Neo-Logicism? An Ontological Reduction of Mathematics to Metaphysics*

Edward N. Zalta†
Center for the Study of Language and Information
Stanford University

Die natürlichen Zahlen hat der liebe Gott gemacht, der Rest ist Menschenwerk.
Leopold Kronecker

It is now well accepted that logicism is false. The primitive notions and proper axioms of mathematical theories are not reducible to primitive logical notions and logical axioms. Even the idea underlying logicism appears to be somewhat problematic, for if the existence claims of mathematics are to be reducible to logical truths, then it would seem that the logicist has to assert the existence of objects of some kind. Though a case can, and has, been made for thinking that logic can contain or imply existence claims, the matter is at least controversial and so there is at least a question as to whether a reduction of mathematics to a logic with existence claims would constitute a reduction of mathematics to logic.

In this paper, however, we defend a philosophical thesis which may preserve some of the spirit of logicism. Our thesis is that mathematical objects just are (reducible to) the abstract objects systematized by a certain axiomatic, *mathematics-free* metaphysical theory. This thesis appears to be a version of mathematical platonism, for if correct, it would make a certain simple and intuitive philosophical position about mathematics much more rigorous, namely, that mathematics describes a realm of abstract objects. Nevertheless, there are two ways in which the present view might constitute a kind of neo-logicism. The first is that the comprehension principle for abstract objects that forms part of the metaphysical theory can be reformulated as a principle that 'looks and sounds' like an analytic, if not logical, truth. Although we shall not argue here that the reformulated principle is analytic, other philosophers have argued that principles analogous to it are. The second is that the abstract objects systematized by the metaphysical theory are, in some sense, *logical objects*. By offering a reduction of mathematical objects to logical objects, the present view may thereby present us with a new kind of logicism.

*This paper was published in *Erkenntnis*, **53**/1-2 (2000): 219–265.

†I am indebted to the Center for the Study of Language and Information and its Director, John Perry, for supporting my research. I would like to thank Mark Colyvan, Bernard Linsky, Chris Menzel, Karl-Georg Niebergall, and an anonymous referee, all of whom carefully read the penultimate draft and offered many suggestions for improvement. Thanks also goes to Sol Feferman, Allen Hazen, Thomas Hofweber, and Brent Mundy for valuable discussions about the ideas contained herein. Finally, I would like to thank Godehard Link for inviting me to the workshop Concepts of Reduction in Logic and Philosophy (in September 1997), where this material was first presented.

To establish our thesis, we need two elements, the first of which is already in place. The first element is the axiomatic, metaphysical theory of abstract objects. We shall employ the background ontology described by the axiomatic theory of abstract objects developed in Zalta [1983] and [1988].[1] The axioms of this theory can be stated without appealing to mathematical primitives or notions of any kind; one of these axioms is a comprehension principle for abstract objects the instances of which explicitly assert the existence of such objects.

The second element required for our thesis is this: we must show that for an arbitrary mathematical theory T, there is a precise interpretation of the terms and predicates of T which (a) analyzes these expressions as denoting abstract objects in the background ontology, and (b) defines a sense in which the theorems of T are true. This second element has only been sketched in previous work. In Zalta [1983] and Linsky & Zalta [1995], a basic analysis of the language of mathematics was developed. The present paper advances the previous work by offering a much more detailed account of reference and truth with respect to mathematical language. In connection with reference, we explicitly identify the steps required to interpret the well-defined terms and relation symbols of an arbitrary mathematical theory so that those expressions denote unique abstract individuals and abstract relations, respectively, in our background ontology. This task is accomplished in Section 4. In connection with

[1]See also Zalta [1993] or [1999] for briefer sketches of the theory and specific applications.

truth, we use the theory of abstract objects to state the conditions under which the theorems of mathematical theories are true. This task is accomplished in Section 6. In the final section of the paper (Section 7), we make several observations about our work and we briefly consider the extent to which the present view constitutes a kind of neo-logicism.

Our background axiomatic theory of abstract objects, henceforth '$\mathcal{O}$', has been developed in both a modal and a type-theoretic setting and these two manifestations of the theory have been applied in numerous ways. The *modal* version of $\mathcal{O}$ axiomatizes abstract individuals, and in the applications of this theory, the laws governing possible worlds, Platonic forms, Leibnizian concepts, and natural numbers (among other things) have been derived and the language of fiction and belief has been given a precise interpretation. The *typed* version of $\mathcal{O}$ axiomatizes the abstract objects of every simple logical type. This version asserts that for each logical type of object (e.g., individual, property of individuals, relation among individuals, property of properties of individuals, relation among properties of individuals, etc.), there are abstract objects of that type (in addition to ordinary objects of that type). In the various applications of this type theory, such things as the Fregean senses of predicates, the fictional properties and relations of rejected scientific theories, mathematical properties (e.g., being prime), and mathematical relations (e.g., set membership) have all been identified as particular abstract properties and abstract relations.[2]

Now it might be thought that all of these applications of $\mathcal{O}$ involve a single kind of ontological reduction. But, in fact, this is not the case. It is very important to recognize that there are two distinct kinds of ontological reduction that can be constructed within $\mathcal{O}$. We shall call these *classical* and *metaphysical* reductions, respectively. This distinction is critical to what follows—we estalish the main thesis of this paper by developing a *metaphysical* reduction of the objects of mathematical theories to the objects of our metaphysics.

In order to distinguish classical and metaphysical reductions in the context of $\mathcal{O}$, it is important to mention first that in *classical* reductions, the reduction is between theories—the axioms of some *theory* T are derived as theorems of $\mathcal{O}$. In the case of *metaphysical* reductions, however, it is more perspicuous to say that the *objects* of a theory T are reduced to the objects of $\mathcal{O}$. (We shall precisely define 'object of theory T' in Section 3.)

[2]See the final chapters of Zalta [1983] and [1988].

Now various philosophers have described the basic idea of a classical reduction as follows: a theory S is reducible to a theory T just in case all of the non-logical notions of S can be explicitly defined in T in such a way that the translations of the theorems of S (via the definitions) are theorems of T.[3] We can recast this definition in language more familiar to logicians by saying that S is reducible to T just in case the theorems of S constitute a subtheory of a definitional extension of T. This definition gives us the basic sense of reduction that the logicist might have used to claim that the proper axioms of mathematics are reducible to the theorems of logic. The logicist idea was that a logic L is defined by a set of analytically true logical axioms and rules of inference. Then, a mathematical theory T is reducible to L just in case the primitive terms and predicates of T are definable in the language of L and the proper theorems of T (when translated into the language of L) become logical theorems of L. Of course, as mentioned earlier, probably no one now believes that the primitive notions or proper axioms and theorems of mathematical theories are reducible in this way to the primitive notions and axioms of logic.

It is important to digress briefly to mention the fact that modern logicians have introduced a variety of much more explicit and fine-grained notions of reduction between theories T and S. They have defined such notions as relative interpretability, proof-theoretic reduction, model-theoretic reduction, and even axiomatized notions of reduction.[4] However, these more fine-grained notions of reducibility will not play a role in what follows. Although the classical reductions available in $\mathcal{O}$ are all instances of relative interpretations, it should become apparent that the above definition of a classical reduction should suffice for the purposes of this paper.

In Section 1, we shall rehearse some classical reductions that are available in $\mathcal{O}$. We'll see that the primitive notions of situation theory, possible world theory, and Dedekind/Peano number theory can be defined in the language of $\mathcal{O}$ and that the proper axioms governing these notions can be (couched in terms of these explicit definitions and) derived as theorems

[3]See Carnap [1967] (p. 6), Quine [1976] (p. 218), and Jubien [1969] (p. 534).

[4]For relative interpretability, see Tarski *et al.* [1953], Feferman [1960], and Visser [1998]. See Feferman [1988] (or [1998a]) for the definition and discussion of 'proof-theoretic reduction'. Finally, see Niebergall [2000], in which the notions of model-theoretic reduction are critically discussed and axioms for the reducibility relation are proposed.

of $\mathcal{O}$. However, as we mentioned earlier, we shall not be appealing to classical reductions of any kind to establish the main thesis of this paper.

Instead, we shall defend our thesis by developing a new kind of ontological reduction which we call *metaphysical* reduction. The additional ontological resources provided by $\mathcal{O}$ make it possible to develop this distinctive and essentially different kind of reduction. We'll spend the preponderance of the paper preparing the ground for, and developing examples of, metaphysical reductions. The examples will show that, for an arbitrary mathematical theory T, a metaphysical reduction identifies both the reference of the well-defined terms and predicates of T and preserves a sense in which the theorems of T are true. More specifically, $\mathcal{O}$ will provide us with a mathematics-free theoretical framework in which we can precisely specify abstract individuals and abstract relations. Certain specifications of abstracta simply objectify the *roles* that mathematical individuals and relations play in a mathematical theory T. So once we extend $\mathcal{O}$ by adding the terms and predicates of T and by adding the analytic mathematical truths which articulate the role that the mathematical objects of T are alleged to play in T, we'll be able to theoretically identify those objects with their objectified roles. The theorems of an arbitrary mathematical theory T will then have compositionally specifiable readings in $\mathcal{O}$ on which they (the theorems) turn out to be true. Moreover, the abstract individuals and relations of $\mathcal{O}$ figure into these readings. Our metaphysical reductions will therefore show that each mathematical theory is about distinctive abstract individuals and abstract relations.

So the reason neither classical reductions nor the other more fine-grained notions of reduction will play a significant role in what follows is that if the view developed here is correct, metaphysical reductions will show that every mathematical theory is about its own distinctive kinds of abstract individuals and/or abstract relations. From the point of view of ontology and the philosophy of language, then, there may be no *metaphysical* reason to investigate these other kinds of reduction of one mathematical theory to another mathematical theory. Of course, there are still *mathematical* reasons to investigate classical and more fine-grained notions of reduction as they apply to mathematical theories, for example, to assess the mathematical power of certain theories in various ways and to understand the various ways in which one mathematical theory might be distinguished as a foundation for the rest of mathematics. But even if one mathematical theory emerges as *the* foundational theory from (in)

which all other mathematical theories can be derived (interpreted), or as that in which all other mathematical theories can be modeled, it doesn't follow that those other mathematical theories are just theories *of* or *about* the objects described by the foundational theory, at least not if we can show that each mathematical theory is about its own distinctive kind of mathematical individuals and relations. Moreover, $\mathcal{O}$ would offer a metaphysical account of truth and reference for the language and theorems of any mathematical theory that emerges as a foundation for mathematics.

Although we shall return to these issues in the last section of the paper, it is important to mention one issue to which we shall not return, namely, our assumption that both classical and metaphysical reductions constitute genuine *ontological* reductions. Although this may be a controversial matter, we shall not spend time in this paper on the matter; instead, we shall assume that the work carried out here offers *some* reason to think that the ontological categories *mathematical individual* and *mathematical relation* are not *sui generis* but rather subcategories of the more fundamental ontological categories *abstract individual* and *abstract relation*, respectively.

We turn, then, to a brief description of some classical reductions in $\mathcal{O}$, so that we will be better prepared to appreciate what is distinctive about the metaphysical reductions that establish the main thesis of this paper. Readers familiar with these applications of object theory may skip ahead to Section 2.

§1: Classical Reductions in $\mathcal{O}$

In order to discuss the classical reductions that have been effected in $\mathcal{O}$, it will be important for the reader to know the language and axioms of the theory. In what follows, we shall presuppose that the reader is familiar with one of the canonical presentations of $\mathcal{O}$ in other publications. In this section, we shall discuss the version of the theory that has been expressed in a syntactically second-order modal (S5 with Barcan formulas) predicate calculus (without identity) which has been modified so as to include a second kind of atomic formula, namely, formulas of the form 'xF^1' (individual x *encodes* property F^1). A single theoretical primitive property '$E!$' ('being concrete') is used to define the property of being abstract ($A!x =_{df} \neg\Diamond E!x$) and the comprehension principle for abstract individuals asserts that for any condition φ (without free xs), there is an abstract individual that encodes just the properties satisfying φ (i.e.,

$\exists x(A!x \,\&\, \forall F(xF \equiv \varphi)))$.[5] Abstract individuals are said to be identical whenever they necessarily encode the same properties, but to show that x and y are the same abstract individual, it suffices to show that x and y encode the same properties, since the logic of encoding is rigid (i.e., $\Diamond xF \rightarrow \Box xF$).

