horsten and Welch, F]. There they develop a view on which classes are to be understood mereologically:

> "We propose to adopt a mereological interpretation of proper classes. We could say that the mathematical universe is a mereological whole and classes, proper as well as improper, are parts of the mathematical universe." ([Horsten and Welch, F], p18)

Classes are then different from sets in that sets are formed *combinatorially*, whereas classes are parts of V formed via *mereological fusion*. According to Horsten and Welch, we are not threatened with problems of proper classes being just like sets (and hence a regress of higher-order class structures) in virtue of this underlying conception:

> "The threat of a hierarchy of super- and hyper-wholes is not looming here. The fusion of the parts of a whole does not create a super-whole, but just the whole itself. So there is no mereological analogue of the creative force of the power set axiom." ([Horsten and Welch, F], p18)

According to Horsten and Welch, the fact that we are conceiving of classes as *mereologically* formed blocks any such problem. Consider an example where we try and create a problem by talking about the class pair $\{V, On\}$. Since classes are determined by the above conception, we understand this talk mereologically and consider the fusion of V and On. But such a fusion will just yield V back once more, and so we fail to increase the type of the objects.

Horsten and Welch's proposal represents a new and interesting development in the philosophy of set theory, not least because it appears to underpin what was largely a defunct analysis of proper classes as objects with a possibly coherent underlying conception. Moreover, Horsten and Welch use this conception to try and motivate some interesting mathematics.[19] However, for several reasons the conception is unsatisfactory for current purposes.

The first problem concerns the extent to which we can interpret non-definable class talk. The amount of discourse interpretable is dependent upon the mechanisms we countenance for forming mereological wholes. One can have more or less restrictive notions of what parts of V exist, corresponding to a greater and smaller amount of non-definable class talk that can be interpreted. However, for all Horsten and Welch have said, their view is consistent with there being no non-first-order definable wholes granted by their theory. While such a conception would be a rather austere view of the nature of mereology, it is nonetheless indicative of a question that needs answering: Exactly what is guaranteed by the conception?[20]

Let us suppose though that there is a good account of what classes exist and the mereological axioms governing them. Horsten and Welch's conception is still problematic in the context of the current discussion. The difficulty lies in its performance with respect to our three constraints, specifically the Foundational Constraint.

Certainly the mereological conception of classes performs well with respect to the Methodological Constraint. Our naive thinking concerning classes is easily represented; though the mechanism underwriting our claims (namely mereological fusion) is different from our naive understanding of class-theoretic operations, the conception does not do too much violence to our modes of thinking concerning classes. For example, if I wish to take a union of two classes A and B, this should be un-

[19]See [Welch, 2014] for presentation and [Barton, Fb] for an evaluation.

[20]In fact there are significant challenges to formulating mereological axioms in combination with set theory. A difficult problem is presented in [Uzquiano, 2006], where it is shown that several plausible principles concerning mereology and set theory are in tension with one another. There are points of Uzquiano's strategy with which Horsten and Welch will disagree (e.g. some of the arguments require that every object can be a member). However, the problems developed are indicative of the general difficulties in formulating a combined theory of mereology and sets.

derstood as mereological fusion, which will result in a whole that is constituted by exactly the same sets as would be members of $A \cup B$.

The Ontological Constraint is possibly slightly more vexed. Classes remain extensional, well-founded objects, and thus look a lot like sets. One may think that Maddy's challenge for making a non-arbitrary distinction has not been answered in a satisfactory manner. However, regarding the Ontological Constraint as violated is simply not to take the conception seriously at all. Horsten and Welch provide an explanation of why (on their view) classes do not form sets and there is no problem of higher-order classes; classes are formed through mereology rather than the combinatorial strength of the Power Set Axiom. Thus, the Ontological Constraint is satisfied on Horsten and Welch's own terms.

Despite this relatively good performance, the Foundational Constraint *is* violated. For now our mathematical discourse is not interpretable solely by sets, but rather by sets and mereological fusions of sets. In this way, Horsten and Welch go substantially beyond sets in terms of the ontology they countenance; none of these mereological fusions are interpretable using sets solely from V. For present purposes then, and the Universist who is motivated by the Foundational Constraint, a different solution is preferable.

Sufficient for resolution of this problem is some species of class *nominalism*. If we wish to conform to this constraint, we must do so through talking *about* the *sets* rather than postulating the existence of additional non-set-theoretic ontology. In this regard, the definable formula interpretation was on the right track; talk of classes was interpreted entirely by talking about sets. We require an interpretation that combines this streamlined view of ontology with the expressive resources granted by non-first-order definable classes.

VI.5.2 Plural resources

Such an interpretation is forthcoming in the literature. One way of interpreting class talk (originally proposed by [Boolos, 1984] and with subsequent development by [Boolos, 1985], [Uzquiano, 2003], and [Burgess, 2004]) is through the use of *plural reference*. Consider the following sentence:

The rocks rained down.

Such a sentence refers to *the rocks* in the plural, and ascribes to them a *non-distributive* plural predicate. The rocks are engaging in raining *together*, no one rock is raining on its own. One might think that such an ascription depends upon a set-theoretic interpretation of the language. Aside from the fact that it is unclear that the set of rocks has any part to play in the raining down (the idea that a set of rocks could rain down has the whiff of a category mistake), further evidence that plural reference is part of common language is available by analysing the following famous example from Boolos:

> "It is haywire to think that when you have some Cheerios you are eating a *set*—what you're doing is: eating THE CHEERIOS." ([Boolos, 1984], p448)

Plural reference thus seems to be part of everyday discourse.[21] Interestingly, it is possible to encode a significant amount of greater than first-order content through the use of plural reference. For example, the following sentence:

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

looks perfectly legitimate, but is similar in syntactic form to:

> There are some numbers such that if a number n is one of them, then n has a predecessor that is also one of them.

Such a sentence implies the existence of an infinite descending sequence of natural numbers, and so can only be true in a non-standard model of arithmetic. By the usual metatheoretic results for first-order theories (such as the Löwenheim-Skolem and Compactness Theorems), one cannot characterise the standard model of arithmetic up to isomorphism using only first-order resources.

[21] Considerations of space prevent a full discussion of the issues surrounding plural reference. Additional discussion is available in [Barton, 2012] and in the published literature in [Boolos, 1984] and recently [Oliver and Smiley, 2013] for (convincing) arguments that plural locutions are part of our ordinary mathematical language.

Indeed, it turns out that, using plural resources, one can provide an interpretation of monadic second-order logic,[22] and if ordering and superplural quantification is permitted this can be extended to full second-order logic.[23] For our purposes though, we shall be primarily interested in a plural interpretation of class talk.

Just such an interpretation was indicated in [Boolos, 1984], discussed in [Uzquiano, 2003], and precisely stated in [Burgess, 2004] but for the sake of clarity we provide our own version here. It will serve first to introduce some formal machinery for making precise plural reference:

Definition 71. The language $\mathscr{L}_{\prec}$ is a two-sorted first-order language comprising the following:

(i) The usual resources of first-order logic: variables x, y, z,... that range over *objects*, quantifiers $\forall$ and $\exists$, and an equality relation $=$.

(ii) Plural variables xx, yy, zz that refer to *some things*.

(iii) Plural quantifiers $\forall xx$, $\exists yy$, read as 'for any things xx', and 'there are some things yy'.

(iv) A two-place relation $x \prec xx$, that holds between singular and plural variables to denote that x is one of the xx.

(v) One's favourite propositional connectives.

We shall not lay down the rules of plural logic just yet, as this discussion will be instructive in the next section. For now, we provide exposition of how classes may be interpreted plurally, and argue that the interpretation is amenable to the Universist's current position.

Given the use of the mechanisms of plural reference, we wish to interpret class talk. We will, from this point on, assume that plural reference and quantification is well understood in the sense that our plural variables have determinate interpretation. We can then consider the language $\mathscr{L}_{\prec,\in} = \mathscr{L}_{\in} \cup \mathscr{L}_{\prec}$ in order to talk about the sets through the mechanisms of plural reference. Let us start with a simple example. The following statement is obviously *true* for the Universist:

[22] See [Boolos, 1984] for details.
[23] See [Hewitt, 2012].

There are some things rr such that x is one of them iff $x \notin x$.

Such a locution serves to pick out all and only the non-self-membered sets. We do *not* thereby commit ourselves to the existence of a class: we are merely talking about the non-self-membered sets, of which there are many. Given the above, it is natural to start comparing what we can say about the sets:

The non-self-membered things are the same things as the sets.

Such a statement provides a plurally rendered, nominalistically acceptable, interpretation of the claim that the Universal class and Russell Class are the one and the same. Again, it is obviously true for a Universist.

We now need to start putting some mathematical meat on the bones of our interpretation. We shall start by stating that the ordinals are well ordered by $\in$. We first introduce the following defined symbol:

Definition 72. The xx are *some of* (or *among*) the yy (written $xx \prec\prec yy$) iff (If x is one of the xx then x is one of the yy).

We can then characterise the ordinals being well-ordered by $\in$ as follows:

There are some things $\alpha\alpha$ such that x is one of the $\alpha\alpha$ iff x is an ordinal. Moreover, for any xx that are some of[24] the $\alpha\alpha$, there is a β that is one of the xx such that there is no $\gamma \neq \beta$ that is also one of the xx with $\gamma \in \beta$, and for any $\delta, \chi \prec xx$ either (i) $\delta \in \chi$, (ii) $\chi \in \delta$, or (iii) $\chi = \delta$.

Moving up in complexity, we can use plurals to talk about elementary embeddings:

There are some ordered pairs jj such that:

(1) If $\langle x, y \rangle$ is one of the jj and $\langle x, z \rangle$ is also one of the jj, then $y = z$ (i.e. the jj code a function-like embedding).

(2) If $\langle x, y \rangle$ is one of the jj, and $\langle z, y \rangle$ is one of the jj, then $x = z$ (i.e. the jj code a one-to-one embedding).

[24]Here we assume that for any things yy there is at least one thing that is one of the yy. We will see some discussion of the possibility of an 'empty plurality' later in the Chapter.

(3) If $\langle x_0, y_0\rangle$,..,$\langle x_n, y_n\rangle$ are each among the jj, then for first-order ϕ, $\phi(x_0, .., x_n)$ iff $\phi(y_0, ..., y_n)$ (i.e. the jj code an elementary embedding).

(4) There is an $\langle x, y\rangle$ that is one of the jj such that $x \neq y$ (i.e. the jj code a non-trivial embedding).

(5) There is a least ordinal κ such that $\langle \kappa, y\rangle$ is one of the jj, and $\kappa \neq y$ (i.e. the jj identify a critical point of the coded embedding).

The jj (should there be such objects) provide a coding in purely plural terms of the existence of a non-trivial elementary embedding $j : V \longrightarrow \mathfrak{M}$. With this interpretation in place, and with suitable abbreviations being made, we can take Kunen's Theorem to establish the following:

> There are no things jj coding a non-trivial elementary embedding from the sets to the sets (i.e. there are no things jj coding a non-trivial elementary $j : V \longrightarrow V$).

The Vickers-Welch embedding is a little more difficult to interpret. Recall the statement of the theorem:

> [Vickers and Welch, 2001] Suppose $I \subseteq On$ witnesses that the ordinals are Ramsey. Then, definably over $(V, \in, I)$, there is a transitive model $\mathfrak{M} = (M, \in)$, and an elementary embedding $j : (M, \in) \longrightarrow (V, \in)$ with a critical point.

We can begin with the following characterisation:

> Suppose that there are some things ii that are (collectively) good indiscernibles for the ordinals. Then there are some things jj coding a non-trivial elementary embedding from some things mm that (collectively) satisfy **ZFC** to the sets.

One aspect of the above characterisation that is left out is the definability of j using I. We will discuss in the next Chapter exactly how to code ordered pairs on classes and hence structures. For the time being it suffices to note that if we are interested in some things ii that are collectively good indiscernibles for On, then there is no barrier to introducing a predicate '$I(x)$' into the language $\mathscr{L}_{\in}$, such that $I(x)$ holds iff $x \prec ii$. We can then use this expanded language to define the embedding over V.

There thus seems to be no barrier to talking about non-definable classes through the use of plural reference. However, we would like to know whether the Universist *should* countenance talk of non-definable classes on this interpretation.

Insofar as the Universist has already countenanced the use of plurals in interpreting class talk, the answer has to be "yes". Consider the case where we restrict attention to some V_κ for uncountable κ. Uncontroversially for the Universist, there are some things that are some of the sets in V_κ, do not form a set in V_κ, and are not all and only the satisfiers in V_κ of some first-order formula. This is made especially perspicuous by the existence of non-definable sets in $V_{\kappa+1}$. Now, in the context of V there are no further stages to make this quite so clear. However, V is patently the richer structure. Therefore, if there are some things xx that are some of the sets in V_κ, do not form a set in V_κ, and are not all and only the satisfiers of some first-order formula ϕ (relative to V_κ), then there should be some things yy that are some of the sets, do not form a set, and are not all and only the satisfiers of some first-order formula.

Further, the interpretation performs very well with respect to the constraints outlined in previous sections. The Ontological Constraint is clearly satisfied; class talk should be understood via plural reference to sets, and so our interpretation of class talk is very different from our interpretation of set talk. The latter refers to objects whereas the former does not refer to objects, rather it refers to sets (in the plural) within V. This further shows why the threat of higher-order classes does not bite on this understanding of class talk. Classes cannot be members because some things xx are not an object that can be a member of anything.

The Methodological Constraint is also nicely satisfied. Our naive thinking about classes carries over well to the plural case. Turning again to the basic example of Union, it is highly plausible that if there are some things xx and some other things yy, then there are some things zz such that an object is one of them iff it is either one of the xx or one of the yy. As we shall see in the next section, this carries over directly to formal theories of classes, and we are able to interpret **MK** plurally.

However, the most important point (given the discussion of Horsten and Welch's proposal) is that the plural interpretation also satisfies the Foundational Constraint. By adopting a class nominalism with reference understood plurally, we make our discussion of classes *about* sets, rather than any objects over and above V. It is still the case that *the sets* provide the standard for correctness and coherence in one foundational mathematical arena. Where Horsten and Welch's proposal accounts for class talk by having it interpreted as *about* a new kind of object, we interpret classes as a providing a new way of talking *about* the sets.

VI.6 Justifying MK

We now provide a justification for **MK** class theory. First, however, we make a remark about methodology, specifically what we will be taking for granted. The above section provides good reason for the Universist to accept that plural reference to sets is well-understood in the sense that we know what it means for any object x to be one of some things xx, and whether or not there are some things yy. In particular, we will refrain from providing a detailed *semantics* for plural reference. This has some precedent in the literature:

> "Throughout philosophical logic, much mischief is caused by a double usage of the word 'semantics'. It is used on the one hand for models, like those provided by Tarski for singular or first-order logic, or by Kripke for modal logic; and it is used on the other hand for a theory of meaning. Confusion between these two usages is manifested in the literature in two different, complementary, ways. On the one hand, if a model theory has not yet been developed for a given logical notion, it may be alleged that the notion is 'meaningless' because it lacks a 'semantics'. On the other hand, once a model theory *has* been developed for a given logical notion, it may be alleged that problematic 'ontological commitments' are implicit in the use of the notion...Both types of objections could be raised against

plural logic. On the one hand, I have not yet presented a model theory for plural logic...On the other hand, when I do present a model theory and an argument that it is satisfactory...then since the model theory will involve an apparatus of sets, it might be claimed that this shows an 'ontological commitment' to sets is implicit in the use of the plural." ([Burgess, 2004], pp216-217)

As Burgess identifies, opponents of the use of plural logic put its adherents in a tricky spot. Either, (i) the plural logician refrains from giving an explicit set-theoretic semantics for the logic of plurals and is criticised for not telling us what the ranges of the plural variables are, or (ii) she provides a (set-theoretic) semantics for a logic of plurals and hence is criticised for using 'set theory in sheep's clothing' or similar. This latter objection is especially damaging in the present context if taken seriously; given that our domains are often proper-class-sized, we do not have the option of using a set-theoretic semantics there. A third option, not mentioned by Burgess in the above quotation, is (iii) use an expanded language in providing the semantics for plurals. This is usually done either via the use of a plural satisfaction predicate ([Boolos, 1984]) or a move to third-order resources ([Linnebo and Rayo, 2012]). This, however, results in the immediate difficulty that the extra resources are plausibly less intelligible than the simple relationship of an object being among some things. There are possible responses to be given here (possibly using *super*plural reference[25]), but even if an answer can be provided, the semantic question can be pushed up another level. Better, instead, to refrain from providing an explicit semantics, and instead axiomatise the rules of logic that govern the use of the language. Indeed, Burgess harbours a similar thought:

"Against the first objection I maintain that even if no one ever did present a satisfactory model theory for plural logic, the plural was in systemic use in natural languages long before model theory for anything had been born or thought of, and such long-standing systemic usages are meaningful if anything is. Against the second objection I maintain...that the tran-

[25]In addition to [Linnebo and Rayo, 2012], see [Linnebo and Nicolas, 2008].

sition from plural language to set-theoretic language in the work of Cantor and his followers involved an intellectual struggle more difficult than would have been called for if the task had been merely one of making explicit something already implicit in ordinary language." ([Burgess, 2004], p217)

We sympathise with these points. It is a curious fact that some philosophers desire a formal set-theoretic semantics for almost every use of language. A set-theoretic semantics is needed when we wish to make precise the use of terms which are not clear. For example, a precise formal semantics to deal with cases of vagueness is often useful to explore the logical space of assertions that can be made regarding vague statements. We contend that it is perfectly clear to (non-contrarian) speakers of English familiar with plural quantification and set theory what the sentences "x is one of some things xx" or "there are some things xx such that ϕ" mean, or indeed for that matter what "there are some sets such that they do not form a set and are not all and only the satisfiers of some first-order formula" means.[26]

What we *can* do, however, is to lay down the logical rules that govern our use of plural quantification. We take them to be the following:

Definition 73. *Plural First-Order Logic* or $\mathbf{PFO}$ is a theory in $\mathscr{L}_{\prec}$ with the following axioms:[27]

(i) All axioms of first-order logic.

(ii) $\forall xx \exists y(y \prec xx)$ (i.e. for any things something is one of them).

(iii) Plural Indiscernibility Scheme:
$\forall xx \forall yy[\forall z(z \prec xx \leftrightarrow z \prec yy) \rightarrow (\phi(xx) \leftrightarrow \phi(yy))]$
(i.e. if some things xx are the same things as some things yy then whatever holds of the xx also holds of the yy, and vice versa).

(iv) Plural Comprehension Scheme: $\exists x \phi(x) \rightarrow \exists xx \forall y[y \prec xx \leftrightarrow \phi(y)]$
(i.e. If there is a ϕ then there are some things that are all and only the satisfiers of ϕ)

[26] If we were pressed to, we would opt for a formally adequate but wholly uninformative Tarski-style semantics with a truth predicate and the relevant T-axioms.