The canonical formulation of the theory of abstract individuals also includes two kinds of complex term. There is a complex way of denoting individuals, namely, rigid definite descriptions of the form $\imath x\varphi$ (for any formula φ). These definite descriptions are axiomatized in the usual way, namely, by a principle which asserts that Russell's analysis of descriptions applies to any atomic formula that contains a description.[6] There is also a complex way of denoting relations, namely, λ-expressions of the form $[\lambda y_1 \ldots y_n\ \varphi]$ (where φ has no free Fs, no encoding subformulas and no descriptions). These λ-expressions are axiomatized by the usual principle λ-Conversion (i.e., λ-abstraction), and by a principle which ensures that exchange of bound variables makes no difference to the relation denoted by the λ-expression.[7] λ-Conversion immediately yields a comprehension principle for relations.[8] The theory of relations is completed by a definition of when relations are said to be identical.[9] Substitution of identicals (whether identical abstract individuals or identical relations) is stipulated to work in all contexts.

Now, in terms of this language and theory, several classical reductions have been effected. These reductions are achieved in precisely the way one would expect, with the exception that it is often the case that the target theory has not been given a canonical presentation. So, for example, although situation theory does not have a canonical axiomatization, we defined (in Zalta [1993]) the following basic notions of situation theory in the language of $\mathcal{O}$: x is a situation, situation s makes state of affairs p true, and situation s is a part of situation t. From these definitions, the *usual* axioms of situation theory are derivable as theorems of $\mathcal{O}$. Similarly, although the theory of possible worlds does not have a canonical axiomatization, we defined (in Zalta [1983] and [1993]) the following notions of world theory: object x encodes proposition p, x is a possible world, proposition p is true at world w, w is maximal, w is consistent, w is modally closed, and w is actual.[10] Then we derived the *usual* principles of world theory: (a) every world is maximal, (b) every world is consistent, (c) every world is modally closed, (d) there is a unique actual world, (e) a proposition is necessarily true iff it is true in all possible worlds,

---

[5]We call individuals x that might have been concrete 'ordinary objects'. In formal terms: $O!x =_{df} \Diamond E!x$. It is axiomatic that ordinary individuals necessarily fail to encode properties.

[6]More specifically, the following is an axiom:

$\psi^{\imath x\varphi}_y \equiv \exists x(\varphi \,\&\, \forall z(\varphi^z_x \rightarrow z\!=\!x) \,\&\, \psi^x_y)$, for any atomic or identity formula $\psi(y)$ in which y is free.

To accomodate descriptions, the classical quantification theory is modified only so as to be 'free' with respect to formulas containing descriptions. Moreover, the above axiom governing descriptions is a logical truth that is not a necessary truth (for the descriptions denote rigidly what they denote at the actual world). So the classical S5 modal logic is modified only to admit the presence of contingent logical truths (the Rule of Necessitation may not be applied to any line that depends on the above axiom governing descriptions).

[7]More specifically, the following are axioms:

$[\lambda y_1 \ldots y_n\ \varphi]x_1 \ldots x_n \equiv \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$

$[\lambda y_1 \ldots y_n\ \varphi] = [\lambda y'_1 \ldots y'_n\ \varphi']$,
where the two λ-expressions are alphabetic variants.

It is also an axiom that:

$[\lambda y_1 \ldots y_n\ F^n y_1 \ldots y_n] = F^n$

Thus, an 'elementary' λ-expression is intersubstitutable for the relation symbol that appears in that expression.

[8]More specifically, the following is a theorem:

$\exists F^n \Box \forall x_1 \ldots \forall x_n (F^n x_1 \ldots x_n \equiv \varphi)$,
where φ has no free F^ns, no encoding subformulas, and no descriptions.

Of course, a relation can be specified in terms of a formula φ containing a definite description $\imath x\varphi$ if it is first proved that $\exists y(y\!=\!\imath x\varphi)$.

[9]More specifically, the definition of identity proceeds by first defining identity for properties F^1 and G^1:

$F^1\!=\!G^1 =_{df} \Box\forall x(xF^1 \equiv xG^1)$

In terms of this definition, we employ λ-expressions with vacuously bound variables to define identity for propositions. Using 'p' and 'q' instead of 'F^0' and 'G^0', we define:

$p\!=\!q =_{df} [\lambda y\ p]\!=\![\lambda y\ q]$

Finally, a definition of relation identity for n-place relations ($n \geq 2$) is constructible in terms of identity for properties. Interested readers may consult one of the presentations of $\mathcal{O}$ cited in the text.

[10]We provide the formal definitions here for those readers unfamiliar with this work. In these definitions, we give the symbols Σ and $\models$ the narrowest possible scope. (For example, '$\Sigma_x p \equiv p$' is to be read as '$(\Sigma_x p) \equiv p$' and '$w \models p \rightarrow p$' is to be read as '$(w \models p) \rightarrow p$'.) We define: (1) x *encodes* p ('$\Sigma_x p$') $\equiv x[\lambda y\ p]$; (2) *World*$(x) \equiv \Diamond\forall p(\Sigma_x p \equiv p)$; (3) p *is true at* w ('$w \models p$') $\equiv \Sigma_w p$; (4) *Maximal*$(w) \equiv \forall p(w \models p \vee w \models \neg p)$; (5) *Consistent*$(w) \equiv \neg\exists p(w \models p \,\&\, w \models \neg p)$; (6) *Modally-closed*$(w) \equiv [w \models p \,\&\, \Box(p \rightarrow q)] \rightarrow w \models q$; and (7) *Actual*$(w) \equiv \forall p(w \models p \rightarrow p)$. See Zalta [1993] for further discussion.

(f) a proposition is possibly true iff it is true in some possible world, (g) whenever worlds have the same propositions true at them, those worlds are identical.[11] Although world theory has not had a canonical axiomatization, it seems clear that any attempt to axiomatize this theory would employ some subset of these principles as axioms. If this is right, then we have a classical ontological reduction of world theory to $\mathcal{O}$.

$\mathcal{O}$ has also been applied so as to reduce Leibniz's theory of (complete individual) concepts and Plato's theory of forms.[12] But as an example where the theory being reduced does have a canonical axiomatization, we note the following. When $\mathcal{O}$ is extended with the logic of actuality and two *a priori* and plausible axioms, Frege's definitions of *predecessor* and 0 can be constructed and the Dedekind/Peano axioms for number theory become theorems (in addition to many Fregean principles about natural numbers). In Zalta [1999], we defined the following notions in the language of $\mathcal{O}$: x is a predecessor of y, x is the number of Gs, x is a natural number, and zero. We then added the formal versions of the following axioms to $\mathcal{O}$: (a) predecessor and its weak ancestral are relations, and (b) if there is a natural number n which numbers the property G, then there might have been a concrete individual distinct from all of the concrete individuals that actually exemplify G.[13] As a result, the Dedekind/Peano axioms for the theory of natural numbers become provable in $\mathcal{O}$.

From a *logical* point of view, there is nothing unusual or distinctive about these classical ontological reductions available in $\mathcal{O}$. They may, however, hold *philosophical* interest for the metaphysician or logician interested in minimizing the number of ontological categories or concerned to find a system in which there are proofs of metaphysical claims that philosophers typically have to stipulate.[14]

[11]Using the formal definitions supplied in the previous footnote, principles (a) – (g) in the text become following the theorems of $\mathcal{O}$: (a) $\forall w Maximal(w)$; (b) $\forall w Consistent(w)$; (c) $\forall w Modally\text{-}closed(w)$; (d) $\exists! w Actual(w)$; (e) $\Box p \equiv \forall w(w \models p)$; (f) $\Diamond p \equiv \exists w(w \models p)$; and (g) $\forall p(w \models p \ \equiv \ w' \models p) \ \rightarrow w = w'$. See Zalta [1993] for the proofs.

[12]See Zalta [1983] for the initial sketches. However, Zalta [2000] has a comprehensive treatment of Leibniz, and Pelletier & Zalta [2000] has a comprehensive treatment of Plato.

[13]We rendered these axioms formally by first defining the technical notions $Precedes(x, y)$, (its weak ancestral) $Precedes^+(x, y)$, $\#_F$ ('the number of Fs'), and $NaturalNumber(x)$ in the language of $\mathcal{O}$. Then the following are axioms:

- $\exists F \forall x \forall y(Fxy \equiv Precedes(x, y))$
  $\exists F \forall x \forall y(Fxy \equiv Precedes^+(x, y))$
- $\exists x(NaturalNumber(x) \ \& \ x = \#_G) \rightarrow \Diamond \exists y(E!y \ \& \ \forall u(\mathcal{A}Gu \rightarrow u \neq_E y))$

The variable 'u' here ranges over possibly concrete objects. See Zalta [1999], Section 5.

§2: Metaphysical Reductions

The classical reductions just discussed are now to be contrasted with the new and distinctive *metaphysical* reductions that are available in $\mathcal{O}$. The point of these metaphysical reductions is not to define the primitive nonlogical notions of mathematics in the language of $\mathcal{O}$ or to derive the axioms of mathematical theories as theorems of $\mathcal{O}$, but rather to: (1) interpret (i.e., formulate denotational truth conditions for) the language and axioms of mathematical theories so as to reveal both that the individual terms of those theories denote abstract individuals and that the relation symbols of those theories denote abstract relations, and (2) develop readings in $\mathcal{O}$ of the theorems of mathematical theories on which those theorems turn out true. Our metaphysical reductions will also show that the mathematical theories themselves can be identified as abstract individuals. In order to make these ideas perfectly clear and precise, we shall need to introduce some technical machinery. This technical machinery will be developed in Section 3. Then, in Section 4, we construct the theoretical descriptions that reduce mathematical objects to abstract objects. In the meantime, however, it is important to prepare the reader for the material in the following sections by briefly outlining the philosophy of mathematics on which this material is based.[15]

Our philosophy of mathematics assumes that the primary data that requires a philosophical analysis are the *true* ordinary mathematical sentences of the form 'In mathematical theory T, p (is true)'. So, for example, we shall try to systematically interpret such statements as 'In real number theory, π is greater than 3' and 'In Zermelo-Fraenkel set theory, no set is a member of the empty set'. Note that these statements, when stripped of the prefix 'In mathematical theory T', are frequently expressed in the formal languages of mathematics. We shall want to show that both

[14]The reductions may also hold some interest for the epistemologist concerned to justify the foundational axioms of metaphysics, especially if they assert the existence of abstract individuals. Linsky and Zalta [1995] has something to offer this project, however.

[15]This philosophy of mathematics was first sketched in a kind of 'fictionalist' form in Zalta [1983] (Chapter VI) and then articulated in some detail in its 'platonist' guise in Linsky and Zalta [1995]. More will be said about the fictionalist and platonist interpretations of the theory in Section. 7.

the mathematical expressions of English and the formal expressions of mathematics denote distinctive abstract objects. (Both the English and the formal symbols require a philosophical interpretation and analysis.) So, on *some* occasions in what follows, we will refer to both the unprefixed English sentences (e.g., 'No set is a member of the empty set') and their standard formal renditions (e.g., '$\neg\exists x(Set(x) \,\&\, x \in \emptyset)$') as *ordinary* statements of mathematical language. We shall call the ordinary language prefix 'In mathematical theory T, ...' the 'theory operator' and, in what follows, we will define a formal notion of $\mathcal{O}$ which will be used to precisely translate this theory operator. The translation thereby yields well-defined, compositional truth conditions for the prefixed statements of mathematics. By contrast, the unprefixed statements become subject to an ambiguity that will be resolved in the framework.

The idea underlying this ambiguity is that predication in ordinary language (including that of mathematics) is subject to a structural ambiguity that is disambiguated by our two modes of predication 'Fx' and 'xF'. The unprefixed statements of mathematics (e.g., 'π is irrational' and 'No set is a member of the empty set') are subject to this ambiguity. They have a reading on which they turn out *true* and a reading on which they turn out false. The true readings of the unprefixed sentences will be analyzed as encoding predications. For example, on one reading, the statement 'π is irrational' (made in connection with real number theory $\Re$) is true if and only if a certain abstract individual, namely $\pi_{\Re}$ (which can be precisely identified), encodes a certain abstract property, namely, *being irrational*$_{\Re}$ (which can also be precisely identified). But on the second reading, 'π is irrational' is true iff $\pi_{\Re}$ exemplifies *being irrational*$_{\Re}$. We take these second 'exemplification' readings of ordinary mathematical statements to be false. $\pi_{\Re}$ does exemplify properties such as being abstract, being non-round, being non-red, being thought about by the reader at this moment, etc. But the present view is that the mathematical properties of $\pi_{\Re}$ are properties that it encodes, not exemplifies.[16] Although the standard exemplification readings of unprefixed mathematical sentences are false, the encoding readings *recover* the mathematical and philosophical intuition that there is a sense in which these unprefixed sentences are true.[17]

[16] $\pi_{\Re}$ encodes only its mathematical properties and will therefore be 'incomplete' in the sense that there are properties F such that $\pi_{\Re}$ neither encodes F nor the negation of F. However, $\pi_{\Re}$, like *all* other objects, is complete in the sense that for any property F, either $\pi_{\Re}$ exemplifies F or $\pi_{\Re}$ exemplifies the negation of F.

[17] Note the similarities and differences with Field [1980] and [1989]. We agree with Field that the standard (exemplification) readings of unprefixed mathematical sentences are false. However, unlike Field, we shall offer a reading on which these unprefixed mathematical sentences are true! We also agree with Field that the theory-prefixed statements are true (he accepts that 'In number theory, 2+2=4' is true), but unlike Field, we shall offer compositional truth conditions for these claims in which the denotations of the constants and predicates play a role!

We shall postpone further discussion and defense of the myriad of interesting issues that arise in connection with this philosophy of mathematics until Sections 6 and 7. But the reader should now be in a position to appreciate the preliminary series of definitions, rules and theorems which follow.

§3: Preliminary Theoretical Principles

We now focus solely on the language and axioms of $\mathcal{O}$, leaving the question of how to represent the data from ordinary mathematics to Sections 4 and 6. Now recall that two paragraphs back, we described a reading for 'π is irrational' (a theorem of $\Re$) in terms of the *abstract* property *being irrational*$_{\Re}$. Just as abstract individuals encode properties, abstract properties encode properties of properties and abstract relations encode properties of relations, etc. To represent such claims precisely, we employ the type-theoretic version of $\mathcal{O}$. This theory is stated in a typed language governed by the following definition of 'logical type':

- i is a logical type.
- Where $t_1, \ldots, t_n$ are any types, $\langle t_1, \ldots, t_n \rangle$ is a logical type.

Our language includes (constants and) variables $x^t, y^t, \ldots$ for each type. Intuitively, i is the type for individuals and so x^i will be a variable ranging over individuals. $\langle t_1, \ldots, t_n \rangle$ is the type for relations that hold among objects having types $t_1, \ldots, t_n$, respectively. Instead of $x^{\langle t_1, \ldots, t_n \rangle}$, we frequently use the variable $F^{\langle t_1, \ldots, t_n \rangle}$ to range over relations of this type, so as to make it clearer that the object in question is a relation. For each type t, there is a distinguished predicate '$E!^{\langle t \rangle}$' ('concrete^t') that applies to things of type t. In terms of this predicate, we define a predicate that characterizes the ordinary objects ('$O!$') and abstract objects ('$A!$') of type t as follows:

$$O!^{\langle t \rangle} x^t =_{df} \Diamond E!^{\langle t \rangle} x^t$$

$$A!^{\langle t \rangle} x^t =_{df} \neg\Diamond E!^{\langle t \rangle} x^t \qquad (1)$$

Finally, since n may be 0 in $\langle t_1, \ldots, t_n \rangle$, we shall use '$p$' as a variable ranging over objects of the empty type $\langle\,\rangle$. Intuitively, this is the type for propositions.

With this typing scheme, we may assume that the formulas and complex terms of the language of our type theory can be specified in the usual way. Note that we still have two kinds of atomic formula:

$$F^{\langle t_1,\ldots,t_n \rangle} x^{t_1} \ldots x^{t_n}$$

$$x^t F^{\langle t \rangle}$$

Since it is straightforward to specify the well-formed formulas and complex terms, we will omit the definition here. The definition can be inferred by examining some of the main principles of $\mathcal{O}$, which now operate at each type (we suppress types on the reoccurrences of a term whose type has already been specified in the formula):

$$O!^{\langle t \rangle} x^t \rightarrow \Box\neg\exists F^{\langle t \rangle} xF \tag{2}$$

$$\exists x^t (A!^{\langle t \rangle} x \;\&\; \forall F^{\langle t \rangle}(xF \equiv \varphi)), \text{ where } \varphi \text{ has no free } x^t\text{s} \tag{3}$$

$$\begin{aligned} x^t = y^t \;\; &=_{df} \\ &O!^{\langle t \rangle} x^t \;\&\; O!y^t \;\&\; \Box\forall F^{\langle t \rangle}(Fx \equiv Fy) \;\vee \\ &A!^{\langle t \rangle} x^t \;\&\; A!y^t \;\&\; \Box\forall F^{\langle t \rangle}(xF \equiv yF) \end{aligned} \tag{4}$$

$$[\lambda y^{t_1} \ldots y^{t_n}\, \varphi] x^{t_1} \ldots x^{t_n} \equiv \varphi^{x^{t_1},\ldots,x^{t_n}}_{y^{t_1},\ldots,y^{t_n}}, \text{ where } \varphi \text{ has no encoding subformulas and no definite descriptions} \tag{5}$$

$$\Diamond x^t F^{\langle t \rangle} \rightarrow \Box xF \tag{6}$$

(2) asserts that ordinary objects, of whatever type, do not encode properties. (3) is the comprehension principle for abstract objects and asserts that when φ is a condition on properties $F^{\langle t \rangle}$ (i.e., the Fs characterize objects of type t), there is an abstract object of type t that encodes all and only the Fs satisfying φ. (4) defines identity conditions for all objects: objects x^t and y^t are identical whenever either they are both ordinary objects of type t and necessarily exemplify the same properties or they are both abstract objects of type t and necessarily encode the same properties. (Substitution of identicals governs this defined notion.) (5) is the λ-Conversion principle that governs λ-expressions. It asserts that objects $x^{t_1}, \ldots, x^{t_n}$ exemplify the complex relation *being a* $y^{t_1}, \ldots, y^{t_n}$ *such that* φ if and only if $x^{t_1}, \ldots, x^{t_n}$ satisfy φ.[18] (6) is a logical axiom which asserts that encoding is not relative to any circumstance—encoded properties are rigidly encoded.

Notice that (3) and (4) jointly guarantee that for any formula φ (with no free x^ts), there is a unique abstract object of type t that encodes all and only the properties satisfying φ. (There couldn't be two distinct abstract objects that encode exactly the properties satisfying φ if distinct abstract objects have to differ with respect to at least one encoded property.) So that means the following canonical description of an abstract object is always well-defined:

$$\imath x^t (A!^{\langle t \rangle} x \;\&\; \forall F^{\langle t \rangle}(xF \equiv \varphi))$$

Moreover, such canonical descriptions are governed by a straightforward consequence of the logic of descriptions, namely, the abstract object x^t that encodes exactly the properties satisfying φ encodes a property $G^{\langle t \rangle}$ iff $G^{\langle t \rangle}$ satisfies φ:

$$\imath x^t (A!^{\langle t \rangle} x \;\&\; \forall F^{\langle t \rangle}(xF \equiv \varphi)) G^{\langle t \rangle} \equiv \varphi^{G^{\langle t \rangle}}_{F^{\langle t \rangle}} \tag{7}$$

We shall appeal to this theorem on occasion in what follows.

In Section 1 of this paper, we discussed only abstract individuals. However, we can now assert the existence of abstract properties of individuals and abstract relations that individuals may exemplify:

$$\exists x^{\langle i \rangle} (A!^{\langle\langle i \rangle\rangle} x \;\&\; \forall F^{\langle\langle i \rangle\rangle}(xF \equiv \varphi)), \text{ where } \varphi \text{ has no free } x^{\langle i \rangle}\text{s}$$

$$\exists x^{\langle i,i \rangle} (A!^{\langle\langle i,i \rangle\rangle} x \;\&\; \forall F^{\langle\langle i,i \rangle\rangle}(xF \equiv \varphi)), \text{ where } \varphi \text{ has no free } x^{\langle i,i \rangle}\text{s}$$

We plan to show that mathematical properties (relations) can be found among the abstract properties which are asserted to exist by the first (second) of the above principles.

Now it will be useful, in what follows, to have a formula of $\mathcal{O}$ that explicitly asserts that something is a mathematical theory. It would suffice for our purposes to just add '*MathTheory*' as a primitive relation symbol of $\mathcal{O}$. Such a predicate would allow us to introduce axioms and specify identity conditions for mathematical theories. However, instead of this expedient of adding a single new primitive notion, it is philosophically more perspicuous to define the notion of 'mathematical theory' in terms of two

[18] It is important to remember that λ-Conversion does *not* guarantee the existence of complex relations and propositions definable in terms of encoding subformulas.

other primitive, but reasonably well-understood, notions. Before we say what these are, it is important to point out that we shall ultimately rely on our pretheoretic ability to recognize mathematical theories when we come across them. We do not intend to use the definition of '$MathTheory(x)$' to *prove* that anything is a mathematical theory. Instead, the purpose of the definition is to tell us, in theoretical terms, what it is that we know when we pretheoretically identify something as a mathematical theory and then assert this in $\mathcal{O}$.

So to define the notion of a 'mathematical theory' in $\mathcal{O}$, we need the following two primitive notions. First, we need the notion of a 'purely mathematical proposition' ('$Math(p)$'). Fortunately, we have a pretty good pretheoretic grasp on what this notion amounts to. It is reasonably clear which primitive constants and predicates are mathematical. Mathematicians and logicians certainly have no trouble identifying which constants and predicates have to be added to a predicate calculus in order to state the proper axioms of some mathematical theory. Though some philosophers have supposed that (primitive) sets are logical objects[19] or that set membership is a logical rather than a mathematical relation,[20] pretty much everyone now agrees that when taken as a primitive notion, 'membership' is a non-logical, theoretical (i.e., mathematical) notion and that the axioms of set theory are non-logical axioms. So, I'll assume that we can judge pretheoretically which predicates and constants are mathematical, and that this ability justifies taking $Math(p)$ as a primitive notion.[21] The other notion we'll need to define a mathematical theory is the notion of 'authorship'. Authorship is a relation that holds between two individuals x^i and y^i and we use '$A^{\langle i,i\rangle}x^iy^i$' to assert that x^i authors y^i. We'll explain why this notion is important in just a moment.

Now in terms of our two primitive notions, we may define a mathematical theory to be any abstract individual (i.e., object of type i) which

[19]See, for example, Frege [1893/1903].

[20]See, for example, Hempel [1945].