[27] For details, see the excellent introduction to plural logic available in [Linnebo, 2014].

We shall use this conception of plural logic to motivate the axioms of **MK**. As an axiomatisation of our use of plurals, **PFO** is reasonably uncontroversial. The only substantive question is what we allow for ϕ in the Plural Comprehension Scheme, this issue is given consideration below. For the moment, we move on to justifying each axiom of **MK** in turn. We first need to specify the content of **MK**:

Definition 74. $\mathscr{L}_{\in,\eta}$ consists of the following:

(i) All apparatus of $\mathscr{L}_{\in}$.

(ii) A class membership predicate η.

(iii) Variables and quantifiers for classes (denoted by uppercase Roman letters, X, Y, Z, X_1, X_2, and $\forall X_1$, $\forall X_2$ and so on).

In addition to the well-formed formulas of $\mathscr{L}_{\in}$, we have:

(iv) If X and Y are class variables, then the formula $X = Y$ is well-formed.

(v) If x is a set variable and Y is a class variable, then the formula $x\eta Y$ is well-formed.

(vi) The usual formulas constructed from atomic sentences using connectives and quantifiers are well-formed.

We now wish to provide an analysis of the class variables and membership using plurals. However, we must first make the following:

Remark 75. A problem with using plural reference to provide an interpretation of the class variables and η is that, on common understandings of plurals, some things must always include at least one thing (and in some accounts at least two things), even though we wish to have an empty class. This is easily remedied, [Burgess and Rosen, 1997] show that different readings of the plural quantifiers (to admit zero or more, one or more, or two or more things) are interdefinable. Thus, for the moment, we let $\prec'$ denote the relation that corresponds to the zero or more reading of plurals.

Remark 76. The reader may, at points, feel like the exposition could flow more quickly in this section. Indeed, many expositions (such as [Uzquiano, 2003]) are quick to move from plurals to classes. It is important in the current foundational context, however, to take things slowly and see exactly how the formalisation of ordinary plural language results in a mature class theory.

We can now provide the following analysis of the variables. The set variables denote sets, and $\in$ denotes set membership. A class variable Y denotes some things yy, and $x\eta Y$ iff $x \prec' yy$.[28]

Remark 77. The use of a distinction between η and $\in$ is formally redundant. We could work in a language that blurs these distinctions, admitting sentences stating that one class is a member of another as well-formed and using a Sethood predicate '$Set(X)$', such that $Set(X)$ holds iff $\exists Y(X\eta Y)$ (this was, in fact, how Gödel formulated his version of $\mathbf{NBG}$[29] class theory). If one is motivated by parsimony in language, this would be the sensible route. On the other hand, we should be mindful that our interpretation of classes through *plural reference* means that, conceptually speaking, it makes more sense to speak of a separate membership relation η between sets and classes, and talk of classes (understood as some things xx) *forming* sets.[30]

Despite this, *some* formal fluidity is desirable, especially when we get on the the trickier class-theoretic results in Chapter VII. What we would like is to keep the two-sorted nature of $\mathscr{L}_{\in,\eta}$ so as to retain our conceptual distinction, but not have to constantly keep track of whether we are using $\in$ or η. We thus provide the following:

Definition 78. The *language of* $\mathbf{MK}$ (denoted by '$\mathscr{L}_{\in^*}$') comprises the following:

(i) The variables, connectives, and quantifiers of $\mathscr{L}_{\in,\eta}$.

[28] This is essentially Boolos' translation given in [Boolos, 1984] and [Boolos, 1985], with subsequent development in [Uzquiano, 2003].

[29] $\mathbf{NBG}$ is a theory, substantially weaker than $\mathbf{MK}$, in which Class Comprehension is restricted to predicative formulas only (i.e. all class quantification in ϕ must be bounded).

[30] See [Bernays, 1958], Part I, Chapter 7 for discussion.

(ii) A dyadic membership relation denoted by $\in^*$.

The well-formed formulas are constructed as follows:

(iii) The equality statements of $\mathscr{L}_{\in,\eta}$ are well-formed.

(iv) If x is a set variable and Y is a class variable, then $x \in^* Y$ is well-formed.

(v) If x is a set variable and y is a set variable then $x \in^* y$ is well-formed.

(vi) The usual formulas constructed from atomic sentences using connectives and quantifiers are well-formed.

$\mathscr{L}_{\in^*}$ keeps the two-sorted conceptual distinction between sets and classes, but allows us to reason with a single membership relation, streamlining proof somewhat.[31] We can now define **MK** proper:

Definition 79. **MK** consists of the following axioms:

(A) Set Axioms:

- (i) Set Extensionality
- (ii) Pairing
- (iii) Infinity
- (iv) Union
- (v) Power Set

(B) Class Axioms:

- (i) Class Extensionality: $\forall X \forall Y \forall z[(z \in^* X \leftrightarrow z \in^* Y) \rightarrow X = Y]$ (i.e. Classes with the same members are identical).
- (ii) Foundation: Every non-empty class has an $\in^*$-minimal element.
- (iii) Scheme of Impredicative Class Comprehension: $\forall X_1, ..., \forall X_n \exists Y (Y = \{x | \phi(x, X_1, ..., X_n)\})$, where ϕ is a formula of $\mathscr{L}_{\in^*}$ that may contain both set and class parameters, and which unrestricted quantification over classes and sets is allowed.

[31]If we did not make such a move, any time we wanted to talk about a class that formed a set which was in turn a member of a class, we would have to switch relation.

(iv) Class Replacement: If F is a (possibly proper-class-sized) function, and x is a set, then $ran(F \upharpoonright x)$ is a set (i.e. $\{F(y)|y \in^* x\}$ is a set).

(v) Global Choice: There is a class function F such that for every non-empty x $\exists y \in^* x\, F(x) = y$. Equivalently, there is a class that well-orders V.[32]

One immediate question is the following: Why bother? Why do we feel the need to move from the two-sorted first-order **PFO** to the class theory **MK**, especially when the former is ontologically honest where the latter appears to misleadingly singularise class talk? There are two reasons for this. The first is nicely summed up by Uzquiano:

> "This [the plural interpretation of proper class discourse] is of course not an invitation to relinquish the vocabulary of classes. For as it will soon become plain, plural paraphrases quickly become unwieldy and difficult to parse. The proposal is rather that we continue to use the vocabulary of classes in the context of set theory, but warn that its grammatical form is not ontologically transparent." ([Uzquiano, 2003], p73)

Simply put, the language of $\mathscr{L}_{\in^*}$ is more easy to understand and work with in a mathematical context. If we kept things plural, when considering a given statement, we would have to be constantly book-keeping what we said to avoid error (such as when dealing with the 'empty plurality' or keeping track of the use of $\prec$ and $\in$). This is especially apparent when we start increasing the complexity of the statements in question and communicating via a mix of formal and natural language. For example, even for a comparatively simple (true) statement such as:

If $X \neq V$, then there is a Y such that $X \subset Y$.

the plural rendering becomes awkward to parse:

[32]Note that since Global Choice and Class Replacement imply their set-sized incarnations, we do not need to include Choice and Replacement in our Set Axioms.

> For any things xx if there is something y such that y is one of the sets but not one of the xx, then there are some things zz such that the xx are among the zz but the zz are not among the xx.

and this is for a very simple claim about classes and V. As we increase complexity through the study of proper-class-sized models and embeddings, the strict plural rendering is likely to become highly intractable.

The second reason for the move is more pragmatic still. Simply put, $\mathbf{MK}$ has received more mathematical attention, and so we have more theorems we can use compared to $\mathbf{PFO}$. Given then that we have a translation between $\mathscr{L}_{\in^*}$ and $\mathscr{L}_{\in,\prec}$, the move is justified, even if only for reasons of expediency and proof.

Since $\mathbf{ZFC}$ is presupposed by this thesis, we already have all the $\mathbf{MK}$ axioms concerning only sets (i.e. part (A) of the definition of $\mathbf{MK}$), as well as Choice and Foundation for sets. We now proceed to motivate each of the class-theoretic axioms in turn.

Class Extensionality is justified by the observation that for any things xx and any things yy, the xx are the same things as the yy just in case every z that is one of the xx is one of the yy (and vice versa). More formally, we note that if we have two classes X and Y such that $z \in^* X$ iff $z \in^* Y$, then there are some things xx and some things yy such that $z \prec xx$ iff $z \prec yy$. Class Extensionality then follows from Plural Indiscernibility.

Foundation. For any Y, either (i) the yy denoted by 'Y' form a set y, or (ii) the yy denoted by 'Y' do not form a set. In case (i) $\in^*$ reduces to $\in$ and $\in^*$-Foundation is immediate from $\in$-Foundation. For (ii) every $y \in^* Y$ is a set, and hence $\in^*$-Foundation follows from $\in$-Foundation.

Scheme of Impredicative Class Comprehension. The next axioms are more difficult to justify. Our methodology will be much like the one we adopted in considering characterisations of proper classes; we will consider both first and second philosophical considerations.

We begin with *first* philosophy. Since our understanding of the class variables and $\in^*$ is given by plural quantification and reference, the truth of Impredicative Class Comprehension reduces to the question of whether or not we can have impredicative definitions in the Plural Comprehension Scheme. Some have regarded this question as obvious. Uzquiano, for example, writes:

> "To the extent to which one accepts unrestricted plural quantification over sets as unproblematic, one will be moved by what David Lewis refers to as the evident triviality of plural comprehension, and thus one will accept all instances of plural comprehension as true. After all, one may explain, in order for an instance of comprehension to be false, there must be a formula ϕ such that it is neither the case that no sets satisfy it nor is it the case that some sets satisfy it. But this could never be the case." ([Uzquiano, 2003], pp76-77)

Thus, insofar as one accepts the determinacy in the range of the plural variables, Impredicative Class Comprehension seems inevitable. For as Uzquiano points out, if impredicative definitions were prohibited from featuring in the Plural Comprehension Scheme, we would be able to state (in our plural theory) a formula that is neither satisfied by no sets nor satisfied by some sets; an apparent contradiction. Furthermore, worries of contradiction are unlikely to threaten; **MK** is consistent relative to the existence of an inaccessible cardinal. There are, however, those who do not agree that the range of plural variables need be determinate.[33] Such a position would undermine the above argument; a formula could be satisfied on some (Henkin-style) assignments but not others. Though we take the ranges of plural variables to be determinate, it is nonetheless desirable to bolster this intuitive plausibility with *second* philosophical considerations.

From a second philosophical perspective, we note that the definition of a **ZFC** *truth predicate* for V has to come from the Scheme of Impredicative Comprehension. To observe this, note that restriction of Class Comprehension to predicative formulas results in **NBG**. Since **NBG** is a conservative extension of **ZFC** it cannot prove the existence of a **ZFC** truth predicate for V (by Tarski's Theorem). Thus, Impredicative Class Comprehension receives second philosophical support; it opens the door to new and plausibly fruitful mathematics.[34]

[33] For a dissenting voice see [Florio and Linnebo, 2015].

[34] For discussion, see [Gitman et al., U]. For a concrete example of how a truth predicate can be used in

Class Replacement and *Global Choice.* We deal with Class Replacement and Global Choice in tandem since they are intimately linked. We first remark that Global Choice has substantial independent merit. From a second philosophical perspective, many constructions in class theory require Global Choice (good examples being class ultrapowers).[35]

From a first philosophical stance, we note that Global Choice is a very robust principle in that if it holds in a model, it is very hard to destroy. Two examples are of interest here. First, Global Choice is consistent with all known large cardinals consistent with AC (so, as far as we know, we cannot destroy it by moving to a model with a new kind of cardinal). Second, Global Choice is preserved by set forcing[36]. Given then that Global Choice is hard to destroy once we have it, we might think that it is more likely to be true; insofar as the model theory of sets indicates how sets behave, Global Choice seems well-entrenched.

From a first *and* second philosophical point of view, however, it is important to note that Global Choice and Class Replacement are intimately linked to limitation of size. We noted in the Introduction that it was at least very unclear if and how Replacement can be justified on the basis of the Iterative Conception of Set. First-order reflection was considered, but is far from uncontroversial in this respect. Instead, many theorists (such as Cantor, Russell, Jourdain, Mirimanoff, and more recently Burgess)[37] have appealed to some principle of a limitation of size; in addition to sets being formed iteratively, sets also exist when given by an operation on a set that does not increase its size. Recall from Chapter I that there were two main ways of expressing this idea:

> *Weak Limitation of Size Principle.* If some objects can be put in a one-to-one correspondence with a set, then there is a set of those objects.
>
> *Strong Limitation of Size Principle.* Some objects form a set iff they are not

deriving the existence of sets, see [Welch, 2014].

[35] Again, see [Gitman et al., U] for details. In fact, far stronger choice principles are required. Interestingly, the justification of many such strong choice principles would make the coding of the next section both easier and more powerful (see [Antos and Friedman, S] for details). However, since $\mathbf{MK}$ is the much more widely accepted theory and is already regarded as a strong class theory by many philosophers, we find it dialectically effective to show that we can accomplish the necessary coding for extensions in $\mathbf{MK}$.

[36] See [Hamkins, 2015] for details.

[37] See [Hallett, 1984] for an excellent exposition of the history and analysis, as well as a more comprehensive list. Burgess' appeals to limitation of size appear in [Burgess, 2004].

bijective with the universe.[38]

The former is only sufficient to guarantee Replacement for sets. However, if the latter is taken as our interpretation of the limitation of size idea, and if we are happy with there being more than one thought underlying set theory we can prove the following:

Theorem 80. The Strong Limitation of Size Principle is equivalent to the conjunction of Global Choice and Class Replacement (modulo the other axioms of **MK**).[39]

Proof. Assume the Strong Limitation of Size Principle. To show Class Replacement, take any set x, function F, and consider $ran(F \restriction x)$. Since $ran(F \restriction x)$ can be mapped into x, it is not proper, and hence there is a set $y = ran(F \restriction x)$. For Global Choice, note than On is a proper class. By the Strong Limitation of Size Principle, On can be mapped onto V, and hence there is a global well-order of V.

In the other direction, assume Class Replacement and Global Choice. We show each conditional of the Strong Limitation of Size Principle in turn. Take any class X and assume that it can be mapped onto V. Then, we know that X does not form a set (otherwise V would be a set by Class Replacement). Conversely, assume that X does not form a set. Map X into an initial segment S of the ordinals (using Global Choice). We know (by Class Replacement) that S cannot form a set, and hence $S = On$. By Global Choice (and composing the relevant functions) X is thus mappable onto V. ■

There are a number of points to be made about this fact. We might simply take the equivalence between these two (reasonably natural) principles to confer extrinsic[40] evidence upon each, much as the wide range of natural principles equivalent

[38]The formal class-theoretic is a little impenetrable but reads as follows:
$\forall C(\neg\exists x\forall y(y \in^* C \leftrightarrow y \in^* x) \leftrightarrow$
$\exists F(\forall x(\exists W(x \in^* W) \rightarrow \exists s(s \in^* C \wedge \langle s,x\rangle \in^* F)) \wedge \forall x\forall y\forall s((\langle s,x\rangle \in^* F \wedge \langle s,y\rangle \in^* F \rightarrow x = y)))$.

[39]I am very grateful to Victoria Gitman and Sam Roberts for discussion of the technical issues surrounding Limitation of Size and Global Choice.

[40]We move to using the terms 'intrinsic' and 'extrinsic' rather than 'first philosophical' and 'second

to AC or the equivalence between **ZC**$-$Infinity with first-order reflection and **ZFC** are each taken to confer extrinsic evidence on their respective axioms. The Strong Limitation of Size Principle, Global Choice, and Class Replacement are all natural enough principles, so any equivalences between them only serves to bolster their case.

For the purposes of intrinsic justification, it is important to note that the equivalence shows that insofar as the Strong Limitation of Size Principle is justified, then Global Choice and Class Replacement are also essential (and vice versa). Further, heuristic justification can be given to link limitation of size and iterativity. Clearly, if some objects form a set, then they are not bijective with V. For the other direction, we note that there are two ways that a class X could fail to be bijective with V. Either (i) it could be bounded in some V_α (in which case X immediately forms a set), or (ii) we might lack the relevant bijection. However, in case (ii) we note that since the ordinals are a proper class, the claim that every proper class is bijective with V is equivalent to the claim that every proper class is bijective with the ordinals. If Strong Limitation of Size fails, we would thus have the rather strange situation where $X \cap V_\alpha$ is well-orderable for every V_α, but X as a whole is not well-orderable. In other words, given that each initial segment of V can be well-ordered, we might think that if some things are not bijective with V then they are unlikely to be unbounded in the V_α (and vice versa).

These arguments are not meant to be conclusive, in particular the previous heuristic consideration concerning unboundedness and limitation of size assumes the principle. Certainly, there are worries here. First, the Strong Limitation of Size Principle is far from obvious, and one might instead opt for the Weak Limitation of Size Principle in justifying Set Replacement. Second, one might not accept that it is possible to adjoin a notion of limitation of size to our iterative conception (though then one must explain how Replacement is justified). Despite these worries, we should be mindful from the Introduction that the justification of **ZFC** itself admitted of sim-

philosophical', because this is one case where it is unclear whether the equivalence is more naturally amenable to a first or second philosophical outlook. Does the equivalence confer justification because it is fruitful mathematics, or rather because it shows a deep fact about our notion of set, amenable to first philosophical analysis? We take no stand on the issue here, but note that the exact relationships between first philosophy, second philosophy, intrinsic justification, and extrinsic justifications are far from obvious.

ilar problems. The important point is that *if* one accepts the Strong Limitation of Size Principle, then one must accept Class Replacement and Global Choice too. The justification for **MK**, it seems, is not that much weaker (if at all) than **ZFC** itself.

Chapter VI: Conclusions

We conclude by making some points about exactly what we have achieved, especially in the context of the current dialectic. In this chapter we began by revisiting the problem of proper classes for the Universist, and noted that there was a convenient way of paraphrasing class talk in terms of definable formulas. We then identified first and second philosophical reasons to accept the use of non-definable classes in set theory. Next, we considered two characterisations of proper classes, one as 'heavyweight' ontological objects formed through mereology, and another as understood through plural quantification, and argued that the latter performed far better with respect to the three constraints. We then provided a justification of **MK** class theory, based on the plural interpretation of proper class discourse.

A few further comments are in order, however. We do not take the above arguments to be in any sense *definitive*. Rather, we have simply shown that there are reasons (both philosophical and mathematical) to accept the legitimacy of discourse involving non-definable classes, and that such an interpretation suggests **MK** class theory as a natural choice of first-order theory to codify such talk in a familiar mathematical language. We do *not* claim that we have provided conclusive reasons to accept **MK**, and indeed philosophical research in this area is ongoing. However, there is no apparent *obstacle* to using **MK**, we have reasons to want to use it, and have a characterisation that suggests it can be interpreted in a philosophically satisfactory manner. As we shall see in the Conclusion, this is important for assessing the philosophical upshot of the coding discussed in this thesis. For now, however, we move on to putting our theory of classes to work.