[21]This pretheoretic ability to judge which predicates and constants are mathematical suggests that we might define 'purely mathematical proposition' as any proposition denoted by a sentence whose non-logical vocabulary consists only of mathematical predicates and, possibly, mathematical constants. But such a definition obviously involves both (a) devices for mentioning pieces of language and (b) semantic notions. We won't officially introduce these devices and notions into $\mathcal{O}$, for then we would have to worry about semantic paradoxes. Moreover, it may be that the notion of a purely mathematical proposition can be defined without introducing these devices and notions into $\mathcal{O}$. But I will not pursue the question here.

(i) encodes only propositional properties constructed out of mathematical propositions, and (ii) is authored by some concrete individual:

$$MathTheory(x^i) =_{df} \forall F^{\langle i\rangle}(xF \rightarrow \exists p(Math(p) \,\&\, F = [\lambda y^i\, p])) \,\& \\ \exists y^i(E!^{\langle i\rangle}y \,\&\, A^{\langle i,i\rangle}yx) \qquad (8)$$

The authorship relation is used so that we can talk primarily about the mathematical theories that have actually been constructed. Of course, we shall want our analysis to apply to any possible mathematical theory, and this is easily done—we simply add the modal operator 'possibly' in front of the second conjunct of (8) (i.e., so that it reads: it *might* be the case that x^i is authored by some concrete individual). However, in what follows, we need not concern ourselves with this subtlety. Henceforth, we use the variables T and S to range over actual theories.

We next say that a proposition p is *true in* theory T ('$T \models p$') if and only if T encodes the property $[\lambda y\, p]$:

$$T \models p =_{df} T[\lambda y\, p] \qquad (9)$$

In the next section, we will use this defined notion to translate the ordinary language theory-prefix 'In mathematical theory T, ...'. But for now, let us note that given this definition, we may extend our central notion, x^t *encodes* $F^{\langle t\rangle}$, so that we may say that certain individuals (in particular, theories) encode propositions. We will say that a theory T encodes proposition p just in case p is true in T, i.e., just in case T encodes $[\lambda y\, p]$.

It is important next to stipulate that mathematical theories are closed under proof-theoretic consequence. If we utilize the notion $\varphi_1, \ldots, \varphi_n \vdash \psi$, which is defined proof-theoretically in $\mathcal{O}$ in the usual way (and which has been so defined in previous work), then we may stipulate that whenever proposition q is a proof-theoretic consequence of propositions $p_1, \ldots, p_n$ and the p_i are all true in mathematical theory T, then q is true in T:

Rule of Closure for Mathematical Theories T:
If $p_1, \ldots, p_n \vdash q$ and $T \models p_1$ and ... and $T \models p_n$, then $T \models q$. (10)

In what follows, therefore, we assume that mathematical theories are closed according to this rule. In addition, we shall often refer to the proof-theoretic consequences of a proposition as its 'logical' consequences. So the reader is hereby cautioned to remember both (a) that '$T \models p$' does *not* assert that p is a logical consequence of T but rather is defined as in (9), and (b) that '$p \vdash q$' asserts that q is a logical (proof-theoretic) consequence of p.

Now we are in a position to appreciate the significance of two simple theorems, the first of which gives us a theoretical description of the abstract individual with which a mathematical theory T is to be identified. It is a simple consequence of our definition of 'mathematical theory' that if T is a mathematical theory, then T is identical with *the* abstract individual that encodes all and only (the properties constructed out of) the propositions that are true in T. In formal terms:[22]

$$\begin{aligned}&\mathit{MathTheory}(T) \rightarrow \\ &\quad T = \imath x^i(A!x \;\&\; \forall F(xF \equiv \exists p(T \models p \;\&\; F = [\lambda y\, p])))\end{aligned} \tag{11}$$

There are two important observations to make about (11). The first is that the definite description that appears in (11), namely:

$$\imath x^i(A!x \;\&\; \forall F(xF \equiv \exists p(T \models p \;\&\; F = [\lambda y\, p]))),$$

[22]The proof depends on theorem (7) governing canonical descriptions of abstract objects: Suppose $\mathit{MathTheory}(T)$. Then, by the definition of identity for abstract objects (4) and the logical axiom (6) that encoded properties are rigidly encoded, we simply have to show that T encodes exactly the properties encoded by:

$$\imath x^i(A!x \;\&\; \forall F(xF \equiv \exists p(T \models p \;\&\; F = [\lambda y\, p]))).$$

($\rightarrow$) So assume that T encodes P (to show that the individual just described encodes P). It then follows from the definition of a mathematical theory (8) that $\exists q(P = [\lambda y q])$. So suppose $P = [\lambda y\, q_1]$. Then T encodes $[\lambda y\, q_1]$. So, by the definition of truth in a theory (9), we therefore know:

$$T \models q_1 \;\&\; P = [\lambda y\, q_1]$$

From which it follows that

$$\exists p(T \models p \;\&\; P = [\lambda y\, p])$$

So by the theorem governing canonical descriptions of abstracta (7), it follows that:

$$\imath x^i(A!x \;\&\; \forall F(xF \equiv \exists p(T \models p \;\&\; F = [\lambda y\, p]))) \text{ encodes } P,$$

which is what we had to show.

($\leftarrow$) Assume that:

$$\imath x^i(A!x \;\&\; \forall F(xF \equiv \exists p(T \models p \;\&\; F = [\lambda y\, p]))) \text{ encodes } P.$$

(to show that T encodes P). Then, by the theorem about canonical descriptions (7), it follows that:

$$\exists p(T \models p \;\&\; P = [\lambda y\, p])$$

Let q_2 be an arbitrary such proposition. So we know:

$$T \models q_2 \;\&\; P = [\lambda y\, q_2]$$

But, by (9), the first conjunct just means that T encodes $[\lambda y\, q_2]$. So, it follows that T encodes P, which is what we had to show.

is well-defined. We established earlier that any such canonical description of an abstractum is well-defined, as a consequence of our comprehension principle (3) and identity principle (4) for abstracta.

The second thing to note about (11) is that its consequent does not offer a *definition* of mathematical theory T, but rather a theoretical description of T. Once we extend the language of $\mathcal{O}$ (in the next section) by introducing names of actual mathematical theories, the above theorem can be instantiated to those names, resulting in specific identity claims which identify the named theories as well-defined abstract individuals. But it is important to remember that we haven't yet introduced any names of actual mathematical theories into our language and so we can't yet instantiate the variable T in (11) in any interesting way. Nor do we have available specific true sentences of the form '$\tau \models \varphi$'. Such sentences will become available in the next section, where we show how to translate the data into our formal system. At this point, we have simply defined some technical notions of $\mathcal{O}$ and have produced a simple theorem that is statable in terms of these notions.

The second simple theorem we are in a position to appreciate tells us identity conditions for mathematical theories. It is a consequence of our identity conditions for abstracta (4) and the definition of mathematical theories (8) that theories T and S are identical if and only if all and only the propositions true in T are true in S:

$$T = S \equiv \forall p(T \models p \equiv S \models p)$$

This, presumably, is exactly what one would expect as identity conditions for theories.

We can complete the specification of the machinery needed for metaphysical reductions with the following series of definitions. First we say that an object x^t (of type t) is an *object of* theory S iff there is a property $F^{\langle t \rangle}$ such that it is true in S that x^t exemplifies F:

$$\mathit{ObjectOf}(x^t, S) =_{df} \exists F^{\langle t \rangle}(S \models F x^t) \tag{12}$$

In what follows, when $t = i$ we say that x^t is a mathematical individual of S; when $t = \langle i \rangle$ we say that x^t is a mathematical property of S; and when $t = \langle i, i \rangle$ we say that x^t is a mathematical relation of S.

We can now formulate a quite general axiom for identifying the objects of theory S as abstract objects. This statement of the axiom is simplified by the fact that we are considering only *pure* mathematical theories.[23] We assert the following as axiomatic:

Reduction Axiom:
$$ObjectOf(x^t, S) \rightarrow x = \imath y^t(A!^{\langle t \rangle} y \,\&\, \forall F^{\langle t \rangle}(yF \equiv S \models Fx)) \tag{13}$$

In other words, if x^t is a type-t object of theory S, then x is the abstract object that encodes exactly the properties true of x in S.

By our theorem (7) governing canonical descriptions, it is an immediate Corollary of the Reduction Axiom that if x^t is an object of S, then x encodes a property $F^{\langle t \rangle}$ iff it is true in S that x exemplifies F:

$$ObjectOf(x^t, S) \rightarrow (xF^{\langle t \rangle} \equiv S \models Fx) \tag{14}$$

We shall see some specific instances of our Reduction Axiom and Corollary in the next section. In the meantime, we note that the principles allow us to identify the objects of a mathematical theory no matter what the logical type of the object. It is to be stressed here that from the point of view of a foundational metaphysics, there is no distinguished 'model-theoretic' perspective to tell us what the 'objects of' a theory T are. From a metaphysical point of view, the objects of a theory are the ones described by its *de re* claims, for these attribute properties to objects. Note that the statement '$\exists x^i P^{\langle i \rangle} x$' counts as a *de re* claim about the property P, but that it doesn't count as a *de re* claim about mathematical individuals. From $T \models \exists x^i P^{\langle i \rangle} x$, we can validly infer $\exists F^{\langle i \rangle}(T \models \exists x^i Fx)$, but we can't validly infer $\exists x^i(T \models Px)$. This tells us that from a logical standpoint, we cannot validly export the quantifiers inside the scope of the theory operator.[24] The implications of this fact will become clearer in Section 6.

§4: The Reduction of Mathematics to Metaphysics

In this section we'll reduce the objects of an arbitrary mathematical theory to the objects of our formal metaphysics. We'll begin by introducing

[23]If we were to consider applied mathematical theories, we would have to distinguish between the abstract, mathematical objects of a theory and the ordinary, non-mathematical objects, since ordinary individuals and ordinary properties may be objects of the applied theory, in our defined sense. When dealing with applied mathematical theories, we would apply the following axiom only to identify the *abstract* objects of the theory. Fortunately, we need not worry about this subtlety here.

[24]Of course, if it is true in theory T that there exists a *unique* object of a certain sort, the theory can be extended to include a well-defined term which denotes the object in question. Such a term would be subject to existential generalization by a quantifer outside the scope of the theory operator.

the expressions of ordinary mathematical language into the language of our formal system. Though it should be clear what the resulting relationship is between the statements of ordinary mathematical language and those of our system, we will leave the explicit discussion of the interesting features of this relationship until the next section.

Now to actually carry out our analysis and reduction, let τ range over names of mathematical theories and suppose that we pretheoretically judge that a group of sentences constitute the proper axioms of a mathematical theory named τ. Let us simply refer to these sentences as 'the axioms of τ'. We shall assume that the axioms of τ have been. or can be, formalized in a first- or second-order predicate calculus (with identity). So whenever sentence s is an axiom of τ, we will also say that its formal rendition φ is an axiom of τ.

So as to reduce the amount of work we shall have to do in what follows, we make the following three simplifying metatheoretical assumptions: (I) the axioms of τ that are instances of a first-order axiom schema can all be replaced by a single second-order axiom which employs quantifiers over relations, (II) whenever the axioms of τ involve a primitive n-place function symbol, they can be replaced by axioms involving an $n+1$-place predicate symbol, and (III) whenever the axioms of τ involve '$=$' as a logical primitive, they can be replaced by axioms involving '$=$' as a distinguished *non-logical* relation symbol (so that the standard two logical axioms for identity become *proper* axioms). We'll discuss simplification (III) in the next section. For the meantime, it should help if we remind the reader that when a sentence of τ asserts an identity, the individuals asserted to be identical exemplify the same properties (or conditions) *expressible in* the language of τ. (For each mathematical theory τ, we shall employ a relation symbol '$=_\tau$' by which one can assert that τ-identical individuals exemplify the same τ-expressible properties.) Moreover, in Section 5, we'll demonstrate that nothing important is lost by appealing to assumption (III).