Chapter VII

V-Logic and Resolution

In previous chapters we have seen that for the purposes of executing Gödel's Programme, there is the possibility of using extensions of V in formulating new axioms and proving consequences about V. We then examined extant attempts to interpret height and width extensions of V and found them wanting in certain respects. In the preceding chapter, we saw the first part of our positive proposal; a motivation (on both first and second philosophical grounds) of the use of **MK** based on a plural conception of proper class discourse.

The time has come to put this theory of classes to work. In this chapter, we show how, using **MK** combined with a strengthening of logical resources (V-logic), we can provide a coding of extensions of V that captures the desired properties in V whilst satisfying various constraints.

Before we provide our coding, we refresh the reader's memory concerning the three challenges we set up for coding of extensions of V. It will be helpful to keep these in mind, as we shall argue towards the end of the chapter that they are satisfied under the coding proposed.

> **The Hilbertian Challenge.** Provide philosophical reasons to legitimise the use of extra-V resources for formulating axioms and analysing intra-V consequences.
>
> **The Foundational Constraint.** In responding to the Hilbertian Chal-

lenge, do so in a way that does not necessitate the use of resources that cannot be represented by sets within V.

The Ontological Constraint. Any interpretation of extra-V resources should make clear the *ontological difference* between the interpretation of extensions and normal sets within V. In other words, any interpretation must make clear in what sense the interpretation does not *literally* refer to extra-V sets.

The Methodological Constraint. In responding to the Hilbertian Challenge, do so in a way that accounts for our naive thinking about extensions and links them to structural features of V.

We shall show, in the remainder of this chapter, that it is possible to provide a codification answering these challenges using $\mathbf{MK}$ class theory (plurally rendered) over V. The structure of the chapter is as follows: After these introductory remarks, §1 provides an exposition of the mechanisms of V-logic. §2 then presents some examples of how we can use V-logic to code extensions. §3 then shows how, if we were allowed to 'go past' V, we would be able to code a sufficient amount of V-logic to interpret extensions of V in $Hyp(V)$—the least admissible set containing V as an element. §4 provides a coding of $Hyp(V)$ with a single class in $\mathbf{MK}$, showing that extensions of V can be coded using class talk. Finally, §5 evaluates the coding from a philosophical perspective. In particular, we argue that the coding performs well with respect to the Hilbertian Challenge under the three constraints.

VII.1 Explaining V-logic

We begin by outlining the logical resources we shall use. The techniques of [Barwise, 1975] will be instrumental, as will the applications noticed by [Antos et al., 2015], [Friedman, S], and [Friedman and Ternullo, S].

VII.1.1 $\mathfrak{M}$-logic

We shall start with the general case. [Barwise, 1975] provides the following logical resources for capturing the properties of a structure. We start with a structure $\mathfrak{M} = (M, R)$ in the language $\mathscr{L}_\in$ and expand to a language $\mathscr{L}_\in^{\mathfrak{M}}$ as follows:

Definition 81. The language $\mathscr{L}_\in^{\mathfrak{M}}$ is $\mathscr{L}_\in$ together with:

(i) A predicate $\bar{\mathfrak{M}}$ to denote $\mathfrak{M}$.

(ii) A constant $\bar{x}$ for every $x \in M$.

We then define a satisfaction relation for structures of $\mathscr{L}_\in^{\mathfrak{M}}$ as follows:

Definition 82. A $\mathfrak{M}$-*structure* for $\mathscr{L}_\in^{\mathfrak{M}}$ is a structure $\mathfrak{N} = (N, E_{\mathfrak{N}})$ satisfying:

(i) The interpretation of $\bar{\mathfrak{M}}$ is $\mathfrak{M}$.

(ii) The interpretation of $\bar{x}$ in $\mathfrak{N}$ is x for every $x \in M$.

(iii) $\mathfrak{M}$ is a substructure of $(N, E_{\mathfrak{N}} \restriction \mathscr{L}_\in)$.

Definition 83. We write $\models_{\mathfrak{M}} \phi$ iff ϕ is true in all $\mathfrak{M}$-structures, and (letting $\mathbf{T}$ be a set of sentences) we write $\mathbf{T} \models_{\mathfrak{M}} \phi$ if ϕ is true in all $\mathfrak{M}$-structures satisfying $\mathbf{T}$.

Remark 84. Officially, we treat $\bar{\mathfrak{M}}$ as a *predicate* symbol, as it is not an object of $\mathfrak{M}$. However, it will be convenient for our purposes (as well as in line with relevant parts of the existing literature) if we render this talk in terms of membership (so $\bar{x} \in \bar{\mathfrak{M}}$ holds iff $\bar{\mathfrak{M}}(\bar{x})$).

We then consider $\mathfrak{M}$-*logic*, which has the following axioms:

Definition 85. $\mathfrak{M}$-*logic* is a system in $\mathscr{L}_\in^{\mathfrak{M}}$, with consequence relation $\vdash_{\mathfrak{M}}$ that consists of the following axioms:

(i) $\bar{x} \in \bar{\mathfrak{M}}$ for every $x \in M$.

(ii) Every atomic or negated atomic sentence of $\mathscr{L}_\in \cup \{\bar{m} | m \in \mathfrak{M}\}$ true in $\mathfrak{M}$ is an axiom of $\mathfrak{M}$-logic.

(iii) The usual axioms of first-order logic.

The rules of inference of $\mathfrak{M}$-logic are given as follows. For a set of sentences $\mathbf{T} \subseteq L^{\mathfrak{M}}_{\in}$:

(a) *Modus ponens*: From $\mathbf{T} \vdash_{\mathfrak{M}} \phi$ and $\mathbf{T} \vdash_{\mathfrak{M}} \phi \to \psi$ infer $\mathbf{T} \vdash_{\mathfrak{M}} \psi$.

(b) *The $\mathfrak{M}$-set rule*: From $\mathbf{T} \vdash_{\mathfrak{M}} \phi(\bar{b})$ for all $b \in a$ infer $\mathbf{T} \vdash_{\mathfrak{M}} \forall x \in \bar{a}\phi(x)$.

(c) *The $\mathfrak{M}$-rule*: From $\mathbf{T} \vdash_{\mathfrak{M}} \phi(\bar{b})$ for all $b \in M$, infer $\mathbf{T} \vdash_{\mathfrak{M}} \forall x \in \bar{\mathfrak{M}}\phi(x)$.

We furthermore say that a set of sentences $\mathbf{T}$ is *consistent in $\mathfrak{M}$-logic* iff $\mathbf{T} \vdash_{\mathfrak{M}} \phi \wedge \neg\phi$ is false for all formulas of $\mathscr{L}^{\mathfrak{M}}_{\in}$.

$\mathfrak{M}$-logic allows us to focus on models where $\mathfrak{M}$ appears as *standard*. We effectively introduce constants and axioms to rigidify the structure of $\mathfrak{M}$, which we can then talk about *syntactically* using $\mathfrak{M}$-logic. We then have the following:

Theorem 86. [Barwise, 1975] *The $\mathfrak{M}$-Soundness Theorem.* Let $\mathbf{T}$ be a set of sentences in $\mathscr{L}^{\mathfrak{M}}_{\in}$. Then if ϕ is a sentence of $\mathscr{L}^{\mathfrak{M}}_{\in}$:

$\mathbf{T} \vdash_{\mathfrak{M}} \phi$ implies that $\mathbf{T} \models_{\mathfrak{M}} \phi$.

Consistency of theories in $\mathfrak{M}$-logic is linked to the existence of models by the following theorem:

Theorem 87. [Barwise, 1975] *The $\mathfrak{M}$-Completeness Theorem.* Let $\mathfrak{M}$ be countable and $\mathbf{T}$ be a set of sentences of $\mathscr{L}^{\mathfrak{M}}_{\in}$. Then if ϕ is a sentence of $\mathscr{L}^{\mathfrak{M}}_{\in}$:

$\mathbf{T} \models_{\mathfrak{M}} \phi$ implies that $\mathbf{T} \vdash_{\mathfrak{M}} \phi$

Thus, if $\mathfrak{M}$ is *countable*, then if a theory $\mathbf{T}$ is consistent in $\mathfrak{M}$-logic, there is a model of $\mathbf{T}$. Of interest for us will be cases where $\mathfrak{M}$ is *uncountable*. Here, the $\mathfrak{M}$-logic completeness theorem radically fails:

Theorem 88. [Barwise, 1975][1] Let $\mathfrak{M}$ be an uncountable structure. Then $\mathfrak{M}$-logic is not even Σ_1-complete.

[1]For details, see discussion in Ch VII of [Barwise, 1975].

The theorem shows that, in terms of metalogical properties, $\mathfrak{M}$-logic is far closer to second-order logic than first-order logic when $\mathfrak{M}$ is uncountable. In this way, the consistency of a theory in $\mathfrak{M}$-logic is no guarantee of there being a model with those properties.

As we shall see, $\mathfrak{M}$-logic becomes exceptionally useful for interpreting extensions of V when we let $\mathfrak{M} = V$.

VII.1.2 V-logic

Certain applications of $\mathfrak{M}$-logic may be familiar to the reader, the example of ω-logic is precisely one such.[2] We now explain V-logic.[3] Quite simply, V-logic is the special case of $\mathfrak{M}$-logic where $\mathfrak{M} = V$. So, exactly as before, we have:

Definition 89. $\mathscr{L}_{\in}^{V}$ is the language consisting of:

(i) A predicate $\bar{V}$ to denote V.

(ii) A constant $\bar{x}$ for every $x \in V$.

Again, as before, we can then define V-logic:

Definition 90. *V-logic* is a system in $\mathscr{L}_{\in}^{V}$, with consequence relation $\vdash_V$ that consists of the following axioms:

(i) $\bar{x} \in \bar{V}$ for every $x \in V$.

(ii) Every atomic or negated atomic sentence of $\mathscr{L}_{\in} \cup \{\bar{x} | x \in V\}$ true in V is an axiom of V-logic.

(iii) The usual axioms of first-order logic in $\mathscr{L}_{\in}^{V}$.

For a set of sentences $\mathbf{T} \subseteq \mathscr{L}_{\in}^{V}$, V-logic contains the following rules of inference:

(a) *Modus ponens*: From $\mathbf{T} \vdash_V \phi$ and $\mathbf{T} \vdash_V \phi \rightarrow \psi$ infer $\mathbf{T} \vdash_V \psi$.

[2] ω-logic adds a predicate $\bar{\mathbb{N}}$ for $\mathbb{N}$ and constants $\bar{n}$ for every standard natural number to the language of **PA**, and then permits inference by the ω-rule: that from $\phi(\bar{n})$ for every $n \in \mathbb{N}$, one may infer $\forall x \bar{\mathbb{N}}(x) \rightarrow \phi(x)$. Essentially, ω-logic is $\mathfrak{M}$-logic for the case of $\mathbb{N}$. See [Barwise, 1975] and [Shapiro, 1991] for an examination of the metalogical properties of ω-logic.

[3] I am very grateful to Sy Friedman for explaining to me the mechanisms of V-logic, and its role in interpreting forcing extensions. See, for more discussion, [Antos et al., 2015], [Friedman, S], and [Friedman and Ternullo, S].

(b) *The Set-rule*: From $\mathbf{T} \vdash_V \phi(\bar{b})$ for all $b \in a$ infer $\mathbf{T} \vdash_V \forall x \in \bar{a} \phi(x)$.

(c) *The V-rule*: From $\mathbf{T} \vdash_V \phi(\bar{b})$ for all $b \in V$, infer $\mathbf{T} \vdash_V \forall x \in \bar{V} \phi(x)$.

We furthermore say that a set of sentences $\mathbf{T}$ is *consistent in V-logic* iff $\mathbf{T} \vdash_V \phi \wedge \neg\phi$ is false for all formulas of $\mathscr{L}_{\in}^{V}$.

Proof codes in V-logic are thus (possibly infinite) well-founded trees with root the conclusion of the proof. Whenever there is an application of the V-rule, we get proper-class-many branches extending from a single node. More formally, we define the notion of a *proof code* in V-logic (an example of which is visually represented in Figure VII.1) as follows:

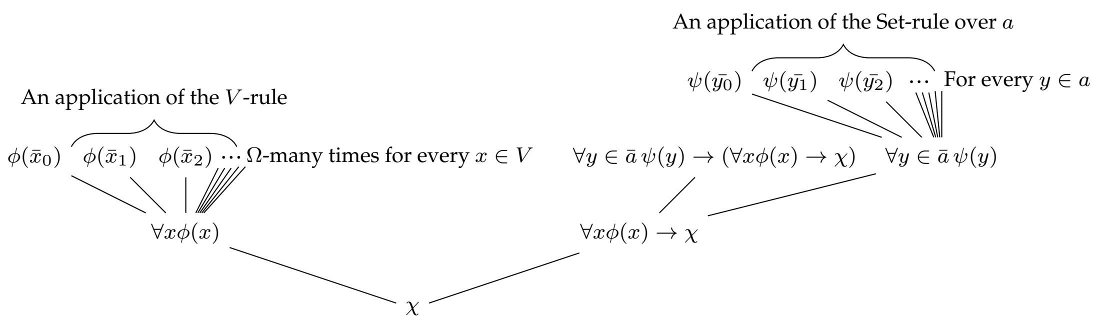
Definition 91. A *proof of χ in V-logic* is a (possibly infinite) well-founded tree, with root the conclusion of the proof (i.e. χ) and where previous nodes are either axioms of V-logic or follow from one of its inference rules.[4]

Definition 92. We furthermore say that a set of sentences $\mathbf{T}$ is *consistent in V-logic* iff $\mathbf{T} \vdash_V \phi \wedge \neg\phi$ is false for all formulas of $\mathscr{L}_{\in}^{V}$.

With the mechanisms of V-logic set up, we now describe how its use is relevant for interpreting extensions of universes.

[4]The eagle-eyed reader may notice that there will be choices to be made about how to code these trees. We shall see discussion of this later.

Figure VII.1: Visual representation of a proof of χ in V-logic



VII.2 Interpreting extensions in V-logic

We have provided explanation of a logic that rigidifies the structure of V; adding constants and axioms to fix V's properties within the syntax of V-logic. How might this allow us to interpret extensions of V? The key fact (developed in [Antos et al., 2015], [Friedman, S], and [Friedman and Ternullo, S]) is that *consistency* in V-logic of theories in $\mathscr{L}_{\in}^{V}$ corresponds to those intra-V consequences holding. We now provide some exposition of how this proceeds:

VII.2.1 Height extensions

We first deal with direct extensions. We wish to develop a theory in V-logic that allows us to talk about a height extension of V. We do this by first adding a constant $\bar{W}$ to our language $\mathscr{L}_{\in}^{V}$. We then add the following axiom:

> $\bar{W}$ is a universe of ZFC with $\bar{V}$ as an initial segment such that Φ.

In particular, in the context of formulating the Extended Reflection Axiom, we have the following axioms in the language of V-logic:

(i) *$\bar{W}$-Height Axiom.* $\bar{W}$ is a universe of ZFC that has $\bar{V}$ as an initial segment.

(ii) *The $\bar{W}$-ERA.* For all first-order formulas ϕ and subclasses $A \subseteq \bar{V}$, if $\phi(A)$ holds in $\bar{W}$ then $\phi(A \cap V_{\bar{\alpha}}^{\bar{V}})$ holds in $V_{\bar{\beta}}^{\bar{V}}$ for some pair of ordinals $\bar{\alpha}$ and $\bar{\beta}$ in $\bar{V}$.

We can then capture the content of the ERA with the following claim:

> *Extended Reflection Axiom** The theory of V-logic with the $\bar{W}$-Height Axiom and $\bar{W}$-ERA added is consistent under $\vdash_V$.

If such a theory is consistent in V-logic, then we can speak as if there is such an extension of V. This is because, in speaking of $\bar{W}$ *syntactically*, we prove *syntactic* facts about $\bar{V}$ and particular $\bar{x}$. However, since the syntax of V-logic keeps V fixed (in the sense that the interpretation of $\bar{V}$ is V and every $\bar{x}$ is x) any syntactic consequence derived from the existence of $\bar{W}$ is mirrored in the *actual* structure of sets in V. Thus, for the purposes of speaking about V, we can speak *as if* there actually is such a $\bar{W}$.

VII.2.2 Width extensions

This also works for width extensions. Again we introduce a constant $\bar{W}$. This time, however, the relevant axioms are of the form:

(i) $\bar{W}$-*Width Axiom.* $\bar{W}$ is a universe satisfying ZFC (or possibly simply KP)[5] with the same ordinals as $\bar{V}$ and containing $\bar{V}$ as a proper subclass.

(ii) $\bar{W}$-Φ-*Width Axiom.* $\bar{W}$ is such that Φ.

We can then have the following axiom to capture the intra-V consequences of an extension such that Φ:

> Φ^*-*Axiom.* The theory in V-logic with the $\bar{W}$-Width Axiom and $\bar{W}$-Φ-Width Axiom is consistent under $\vdash_V$.

This captures the intra-V consequences of any such axiom. Just as with height extensions, any syntactic consequence concerning either some $\bar{x}$ or $\bar{V}$ derived from the axioms mentioning $\bar{W}$ will hold of the respective *actual* structures.

To see this more clearly, let us examine V-logic in action with respect to some of the examples we outlined in Chapter III.

In the case of a set forcing we could have the following:

> $\bar{W}$-*G*-*Width Axiom.* $\bar{W}$ is such that it contains some $\bar{V}$-$\bar{\mathbb{P}}$-generic G.

If the resulting V-logic theory is consistent, then any syntactic consequence of the existence of $\bar{W}$ concerning $\bar{V}$ will then be true in V. In particular, formulations of both generic embeddings and remarkable cardinals are thereby dealt with: the behaviour of the relevant objects in the extension can be formulated via consideration of appropriate $\bar{W}$. The situation with class forcing is similar, but with a small twist. For, in the case of class forcing using some class poset $\mathbb{P}^C$, the existence of a V-$\mathbb{P}^C$ generic G^C is not a *first-order* property of $\bar{W}$. Despite this, in V-logic we have the ability to add predicates (as we did with $\bar{V}$ and $\bar{W}$). Thus, we can add additional predicates $\bar{\mathbb{P}^C}$ and $\bar{G^C}$ for $\mathbb{P}^C$ and G^C into the usual syntax of V-logic, and state the following axiom:

[5] KP is a significant weakening of ZFC. We shall provide an exposition of KP later in the chapter and see some of the theory's interesting features with respect to V-logic. For now, we merely note that KP satisfaction is a minimal requirement on $\bar{W}$.

> $\bar{W}$-G^C-*Width Axiom.* $\bar{W}$ is such that $\bar{G^C} \subseteq \bar{W}$ and $\bar{G^C}$ is $\bar{\mathbb{P}^C}$-generic over V.

and then examine whether the resulting theory is consistent in V-logic. Any intra-V consequence of such a (consistent) theory would, for exactly the same reasons as in the case of set forcing, naturally transfer to truths concerning V.

Note here that a problem regarding interpreting class forcing is immediately resolved. Earlier (III.1.4) we remarked that there are class forcings that add only classes to the model without adding sets, and that this represented an additional challenge in interpreting extensions. Moreover, given that we have argued earlier (Chapter VI) that we should interpret quantifiction over classes via plural quantification, one might think that there is the difficulty that the addition of a plural parameter (such as G^C) would necessitate the addition of sets (since one cannot add some sets without adding a set). Reflection on the method of interpretation dissolves this problem. Since all reference to $\bar{G^C}$ and $\bar{W}$ is being interpreted as purely *ideal* and *syntactic* in V-logic, we can perfectly well consider $\bar{W}$ with additions of classes but no extra sets. Since *every* 'extension' is being interpreted in a purely syntactic manner, *no* 'extension' of V literally adds sets (or any other object for that matter), and so we can use any theory whatsoever that is consistent in V-logic without worrying about specific puzzles about ontology within the extension.

This liberation of methods via syntactic means also use allows us to formulate axioms that capture *non*-forcing extensions. For example:

> $\bar{W}$-$\sharp$-*Width Axiom.* $\bar{W}$ is such that it contains a sharp that generates $\bar{V}$.

This then allows us to express the claim that V is sharp generated:

> *Axiom*-$(\sharp)^*$. The theory in V-logic with the $\bar{W}$-Width Axiom and $\bar{W}$-$\sharp$-Width Axiom is consistent under $\vdash_V$.

For exactly the same reasons as before, this allows us to interpret the claim that V is sharp-generated, with all the consequences of sharp-generation that we would like.

So far, we have only showed that the interpretation of extensions of V using V-logic performs at least as well as the countable transitive model strategy and use of $\mathfrak{V}$. Recall, however, that the problem with that strategy was that it could not account for axioms that were greater than first-order over V. We now show that V-logic can be put to this task.

VII.2.3 Inner model density: redux

We noted earlier (in Chapter V) that all the above could be done with a conservative extension of **ZFC**, by conducting the construction over a countable transitive model $\mathfrak{V}$ that satisfied exactly the same first-order sentences as V. It was noted there, however, that the production of such a model provided no guarantee that the model would respect greater than first-order features of V, in particular the density of inner models provided by the IMH. The key fact here is that now we have the notion of interpreting extensions via consistency in V-logic, we are able to simulate statements about the existence of models.

Again, we add a constant $\bar{W}$ to our language and formulate axioms concerning width extensions represented syntactically by the relevant $\bar{W}$. We can then express the intended content of the IMH as follows:

> IMH^*. Suppose that ϕ is a first-order sentence. Let $\mathbf{T}$ be a V-logic theory containing the $\bar{W}$-Width Axiom and also the $\bar{W}$-ϕ-Width Axiom (i.e. $\bar{W}$ satisfies ϕ). Then if $\mathbf{T}$ is consistent under $\vdash_V$, there is an inner model of V satisfying ϕ.

Thus, by interpreting the existence of outer models through the consistency of theories, we can now make claims concerning consequences (about V) of the existence of outer models. In particular, we can say that if ϕ is satisfiable in an extension of V (syntactically formulated as $\bar{W}$) then it is satisfied in an inner model of V. So, the IMH^* holds iff, (if the mathematical structure of V does not preclude the V-logic consistency of an outer model satisfying ϕ, then V has an inner model satisfying ϕ). In this way, we make claims concerning greater than first-order properties of V needed to express the IMH in its maximal sense.

Thus, if we allow the use of V-logic, we are able to syntactically code the effects of extensions of V on V. However, it is one thing to have provided a system for interpreting extensions, and quite another to argue that it is acceptable on the Universist's philosophical picture. This will be our focus for the rest of the chapter.

VII.3 V-logic and admissibility

Thus far, we have provided exposition of a logical system that can simulate talk of arbitrary extensions of V syntactically. An immediate question is whether or not it is possible to get a handle on this system through consideration of V. Our methodology will be slightly unusual. We will drop, for a moment, the Universist perspective and assume that height extensions of V are available.[6] We will show that if ϕ is a consequence of a V-logic theory $\mathbf{T}$, then a proof of ϕ appears in a mild lengthening of V known as $Hyp(V)$ (i.e. the least admissible set containing V). We then argue that in fact $Hyp(V)$ can be coded within the class theory set up in the previous chapter, allowing us to code extensions using sets from V.[7]

We first set up the system of *Kripke-Platek* set theory KP.

Definition 93. *Kripke-Platek Set Theory* (or simply 'KP') comprises the following axioms:

(i) Extensionality

(ii) Union

(iii) Pairing

(iv) Foundation

(v) Δ_0-Separation: If ϕ is a Δ_0 formula in which b does not occur free, then:

$\forall a \exists b \forall x [x \in b \leftrightarrow (x \in a \wedge \phi(x))]$

Intuitive characterisation: If ϕ is a Δ_0 formula then, given a set a, we can separate those members of a which satisfy ϕ into a new set b.

[6] In this, we begin by following the methodology of [Antos et al., 2015], [Friedman, S], and [Friedman and Ternullo, S].

[7] I am especially grateful to Sy Friedman for his help with the technical exposition here.

(vi) Δ_0-Collection: If ϕ is a Δ_0 formula in which b does not occur free:

$\forall a[\forall x \in a \exists y \phi(x, y) \to \exists b \forall x \in a \exists y \in b \phi(x, y)]$

Intuitive characterisation: If ϕ is a Δ_0 formula then, given a set a, if for every x in a there is some y that is ϕ-related to x, then there is some b such that everything in a has a ϕ-related object in b.

We make a further pair of definitions:

Definition 94. A set $\mathfrak{N}$ is *admissible over* $\mathfrak{M}$ iff $\mathfrak{N}$ is a model of KP containing $\mathfrak{M}$ as an element.

Definition 95. $Hyp(\mathfrak{M})$ is the smallest transitive x such that x is admissible over $\mathfrak{M}$.

Our interest will be in $Hyp(V)$; the least admissible set containing V as an element. We wish to show that if there is a proof of ϕ in V-logic, then there is a proof code of ϕ in $Hyp(V)$. First, however, we must be precise about how we interpret the extended syntax of V-logic within this height extension:

Definition 96. The language and proofs of V-logic are interepreted as follows:

(i) Every set x is named by $\langle x, 3 \rangle$ (so, for example, if $x = \omega$, then $\bar{\omega} = \langle \omega, 3 \rangle$ (to avoid the double use of names for natural numbers and the Gödel coding of the connectives).

(ii) $\in^V$ and V name $\in$ and $\bar{V}$ respectively (remembering that we are currently allowing ourselves height extensions so that these are sets).

(iii) The relevant $\bar{W}$ (and possibly $\bar{\mathbb{P}^C}$, $\bar{G^C}$, or any other required predicates) can be represented by any object not otherwise required for the syntax of V-logic, so we may use $\{V\}, \{\{V\}\}, \{\{\{V\}\}\}, ...$ and so on (for any Zermelo-style construction of singletons derived from V).[8]

[8]For most axioms we only need one (and at most three) extra predicates, but we make room for the use of several different outer models in case others wish to talk about relationships between *incompatible* extensions using the same axiom.

(iv) After a suitable Gödel coding for the connectives and quantifiers has been chosen, we represent well-formed formulas of V-logic with sequences of symbol codes.

(v) Proofs are represented by the appropriate trees comprising codes of the relevant sentences as nodes.

We can now state and prove the following:

Theorem 97. [Barwise, 1975] Suppose that there is a proof of ϕ in V-logic. Then there is a proof code of ϕ in $Hyp(V)$.

Proof. We first need to identify the following:

Lemma 98. [Barwise, 1975] For any $\mathfrak{M}$, $Hyp(\mathfrak{M})$ is of the form $L_\alpha(\mathfrak{M})$ for some ordinal α.

As noted earlier, since we will be going 'above' V (and then showing how to code this acceptably later), we will refer to ideal ordinals 'past' V using variants of the Greek letters (such as 'ϑ'). Since we are currently working from a perspective on which we assume that height extensions are available, we are now interested in a structure $Hyp(V) = L_\vartheta(V)$.

We proceed by induction on the complexity of our proof P. Suppose that there is a proof P of ϕ in V-logic. Then either:

(a) P is one line.

(b) P is more than one line.

We deal with (a) first. Suppose that P is one line. Then either (i) ϕ is of the form $\bar{x} \in \bar{V}$, (ii) ϕ is an atomic or negated atomic sentence of $\mathscr{L}_\in \cup \{\bar{x}|x \in V\}$, (iii) ϕ is an axiom of first-order logic in $\mathscr{L}_\in^V$, or (iv) ϕ is an additional axiom containing some extra predicate (such as $\bar{W}$, $\bar{\mathbb{P}^C}$, or $\bar{G^C}$).

For (i), suppose ϕ is of the form $\bar{x} \in \bar{V}$ for $x \in V$. Then, the result is immediate: the required sentence is in $Hyp(V)$ by Pairing, and hence so is the tree coding its proof (i.e. $\{\{\phi\}, \emptyset\}$). In the case of (ii), ϕ appears in V, and so it is immediate that ϕ is in $Hyp(V)$, with the relevant proof tree. For (iii), we note that that all constructions of first-order axioms from simpler formulas ψ and χ (that are assumed, for induction, to be in $Hyp(V)$) can be chained together through Pairing. For (iv), since we represent the various extra predicates by objects that are not pieces of syntax in other parts of V-logic but are in $Hyp(V)$, any axiom of the form "$\bar{W}$ is such that Φ" is simply a finite sequence of sets already present in $Hyp(V)$ (and similarly when $\bar{\mathbb{P}^C}$ or $\bar{G^C}$ are present). Again, repeated application of Pairing ensures that ϕ is in $Hyp(V)$, as well as the relevant proof tree.

(b) Suppose then that P is more than one line. Assume for induction that all prior steps to the final inference to ϕ have proofs in $Hyp(V)$. Then either (i) ϕ is an axiom, or (ii) ϕ follows from ψ, $(\psi \rightarrow \phi)$ via modus ponens, or (iii) ϕ is of the form $\forall x \in \bar{a}\psi(x)$ and follows from $\psi(\bar{b})$ for all $b \in a$ by the Set-rule, or (iv) ϕ is of the form $\forall x \in \bar{V}\psi(x)$ and follows from $\psi(\bar{x})$ for all $x \in V$ by the V-rule.

For each of the steps we need to construct, from the given proof trees, a new proof tree coding a proof of ϕ. We already know that the relevant pieces of syntax exist (by part (a)) and so the challenge is simply in the construction of the trees in $Hyp(V)$.

(i) has already been dealt with in part (a). (ii) Suppose for induction that ψ and $(\psi \rightarrow \phi)$ have proofs coded in $Hyp(V)$ by $\mathbb{T}_\psi = \langle T_\psi, <_\psi \rangle$ and $\mathbb{T}_{(\psi\rightarrow\phi)} = \langle T_{(\psi\rightarrow\phi)}, <_{(\psi\rightarrow\phi)} \rangle$. Since we know that $Hyp(V)$ satisfies finite iterations of Pairing, we only need to construct T_ϕ and $<_\phi$. We can easily construct $T_\phi = T_\psi \cup T_{(\psi\rightarrow\phi)} \cup \{\phi\}$. Next, we define $<_\phi$ as follows:

$x <_\phi y$ iff:

(i) $x <_\psi y$, or

(ii) $x <_{(\psi\rightarrow\phi)} y$, or

(iii) $y = \phi$.

Since we have $<_\psi$ and $<_{(\psi\to\phi)}$ already ($Hyp(V)$ is transitive), we just need to construct $\{\langle x, \phi\rangle | x \in T_\psi \vee x \in T_{(\psi\to\phi)}\}$. We have that $\phi \in Hyp(V)$, and also for any object $y \in Hyp(V)$, $\langle y, \phi\rangle \in Hyp(V)$. We then (working with $Hyp(V)$) define the following formula:

$\chi(x, y) =_{df} x \in T_\psi \cup T_{(\psi\to\phi)} \wedge y = \langle x, \phi\rangle$

$\chi(x, y)$ is clearly a Δ_0 formula defining a function that maps any particular $x \in T_\psi \cup T_{(\psi\to\phi)}$ to $\langle x, \phi\rangle$. We also have the following lemma:

Lemma 99. [Barwise, 1975] Σ_1-*Replacement.* $Hyp(V)$ satisfies Replacement for Σ_1 formulas.

We thus have $\{\langle x, \phi\rangle | x \in T_\psi \wedge x \in T_{(\psi\to\phi)}\} \in Hyp(V)$ as desired. $\langle T_\phi, <_\phi\rangle$ clearly codes a proof of ϕ from ψ and $(\psi \to \phi)$, the proof steps are inherited from the previous trees, and each proof step prior to ϕ is $<_\phi$-related to ϕ.

We deal with (iii) and (iv) in tandem. As the strategy is the same for both, we give only the proof of (iv). Suppose then that ϕ is of the form $\forall x \in \bar{V}\psi(x)$ and follows from $\psi(\bar{b})$ for all $b \in V$ by the V-rule. Assume for induction that every $\psi(\bar{b})$ has a proof code $\mathbb{T}_{\psi(\bar{b})} = \langle T_{\psi(\bar{b})}, <_{\psi(\bar{b})}\rangle \in Hyp(V)$. Since $Hyp(V) = L_\vartheta(V)$ and contains a well-ordering of V (by Global Choice in V) $Hyp(V)$ has a Σ_1-definable well-order R. Using this well-order, we define the following function:

$\chi(x, y) =_{df} x \in V \wedge$ "y is the R-least code of a proof tree of $\phi(\bar{x})$"

By Σ_1-Replacement, we then have a set $\{\mathbb{T}_{\psi(\bar{b})} | b \in V\} \in Hyp(V)$. By transitivity of $Hyp(V)$, a further application of Σ_1-Replacement, and Paring with $\{\phi\}$, we obtain T_ϕ (and the union of the $<_{\psi(\bar{b})}$). The argument for the existence of $<_\phi$ is then exactly the same as in (ii). Since we have T_ϕ, we can map each $x \in T_\phi$ such that $x \neq \phi$ to $\langle x, \phi\rangle$ to obtain (by Σ_1-Replacement) the set of all $\langle x, \phi\rangle$. Again, by Transitivity, Union, and Pairing, $<_\phi \in Hyp(V)$. This concludes the proof.

■

The above result, translated from the work of [Barwise, 1975] to the context of V, shows that if height extensions of V are available, we can code extensions of V. However, for the Universist, we do not yet have a satisfactory coding. If we wish to make use of the above result, we have to show how the 'ideal' $Hyp(V)$ can be coded using only sets from within V.

VII.4 Coding $Hyp(V)$ using MK

We now are in a position where:

(1) Extensions of V can be coded syntactically using V-logic.

(2) If ϕ is provable in V-logic, then ϕ has a proof code in $Hyp(V)$.

Of course, we are not yet able to say that we have coded extensions of V using sets from V; the current coding depends upon going 'past' all the ordinals to a level $L_\vartheta(V) = Hyp(V)$. What we want to show now is that we can code $Hyp(V)$ in the theory of **MK**. We shall in fact do a little more. We code a model $(V)^+$, built up from V, that satisfies Δ_n-Separation for every n and in which $Hyp(V)$ appears as a single ideal set (and hence is coded by a some things hh).

We first provide a sketch of how this can be done. The bulk of the work occurs in [Antos and Friedman, S] and [Antos, 2015]. There, however, additional axioms were required for certain applications[9] of the coding, and since we have only justified **MK** thus far we proceed in plain **MK**.[10]

We first need to code the notion of an *ordered pair* of classes. Initially, this seems problematic; normal pairing functions on sets are type-raising in the sense that they have the things they pair as members. This is not available on the current interpretation; we adopt a class nominalism with class theory interpreted *plurally*, and so cannot speak of classes being members. Despite this, we can use the abundance of sets within V (in particular the closure of V under pairing) to code pairs:

[9]Namely *hyper*class forcing.

[10]Antos and Friedman use a version of *Class Bounding*, equivalent (modulo **MK**) to AC_∞; a particular kind of choice principle for classes. It is an interesting question (though one I lack the space to address here) whether or not a Universist should accept that such a principle holds of V, not least because the principle is necessary for a good deal of mathematical work (see, for example, [Gitman et al., U]). The coding goes through far more easily with the principle, but since we do not need it for the purpose of representing extensions of V, we show the coding works without it.

Definition 100. Let X and Y be classes. We define those things that *represent the ordered pair of X and Y*, or '$REP(\langle X, Y\rangle)$' as follows:

$REP(\langle X, Y\rangle) = \{\langle z, i\rangle | (z \in^* X \wedge i = 1) \vee (z \in^* Y \wedge i = 2)\}$

Effectively, we talk about coding an ordered pair of classes by tagging all the members of X with 1 and all the members of Y with 2, and referring to the resulting class. A possible problem here is that we will have certain classes that we cannot distinguish from pairs of classes. For example, under the above coding the class:

$Z = \{\langle z, i\rangle | (z \in^* X \wedge i = 1) \vee (z \in^* Y \wedge i = 2)\}$

will refer to the same things as the class representing $\langle X, Y\rangle$, which is undesirable. The problem is simply avoided, we just let each class be represented as follows:

$REP(Z) = \{\langle z, 0\rangle | z \in^* Z\}$

The raising of type of each the members of Z avoids the previous problem: by tagging the members of a class with 0, we uniformly increase the type of members of Z. This provides the means to distinguish between our original classes and their pairs. Effectively, this tagging procedure produces a 'copy' of V (and its classes) for each tag, allowing us to code relationships between them.

Remark 101. The above coding gives us a natural way of coding proper-class-sized structures. Letting, for the moment, E_V denote those ordered pairs that comprise the $\in$-relation, we can now represent (V, E_V) as follows:

$REP((V, E_V)) = \{\langle z, i\rangle | (z \in^* V \wedge i = 1) \vee (z \in^* E_V \wedge i = 2)\}$

This is in turn, distinct from:

$REP(\{\langle z, i\rangle | (z \in^* V \wedge i = 1) \vee (z \in^* E_V \wedge i = 2)\}) =$

$\{\langle x, 0\rangle | x \in^* \{\langle z, i\rangle | (z \in^* V \wedge i = 1) \vee (z \in^* E_V \wedge i = 2)\}\}$

We now have a method for coding pairs of classes (moreover, short reflection on the above coding shows that it is easy to generalise this to ordered α-tuples by letting α tag α-many copies of V). We are now able to make use of the following definition (with visual representation[11] of some examples provided in Figures VII.2 and VII.3):

[11] I am grateful to Carolin Antos and Sy Friedman for their kind permission to include these diagrams.

Figure VII.2: An example of a tree corresponding to a coding pair showing membership.