§4.1 Extending $\mathcal{O}$

In this subsection, we extend $\mathcal{O}$ by adding new expressions to the language and by adding certain analytic truths and certain obvious facts as new axioms. Now, whenever we pretheoretically judge that a group of sentences constitute the axioms for a mathematical theory named τ, we *extend* the language of $\mathcal{O}$ as follows: (a) we add the name τ as a new constant of type

i to our language, (b) for each primitive constant κ (if there are any) that appears in an axiom of τ, we add κ_τ as a new constant of type i to our language, and (c) for each n-place primitive relation symbol Π appearing in an axiom of τ, we add Π_τ as a new relation symbol of type $\langle i, \ldots, i\rangle$ (with n occurrences of 'i') to our language. Note that clause (c) and simplifying assumption (III) together ensure that if an axiom of τ involves '$=$', we shall be adding '$=_\tau$' as a new 2-place relation symbol of $\mathcal{O}$.

Next, for each proper axiom φ of theory τ, let φ^* designate the formula of $\mathcal{O}$ that results when each primitive constant κ in φ is replaced by κ_τ and each primitive predicate Π in φ is replaced by Π_τ. Then, for each proper axiom φ of τ, we take the following as a *new axiom* of $\mathcal{O}$:

$$\tau \models \varphi^* \qquad \text{(whenever } \varphi \text{ is an axiom of } \tau) \tag{15}$$

In the remainder of this essay, we adopt the following convention. In any displayed line of the form $\tau \models \psi$, the scope of '$\models$' will extend over ψ, no matter how complex ψ may be. So this is a special case where the convention (described in footnote 10) of giving '$\models$' the narrowest possible scope within a formula is to be overridden.

Notice that these new axioms of $\mathcal{O}$ are, in a real sense, *analytic* truths. They explicitly represent the ordinary language claims of the form 'In mathematical theory τ, ...' in terms of our formal machinery. Note also that if the axioms of τ involve '$=$', then given simplifying assumption (III), (15) requires us to add the following as new axioms of $\mathcal{O}$ (in which x, y are variables of type i and F is a variable of type $\langle i\rangle$):

$$\tau \models x =_\tau x \tag{16}$$

$$\tau \models x =_\tau y \rightarrow \forall F(Fx \equiv Fy) \tag{17}$$

Notice that the quantifier '$\forall F$' in (17) lies within the scope of the theory operator '$\tau \models \ldots$'. So the indiscernibility of x and y (with respect to exemplification) is conditioned on the τ-identity of x and y only relative to τ itself. From the (17) and the fact that $\tau \models x =_\tau y$, our Rule of Closure lets us conclude only that $\tau \models \forall F(Fx \equiv Fy)$ and so we have to show that $\tau \models Px$ if we want to conclude that $\tau \models Py$. The quantifier '$\forall F$' therefore governs the τ-relative properties of x and y.

Finally, for each (pretheoretically determined) mathematical theory τ, we also add the following obvious fact as a new assumption of $\mathcal{O}$:

$$\mathit{MathTheory}(\tau) \tag{18}$$

Notice that given (15) and (18), we know metatheoretically that our Rule of Closure ensures that the translation φ^* of every ordinary theorem φ of τ becomes derivable in $\mathcal{O}$ as an explicit τ-relative truth:

$$\tau \models \varphi^* \qquad \text{(whenever } \vdash_\tau \varphi) \tag{19}$$

Reasoning behind the theory operator, therefore, is classical whenever the theory in question employs classical logic.[25]

§4.2 Reducing the Objects of Set Theories

We now examine how the foregoing facilitates a metaphysical reduction of any set described by any set theory. As a particular example, we shall consider the sets of a simple 'adjunctive' set theory. This simple theory is representative and it should be clear how the techniques used can be applied to reduce the objects of Zermelo-Fraenkel set theory and other set theories. Let us designate the following as the axioms of the theory named 'ST':

- Sets which have the same members are identical. (ST1)
- The empty set is a set. (ST2)
- No set is a member of the empty set.[26] (ST3)
- For any two sets, there is a set having the first set and the members of the second set as members. (ST4)
- For any property F and set x, there is a set which has as members all and only those members of x which exemplify F. (ST5)

These are familiar axioms—(ST1) is the Axiom of Extensionality, (ST2) and (ST3) describe the empty set, (ST4) is the Axiom of Adjunction, and (ST5) is the Axiom of Separation. It is straightforward to formalize these axioms in classical exemplification logic by assuming that the primitive,

[25]For mathematical theories involving non-classical logic, we have to adjust our Rule of Closure, so that we add to $\mathcal{O}$ only those claims derivable using the non-classical logic in question.

[26]Of course, if we were to allow urelements in the formulation of ST, we would revise this axiom so that it asserts that *nothing* whatsoever is an element of the empty set. But though this would simplify the statement of (ST3), we would no longer have a reasonably simple truth that involves *all* of the primitive non-logical notions of ST. Having such a sentence proves to be useful in what follows.

non-logical expressions are the singular term 'the empty set' ('$\emptyset$') and the non-logical predicates 'is a set' ('S'), 'is a member of' ('$\in$') and 'is the same as' ('$=$'). The formalization would go as follows:

$$\forall x \forall y[Sx \,\&\, Sy \rightarrow [\forall z(z \in x \equiv z \in y) \rightarrow x = y]]$$

$$S\emptyset \qquad \text{(ST2}'\text{)}$$

$$\neg\exists x(Sx \,\&\, x \in \emptyset) \qquad \text{(ST3}'\text{)}$$

$$\forall x \forall y[Sx \,\&\, Sy \rightarrow \exists z \forall w(w \in z \equiv w = x \vee w \in y)]$$

$$\forall F \forall x[Sx \rightarrow \exists y(Sy \,\&\, \forall z(z \in y \equiv z \in x \,\&\, Fz))] \qquad \text{(ST5}'\text{)}$$

Note that we have invoked simplifying assumption (I) so as to formulate the Separation axiom (ST5′) in its second-order guise. Since we have such quantifiers in the language of $\mathcal{O}$, we need not bother with the instances of the first-order Separation Schema.

We now translate these axioms into analytic truths of $\mathcal{O}$ as follows. According to the procedure outlined above, we first extend the language of $\mathcal{O}$ with the new non-logical constants 'ST' and '$\emptyset_{\text{ST}}$' and with the new non-logical relation symbols 'S_{ST}', '$\in_{\text{ST}}$', and '$=_{\text{ST}}$'. (It should be clear that 'ST' and '$\emptyset_{\text{ST}}$' are expressions of type i, that 'S_{ST}' is an expression of type $\langle i \rangle$, and that '$\in_{\text{ST}}$' and '$=_{\text{ST}}$' are expressions of type $\langle i, i \rangle$.) Then, using x, y, z, w as variables of type i and F as a variable of type $\langle i \rangle$, we add the following analytic truths as new axioms of $\mathcal{O}$ (in addition to the new axioms for identity discussed above):

$$\text{ST} \models \forall x \forall y[S_{\text{ST}}x \,\&\, S_{\text{ST}}y \rightarrow [\forall z(z \in_{\text{ST}} x \equiv z \in_{\text{ST}} y) \rightarrow x =_{\text{ST}} y]] \qquad (20)$$

$$\text{ST} \models S_{\text{ST}}\emptyset_{\text{ST}} \qquad (21)$$

$$\text{ST} \models \neg\exists x(S_{\text{ST}}x \,\&\, x \in_{\text{ST}} \emptyset_{\text{ST}}) \qquad (22)$$

$$\begin{aligned}\text{ST} \models \forall x \forall y[&S_{\text{ST}}x \,\&\, S_{\text{ST}}y \rightarrow \\ &\exists z \forall w(w \in_{\text{ST}} z \equiv w =_{\text{ST}} x \vee w \in_{\text{ST}} y)]\end{aligned} \qquad (23)$$

$$\text{ST} \models \forall F \forall x[S_{\text{ST}}x \rightarrow \exists y(S_{\text{ST}}y \,\&\, \forall z(z \in_{\text{ST}} y \equiv z \in_{\text{ST}} x \,\&\, Fz))]$$

Finally, we also add the assumption

$$\mathit{MathTheory}(\text{ST}) \qquad (24)$$

Now given this last fact, we know from (19) that the translation φ^* of every ordinary theorem φ of ST becomes derivable in $\mathcal{O}$ as an explicit ST-relative truth:

$$\text{ST} \models \varphi^* \qquad (\text{whenever} \vdash_{\text{ST}} \varphi)$$

With this group of theorems in $\mathcal{O}$, we are now ready to metaphysically reduce the sets described by ST. However, it is to be emphasized that there is nothing implied by the order of presentation in what follows. We are not constructing the objects 'in stages'. We are simply showing which sequences of formulas in $\mathcal{O}$ constitute proofs.

First we identify the theory ST as a particular abstract individual. To do this, we instantiate (11) to ST and then derive the consequent of the result from our assumption (24) that *MathTheory*(ST). The result is:

$$\text{ST} = \imath x^i(A!x \,\&\, \forall F(xF \equiv \exists p(\text{ST} \models p \,\&\, F = [\lambda y\, p]))) \qquad (25)$$

It is essential to recognize that this is not a definition of 'ST', but rather an exact theoretical description of a particular abstract individual. Given that sentences of the form '$\text{ST} \models p$' are well-defined and that the ordinary theorems of ST appear in this form as theorems of $\mathcal{O}$, we know *in principle* which abstract individual ST is.

Second, we identify $\emptyset_{\text{ST}}$. To do this, recall that (21) is a new axiom of $\mathcal{O}$:

(21) $\text{ST} \models S_{\text{ST}}\emptyset_{\text{ST}}$

It therefore follows that

$$\exists F^{\langle i \rangle}(\text{ST} \models F\emptyset_{\text{ST}}) \qquad (26)$$

So, by (12), it follows that $\emptyset_{\text{ST}}$ is an *object of* ST:

$$\mathit{ObjectOf}(\emptyset_{\text{ST}}, \text{ST}) \qquad (27)$$

We may therefore instantiate our Reduction Axiom (13) and detach the consequent to yield the following theorem of $\mathcal{O}$:

$$\emptyset_{\text{ST}} = \imath x^i(A!x \,\&\, \forall F(xF \equiv \text{ST} \models F\emptyset_{\text{ST}})) \qquad (28)$$

We have therefore identified $\emptyset_{\text{ST}}$ as an abstract individual. Notice that by the Corollary (14) to the Reduction Axiom, it also follows from the fact that $\emptyset_{\text{ST}}$ is an object of ST that $\emptyset_{\text{ST}}$ encodes a property F if and only if it is a truth of ST that $\emptyset_{\text{ST}}$ exemplifies F:

$$\emptyset_{\text{ST}}F \equiv \text{ST} \models F\emptyset_{\text{ST}} \tag{29}$$

This fact will prove instrumental in Section 6, when we look at the relationship between our formal theorems and their counterparts in ordinary mathematical language.