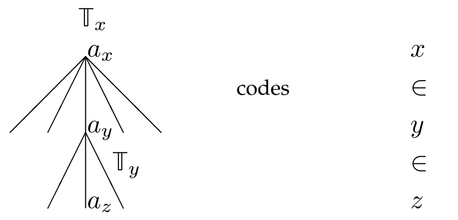
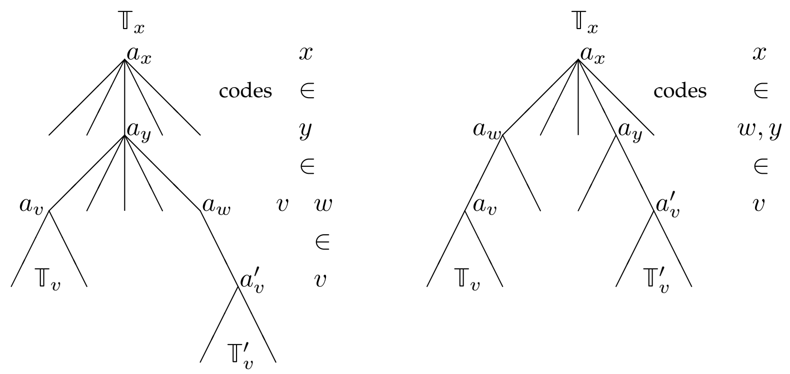


Figure VII.3: More complicated coding pair tree structures.



Definition 102. [Antos, 2015], [Antos and Friedman, S] A pair $\langle M_0, R\rangle$ is a *coding pair* iff M_0 is a class with distinguished element[12] a, and R is a class binary relation on M_0 such that:

(i) $\forall z \in^* M_0 \exists! n$ such that z has R-distance n from a (i.e. for any element z of M_0, z is a single finite R-distance away from a), and

(ii) let $\langle M_0, R\rangle \restriction x$ denote the R-transitive closure below x. Then if $x, y, z \in^* M_0$ with $y \neq z$, yRx, and zRx, then $\langle M_0, R\rangle \restriction y$ is not isomorphic to $\langle M_0, R\rangle \restriction z$ (and respectively for z), and

(iii) if $y, z \in^* M_0$ have the same R-distance from a, and $y \neq z$, then for all v, $vRy \to \neg vRz$, and

(iv) R is well-founded.

These coding pairs shall be essential in coding the structure of ideal sets that would have to be 'above' V were they to exist. One can think of the coding pair as a

[12]Such an element exists by Global Choice.

tree $\mathbb{T}$ which has as its nodes the elements of M_0, a top node of a, and R the extension relation of $\mathbb{T}$. For each tree there are only countably many levels, but each level can have proper-class-many nodes.

Next, we code the ideal objects of $(V)^+$ using coding pairs. For any particular ideal set x in $(V)^+$, we will code the transitive closure of $\{x\}$. A fact of the above coding is that any tree will have many isomorphic subtrees, and hence will not be isomorphic to $TC(\{x\})$.[13] We therefore need to form *quotient* pairs that provide a coding of ideal sets in $(V)^+$ (we provide a visual representation[14] of an example in Figure VII.4).

Definition 103. [Antos, 2015], [Antos and Friedman, S] *Quotient Pairs.* Let $\langle M_0, R\rangle$ be a coding pair and a be a set in M_0. We then define the equivalence class of a (denoted by '$[a]$') of all top nodes of the associated coding tree isomorphic to the subtree $\mathbb{T}_a$:

$[a] = \{b \in^* M_0 |$"$\langle M_0, R\rangle \restriction b$ is isomorphic to $\langle M_0, R\rangle \restriction a$"$\}$

Since we have Global Choice, we let $\tilde{a}$ be a fixed representative of $[a]$. We then define the *quotient pair* $\langle \tilde{M}_0, \tilde{R}\rangle$ as follows:

$\tilde{M}_0 = \{\tilde{a}|$ "$\tilde{a}$ is the representative of the class $[a]$ for all $a \in^* M_0\}$

$\tilde{a}\tilde{R}\tilde{b}$ iff "There is an $a_0 \in^* [a]$ and a $b_0 \in^* [b]$ such that $a_0 R b_0$."

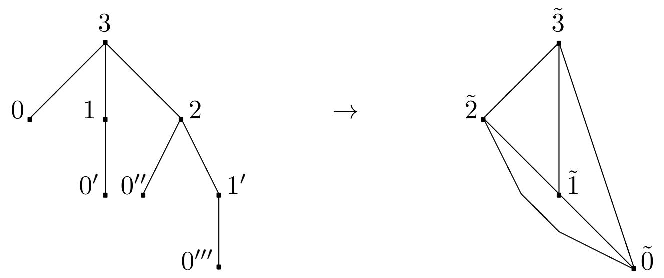
Remark 104. The quotient pairs work by taking fixed representatives of the equivalence class of top nodes of isomorphic subtrees. We then define the relation on these representatives by searching through the equivalence classes to find a relevant subtree in which two members of the equivalence class are R-related.

We now have a quotient structure for the coding pairs. Next, we begin to verify some of the properties of these coding pairs for the purposes of showing that they code ideal sets:

[13]See [Antos, 2015] and [Antos and Friedman, S] for details and further explanation. We would like subtrees to correspond to elements in the transitive closure. However, isomorphic subtrees would code the same element of $(V)^+$, and so as it stands our coding pairs are not extensional.

[14]Again, thanks are due to Carolin Antos and Sy Friedman for their kind permission to include the diagrams.

Figure VII.4: The quotient process for a coding pair of 3.



Lemma 105. [Antos, 2015], [Antos and Friedman, S] Let $\langle M_0, R\rangle$ be a coding pair. Then the quotient pair $\langle \tilde{M_0}, \tilde{R}\rangle$ is extensional and well-founded.

Proof.

Well-foundedness. By Class Comprehension $\tilde{R}$ refers to some things in V (we have provided its definition above). Take any things $\tilde{x}\tilde{x}$ that are some of $\tilde{M_0}$. We claim that $\tilde{x}\tilde{x}$ have a $\tilde{R}$-minimal element $\tilde{a}$. In particular, the $\tilde{x}\tilde{x}$ that are some of $\tilde{M_0}$ can be traced back to some things xx that are some of M_0, and $\tilde{a}$ is then the representative of $[a]$ for the R-minimal object $a \prec xx$. Suppose $\tilde{a}$ is not $\tilde{R}$-minimal. Then there is some $\tilde{a}'$ such that $\tilde{a}'\tilde{R}\tilde{a}$. Therefore, there is an $a_0' \in^* [a_0']$ such that $a_0'Ra$, contradicting the claim that a was R-minimal, $\perp$.

Extensionality. Assume for contradiction that $\tilde{y} \in^* \tilde{M_0}$ and $\tilde{z} \in^* \tilde{M_0}$, with $\tilde{y} \neq \tilde{z}$ and $\{\tilde{x}|\tilde{x}\tilde{R}\tilde{y}\} = \{\tilde{x}|\tilde{x}\tilde{R}\tilde{z}\}$. By the definition of $\tilde{R}$, this means that within the equivalence classes $[y]$ and $[z]$, we have for every x_0, y_0, $z_0 \in^* M_0$ such that $x_0 R y_0$, with $y_0 \in^* [y]$ and $z_0 \in^* [z]$ there is an x_1 with $x_1 R z_0$ such that $x_0, x_1 \in^* [x]$. By property (ii) of the definition of the coding pair $\langle M_0, R\rangle$, we then know that $[y] = [z]$ as there are no isomorphic subtrees connected to the same R-predecessor (by (ii) in the definition of coding pairs). Thus $\tilde{y} = \tilde{z}$, $\perp$. ■

Let us take stock. We have quotient structures of coding pairs that behave extensionally and in a well-founded manner, and are coded by individual classes. We are now ready to establish the main theorem, showing that these quotient coding pairs obey certain operations, and thus we have a code for $Hyp(V)$. We first, however, explain how we will state what we wish to say. We will be interested in the quotient coding pairs coding $(V)^+$ and satisfying a certain theory *together*. However, we have a challenge here, as $(V)^+$ is not even represented by a second-order class, but rather appears to be a *collection* of $\langle \tilde{M}_0, \tilde{R} \rangle$. This is *not* going to be countenanced by the Universist on the current way of understanding classes; understood through plural reference, classes cannot be members. We therefore need to modify how we express the satisfaction of different axioms by the quotient pairs.

We begin with the following:

Remark 106. We will talk about 'a structure' $(V)^+$ under a 'relation' $\widehat{\in}$ that we will show satisfies certain *first-order* axioms. We will therefore use locutions like "$x \widehat{\in} y$", "$x \widehat{\in} (V)^+$", and more generally "$\phi(x)$" for some first-order ϕ. Really, however, since the objects of $(V)^+$ are *ideal*, this should all be paraphrased in terms of quotient coding pairs. Though we *deny* that we are using singular reference to refer to *objects*, there is no obstruction to using a first-order language to represent the theory of quotient coding pairs (just as there is no contradiction in using a two-sorted first-order language to talk plurally). We therefore make the following definitions:

Definition 107. Suppose there are some things xx coding a quotient pair $\langle \tilde{M}_0, \tilde{R} \rangle$, which in turn codes some $x \widehat{\in} (V)^+$. Since it will often be useful to talk about the xx coding $\langle \tilde{M}_0, \tilde{R} \rangle$ under their presentation as a tree coding x, we say that:

$\mathbb{T}_x$ is a *quotient pair tree for* x iff there are some things xx representing $\langle \tilde{M}_0, \tilde{R} \rangle$ coding x, and $\mathbb{T}_x$ is the tree structure that the xx collectively exemplify.

Definition 108. Say there are some things xx coding a quotient pair tree $\mathbb{T}_x$. Then $\mathbb{T}_y$ (coded by some things yy) is a *direct subtree* of $\mathbb{T}_x$ iff $\mathbb{T}_y$ a proper subtree of $\mathbb{T}_x$ and the top node a_y of $\mathbb{T}_y$ is in the level immediately below the top node a_x of $\mathbb{T}_x$.

Definition 109. Let xx and yy code quotient pair trees $\mathbb{T}_x$ and $\mathbb{T}_y$ respectively. Then *the xx bear the E_T relation to the yy* iff $\mathbb{T}_y$ is isomorphic to a direct subtree of $\mathbb{T}_x$. We shall also write $\mathbb{T}_y E_T \mathbb{T}_x$ to represent this relation between some things xx and some other things yy, and will then speak of $y \widehat{\in} x \widehat{\in} (V)^+$.

Definition 110. Let xx and yy code quotient pair trees $\mathbb{T}_x$ and $\mathbb{T}_y$ respectively. Then we say that xx and yy are *quotient pair equivalent* (or $\mathbb{T}_x =_T \mathbb{T}_y$ iff $\mathbb{T}_x$ and $\mathbb{T}_y$) are isomorphic.

Definition 111. $(V)^+$ is the structure comprised of quotient coding pairs over V under a relation $\widehat{\in}$. If x is an object in this language, then it corresponds to a coding pair $\langle \tilde{M}_x, \tilde{R}_x \rangle$, that in turn codes a tree $\mathbb{T}_x$. Membership $\widehat{\in}$ and equality $\widehat{=}$ in $(V)^+$ correspond to E_T and $=_T$ respectively.

Theorem 112. [Antos, 2015], [Antos and Friedman, S], [Barton, 2016][15] $(V)^+$ satisfies Infinity, Extensionality, Foundation, Pairing, Union, Δ_n-Separation for every n, and Σ_n-Collection for every n. Rendered in a manner acceptable to the Universist, this states:

(1.) *Infinity*$^+$: There are some things $\omega\omega$ representing a tree $\mathbb{T}_\omega$ for ω.

(2.) *Transitivity*$^+$: For any things xx representing a quotient coding pair tree $\mathbb{T}_x$, and for any things yy representing a quotient coding pair tree $\mathbb{T}_y E_T \mathbb{T}_x$, then if a_z is a node directly below a_y in $\mathbb{T}_y$, then there are some things zz coding a direct subtree $\mathbb{T}_z$ of $\mathbb{T}_y$.

(3.) *Pairing*$^+$: For any things xx coding a tree $\mathbb{T}_x$ and any things yy coding a tree $\mathbb{T}_y$, there are some things zz coding a tree $\mathbb{T}_z$ such that $\mathbb{T}_x E_T \mathbb{T}_z$ and $\mathbb{T}_y E_T \mathbb{T}_z$.

(4.) *Union*$^+$: Suppose that there are some things xx coding a quotient pair tree $\mathbb{T}_x$. Then for any things yy coding a quotient pair tree $\mathbb{T}_y$

[15]The bulk of the work here is philosophical, in that I show how the coding works in the setting of V with plural reference added. However, some modifications were required. In particular, the absence of Class Bounding was problematic for proving Σ_n-Collection$^+$, and so we have provided a separate proof. Both this result and modifications to the original proofs were developed in correspondence with Antos and Friedman.

such that $\mathbb{T}_y E_T \mathbb{T}_x$, and any things zz coding trees $\mathbb{T}_z E_T \mathbb{T}_x$, there are some things $\cup\cup$ coding a quotient pair tree $\mathbb{T}_{\cup x}$ such that each $\mathbb{T}_z E_T \mathbb{T}_{\cup x}$.

(5.) Δ_n-*Separation*$^+$: Suppose that there are some things xx coding a quotient pair tree $\mathbb{T}_x$, in turn coding some $x\widehat{\in}(V)^+$. Let $\phi(y)$ be a Δ_n formula in the language of $(V)^+$. Then there are some things zz coding a tree $\mathbb{T}_z$ such that for any things yy coding a tree $\mathbb{T}_y$ (with $y\widehat{\in}(V)^+$) such that $\mathbb{T}_y E_T \mathbb{T}_x$, if $\phi(y)$ holds in the theory of $(V)^+$ then $\mathbb{T}_y E_T \mathbb{T}_z$.

(6.) Σ_n-*Collection*$^+$. Suppose that there are some things xx coding a quotient pair tree $\mathbb{T}_x$, in turn coding some $x\widehat{\in}(V)^+$. Let $\phi(p,q)$ be a Σ_n formula in the language of $(V)^+$. Suppose further that for every $\mathbb{T}_y$ coding some $y\widehat{\in}(V)^+$ such that $\mathbb{T}_y E_T \mathbb{T}_x$, there is a $z\widehat{\in}(V)^+$ (represented by some things zz coding a tree $\mathbb{T}_z$), such that $\phi(y,z)$. Then there are some things aa coding a tree $\mathbb{T}_a$ and $a\widehat{\in}(V)^+$ such that for any things *bb* coding a tree $\mathbb{T}_b E_T \mathbb{T}_x$ and $b\widehat{\in}x\widehat{\in}(V)^+$, there are some things cc coding a tree $\mathbb{T}_c$ and $c\widehat{\in}(V)^+$ such that $\mathbb{T}_c E_T \mathbb{T}_a$, $c\widehat{\in}a$, and $\phi(b,c)$.

Proof. Most of the work has already been accomplished in [Antos, 2015] and [Antos and Friedman, S]. There, they use an axiomatisation MK^* that also includes the following Class Bounding Axiom:

$\forall x \exists A \phi(x,A) \rightarrow \exists B \forall x \exists y \phi(x,(B)_y)$

where $(B)_y$ is defined as follows:

$(B)_y = \{z | \langle y,z\rangle \in^* B\}$.

They also show that, when working over a countable transitive model $\mathfrak{M} = (M, \in, \mathcal{C})$ such that $\mathfrak{M} \models \mathbf{MK}$, the model $(M^+, \in)$, where M^+ is defined as follows:

$M^+ = \{x|$ "There is a coding pair $\langle M_x, R_x\rangle$ for x in $\mathcal{C}$"$\}$

satisfies $\mathbf{SetMK}^*$; a version of $\mathbf{ZFC}-$Power Set with a Set Bounding Axiom and some constraints on the cardinal structure of M^+. We show that the proofs we require are (a) amenable to the current context, (b) realisable using the mechanisms of plural reference, and (c) can be accomplished avoiding the use of Bounding. For the purposes of the proof, it will be much easier to speak of the tree structures $\mathbb{T}_x$, $\mathbb{T}_y$, and $\mathbb{T}_z$, rather than constantly paraphrasing in terms of those things representing the quotient coding structures.

The proofs in [Antos, 2015] and [Antos and Friedman, S] rely on the following two lemmas:

[Antos, 2015], [Antos and Friedman, S] *First Coding Lemma.* Let $\mathfrak{M} = (M, \mathcal{C})$ be a transitive β-model of $\mathbf{MK}^*$. Let $\langle N_1, R_1\rangle$ and $\langle N_2, R_2\rangle$ be coding pairs. Then if there is an isomorphism between $\langle N_1, R_1\rangle$ and $\langle N_2, R_2\rangle$ then there is such an isomorphism in $\mathcal{C}$.

[Antos, 2015], [Antos and Friedman, S] *Second Coding Lemma.* For all $x \in M^+$ there is a one-to-one function $f \in M^+$ such that $f : x \longrightarrow M_x$, where $\langle M_x, R_x\rangle$ is a coding pair for x.

Immediately we need to make the following:

Remark 113. [Antos, 2015] and [Antos and Friedman, S] are concerned with higher-order forcing (i.e. using forcing posets that have some *proper classes* of a model as conditions) over models of $\mathbf{MK}^*$. In order to deal with the obvious metamathematical difficulties, they *explicitly* define the construction over models satisfying $\mathbf{MK}^*$ that are *countable*, *transitive*, and are β-models (i.e. the models are correct about which relations are well-founded). Their strategy is to code a model $(\mathfrak{M})^+$ of $\mathbf{SetMK}^*$ in the original model of $\mathbf{MK}^*$, and perform a definable class forcing over $(\mathfrak{M})^+$. This then corresponds to a hyperclass forcing over $\mathfrak{M}$. The work of the current thesis shows that the coding outlined is not dependent upon the countability of the models, nor the extra assumption of Class Bounding. Instead, we can take the coding over V using the interpretation of $\mathbf{MK}$

through plural reference, and show how to code the theory of $(V)^+$ using these resources. For the above lemmas then, a few remarks are in order.

(I) The assumption that the model over which we code is a β-model is trivial in the present setting; we are working over the Universist's V. For the Universist, V *sets the standard* for what a β-model is, and so is trivially a β-model.

(II) The First Coding Lemma is a non-trivial result when we are concerned with a countable transitive model $\mathfrak{M} \models \mathbf{MK}^*$. Since $\mathfrak{M}$ has an impoverished view of what classes there are we need to establish that $\mathbf{MK}^*$ satisfaction *alone* ensures that there is a class of the relevant kind in $\mathcal{C}$. For the purposes at hand, however, the result is again trivial; the relevant class theory over V *sets the standard* for when two trees representing classes are isomorphic, and so we *cannot* have isomorphic trees $\mathbb{T}_x$ and $\mathbb{T}_y$ for which there is not a class coding an isomorphism. If there are no things coding an isomorphism between $\mathbb{T}_x$ and $\mathbb{T}_y$ then they are simply *not* isomorphic.

(III) The Second Coding Lemma will be required for showing certain propeties of $(V)^+$. We deal with it in due course.