To complete our metaphysical reduction, we identify the mathematical relations of ST. We need to identify $S_{\text{ST}}, \in_{\text{ST}}$, and $=_{\text{ST}}$. It will suffice to show how to identify the first two (since the reduction of $=_\tau$ is carried out exactly like that for $\in_{\text{ST}}$). Now recall that (22) is a new axiom of $\mathcal{O}$:

(22) $\text{ST} \models \neg\exists x(S_{\text{ST}}x \;\&\; x \in_{\text{ST}} \emptyset_{\text{ST}})$

From this, our Rule of Closure, and λ-Conversion, we can 'abstract out' a property of properties that (the property) *being a set*$_{\text{ST}}$ exemplifies (in ST) and a property of relations that the *membership*$_{\text{ST}}$ relation exemplifies (in ST):

$$\text{ST} \models [\lambda F^{\langle i \rangle}\ \neg\exists x(Fx \;\&\; x \in_{\text{ST}} \emptyset_{\text{ST}})]S_{\text{ST}} \tag{30}$$

$$\text{ST} \models [\lambda F^{\langle i,i \rangle}\ \neg\exists x(S_{\text{ST}}x \;\&\; Fx\emptyset_{\text{ST}})] \in_{\text{ST}} \tag{31}$$

We can now generalize on each of the above λ-expressions, remembering that expressions of the form $[\lambda F^{\langle i \rangle}\ \psi]$ are expressions of type $\langle\langle i \rangle\rangle$ and that expressions of the form $[\lambda F^{\langle i,i \rangle}\ \psi]$ are expressions of type $\langle\langle i,i \rangle\rangle$:

$$\exists F^{\langle\langle i \rangle\rangle}(\text{ST} \models FS_{\text{ST}}) \tag{32}$$

$$\exists F^{\langle\langle i,i \rangle\rangle}(\text{ST} \models F \in_{\text{ST}}) \tag{33}$$

So, by the definition of *ObjectOf*, we may derive the following two facts:

$$\mathit{ObjectOf}(S_{\text{ST}}, \text{ST}) \tag{34}$$

$$\mathit{ObjectOf}(\in_{\text{ST}}, \text{ST}) \tag{35}$$

We may therefore instantiate our Reduction Axiom (13) and detach the consequent to yield the following theorems of $\mathcal{O}$:

$$S_{\text{ST}} = \imath x^{\langle i \rangle}(A!^{\langle\langle i \rangle\rangle}x \;\&\; \forall F^{\langle\langle i \rangle\rangle}(xF \equiv \text{ST} \models FS_{\text{ST}})) \tag{36}$$

$$\in_{\text{ST}} = \imath x^{\langle i,i \rangle}(A!^{\langle\langle i,i \rangle\rangle}x \;\&\; \forall F^{\langle\langle i,i \rangle\rangle}(xF \equiv \text{ST} \models F \in_{\text{ST}})) \tag{37}$$

Theorems (36) and (37) identify S_{ST} and $\in_{\text{ST}}$ as an abstract property and abstract relation, respectively. (Note that if we had begun with axiom (20) or (23) instead of (22), we could have reconstructed the above deduction to prove that $=_{\text{ST}}$ is an abstract relation.) Moreover, by our Corollary (14) to the Reduction Axiom, the following are also consequences of the fact that S_{ST} and $\in_{\text{ST}}$ are objects of ST:

$$S_{\text{ST}}F^{\langle\langle i \rangle\rangle} \equiv \text{ST} \models FS_{\text{ST}} \tag{38}$$

$$\in_{\text{ST}} F^{\langle\langle i,i \rangle\rangle} \equiv \text{ST} \models F \in_{\text{ST}}$$

In other words, S_{ST} and $\in_{\text{ST}}$ encode precisely the properties they exemplify in ST. These facts will prove instrumental in Section 6, when we look at the relationship between ordinary mathematical language and the formal theorems of $\mathcal{O}$.

By analogy, given the foregoing derivations, as soon as we analyze the mathematical theories Zermelo-Fraenkel set theory (ZF) or ZF + Axiom of Choice (ZFC) and supplement $\mathcal{O}$ with new terms and analytic truths in the manner prescribed above, the following become theorems which identify the primitive individuals and relations of these two theories:

$$\emptyset_{\text{ZF}} = \imath x^i(A!x \;\&\; \forall F(xF \equiv \text{ZF} \models F\emptyset_{\text{ZF}}))$$

$$\emptyset_{\text{ZFC}} = \imath x^i(A!x \;\&\; \forall F(xF \equiv \text{ZFC} \models F\emptyset_{\text{ZFC}}))$$

$$\in_{\text{ZF}} = \imath x^{\langle i,i \rangle}(A!x \;\&\; \forall F^{\langle\langle i,i \rangle\rangle}(xF \equiv \text{ZF} \models F\in_{\text{ZF}}))$$

$$\in_{\text{ZFC}} = \imath x^{\langle i,i \rangle}(A!x \;\&\; \forall F^{\langle\langle i,i \rangle\rangle}(xF \equiv \text{ZFC} \models F\in_{\text{ZFC}}))$$

On this analysis, the empty sets and membership relations of ZF and ZFC are analyzed as distinct abstract individuals and distinct abstract relations, respectively, which are defined by their theoretical roles, i.e., they are defined by the truths of ZF and ZFC.[27]

§4.3 Reducing the Objects of Number Theories

(In this subsection, we describe how the above procedure would be applied to the primitive objects of the Dedekind/Peano axioms for number theory. Since the procedure is almost exactly analogous to the one described in the previous subsection, some readers might wish to skip this subsection. However, such readers should note that the final two paragraphs discuss

[27] When Field and Balaguer use such expressions as 'sets$_{\aleph_1}$', 'sets$_{\aleph_{817}}$', '$\in_{\aleph_1}$', and '$\in_{\aleph_{817}}$', the above theorems can tell us exactly which objects in the plenitude of abstracta that these expressions refer to. See Field [1994] (pp. 420-422), Field [1998a] (p. 293), Balaguer [1995] (pp. 316-317), and Balaguer [1998] (p. 59).

why it is important to distinguish between the metaphysical reduction of the primitive objects of Dedekind/Peano number theory as outlined in this subsection and the classical reduction of the natural numbers we described in Section 1.)

We now metaphysically reduce the numbers described by various number theories. As a representative example, we'll focus on a classic number theory. Let us designate the following as the axioms of 'Number Theory' ('NT'):

- Zero is a number.
- Zero doesn't succeed any number.
- No two numbers have the same successor.
- Every number has a successor.
- If (a) 0 exemplifies the property F and (b) every two successive numbers x and y are such that if x exemplifes F then y exemplifies F, then every number exemplifies F.

Now suppose these axioms have been formalized in terms of the non-logical expressions Zero ('0'), 'is a number' ('N'), 'succeeds' ('S'), and 'is the same as' ('='). Then we extend the language of $\mathcal{O}$ with the expressions 'NT', '$0_{\rm NT}$', '$N_{\rm NT}$', '$S_{\rm NT}$', and '$=_{\rm NT}$'. (It should be clear that 'NT' and '$0_{\rm NT}$' are expressions of type i, that '$N_{\rm NT}$' is an expression of type $\langle i \rangle$, and that '$S_{\rm NT}$' and '$=_{\rm NT}$' are expressions of type $\langle i, i \rangle$.) We now add the following analytic truths as new axioms of $\mathcal{O}$ (in addition to the new axioms for identity discussed above):

$$\text{NT} \models N_{\rm NT} 0_{\rm NT} \tag{39}$$

$$\text{NT} \models \neg \exists x (N_{\rm NT} x \;\&\; S_{\rm NT} 0_{\rm NT} x) \tag{40}$$

$$\text{NT} \models \forall x \forall y [N_{\rm NT} x \;\&\; N_{\rm NT} y \;\&\; x \neq_{\rm NT} y \rightarrow \neg \exists z (N_{\rm NT} z \;\&\; S_{\rm NT} zx \;\&\; S_{\rm NT} zy)]$$

$$\text{NT} \models \forall x (N_{\rm NT} x \rightarrow \exists y (N_{\rm NT} y \;\&\; S_{\rm NT} yx))$$

$$\text{NT} \models \forall F [F 0_{\rm NT} \;\&\; \forall x \forall y (N_{\rm NT} x \;\&\; N_{\rm NT} y \;\&\; S_{\rm NT} yx \;\&\; Fx \rightarrow Fy) \rightarrow \forall x (N_{\rm NT} x \rightarrow Fx)] \tag{41}$$

Note that we have invoked simplifying assumption (I) so as to formulate the Induction Axiom (41) in its second-order guise.

Finally, we also add the assumption:

$$\mathit{MathTheory}(\text{NT}) \tag{42}$$

Now given this last fact, we know from (19) that the translation φ^* of every ordinary theorem φ of NT becomes derivable in $\mathcal{O}$ as an explicit NT-relative truth:

$$\text{NT} \models \varphi^* \qquad (\text{whenever} \vdash_{\rm NT} \varphi)$$

With this group of theorems in $\mathcal{O}$, we begin our reduction by identifying the theory NT as a particular abstract individual. To do this, we instantiate (11) to NT and appeal to (42) to conclude:

$$\text{NT} = \imath x^i (A!x \;\&\; \forall F (xF \equiv \exists p (\text{NT} \models p \;\&\; F = [\lambda y\, p]))) \tag{43}$$

Secondly, we identify $0_{\rm NT}$. Beginning with (39), we follow the same steps that we followed in moving from (21) through (26) and (27) to reach (28). That is, beginning with (39), we abstract out a property that $0_{\rm NT}$ exemplifies according to NT, generalize on that property, conclude that $0_{\rm NT}$ is an object of NT, and then instantiate our Reduction Axiom to conclude:

$$0_{\rm NT} = \imath x^i (A!x \;\&\; \forall F (xF \equiv \text{NT} \models F 0_{\rm NT})) \tag{44}$$

We have therefore identified $0_{\rm NT}$ as an abstract individual.

To complete our metaphysical reduction, we identify the mathematical relations of NT. We need to identify $N_{\rm NT}$, $S_{\rm NT}$, and $=_{\rm NT}$. Again, it will suffice to show how to identify the first two. Beginning with (40), we just follow the same steps that we followed in moving from (22) via (30) – (35) to reach both (36) and (37). That is, beginning with (40), we abstract out various properties, generalize, apply the definition of *ObjectOf* and instantiate the Reduction Axiom. We thereby prove the following two theorems of $\mathcal{O}$:

$$N_{\rm NT} = \imath x^{\langle i \rangle} (A!^{\langle\langle i \rangle\rangle} x \;\&\; \forall F^{\langle\langle i \rangle\rangle} (xF \equiv \text{NT} \models F N_{\rm NT})) \tag{45}$$

$$S_{\rm NT} = \imath x^{\langle i,i \rangle} (A!^{\langle\langle i,i \rangle\rangle} x \;\&\; \forall F^{\langle\langle i,i \rangle\rangle} (xF \equiv \text{NT} \models F S_{\rm NT})) \tag{46}$$

We have therefore identified the mathematical relations $N_{\rm NT}$ and $S_{\rm NT}$ as abstract relations.

Before we turn our attention to the objects of arbitrary mathematical theories in the next subsection, an important observation is in order. It is important to recognize that the objects identified by the above metaphysical reduction of NT are completely different from the objects identified in the classical reduction of the natural numbers in $\mathcal{O}$ described in Section 1. In Section 1, we described how $\mathcal{O}$ has the resources to define the natural cardinal 0 and to define the concepts of 'natural number' and 'predecessor'. This classical reduction was the subject of Zalta [1999], and in that paper the natural cardinals were defined so as to encode ordinary non-mathematical properties. For example, the natural cardinal 0 encodes all and only those properties which are exemplified by no ordinary objects (e.g., 0 encodes the property of being a giraffe in the Arctic Circle), and the natural cardinal 9 encodes just those properties that are exemplified by nine ordinary objects (e.g., 9 encodes the property of being a planet in our solar system). Moreover, the predecessor relation and its weak ancestral are asserted to be ordinary relations (since they aren't abstract relations, they don't encode any properties). They were defined in terms of encoding, using a definition similar to Frege's.[28] And the concept of 'natural number' was then defined to be any abstract object to which the natural cardinal 0 bears the weak ancestral of the predecessor relation. Given such definitions, the *unprefixed* Dedekind/Peano axioms become *unprefixed* theorems of $\mathcal{O}$. So $\mathcal{O}$ rules that the basic laws of number theory are true *simpliciter*.

Consequently, it is important to distinguish these *natural* numbers, which are defined in terms of the application of counting the ordinary objects of the natural world, from the *theoretical* numbers of NT (and from the theoretical numbers of every other mathematical theory of numbers). These number systems are different because the individual numbers play different roles in their respective theories. The theoretical numbers of NT encode *only* the properties assigned to them by their theoretical role in NT. As such, they do not encode ordinary properties such as being a giraffe in the Arctic Circle or being a planet. Similarly, the relations $\text{predecessor}_{\text{NT}}$ and $\text{number}_{\text{NT}}$ are both primitive (not defined) in NT. They can be identified as abstract relations that encode only the properties of relations that NT assigns to them. The standard laws of number theory,

[28]Since these relations were defined in terms of encoding subformulas, we had to explicitly assert that they are (ordinary) relations, and prove that when those assertions are added to $\mathcal{O}$, it remains consistent. See Zalta [1999].

as formulated in exemplification logic in terms of the *primitive notions* of zero_{NT}, $\text{number}_{\text{NT}}$, and $\text{predecessor}_{\text{NT}}$, remain true, when translated into $\mathcal{O}$, *only* when prefixed by the appropriate theory operator. So although $\mathcal{O}$ has enough mathematical power (via a definition of 'natural number') to imply the basic laws of number theory as unprefixed (i.e., objective) truths, we rely on its philosophical power rather than its mathematical power to give a *metaphysical* reduction of the objects of arbitrary mathematical theories, as we shall now see.

§4.4 Reducing the Objects of Arbitrary Mathematical Theories

To reduce the objects of an arbitrary mathematical theory, we first identify the theories themselves by proving a theorem similar to (25) and (43). Suppose that we pretheoretically judge that a group of sentences constitute the axioms of a mathematical theory named τ. Suppose further that the axioms of τ have been given some standard first- or second-order formalization in accordance with our simplifying assumptions (I) – (III). Now suppose we have extended $\mathcal{O}$ in the way described above. We can then theoretically identify the mathematical theory τ as follows:

$$\tau = \iota x^i(A!x \,\&\, \forall F(xF \equiv \exists p(\tau \models p \,\&\, F = [\lambda y\, p])))$$

Recall that this is provable from (11) and (18).

We now simply have to identify the objects of τ. Consider any primitive non-logical expression κ^t that appears in (a sentence pretheoretically judged to be) an axiom of τ. Then, given our simplifying assumptions, κ^t is either an individual constant of type i or an n-place relation symbol of type $\langle i, \ldots, i\rangle$ (with n occurrences of i). We therefore add κ^t_τ to the language of $\mathcal{O}$. Now suppose φ is (a sentence pretheoretically judged to be) an axiom of τ and that φ contains κ^t. Then where φ^* is the translation of φ into the language of $\mathcal{O}$, we know that the following is a new axiom of $\mathcal{O}$:

$$\tau \models \varphi^*$$

From this, we can 'abstract out' a property that κ^t exemplifies in theory τ. Let φ^- be the result of substituting the new variable y^t for κ^t_τ in φ^*. Then we may use λ-Conversion and our Rule of Closure (10) to prove that:

$$\tau \models [\lambda y^t\, \varphi^-]\kappa_\tau,$$

And by generalizing on the λ-expression, it follows that:

$$\exists F^{\langle t \rangle}(\tau \models F\kappa^t_\tau)$$

So, by (12), κ^t_τ is a type-t object of τ:

$$ObjectOf(\kappa^t_\tau, \tau)$$

Then, by our Reduction Axiom (13), we can provably identify κ^t_τ in $\mathcal{O}$ as follows:

$$\kappa^t_\tau = \imath x^t(A!x \;\&\; \forall F(xF \equiv \tau \models F\kappa^t_\tau)) \tag{47}$$

The significance of this theorem cannot be overemphasized. It offers a general ontological reduction of mathematical objects (individuals and relations) to the abstract objects of our background ontology. Given (14), it is an immediate consequence of (47) that:

$$\kappa^t_\tau F^{\langle t \rangle} \equiv \tau \models F\kappa_\tau \tag{48}$$

In other words, a mathematical object κ^t_τ encodes exactly the properties it exemplifies in theory τ.

§5: Some Consequences of the Reduction

There are some issues that arise in connection with the foregoing that deserve commentary. In this section, we shall discuss our treatment of identity (§5.1) and describe some interesting consequences of the theorems just proved (§5.2). The consequences discussed in Section 5.2 play an important role in Section 6, where we analyze (ordinary) mathematical language in a way that reveals a correlation between the theorems of arbitrary mathematical theories and theorems of (extended) $\mathcal{O}$.

§5.1 When Identity is Primitive in τ

Recall that simplifying assumption (III) was that whenever a mathematical theory τ is formulated in a language with identity, τ can be reformulated in a language without identity in which (a) the symbol '=' becomes a distinguished binary relation symbol that is a non-logical primitive of the theory, and (b) the standard (two) logical axioms for identity become (reformulated as) *proper* axioms which govern the primitive binary relation symbol '='. Now one might argue, from considerations of model theory, that theories reformulated in this way would not have the same expressive

capacity as the original. The argument would be that the proper axioms governing '=' in the reformulated theory would only guarantee that '=' denotes an equivalence relation and not *identity*.

The response to this argument is straightforward, however. If the italicized use of the word 'identity' at the end of the previous paragraph is supposed to denote some relation that is primitive in model theory, then we simply point out that from the point of view of the present metaphysics, there is no such primitive relation. $\mathcal{O}$ uses both a *defined* notion of identity and a *proper theory* of identity.[29] To defend $\mathcal{O}$, we get to assume that it is true and that, consequently, the facts about identity are as the theory says. So, unless that theory is shown to be defective in some way, its theory of identity trumps the primitive model-theoretic notion. The argument from model-theory simply becomes unpersuasive to the object theorist.

However, if the italicized use of the word 'identity' is supposed to denote the notion of identity defined as 'exemplifying the same properties', then the question becomes whether that definition is correct (i.e., consistent with the unrestricted substitution of identicals). From the point of view of $\mathcal{O}$, this standard definition of identity is *not* correct; the identity of indiscernibles correctly applies *only* to non-abstract (i.e., ordinary) objects. It is a theorem of $\mathcal{O}$ that there are abstract objects x and y that are distinct (in the sense that they *encode* different properties) but which exemplify the same properties![30] There are so many abstract objects that the traditional mode of predication, namely exemplification, cannot always discern abstract objects that encode different properties.

We may also put the model-theoretic concern to rest by showing how our notion of identity for abstract objects, defined in (4) as 'encoding the same properties', does precisely the work it should do. We show that, using simplification (III), whenever x and y are objects of τ and $\tau \models x =_\tau y$, then our metaphysics guarantees that x and y are identical in the sense defined by (4).

To see this, recall that when we have a theory τ expressed in the language of identity that has been reformulated according to simplifying assumption (III), the proper axioms of τ that govern the new relation

[29] It is derivable that $x^t = x^t$, from the definition in (4) of '='. Substitution of identicals is asserted as a proper axiom.

[30] This theorem is proved and explained in Zalta [1999] (Section 2), but we will not take the time to repeat the proof and explanation here.

symbol $=_\tau$ become added to $\mathcal{O}$ as the axioms (16) and (17):

(16) $\tau \models x =_\tau x$

(17) $\tau \models x =_\tau y \rightarrow \forall F(Fx \equiv Fy)$

Note that (17) tells us that it is a truth in τ that, if x and y are τ-identical, then x and y exemplify the same properties. (Note also that the quantification over properties might be sufficient to put the model-theoretic concern to rest.)

Now we want to show that if x and y are objects of τ and $\tau \models x =_\tau y$, then $x = y$. So assume the claims required by the antecedent:

$$\begin{array}{l} \mathit{ObjectOf}(x, \tau) \\ \mathit{ObjectOf}(y, \tau) \\ \tau \models x =_\tau y \end{array}$$

Then by (17) and our Rule of Closure, it follows that:

$$\tau \models \forall F(Fx \equiv Fy) \qquad (\vartheta)$$

Now to show $x = y$, we have to show that x and y encode the same properties. Without loss of generality, we simply prove that if x encodes P, then y encodes P, since the converse uses the same reasoning. So suppose that x encodes P. Since x is an object of τ, we may appeal to the Corollary (14) of our Reduction Axiom to conclude:

$$xF \equiv \tau \models Fx$$

So since x encodes P, it follows that $\tau \models Px$. We can now appeal to (ϑ) and our Rule of Closure to infer that $\tau \models Py$. But y is also an object of τ, and so the Corollary to the Reduction Axiom implies:

$$yF \equiv \tau \models Fy$$

So y encodes P, which is what we had to show.

So, from the point of view of $\mathcal{O}$, whenever x and y are τ-objects that are τ-identifical, our metaphysics concludes that x and y are the same abstract object. Thus, anything true of the one is true of the other.

§5.2 Some Further Theorems

In this subsection, we describe some interesting consequences of the theorems proved in the last section. Consider the ordinary axiom of ST that $\emptyset$ is a set. This simple claim was introduced into $\mathcal{O}$ as the analytic, theory-prefixed axiom (21):

(21) $\mathrm{ST} \models S_{\mathrm{ST}}\emptyset_{\mathrm{ST}}$

The Rule of Closure and λ-Conversion immediately yield that in ST, S_{ST} exemplifies the property of being a property that $\emptyset_{\mathrm{NT}}$ exemplifies:

$$\mathrm{ST} \models [\lambda G^{\langle i \rangle}\, G\emptyset_{\mathrm{ST}}]S_{\mathrm{ST}} \qquad (49)$$

We now proceed to show that from (21) and (49), we can derive two further facts, namely that $\emptyset_{\mathrm{ST}}$ encodes S_{ST}, and that S_{ST} encodes the higher-order property $[\lambda G\, G\emptyset_{\mathrm{ST}}]$. In formal terms, we prove the following:

$$\emptyset_{\mathrm{ST}}S_{\mathrm{ST}} \qquad (50)$$

$$S_{\mathrm{ST}}[\lambda G^{\langle i \rangle}\, G\emptyset_{\mathrm{ST}}] \qquad (51)$$

Now to derive (50), recall that we proved (29) in the previous section:

(29) $\emptyset_{\mathrm{ST}}F \equiv \mathrm{ST} \models F\emptyset_{\mathrm{ST}}$

In light of this, (50) is an immediate consequence of (21). Recall also that we proved (38) in the previous section:

(38) $S_{\mathrm{ST}}F^{\langle\langle i \rangle\rangle} \equiv \mathrm{ST} \models FS_{\mathrm{ST}}$

Now in virtue of this, (51) is an immediate consequence of (49).

We might reflect for a moment on the fact that if there is an atomic relational axiom of theory τ of the form $\Pi\kappa_1\kappa_2$, not only would it be represented as the following axiom of $\mathcal{O}$:

$$\tau \models \Pi_\tau \kappa_{1\tau}\kappa_{2\tau},$$

but it would also have the following as consequences:

$$\kappa_{1\tau}[\lambda x\, \Pi_\tau x \kappa_{2\tau}]$$

$$\kappa_{2\tau}[\lambda x\, \Pi_\tau \kappa_{1\tau} x]$$

$$\Pi_\tau[\lambda F\, F\kappa_{1\tau}\kappa_{2\tau}]$$

All three encoding claims would therefore be theorems of $\mathcal{O}$.