We now proceed to prove (1.)–(5.), transposing the proofs from [Antos, 2015] and [Antos and Friedman, S] to the Universist's framework (and making changes where necessary).

(1.) *Infinity*$^+$ is trivial, $\langle \omega \cup \{\omega\}, \in^* \rangle$ is a coding pair for ω, hence there is a quotient coding tree $\mathbb{T}_\omega$ for ω.

(2.) *Transitivity*$^+$ follows from Class Comprehension. If $\mathbb{T}_y$ is a subtree of $\mathbb{T}_x$ then there is a quotient coding pair for $\mathbb{T}_y$ (the above definitions ensure that we have a formula defining $\mathbb{T}_y$). As the quotient coding pairs behave in a transitive manner, there are no infinite descending sequences of direct subtrees (i.e. the equivalent of Foundation for quotient coding pairs and E_T holds) and if $\mathbb{T}_x$ and $\mathbb{T}_y$ have isomorphic direct subtrees then $\mathbb{T}_x$ is isomorphic to $\mathbb{T}_y$ (i.e. Extensionality for quotient coding pairs holds).

(3.) *Pairing*$^+$. Let $\mathbb{T}_x$ and $\mathbb{T}_y$ be coding trees with top nodes a_x and a_y. Then $\mathbb{T}_{\{x,y\}}$ has top node $\{a_x, a_y\}$ with immediate $\tilde{R}_{\{a_x,a_y\}}$-predecessors a_x and a_y, and the inclusive $\tilde{R}_{\{a_x,a_y\}}$-downclasses of a_x and a_y are $\mathbb{T}_x$ and $\mathbb{T}_y$ respectively. Effectively, we join the trees $\mathbb{T}_x$ and $\mathbb{T}_y$ to the top node $\{a_x, a_y\}$.

(4.) *Union*$^+$. Suppose that $\mathbb{T}_x$ is a tree with top node a_x, with immediate subtrees (i.e. 'members') $\mathbb{T}_y$, $\mathbb{T}_z$,... with their respective top nodes a_y, a_z,..., with subtrees $\mathbb{T}_{y_0}$, $\mathbb{T}_{y_1}$,.. of $\mathbb{T}_y$, and respectively for $\mathbb{T}_z$. Our tree $\mathbb{T}_{\cup x}$ will have some top node $a_{\cup x}$. We cannot simply join the trees as in (3.) as some subtrees may be isomorphic. Instead, we take equivalence classes of top nodes from the second level below a_x, and join the trees to $a_{\cup x}$ using Class Comprehension (and the First Coding Lemma to ensure knowledge of the isomorphism).

(5.) Δ_n-*Separation*$^+$ is slightly more problematic in that it is not clear how to code a *first-order* formula within $(V)^+$. This is dealt with by the following:

Lemma 114. [Antos, 2015], [Antos and Friedman, S] Let ϕ be a Δ_n-formula in the language of $(V)^+$. Then there is a formula in the language of plurals (and the trees they code) such that the theory of $(V)^+$ contains $\phi(x_1, ..., x_n)$ iff $\psi(\mathbb{T}_{x_1}, ..., \mathbb{T}_{x_n})$.

Proof. By induction on the complexity of ϕ. Suppose ϕ is of the form $y\widehat{\in}x$, and let $\mathbb{T}_y$ and $\mathbb{T}_x$ be the associated trees. As $y\widehat{\in}x$, there is a tree $\mathbb{T}_{y'}$ with top node $a_{y'}$ in the level below the top node of $\mathbb{T}_x$, such that $\mathbb{T}_y$ is isomorphic to $\mathbb{T}_{y'}$. By the First Coding Lemma (trivial in the current setting), ψ is then "there are some things yy coding a tree $\mathbb{T}_y$ such that $\mathbb{T}_y$ is isomorphic to some things $y'y'$ coding a direct subtree $\mathbb{T}_{y'}$ of $\mathbb{T}_x$".

Suppose then that ϕ is of the form $y\widehat{=}x$. Then ψ is simply "The things that code $\mathbb{T}_y$ are isomorphic to the things that code $\mathbb{T}_x$", which is again dealt with by the First Coding Lemma.

For the inductive steps where ϕ is of the form $\neg\chi_0$ or[16] $\chi_1 \wedge \chi_2$, the result is immediate, we just either negate or conjoin the class-theoretic correlate of the χ_n provided by the induction step.

Suppose then that ϕ is of the form $\forall x\chi$, where χ is translatable in our class theory by χ'. Then ψ is "For any things xx coding some tree $\mathbb{T}_x$, $\chi'(xx)$". ■

We now can proceed with the proof of Δ_n-Separation$^+$. Let $a, x_1, .., x_n$ be first-order names in the theory of $(V)^+$ and $\phi(x, x_1, ..., x_n, a)$ be any Δ_n-formula in the language of $(V)^+$. We need to show that:

$b = \{x\widehat{\in}a|\phi(x, x_1, .., x_n, a)\}\widehat{\in}(V)^+$

and hence that there are some things bb representing a coding pair for b with a corresponding tree $\mathbb{T}_b$.

Let $\mathbb{T}_{x_1}$,...,$\mathbb{T}_{x_n}$, $\mathbb{T}_a$ be the codes for each of $x_1, ..., x_n$ and a. Further let ψ be the class-theoretic correlate of ϕ. If b is empty the result is immediate as there is a coding pair for the empty set. Assume then that b is non-empty and $(V)^+$ thinks that b contains some c_0 with coding pair tree $\mathbb{T}_{c_0}$. Let $\mathbb{T}_{a(c)}$ be a variable over trees (rendered as restricted plural quantification over the relevant coding pairs), such that each $\mathbb{T}_{a(c)}$ corresponds to a direct subtree of $\mathbb{T}_a$ (i.e. $a(c)$ is a member of a in $(V)^+$). By Class Comprehension, there are some things zz such that if $\psi(\mathbb{T}_{a(c)}, \mathbb{T}_{x_1}, ..., \mathbb{T}_{x_n}, \mathbb{T}_a)$ holds then $\{z|\langle c, z\rangle \prec zz\}$ is the direct subtree $\mathbb{T}_{a(c)}$ of $\mathbb{T}_a$, and if $\psi(\mathbb{T}_{a(c)}, \mathbb{T}_{x_1}, ..., \mathbb{T}_{x_n}, \mathbb{T}_a)$ does not hold then $\{z|\langle c, z\rangle \prec zz\} = \mathbb{T}_{c_0}$.

We then let $\mathbb{T}_b$ be a tree with top node b_0 that has as all its direct subtrees the various $\{z|\langle c, z\rangle \prec zz\}$. The tree $\mathbb{T}_b$ then codes the existence of the necessary $b = \{x\widehat{\in}a|\phi(x, x_1, ..., x_n, a)\}\widehat{\in}(V)^+$.

(6.) Σ_n-*Collection*$^+$. In order to prove Σ_n-Collection$^+$, it is far easier to work from the perspective of $(V)^+$. We first mention some of $(V)^+$'s basic properties. For ascertaining some of these properties, we return (as promised) to the:

[16]The choice of $\wedge$ here was somewhat arbitrary, any suitable connective will do.

Lemma 115. [Antos and Friedman, S] *Second Coding Lemma.* For every $x\widehat{\in}(V)^+$ there is a one-to-one $f\widehat{\in}(V)^+$ such that $f : x \longrightarrow M_x$, where $\langle M_x, R_x\rangle$ is a coding pair for x. i.e. For any $x\widehat{\in}(V)^+$, $(V)^+$ has a function f_x that codes a one-to-one function onto the domain of a coding pair for x.

Proof. Let $\mathbb{T}_x$ be a coding tree for x and each $y \in x$ have a coding tree $\mathbb{T}_y$ with top node a_y. We start by showing that $\{\langle y, a_y\rangle | y \in x\}\widehat{\in}(V)^+$. We have each $\mathbb{T}_{\langle y,a_y\rangle}$ in $(V)^+$ by Pairing. We then get a coding pair for f as follows.

$M_f = \bigcup_{z\in x} M_{\langle z,a_z\rangle} \cup \{a_f\}$, where $a_f \in V$ and $a_f \notin M_{\langle z,a_z\rangle}$ for every $z \in x$.

$R_f = \{\langle v, w\rangle|$ for some $y \in x$ either $\langle v, w\rangle \in R_{\langle y,a_y\rangle}$ or ($v = a_{\langle y,a_y\rangle}$ and $w = a_f$) $\}$.

$\mathbb{T}_f$ is ordered by R_z below every a_z and $a_{\langle y,a_y\rangle}$ below a_f otherwise. Essentially, we have mapped trees to their top notes in constructing $\mathbb{T}_f$. ■

We can now prove:

Lemma 116. $(V)^+$ believes it has a largest cardinal, namely $Ord(V)$. More precisely, any $x\widehat{\in}(V)^+$ can be mapped one-to-one onto an initial segment of the ordinals.

Proof. Since $(V)^+$ believes that there is a bijection from any x to M_x, it suffices to show that $(V)^+$ can see an injection $f : M_x \longrightarrow On$, to obtain an injection from any $x\widehat{\in}(V)^+$ to On^V. The case for set-sized coding pairs is trivial. In the case of class-sized M_x, we know by Class Replacement and Global Choice that every class is bijective with an initial segment of the ordinals, and hence there is a class F_x coding a bijection from M_x to On^V. To then obtain that F_x has a representative in $(V)^+$, we just need to show that any class C has a coding pair and associated element of $(V)^+$. To this end, assume C is a class and $y \in C$. Then $y \in V$ and y is coded by a tree with top node a_y. Then we consider a tree $\mathbb{T}_x$ with top node a_x:

$M_x = \bigcup_{y\in X} M_y \cup \{a_x\}$

and

$R_x = \{\langle v, w\rangle |$ for some $y \in C$ either $\langle v, w\rangle \in R_y$ or $v = a_y$ and $w = a_x\}$.[17]

■

With these properties in place, we proceed with this proof of Σ_n-Collection$^+$. Assume for contradiction that Σ_n-Collection fails in $(V)^+$. Then we have a Σ_n-definable function $\phi(x, y)$ from $Ord(V)$ unbounded in $L_\varkappa(V) = (V)^+$. Since $Ord(V)\widehat{\in}(V)^+$ is the largest cardinal of $(V)^+$, we can extend $\phi(x, y)$ to a Σ_n-definable bijection between $Ord(V)$ and $L_\varkappa(V)$.We then have a Σ_n-definable well-order $<_R$ on the subsets of $Ord(V)$ in $L_\varkappa(V)$. For $i\widehat{\in}I$ indexing $<_R$ on subsets $x_0, x_1, ..., x_i, ...$ of $Ord(V)$ we then diagonalise to produce the following $X \subseteq Ord(V)$:

$X = \{\alpha \in^* On | \neg\alpha\widehat{\in}x_\alpha\}$

This X is Σ_n-definable over $Ord(V)$ by the properties of $<_R$ and the fact that $Ord(V)\widehat{\in}(V)^+$, but cannot be in $(V)^+$. Since $Ord(V)\widehat{\in}(V)^+$, this violates Δ_n-Separation, $\perp$. ■

We now need to finally put the pieces together to provide a coding of $Hyp(V)$:

Theorem 117. There is an $x\widehat{\in}(V)^+$ such that $x\widehat{=}Hyp(V)$. In other words, there are some things hh coding a quotient coding pair for $Hyp(V)$ and associated tree $\mathbb{T}_{Hyp(V)}$.

Proof. Now we have more resources at our disposal, we work in $(V)^+$. Since $(V)^+$ is of the form $L_\varkappa(V)$ and satisfies Δ_n-Separation and Σ_n-Collection, there is some level $L_\vartheta(V)$ satisfying Δ_0-Separation and Σ_1-Collection, with ϑ the least such.

By the results of [Barwise, 1975], $L_\vartheta(V)\widehat{=}Hyp(V)$. Moving back to the coding, this implies the existence of some things hh coding a tree $\mathbb{T}_{Hyp(V)}$ for $Hyp(V)$. ■

The above machinery then provides the resources to code $Hyp(V)$ (and hence

[17] V is a special case here but we can deal with this by representing every set by its singleton and use $\emptyset$ for the top node a_v to obtain a coding pair for V.

consistency in V-logic) by referring to some sets within V. Before we sum up the current section, we make a final:

Remark 118. We mentioned earlier that we would show how long definable well-orders over V could be coded within **MK**. We could code such well-orders in height extensions as we did earlier in V-logic. Equally though, we could note the following:

Theorem 119. Let ϕ be a formula defining a V-definable well-order longer than Ω^V. Then there is a coding pair representing ϕ.

Proof. We proceed by working from the perspective of $Hyp(V)\widehat{\in}(V)^+$. We will show that the set of V-definable well-orders is in $Hyp(V)$. We do this by first noting that V can tell which are the V-definable well-orders, and thus so can $Hyp(V)$. Then the function that assigns to each V-definable wellorder its ordertype is Σ_1-definable in $Hyp(V)$ by a formula $\phi(x, y)$. Thus, the set of all V-definable well-orders is in $Hyp(V)$ (by Σ_1-Collection and the fact that $V\widehat{\in}Hyp(V)$) and thus has a coding pair in **MK** over V. ■

The point here is the following: While we could code long well-orders syntactically using V-logic and a sentence about some height extension $\bar{W}$, definable long well-orders can be coded more directly using coding pairs.

In sum, we have seen thus far that (i) extensions of V can be coded using the consistency of theories in V-logic, and (ii) this can be coded within the class theory licensed by the plural interpretation of proper classes. It remains, however, to show that this coding is *philosophically virtuous*.

VII.5 Evaluation

We now have a coding in the (plurally rendered) class theory **MK** for consistency in V-logic, which in turn can capture the intra-V consequences of extensions of V. This is true even for *higher-order* axioms postulating the existence of inner models (such as the IMH). We now wish to analyse the extent to which this coding is philosophically satisfactory.

Recall the main challenge (outlined in II.4) that we wanted to satisfy:

> **The Hilbertian Challenge.** Provide philosophical reasons to legitimise the use of extra-V resources for formulating axioms and analysing intra-V consequences.

The coding performs well with respect to The Hilbertian Challenge in this raw form. We have provided philosophical reasons to accept the use of **MK** over V using the mechanisms of plural reference. We then showed how to code extensions of V using V-logic and that $Hyp(V)$ can formalise the notion of consistency in V-logic. Since $Hyp(V)$ can be coded using **MK** over V, we have some things in V that collectively code consistency in V-logic, and hence discourse about extensions of V. Given the cogency of the resources of **MK** (argued for in the previous chapter) we can see why talk of extensions will not lead us astray.

The Hilbertian Challenge was, however, tempered by three additional desiderata on any interpretation of extension talk:

> **The Foundational Constraint.** In responding to the Hilbertian Challenge, do so in a way that does not necessitate the use of resources that cannot be represented by sets within V.
>
> **The Ontological Constraint.** Any interpretation of extra-V resources should make clear the *ontological difference* between the interpretation of extensions and normal sets within V. In other words, any interpretation must make clear in what sense the interpretation does not *literally* refer to extra-V sets.
>
> **The Methodological Constraint.** In responding to the Hilbertian Challenge, do so in a way that accounts for our naive thinking about extensions and links them to structural features of V.

As we noted earlier (V.5), these three constraints appear to be in tension with one another. For, how could we be faithful to the *naive* talk, whilst ensuring that what we actually interpret as extensions of V is substantially ontologically *different* from set-talk, without using objects not representable using sets from V? Is it not the case that in order to interpret discourse *naively* we have to use bona fide sets in our interpretation?

We should pay attention to some features of the coding, and its relationship to the countable transitive model strategy outlined in V.4. Since the underlying interpretation of our coding is plural, the Foundational Constraint is immediately satisfied. The coding does not require the use of entities not representable using objects from V: everything is rendered in terms of plural reference to and quantification over the sets.

For the Ontological Constraint, note that if we examine the coding, we can see that our interpretation of extension talk concerning V makes the distinction between sets and extensions clear. We are interpreting a claim concerning an extension of V as a one about the consistency of a particular theory in V-logic, which in turn can be coded as talk about some sets within V. There is little pressure for the Universist to accept this talk as committing her to *actual* extensions of V, in much the same way as talk of non-well-founded models of sets does not commit her to the existence of *actually* non-well-founded sets.[18]

The Methodological Constraint is slightly more involved, however. We do *not* have a naive interpretation of extensions of V: we are interpreting model-theoretic claims about extensions of V as the syntactic consistency of theories in V-logic. We would like to find a place for our naive thinking concerning extensions and relate this discourse to our analysis of truth in V. The key fact here is that $Hyp(V)$ (and hence claims about consistency in V-logic) is coded by a *single* class. Recall that, in VI.4, we discussed the following embedding:

> [Vickers and Welch, 2001] Suppose $I \subseteq Ord$ witnesses Ord is Ramsey. Then, definably over $(V, \in, I)$, there is a transitive model $\mathfrak{M} = (M, \in)$, and an elementary embedding $j : (M, \in) \longrightarrow (V, \in)$ with a critical point.

[18]The case for the actual existence of extensions of V is in fact weaker than that for non-well-founded sets. For, the relevant interpretation of non-well-founded sets is both model-theoretic and syntactic. By contrast, our interpretation of extensions of V is *purely syntactic* in that there are no models corresponding to the relevant $\bar{W}$.

We then noted that the most natural way of interpreting and talking about the classes in this embedding was to talk of some sets that collectively are good indiscernibles for On^V. We then noted that, despite the non-definability of I, we could perfectly well introduce a plural parameter ii and predicate $I(x)$ to talk about the indiscernibles and use them to define j.

In the case of our coding in V-logic, we have a single class H coding $Hyp(V)$. In particular, there are some things hh in the language of plurals corresponding to our use of the term 'H'. But now we can introduce a predicate '$H(x)$', such that $H(x)$ holds iff $x \prec hh$. Now, consider the structure $(V, \in, H)$ in *first-order* ZFC augmented with the predicate $H(x)$. This theory is countable, and so we can construct the same Downward Löwenheim-Skolem argument as in the original countable transitive model strategy (V.4). There, we posited the existence of a set of Skolem functions F^V for V (by 'choosing' a witness for the countably many existential statements in $\mathscr{L}_\in$ satisfied by V with a countable sequence of choices). Then, by Skolemising and Collapsing, we obtained a countable transitive model $\mathfrak{V}$ that satisfies exactly the same parameter-free first-order sentences as V. In our current setting, we are taking (given the determinateness of plural reference) our use of the predicate $H(x)$ to be well-understood. We have explained to what things the predicate applies, and hold ourselves (if we are Universists) to be talking about a determinate range of objects when we discuss V. Thus we augument $\mathscr{L}_\in$ with the predicate $H(x)$ (let the resulting language be denoted by $\mathscr{L}_{\in,H}$), and (since we hold ourselves to understand $H(x)$), admit its use in the Replacement and Comprehension Schema (let the resulting theory be $\mathbf{ZFC}_H$). Again, we may now assume (by 'choosing' a witness for the countably many existential statements in $\mathscr{L}_{\in,H}$ satisfied by V with a countable sequence of choices), that there is a set of Skolem functions F^V_H for the language $\mathscr{L}_{\in,H}$ for V.

Again by Tarski's Theorem on the undefinability of truth, while V can see F^V_H, it does not know that F^V_H provides its own set of Skolem functions. Again, we can formalise this using a modification of Feferman's approach in [Feferman, 1969]:

Definition 120. Let $\mathscr{L}_{\in,H,\mathcal{V}}$ be the language $\mathscr{L}_{\in,H}$ augmented with a single constant symbol $\mathcal{V}$. $\mathbf{ZFC}^{\mathcal{V}}_H$ is then a theory in $\mathscr{L}_{\in,H,\mathcal{V}}$ with the following axioms:

(i) $\mathbf{ZFC}_H$

(ii) $\mathcal{V}$ is countable and transitive.

(iii) For every ϕ in $\mathscr{L}_{\in,H}$, $\phi \leftrightarrow \phi^{\mathcal{V}}$ (by Tarski's Theorem, this is an axiom scheme).

Again, we know by the Reflection, Löwenheim-Skolem, and Mostowski Collapse Theorems that $\mathbf{ZFC}_H^{\mathcal{V}}$ is a conservative extension of $\mathbf{ZFC}_H$. However, $\mathcal{V}$ then satisfies *exactly* the same parameter-free first-order sentences of $\mathbf{ZFC}_H$ as V. The end result is a countable transitive model $\mathcal{V}$ (with its own $\mathcal{V}$-logic) that satisfies *both* the same first-order sentences *and* the same $Hyp(V)$ theory as V. However, since $\mathcal{V}$ *is* countable, Barwise's $\mathcal{V}$-logic completeness theorem *does* hold, and so any consistency in $\mathcal{V}$-logic *is* mirrored in the existence of an actual extension of $\mathcal{V}$ in V. We can then represent our naive reasoning about extensions as concerned with $\mathcal{V}$, exporting the results obtained back up to V as desired. The same was not true of $\mathfrak{V}$, as there we had no guarantee that the structure of $\mathfrak{V}$ mirrored that of V with respect to certain greater than first-order properties (such as the existence of inner models). This is precisely what is delivered by the above coding in combination with a countable $\mathcal{V}$.

Chapter VII: Conclusions

The chapter began by identifying our position within the dialectic. We had a motivation for **MK** based on the devices of plural reference. We then provided an extension of our logical resources to V-logic, and argued that it could be used to interpret extensions of V. Next, we moved to a perspective on which height extensions of V were available, and showed how to code proofs in V-logic into $Hyp(V)$. We then showed how to code $Hyp(V)$ with some things hh within V. Finally, we provided an evaluation of this coding. We first noted that by introducing a predicate for the hh we could reduce to the countable. We then argued that this enabled us to satisfy the Hilbertian Challenge in a manner consonant with the three constraints.

Chapter VIII

Conclusions

We have arrived at a quite satisfying situation, providing a response to the Hilbertian Challenge concerning extensions of V that performs well with respect to the Ontological, Methodological, and Foundational constraints. However, a few concluding remarks are in order to make precise exactly what the dialectic is with respect to the problems discussed, and indicate some directions for future research.

VIII.1 What has been shown

We begin by reviewing what has been shown through the course of the thesis. Chapter I situated the discussion within the current philosophical landscape, and identified the broad concerns with extending V. Chapter II then provided a more thorough analysis of why extensions of V might be interesting for a Universist. In particular, we argued that it is natural for a Universist to be interested in executing Gödel's Programme for formulating and justifying new axioms. The possibility for using extensions of V in this execution was discussed, both with respect to problems of formulation and also issues concerning justification. A Hilbertian Challenge was identified; provide philosophical reasons to legitimise the use of extra-V resources for formulating axioms and analysing intra-V consequences. Chapter III then lent some mathematical precision to these difficulties, outlining several axioms that required extensions of V for non-trivial formulation, and noting that there were cases in which extensions were desirable for proving facts about V. In Chapter IV, we examined attempts to interpret height extensions of V. We argued that the difficulties associated with these attempts highlighted two desirable constraints on responses to the Hilbertian Challenge: the Ontological Constraint that extensions of V should have a substantially different interpretation from sets in V, and the Foundational Constraint that we should only use mathematical objects representable with sets from V in any such interpretation. In Chapter V, we analysed some possibilities for interpreting width extensions of V. Again, we argued that there were difficulties with each interpretation considered, and that this highlighted the Methodological Constraint on resolution of Gödel's Programme: provide an interpretation that finds a place for our naive thinking concerning extensions and link it to truth within V. A tension between the three constraints was then identified; how can we possibly interpret extensions of V using only sets from V, whilst making the ontological character of said extensions substantially different from that of the sets, yet ensuring that our naive (set-like) thinking concerning extensions was linked to truth in V? We then in Chapter VI returned to the issue of proper classes. We argued that there are both first and second philosophical reasons to want to interpret non-definable talk concerning classes. We then provided an interpretation of proper classes via plural quantification and argued that it (i) performed well with respect to our three constraints, and (ii) provided reasons to accept the use of **MK** as a first-order theory of classes. In Chapter VII, we showed then how to interpret extensions of V using V-logic, and how to code proofs in V-logic with classes. Finally, we argued that the coding of $Hyp(V)$ by a single class facilitated a reduction to the countable, securing satisfaction of our three constraints within the codification of extensions of V provided.

VIII.2 What has not been shown

So far so good. However, we should be fully precise as to the philosophical impact of our discussion of codifications. Two points are especially important to clarify what we have *not* shown.

Firstly, it might be claimed that the thesis shows that the statement "V has an extension such that Φ" is perfectly coherent. It is important to be clear that this is not (and can never be) the case for a Universist. V is all the sets there are, and so such a statement (with its natural interpretation), is obviously false. Further, we should pay attention to just how gerrymandered the coding is. We do not claim that talk of extensions of V is *coherent* for the Universist: rather we assert that, given the use of additional expressive resources, she can reinterpret these statements in such a way as to be about objects countenanced by her ontology. This rendering proceeds through the mechanisms of V-logic and plural quantification, and allows us to capture the relevant intra-V consequences and formulations we desire. However a particular claim about an extension of V *really* says for the Universist that it is not the case that there are some things coding a proof in V-logic of $0 = 1$ from a particular theory $\mathbf{T}$. There is no claim that extensions are 'fine after all': such talk should rather be understood as unintended and pathological but nonetheless useful.

The second point is to note how *dialectically conservative* many of the chapters were. In Chapter II we noted that whether or not we can answer the Hilbertian Challenge is an *interesting question*. Despite this, the interest of an interesting question is highly defeasible. We might provide other reasons as to why we should not use extension talk, despite its *prima facie* usefulness for talking about V. Similarly, Chapter VI showed that if one is a Universist, there are *reasons to accept* $\mathbf{MK}$ class theory (plurally rendered) for talking about sets in V. We do not purport to have provided uncontroversial justification for $\mathbf{MK}$, instead we merely claim that its use appears both motivated and desirable. We are perfectly prepared to hold the justification of $\mathbf{MK}$ as defeasible in the face of good arguments to the contrary. Chapter VII showed how to code talk of extensions in class theory using an expansion of logical resources, but again, we might provide reasons to reject such logical profligacy.

This dialectical conservativeness plays out in what the reader should take from the thesis. What we have established is that *if* one is a Universist, and *if* one accepts that extension talk concerning V is worth scrutiny, and *if* one thinks that the three conditions are good constraints on an answer to Hilbert's Challenge, and *if* one finds the motivation of $\mathbf{MK}$ convincing and regards its use as legitimate, and *if* one thinks that we can extend our logical resources in the manner outlined in Chapter VII, *then* one can encode a substantial amount of talk concerning extensions of V using sets from V. That is an awful lot of conditionals, and we should be mindful of the dictum "one person's modus ponens is another's modus tollens". It is perfectly within the remit of the thesis for the conclusion to be taken as absurd and call for a rejection of one (or more) of the antecedents. One might take the thesis to show that there is a problem with talking about extended languages or the use of **MK**. One might think that the constraints are too strict, or not strict enough. Still further, one might think that this shows incoherence in the current practice of set theorists. Further still, one might take the sheer amount of talk of extensions of V that can be coded within the Universist's framework to be evidence of the falsity of her position. We have not taken a stand on any of these issues here, but have shown what can be accomplished given the acceptance of certain positions and resources.

VIII.3 Directions for future research

These observations indicate several possible directions of future research. In particular each of the antecedents in the above conditionals merits further scrutiny. Aside from considering the general questions surrounding the nature of sets, we propose several concrete recommendations:

> **Question 121.** The use of what was effectively an infinitary language was essential to the results of Chapter VII. How should we philosophically view the use of these infinitary languages? Of course such a language cannot be written down by a finite agent such as ourselves (it rather corresponds to a particular way of talking about classes in the coding). However, we might ask what our ability to use such languages in talking about V tells us about the relationship between ourselves, languages, and mathematical objects.
>
> **Question 122.** In Chapter VI, the class theory **MK** was given a plural rendering, and motivated. However, some of the justificatory force was provided by the Strong Limitation of Size Principle, which was noted to

be on somewhat shaky ground. What is the status of these principles, and more generally $\mathbf{MK}$?

Question 123. More generally, the motivation of $\mathbf{MK}$ was flagged as one of the more contentious aspects of the thesis. However, we do not need full $\mathbf{MK}$ in order to capture the proofs of V-logic (and definable well-orders) within $Hyp(V)$. In particular, having Σ_n-Collection and Δ_n-Separation in $(V)^+$ for any n is overkill. For V-logic, we just need to be able to simulate the objects of $Hyp(V)$, all of which are well below $L_{\varkappa}(V) = (V)^+$. This raises two questions, one technical and one philosophical. On the technical side, exactly how much Class Comprehension is required to capture consistency in V-logic and definable long well-orders using coding pairs? On the philosophical side, how does the required amount of Class Comprehension relate to other (less liberal) conceptions of the values of class variables?

Question 124. Further, while motivating $\mathbf{MK}$, we demured from providing a semantics for plural reference. Do our results (showing how much can be coded if plural reference is assumed to be well-understood) have any bearing on the debates in the semantics of plural reference and set theory?

Question 125. Certain kinds of simulations of extensions of V within V have been cited as evidence against the Universist's position (see, for example, the arguments in [Hamkins, 2012]). Do the new simulations presented in this thesis present a different kind of challenge? Does the fact that V is kept standard affect the debate here? On the other hand, does the ability of the Universist to talk about extensions of V provide a response to the charge that her view is overly restrictive?

Question 126. In particular, while it was shown how talk concerning long well-orders could be interpreted as about objects within V, the conceptual problem raised by the Burali-Forti Paradox remains. Do our codings serve to assuage these worries or exacerbate them?

However, assuming that we accept the coding, there are several interesting open questions and problems that spring from the thesis. The first is how, given that the Universist accepts our results, the axioms considered might be motivated. In particular, given the fact that these axioms rely essentially on significant reinterpretation of the statements concerning extensions, and the way that they are intimately connected with the merely countable, we might ask:

Question 127. How might we motivate these axioms? Are there interesting distinctions in motivation between these axioms and more 'standard' axioms of set theory?

and

Question 128. Does the reduction to the countable indicate any new relationships between truth in V and the existence of countable models?

Finally, there were at least two classes of axioms that were perspicuously *absent* from the discussion, despite their relevance. The first concerns the exact manner in which the coding was proved. One can code a greater amount of talk (and with less difficulty) if one accepts the following Class Bounding axiom:

$\forall x \exists A \phi(x, A) \rightarrow \exists B \forall x \exists y \phi(x, (B)_y)$

where $(B)_y$ is defined as follows:

$(B)_y = \{z | \langle y, z \rangle \in^* B\}$.

Such a principle is equivalent (assuming Global Choice) to the following Choice principle for classes:

AC_∞: $\forall x \exists A \phi(x, A) \rightarrow \exists B \forall x \phi(x, (B)_x)$

There are, in fact, many stronger choice principles for classes. These principles have received some attention from the mathematical community, but strengthenings of Global Choice have been largely ignored by philosophers. This is both interesting and represents a gap in philosophical discussion, especially from the point of view of the second philosopher; such principles are necessary for a wide variety of constructions involving classes (a good example being class ultrapowers).[1] A natural question then is:

> **Question 129.** What is the philosophical status of class choice principles stronger than Global Choice?

It was remarked in Chapter IV that one possibility for interpreting extensions of V would be the use of modal notions, partly inspired by the work of [Reinhardt, 1974] and [Reinhardt, 1980]. The suggestion was rejected as problematic for several reasons. However, since we now *have* an interpretation of extensions, a natural question is:

> **Question 130.** How should we think about justification concerning Reinhardt's axioms on a Universist picture?
>
> **Question 131.** More generally, how does the ability to interpret extensions of V facilitate an understanding of modal notions for a Universist?

In addition to all these specific questions though, there is one guiding question that remains unresolved for Universist-inspired mathematics:

> **Question 132.** What is true (in V)?

Precise answers to this question have not been uncovered by this thesis. However, we have shown that with a smidgeon of extra expressive resources, the Universist can utilise far more mathematics than previously thought. The doors are thus open to new and exciting philosophical and mathematical discussions.

[1] See [Gitman et al., U] for discussion.

Bibliography

[Antos, 2015] Antos, C. (2015). *Foundations of Higher-Order Forcing*. PhD thesis, Universität Wien.

[Antos and Friedman, S] Antos, C. and Friedman, S.-D. (S). Hyperclass forcing in Morse-Kelley class theory. Submitted.

[Antos et al., 2015] Antos, C., Friedman, S.-D., Honzik, R., and Ternullo, C. (2015). Multiverse conceptions in set theory. *Synthese*, 192(8):2463–2488.

[Arrigoni and Friedman, 2013] Arrigoni, T. and Friedman, S.-D. (2013). The Hyperuniverse Program. *Bulletin of Symbolic Logic*, 19:77–96.

[Balaguer, 1998] Balaguer, M. (1998). *Platonism and Anti-Platonism in Mathematics*. Oxford University Press.

[Barton, 2012] Barton, N. (2012). Proper classes; understanding ontology through paradox. Master's thesis, University College London.

[Barton, 2016] Barton, N. (2016). *Executing Gödel's Programme in Set Theory*. PhD thesis, Birkbeck, University of London.

[Barton, Fa] Barton, N. (Fa). Multiversism and concepts of set: How much relativism is acceptable? In Boccuni, F. and Sereni, A., editors, *Objectivity, Realism, and Proof. FilMat Studies in the Philosophy of Mathematics*, Boston Studies in the Philosophy and History of Science. Springer. Forthcoming.

[Barton, Fb] Barton, N. (Fb). Richness and reflection. *Philosophia Mathematica*. Forthcoming, DOI: 10.1093/philmat/nkv036.

[Barwise, 1975] Barwise, J. (1975). *Admissible Sets and Structures: An approach to definability theory.* Springer.

[Baumgartner and Hajnal, 1973] Baumgartner, J. and Hajnal, A. (1973). A proof (involving Martin's Axiom) of a partition relation. *Fundamenta Mathematicae*, 78(3):193–203.

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

[Benci et al., 2006] Benci, V., Di Nasso, M., and Forti, M. (2006). An Aristotelian notion of size. *Annals of Pure and Applied Logic*, 143(1):43–53.

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

[Bernays, 1961] Bernays, P. (1961). Zur Frage der Unendlichkeitsschemata in der axiomatische Mengenlehre. In Bar-Hillel, Y., editor, *Essays on the Foundations of Mathematics*, pages 3–49. Magnus Press, Hebrew University of Jerusalem,.

[Boolos, 1971] Boolos, G. (1971). The Iterative Conception of Set. *The Journal of Philosophy*, 68(8):215–231.

[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(8):430–449.

[Boolos, 1985] Boolos, G. (1985). Nominalist platonism. *The Philosophical Review*, 94(3):327–344.

[Boolos, 1989] Boolos, G. (1989). Iteration again. *Philosophical Topics*, 17(2):5–21.

[Boolos, 1998] Boolos, G. (1998). Reply to Charles Parsons' 'sets and classes'. In *Logic, Logic, and Logic*, pages 30–36. Harvard University Press.

[Broughton, 2002] Broughton, J. (2002). *Descartes's Method of Doubt*. Princeton University Press.

[Burali-Forti, 1897] Burali-Forti, C. (1897). Una question sui numeri transfiniti. *Rendiconti del Circolo Matematico di Palermo*, 11:154–164.

[Burgess and Rosen, 1997] Burgess, J. and Rosen, G. (1997). *A Subject with No Object*. Clarendon Press.

[Burgess, 2004] Burgess, J. P. (2004). E pluribus unum: Plural logic and set theory. *Philosophia Mathematica*, 12(3):193–221.

[Burgess, 2015] Burgess, J. P. (2015). *Rigor and Structure*. Oxford University Press.

[Buss, 1994] Buss, S. R. (1994). On Gödel's theorems on lengths of proofs I: Number of lines and speedup for arithmetics. *The Journal of Symbolic Logic*, 59(3):737–756.

[Cantor, 1899] Cantor, G. (1899). Letter to Dedekind. In *[van Heijenoort, 1967]*, pages 113–117. Harvard University Press.

[Cantor et al., 1899] Cantor, G., Hilbert, D., and Dedekind, R. (1897-1899). Cantor's late correspondence with Dedekind and Hilbert. In *[Ewald, 1996a]*, pages 923–940. Oxford University Press.

[Cohen, 1963] Cohen, P. (1963). A minimal model for set theory. *Bull. Amer. Math. Soc*, (69):537–540.

[Cohen, 1966] Cohen, P. (1966). *Set Theory and The Continuum Hypothesis*. W.A. Benjamin, Inc.

[Cottingham and Williams, 1996] Cottingham, J. and Williams, B. (1996). *Meditations on First Philosophy with Selections from the Objections and Replies*. Cambridge University Press.

[Descartes, 1637] Descartes, R. (1637). *Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth in the Sciences*. Penguin. Translated by F.E. Sutcliffe in [Sutcliffe, 1968].

[Descartes, 1641] Descartes, R. (1641). *Meditations on First Philosophy*. Oxford University Press. Translated by J. Cottingham in [Cottingham and Williams, 1996].

[Detlefsen, 1986] Detlefsen, M. (1986). *Hilbert's Program*. Dordrecht: Reidel.

[Detlefsen, 1990] Detlefsen, M. (1990). On an alleged refutation of Hilbert's program using Gödel's first incompleteness theorem. *Journal of Philosophical Logic*, 19:343–377.

[Detlefsen, 2001] Detlefsen, M. (2001). What does Gödel's Second Theorem say? *Philosophia Mathematica*, 9:37–71.

[Devlin, 1979] Devlin, K. (1979). *The Joy of Sets*. Springer-Verlag, 2nd edition.

[Drake, 1974] Drake, F. R. (1974). *Set Theory: An Introduction to Large Cardinals*. North Holland Publishing Co.

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

[Erdős and Hajnal, 1966] Erdős, P. and Hajnal, A. (1966). On a problem of B. Jónsson. *Bull. Acad. Pol. Sci., Sér. Sci. Math. Astron. Phys. 14, 19-23 (1966).*

[Ewald, 1996a] Ewald, W. B., editor (1996a). *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, volume II. Oxford University Press.

[Ewald, 1996b] Ewald, W. B., editor (1996b). *From Kant to Hilbert. A Source Book in the Foundations of Mathematics.*, volume I. Oxford University Press.

[Feferman, 1969] Feferman, S. (1969). Set-theoretical foundations of category theory. In *Reports of the Midwest Category Seminar III, Lecture Notes in Mathematics Volume 106*, pages 201–247. Springer.

[Feferman, 2009] Feferman, S. (2009). Operational set theory and small large cardinals. *Information and Computation*, 207:971–979.

[Feferman et al., 2000] Feferman, S., Friedman, H., Maddy, P., and Steel, J. (2000). Does mathematics need new axioms? *Bulletin of Symbolic Logic*, 6(4):401–446.

[Field, 2003] Field, H. (2003). Do we have a determinate conception of finiteness and natural number? In Schirn, M., editor, *The Philosophy of Mathematics Today*. Clarendon Press.

[Florio and Linnebo, 2015] Florio, S. and Linnebo, Ø. (2015). On the innocence and determinacy of plural quantification. *Noûs*, 49(1):1–19.

[Foreman, 1986] Foreman, M. (1986). Potent axioms. *Transactions of the American Mathematical Society*, 294(1):1–28.

[Foreman, 1998] Foreman, M. (1998). Generic large cardinals: New axioms for mathematics? *Documenta Mathematica*, 2:11–21.

[Foreman, 2010] Foreman, M. (2010). Ideals and generic elementary embeddings. In Kanamori, A. and Foreman, M., editors, *Handbook of Set Theory*, pages 885–1147. Springer.

[Forster, 1995] Forster, T. (1995). *Set Theory With a Universal Set: Exploring an Untyped Universe*. Clarendon Press.

[Fraenkel and Bar-Hillel, 1966] Fraenkel, A. and Bar-Hillel, Y., editors (1966). *Essays on the Foundations of Mathematics*. Magnes Press Hebrew University.

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

[Frege, 1893] Frege, G. (1893). *The Basic Laws of Arithmetic: Exposition of the System*. University of California Press.

[Friedman, 2000] Friedman, S.-D. (2000). *Fine Structure and Class Forcing*. de Gruyter. de Gruyter Series in Logic and its Applications, Vol. 3.

[Friedman, 2006] Friedman, S.-D. (2006). Internal consistency and the inner model hypothesis. *Bulletin of Symbolic Logic*, 12(4):591–600.

[Friedman, 2010] Friedman, S.-D. (2010). Constructibility and class forcing. In Kanamori, A., editor, *Handbook of Set Theory*, pages 557–604. Springer.

[Friedman, S] Friedman, S.-D. (S). Evidence for set-theoretic truth and the Hyperuniverse Programme. Submitted.

[Friedman and Honzik, 2016] Friedman, S.-D. and Honzik, R. (2016). On strong forms of reflection in set theory. *Mathematical Logic Quarterly*, 62(1-2):52–58.

[Friedman and Ternullo, F] Friedman, S.-D. and Ternullo, C. (F). The search for new axioms in the Hyperuniverse Programme. In Boccuni, F. and Sereni, A., editors, *Objectivity, Realism, and Proof. FilMat Studies in the Philosophy of Mathematics*, Boston Studies in the Philosophy and History of Science. Springer. Forthcoming.

[Friedman and Ternullo, S] Friedman, S.-D. and Ternullo, C. (S). Believing the new axioms. Submitted.

[Gettier, 1963] Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23:121–123.

[Gitman and Hamkins, 2010] Gitman, V. and Hamkins, J. D. (2010). A natural model of the multiverse axioms. *Notre Dame Journal of Formal Logic*, 51(4):475–484.

[Gitman et al., U] Gitman, V., Hamkins, J. D., and Johnstone, T. (U). Kelley-Morse set theory and choice principles for classes. Unpublished.

[Gödel, 1931] Gödel, K. (1931). On formally undecidable propositions of *Principia mathematica* and related systems I. In *[Gödel, 1986]*, pages 144–195. Oxford University Press.

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

[Gödel, 1947] Gödel, K. (1947). What is Cantor's continuum problem? In *[Gödel, 1990]*, pages 176–187. Oxford University Press.

[Gödel, 1964] Gödel, K. (1964). What is Cantor's continuum problem? In *[Gödel, 1990]*, pages 254–270. Oxford University Press.

[Gödel, 1986] Gödel, K. (1986). *Collected Works, Volume I: Publications 1929-1936*. Oxford University Press. Edited by: Solomon Feferman (Editor-in-chief), John W. Dawson, Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, Jean van Heijenoort.

[Gödel, 1990] Gödel, K. (1990). *Collected Works, Volume II: Publications 1938-1974*. Oxford University Press. Edited by: Solomon Feferman (Editor-in-chief), John W.

Dawson, Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, Jean van Heijenoort.

[Hale, 2013] Hale, B. (2013). *Necessary Beings: An Essay on Ontology, Modality, and the Relations Between Them*. Oxford University Press.

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

[Hamkins et al., 2012] Hamkins, J., Kirmayer, G., and Perlmutter, N. (2012). Generalizations of the Kunen inconsistency. *Annals of Pure and Applied Logic*, 163(12):1972–1890.

[Hamkins, 2012] Hamkins, J. D. (2012). The set-theoretic multiverse. *The Review of Symbolic Logic*, 5(3):416–449.

[Hamkins, 2015] Hamkins, J. D. (2015). Being *hod*-of-a-set is invariant throughout the generic multiverse. available from http://jdh.hamkins.org/being-hod-of-a-set-is-invariant-throughout-the-generic-multiverse/.

[Hamkins and Seabold, 2012] Hamkins, J. D. and Seabold, D. E. (2012). Well-founded Boolean ultrapowers as large cardinal embeddings. arXiv:1206.6075 [math.LO].

[Hamkins and Yang, 2013] Hamkins, J. D. and Yang, R. (2013). Satisfaction is not absolute. arXiv:1312.0670v1 [math.LO].

[Hauser, 2006] Hauser, K. (2006). Gödel's program revisited part I: The turn to phenomenology. *Bulletin of Symbolic Logic*, 12(4):529–590.

[Hellman, 1989] Hellman, G. (1989). *Mathematics Without Numbers*. Oxford University Press.

[Hewitt, 2012] Hewitt, S. (2012). The logic of finite order. *Notre Dame Journal of Formal Logic*, 53(3):297–318.

[Hilbert, 1922] Hilbert, D. (1922). The new grounding of mathematics. First report. In *[Ewald, 1996b]*, volume 2, pages 1115–1134. Oxford University Press.

[Hilbert, 1925] Hilbert, D. (1925). On the infinite. In *[van Heijnoort, 1967]*, pages 367–392. Harvard University Press.

[Holmes, 1998] Holmes, R. (1998). *Elementary Set Theory with a Universal Set*. Bruylant-Academia.

[Holy et al., F] Holy, P., Krapf, R., Lücke, P., Njegomir, A., and Schlicht, P. (F). Class forcing, the forcing theorem and Boolean completions. Forthcoming in the*Journal of Symbolic Logic*.

[Horsten and Welch, F] Horsten, L. and Welch, P. (F). Reflecting on absolute infinity. *Journal of Philosophy*. Forthcoming.

[Iannucci et al., 2009] Iannucci, A., Tandy, A., and Loader, K. (2009). In the Loop. Motion picture.

[Isaacson, 2011] Isaacson, D. (2011). The reality of mathematics and the case of set theory. In Noviak, Z. and Simonyi, A., editors, *Truth, Reference, and Realism*, pages 1–75. Central European University Press.

[Jech, 1973] Jech, T. (1973). *The Axiom of Choice*. North Holland Publishing Co.

[Jech, 2002] Jech, T. (2002). *Set Theory*. Springer.

[Jensen, 1968] Jensen, R. (1968). On the consistency of a slight (?) modification of Quine's "New Foundations". *Synthese*, 19(1/2):250–264.

[Jensen, 1972] Jensen, R. (1972). The fine structure of the constructible hierarchy. *Annals of Mathematical Logic*, 4:229–308.

[Jensen, 1995] Jensen, R. (1995). Inner models and large cardinals. *The Bulletin of Symbolic Logic*, 1(4):393–407.

[Kanamori, 2009] Kanamori, A. (2009). *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings*. Springer, 2nd edition.

[Koellner, 2009] Koellner, P. (2009). On reflection principles. *Annals of Pure and Applied Logic*, 157:206–219.

[Koellner, 2010] Koellner, P. (2010). On the question of absolute undecidability. In Feferman, S., Parsons, C., and Simpson, S. G., editors, *Kurt Gödel: Essays for his Centennial*, pages 189–222. Association for Symbolic Logic.

[Koellner, 2011] Koellner, P. (2011). Independence and large cardinals. In *The Stanford Encyclopedia of Philosophy*. `http://plato.stanford.edu/archives/sum2011/entries/independence-large-cardinals/`, Summer 2011 edition.

[Koellner, 2013] Koellner, P. (2013). Hamkins on the multiverse. In Koellner, P., editor, *Exploring the Frontiers of Incompleteness*.

[Kunen, 1971] Kunen, K. (1971). Elementary embeddings and infinitary combinatorics. *The Journal of Symbolic Logic*, 36:407–413.

[Kunen, 2013] Kunen, K. (2013). *Set Theory*. College Publications.

[Levy, 1960] Levy, A. (1960). Axiom schemata of strong infinity and axiomatic set theory. *Pacific Journal of Mathematics*, 10:223–238.

[Linnebo, 2006] Linnebo, Ø. (2006). Sets, properties, and unrestricted quantification. In *Absolute Generality*. Oxford University Press.

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

[Linnebo, 2014] Linnebo, Ø. (2014). Plural quantification. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2014 edition.

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

[Linnebo and Pettigrew, 2011] Linnebo, Ø. and Pettigrew, R. (2011). Category theory as an autonomous foundation. *Philosophia Mathematica*, 19:227–254.

[Linnebo and Rayo, 2012] Linnebo, Ø. and Rayo, A. (2012). Hierarchies ontological and ideological. *Mind*, 121(482):269–308.

[Mac Lane, 1971] Mac Lane, S. (1971). *Categories for the Working Mathematician*. Springer.

[Maddy, 1983] Maddy, P. (1983). Proper classes. *Journal of Symbolic Logic*, 48:113–119.

[Maddy, 1988a] Maddy, P. (1988a). Believing the axioms. I. *The Journal of Symbolic Logic*, 53(2):481–511.

[Maddy, 1988b] Maddy, P. (1988b). Believing the axioms II. *The Journal of Symbolic Logic*, 53(3):736–764.

[Maddy, 1997] Maddy, P. (1997). *Naturalism in Mathematics*. Oxford University Press.

[Maddy, 2007] Maddy, P. (2007). *Second Philosophy*. Oxford University Press.

[Maddy, 2011] Maddy, P. (2011). *Defending the Axioms*. Oxford University Press.

[Martin, 2001] Martin, D. (2001). Multiple universes of sets and indeterminate truth values. *Topoi*, 20(1):5–16.

[McGee, 1997] McGee, V. (1997). How we learn mathematical language. *The Philosophical Review*, 106(1):35–68.

[Meadows, 2013] Meadows, T. (2013). What can a categoricity theorem tell us? *The Review of Symbolic Logic*, 6:524–544.

[Menzel, 1986] Menzel, C. (1986). On the iterative explanation of the paradoxes. *Philosophical Studies*, 49:37–61.

[Menzel, 2014] Menzel, C. (2014). $ZFCU$, wide sets, and the iterative conception. *Journal of Philosophy*, 111(2):57–83.

[Montague, 1961] Montague, R. (1961). Fraenkel's addition to the axioms of Zermelo. In Fraenkel, A. and Bar-Hillel, Y., editors, *[Fraenkel and Bar-Hillel, 1966]*, pages 91–114. Magnes Press Hebrew University.

[Moore, 1990] Moore, G. H. (1990). Introductory note to 1947 and 1964. In *[Gödel, 1990]*, pages 154–175. Oxford University Press.

[Muller, 2001] Muller, F. A. (2001). Sets, classes and categories. *British Journal for the Philosophy of Science*, 52:539–573.

[Oliver and Smiley, 2013] Oliver, A. and Smiley, T. (2013). *Plural Logic*. Oxford University Press.

[Paseau, 2007] Paseau, A. (2007). Boolos on the justification of set theory. *Philosophia Mathematica*, 15:30–53.

[Paseau, 2009] Paseau, A. (2009). Reducing arithmetic to set theory. In Bueno, O. and Linnebo, Ø., editors, *New waves in Philosophy of Mathematics*, pages 35–55. Palgrave Macmillan.

[Potter, 2004] Potter, M. (2004). *Set Theory and its Philosophy: A Critical Introduction*. Oxford University Press.

[Quine, 1937] Quine, W. (1937). New foundations for mathematical logic. *The American Mathematical Monthly*, 44:70–80.

[Quine, 1960] Quine, W. V. O. (1960). *Word and Object*. Harvard University Press.

[Quine, 1969a] Quine, W. V. O. (1969a). Epistemology naturalized. In *[Quine, 1969b]*, pages 69–90. Columbia University Press.

[Quine, 1969b] Quine, W. V. O. (1969b). *Ontological Relativity and Other Essays*. Columbia University Press.

[Rathjen, 2012] Rathjen, M. (2012). Constructive Zermelo-Fraenkel set theory, power set, and the calculus of constructions. In Dybjer, P., Lindstrom, S., Palmgren, E., and Sundholm, G., editors, *Epistemology versus Ontology: Essays on the philosophy and foundations of mathematics in honour of Per Martin-Löf*. Springer.

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

[Reinhardt, 1980] Reinhardt, W. (1980). Satisfaction definitions and axioms of infinity in a theory of properties with necessisty operator. In Arruda, A., Chuaqui, R., and Costa, N. D., editors, *Mathematical Logic in Latin America*, pages 267–303. North Holland Publishing Co.

[Russell, 1902] Russell, B. (1902). Letter to Frege. In *[van Heijnoort, 1967]*, pages 124–125. Harvard University Press.

[Russell, 1907] Russell, B. (1907). The regressive method of discovering the premises of mathematics. In *[Russell, 1973]*, pages 272–283. George Allen & Unwin Ltd.

[Russell, 1973] Russell, B. (1973). *Essays in Analysis*. George Allen & Unwin Ltd. Edited by Douglas Lackie.

[Russell and Whitehead, 1910] Russell, B. and Whitehead, A. (1910). *Principia Mathematica to *56*. Cambridge University Press.

[Sargsyan, 2013] Sargsyan, G. (2013). Descriptive inner model theory. *Bulletin of Symbolic Logic*, 19(1):1–55.

[Schimmerling, 2001] Schimmerling, E. (2001). The abc's of mice. *The Bulletin of Symbolic Logic*, 7(4):485–503.

[Schindler, 2000] Schindler, R. (2000). Proper forcing and remarkable cardinals: I. *Bulletin of Symbolic Logic*, 6(2):176–184.

[Schindler, 2001] Schindler, R. (2001). Proper forcing and remarkable cardinals: II. *Journal of Symbolic Logc*, 66(3):1481–1492.

[Schindler, 2014] Schindler, R. (2014). *Set Theory: Exploring Independence and Truth*. Springer.

[Shapiro, 1991] Shapiro, S. (1991). *Foundations without Foundationalism: A Case for Second-order Logic*. Oxford University Press.

[Shapiro and Wright, 2006] Shapiro, S. and Wright, C. (2006). All things indefinitely extensible. In Rayo, A. and Uzquiano, G., editors, *Absolute Generality*. Oxford University Press.

[Shepherdson, 1953] Shepherdson, J. (1953). Inner models for set theory–part iii. *The Journal of Symbolic Logic*, 18(2):145–167.

[Shepherdson, 1951] Shepherdson, J. C. (1951). Inner models for set theory–part i. *Journal of Symbolic Logic*, 16(3):161–190.

[Shepherdson, 1952] Shepherdson, J. C. (1952). Inner models for set theory–part ii. *Journal of Symbolic Logic*, 17(4):225–237.

[Shoenfield, 1967] Shoenfield, J. (1967). *Mathematical Logic*. Addison-Wesley Publishing Co.

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

[Smullyan and Fitting, 1996] Smullyan, R. M. and Fitting, M. (1996). *Set Theory and the Continuum Problem*. Clarendon Press.

[Solovay et al., 1978] Solovay, R., Reinhardt, W., and Kanamori, A. (1978). Strong axioms of infinity and elementary embeddings. *Annals of Mathematical Logic*, 13:73–116.

[Steel, 1996] Steel, J. (1996). The core model iterability problem. In *Lecture Notes in Logic*, volume 8. Springer.

[Steel, 2014] Steel, J. (2014). Gödel's program. In Kennedy, J., editor, *Interpreting Gödel*. Cambridge University Press.

[Sutcliffe, 1968] Sutcliffe, F. (1968). *Discourse on Method and The Meditations*. Penguin.

[Suzuki, 1999] Suzuki, A. (1999). No elementary embedding from V into V is definable from parameters. *The Journal of Symbolic Logic*, 64(4):1591–1594.

[Tait, 2005] Tait, W. (2005). Constructing cardinals from below. In *The Provenance of Pure Reason: Essays in the Philosophy of Mathematics and Its History, Logic and Computation in Philosophy*, pages 133–154. Oxford University Press.

[Uzquiano, 2003] Uzquiano, G. (2003). Plural quantification and classes. *Philosophia Mathematica*, 11(1):67–81.

[Uzquiano, 2006] Uzquiano, G. (2006). Unrestricted unrestricted quantification: The cardinal problem of absolute generality. In *Absolute Generality*. Oxford University Press.

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

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

[Vickers and Welch, 2001] Vickers, J. and Welch, P. (2001). On elementary embeddings from an inner model to the universe. *The Journal of Symbolic Logic*, 66(3):1090–1116.

[Welch, 2014] Welch, P. (2014). Global reflection principles. Isaac Newton Institute pre-print series, No. NI12051-SAS.

[Williamson, F] Williamson, T. (F). Absolute provability and safe knowledge of axioms. In Horsten, L. and Welch, P., editors, *The Limits of Mathematical Knowledge*. Oxford University Press.

[Yablo, 2014] Yablo, S. (2014). *Aboutness*. Princeton University Press.

[Zermelo, 1930] Zermelo, E. (1930). On boundary numbers and domains of sets. In *[Ewald, 1996b]*, volume 2, pages 1208–1233. Oxford University Press.