Finally, we contemplate the consequences of those theorems of a mathematical theory which are expressed by molecular and quantified formulas. For example, consider the axiom of ST which asserts that no set is a member of $\emptyset$. As we noted above, this becomes the following axiom (22) of $\mathcal{O}$:

(22) $\text{ST} \models \neg\exists x(S_{\text{ST}}x \;\&\; x \in_{\text{ST}} \emptyset_{\text{ST}})$

By now familiar reasoning, this axiom of $\mathcal{O}$ implies the following:

$$\emptyset_{\text{ST}}[\lambda y^i \; \neg\exists x(S_{\text{ST}}x \;\&\; x \in_{\text{ST}} y)] \tag{52}$$

$$S_{\text{ST}}[\lambda F^{\langle i \rangle} \; \neg\exists x(Fx \;\&\; x \in_{\text{ST}} 0_{\text{ST}})] \tag{53}$$

$$\in_{\text{ST}}[\lambda G^{\langle i,i \rangle} \; \neg\exists x(S_{\text{ST}}x \;\&\; Gx0_{\text{ST}})] \tag{54}$$

With these consequences in mind, we now reconsider the relationship between the language and theorems of ordinary mathematics and the theorems of our extended $\mathcal{O}$.

§6: Analysis of (Ordinary) Mathematical Language

Recall that in Section 2, we divided the true statements of ordinary mathematical language into the basic ones, which begin with the theory operator and the non-basic ones, which don't. Examples of the basic statements are:

(A) In ST, the empty set is a set.

(B) In ST, no set is a member of the empty set.

Clearly, the analyses (i.e., philosophically correct truth conditions) of these claims are given by their direct translations into our formal system as (21) and (22), respectively:

(21) $\text{ST} \models S_{\text{ST}}\emptyset_{\text{ST}}$

(22) $\text{ST} \models \neg\exists x(S_{\text{ST}}x \;\&\; x \in_{\text{ST}} \emptyset_{\text{ST}})$

These analyses reveal that the truth conditions for these statements are compositionally determined.[31] Note that our truth conditions are accompanied by a philosophical account of the abstract individuals and abstract relations that serve as the denotations of the expressions 'ST', 'set', 'is a member of', and 'empty set' as they occur in (A) and (B).

However, when we consider these same statements without the theory operator, we have several interpretative options. Consider first:

(ST2) The empty set is a set.

as asserted in the context of ST. This simple predication becomes ambiguous in the present theory. On the one hand, the following atomic exemplification reading of this claim is false:

(ST2*) $S_{\text{ST}}\emptyset_{\text{ST}}$

Though $\mathcal{O}$ doesn't assert that this formula is false, we may consistently add the assumption that it is false. Recall that this assumption is grounded in our philosophy of mathematics, on which it is asserted that (i) $\emptyset_{\text{ST}}$ encodes rather than exemplifies its mathematical properties, (ii) $\emptyset_{\text{ST}}$ exemplifies such properties as being non-red, being non-round, being thought about by the reader now, etc., and (iii) $\emptyset_{\text{ST}}$ is complete with respect to the exemplification of properties but not with respect to the encoding of properties. Given these philosophical ideas, the truth conditions of (ST2*), which it wears on its sleeve, do not obtain. So when read as (ST2*), both (ST2) and its traditional formal rendition as (ST2′) turn out to be false.

However, $\mathcal{O}$ offers a reading for (ST2) on which it turns out true. (50) is an atomic encoding claim which is not only true but a theorem of $\mathcal{O}$, as we saw in the previous subsection:

(50) $\emptyset_{\text{ST}}S_{\text{ST}}$

Given the ambiguity in language described in Section 2, (50) becomes a legitimate reading for (ST2). So we have *recovered* a sense in which (ST2) is *true*. (This preserves the intuition of mathematicians that they

[31] As soon as a philosopher of mathematics takes claims such as (A) and (B) to be fundamental, it then becomes important to specify (compositional) truth conditions for these statements. We can therefore provide the truth conditions needed to complete the position described in Field [1989] (p. 3). Without such truth conditions, Field's position has an important explanatory gap. Moreover, we shall offer, in just a moment, a reading on which the unprefixed version of (A) and (B) turn out true.

are saying something true.[32]) In this sense, our analysis says that (ST2) is *about* the set $\emptyset_{\text{ST}}$.

Note that our analysis suggests that we might equally well have regarded (ST2) as a statement about the abstract property of being an ST-set. Our work in the previous subsection suggests that theorem (51) *also* offers a true reading of the ordinary statement (ST2):

(51) $S_{\text{ST}}[\lambda G\, G\emptyset_{\text{ST}}]$

This asserts that the property of being an ST-set encodes the property of being a property that $\emptyset_{\text{ST}}$ exemplifies. On this reading, our analysis says that (ST2) is about the abstract property of being an ST-set.

We'll discuss the fact that (ST2) has alternative true readings in just a moment. But first, consider (ST3):

(ST3) No set is a member of the empty set.

We assert that (ST3) is false when represented as the formal claim:

(ST3*) $\neg\exists x(S_{\text{ST}}x \,\&\, x \in_{\text{ST}} \emptyset_{\text{ST}})$

However, any of the formal representations (52), (53), or (54), which turned up as theorems in the previous subsection, provide us with a true reading of (ST3):

(52) $\emptyset_{\text{ST}}[\lambda y^i\, \neg\exists x(S_{\text{ST}}x \,\&\, x \in_{\text{ST}} y)]$

(53) $S_{\text{ST}}[\lambda F^{\langle i\rangle}\, \neg\exists x(Fx \,\&\, x \in_{\text{ST}} \emptyset_{\text{ST}})]$

(54) $\in_{\text{ST}}[\lambda G^{\langle i,i\rangle}\, \neg\exists x(S_{\text{ST}}x \,\&\, Gx\emptyset_{\text{ST}})]$

(ST3) is not only about $\emptyset_{\text{ST}}$ but also about the property of being a set_{ST} and about the relation of $\text{membership}_{\text{ST}}$. In some sense, it doesn't matter which of theorems (52) – (54) we assign to (ST3) as the disambiguated condition under which it is true. From any one of these statements, we can recover the other two, by appealing to the Corollary to the Reduction Axiom and λ-Conversion.

However, we can take our analysis of (ST3) one step further. Let us define an *extended* sense of 'encodes' in terms of which we can say that the abstract objects $\emptyset_{\text{ST}}$, S_{ST}, and $\in_{\text{ST}}$ *encode* the following complex relation,

[32]Here again, this fills another important gap in Field's theory, for the latter doesn't offer any reading on which *unprefixed* theorems of mathematics are true. Without such a reading, the beliefs of mathematicians become something of a mystery.

namely, being an individual y^i, property $F^{\langle i\rangle}$, and relation $G^{\langle i,i\rangle}$ such that $\neg\exists x(Fx \& Gxy)$. The intuitive idea here is to define 'x, y, and z encode R' ('$xyzR$') as the conjunction of $x[\lambda u\, Ruyz]$, $y[\lambda u\, Rxuz]$, and $z[\lambda u\, Rxyu]$. To employ this idea in the case at hand, we can let the following notation be defined as the conjunction of (52), (53), and (54):

$$\emptyset_{\text{ST}} S_{\text{ST}} \in_{\text{ST}} [\lambda y^i F^{\langle i\rangle} G^{\langle i,i\rangle}\, \neg\exists x(Fx \,\&\, Gxy)] \tag{55}$$

(In the above notation, the λ-expression is a relational expression of the form $[\lambda yFG\, \psi]$, in which y, F and G are all bound by the λ.) We can use this newly defined statement of $\mathcal{O}$, and the truth conditions it encapsulates, as the reading of (ST3) on which it is true. Similarly, whereas we take the straightforward translation (ST3*) of (ST3′) to be false, (55) offers a way to understand (ST3′) as representing a truth.

Of course, this leads to a very general technique for constructing an encoding condition that expresses the reading under which a complex ordinary sentence S of a mathematical theory τ is true. Suppose that τ has been formulated in classical exemplification logic and that φ is the *traditional* formal exemplification statement which precisely renders S. (As an example, let S be (ST3) and let φ be (ST3′).) Let the primitive non-logical constants and predicates of φ be listed as $\kappa^{t_1}, \ldots, \kappa^{t_n}$. Then where $\kappa^{t_1}_{\tau}, \ldots, \kappa^{t_n}_{\tau}$ are the new corresponding symbols of $\mathcal{O}$, let φ^* be the sentence of $\mathcal{O}$ which results when we substitute $\kappa^{t_i}_{\tau}$ for κ^{t_i} in φ ($1 \leq i \leq n$). (Continuing with our example, when φ is (ST3′), φ^* is (ST3*).) Then, as we've seen, whether or not φ and S are theorems of τ, φ^* is to be regarded as false (though if φ is a theorem of τ, we know that $\tau \models \varphi^*$ is true). However, there is a statement of $\mathcal{O}$ which expresses a metaphysical truth if and only if φ and S are theorems of τ. To specify this statement in complete generality, let φ^- be the result of substituting new variables $y^{t_1}, \ldots, y^{t_n}$ for all the occurrences of the non-logical expressions $\kappa^{t_1}_{\tau}, \ldots, \kappa^{t_n}_{\tau}$, respectively, in φ^*, and let $\psi(\alpha^{t_i}/\kappa^{t_i})$ be the result of substituting the variable α^{t_i} for all the occurrences of the (constant or predicate) symbol κ^{t_i} in ψ. We may then use the definiendum in the following definition as the reading which captures the mathematical truth underlying φ:

$$\begin{aligned}&\kappa^{t_1}_{\tau} \ldots \kappa^{t_n}_{\tau} [\lambda y^{t_1} \ldots y^{t_n}\, \varphi^-] \;=_{df} \\ &\quad \kappa^{t_1}_{\tau}[\lambda y^{t_1}\, \varphi^*(y^{t_1}/\kappa^{t_1}_{\tau})] \;\&\; \ldots \;\&\; \kappa^{t_n}_{\tau}[\lambda y^{t_n}\, \varphi^*(y^{t_n}/\kappa^{t_n}_{\tau})]\end{aligned} \tag{56}$$

It should be clear that when (55) is taken as an example of the definiendum in (56), the conjunction of (52), (53), and (54) is an example of the definiens. (56) allows us to represent the truth conditions for ordinary mathematical statements (e.g., (ST3) or (ST3′)) of arbitrary complexity in terms of a single defined formula of $\mathcal{O}$. It should also be clear that the definiens of (56) is derivable as a theorem of $\mathcal{O}$ whenever φ is a theorem of τ. This sets up a correlation—each theorem of an arbitrary mathematical theory can be correlated with a theorem of $\mathcal{O}$ that is *unprefixed* by the theory operator! At this point, it may be that enough has been said to give the reader a sense of how the (ordinary) language of mathematics is to be analyzed.

Before we turn to the final section, it is important to address one objection that might be raised against our reduction of mathematics to metaphysics. The objection criticizes the reduction from a model-theoretic perspective. It might go as follows:[33]

> In various mathematical theories, many of the objects are not uniquely identifiable by descriptions expressible in the language of the theory. An example might be real number theory, where only countably many reals are nameable in a standard language. Indeed, in some mathematical theories, *none* of the objects are identifiable, for reasons of symmetry. Examples are classical geometries and Cantor's theory of dense linear orderings without endpoints. In models of these homogeneous theories, every element of the domain possesses exactly the same properties meaningful for the theory T. So you can't reduce all the distinct objects of these theories to distinct abstract objects.

There are actually two separate questions raised here, namely, what to do about theories which assert the existence of objects that are not uniquely identifiable by descriptions expressible in the theory, and what to do about theories which assert the existence of distinct symmetrical objects. The two questions are related, however. They both arise because the model-theoretic conception of the 'objects of' a theory is rather different from the metaphysical conception, which we defined above as (12). On the model-theoretic conception, an 'object of' a theory is any element of any domain of quantification that is part of the intended model of the theory.

[33]I am quoting and paraphrasing here from an unpublished paper by Brent Mundy. I think he states the objection nicely. Allen Hazen raised a similar concern in oral presentations at the Australian National University and the University of Alberta. And Thomas Hofweber raised a variant of the objection in a recent conversation about the present paper.

The model-theoretician uses this definition to claim that there can be objects of a mathematical theory that are inexpressible, since the theory has no well-defined terms to denote them. What analysis does $\mathcal{O}$ offer when this is the case?

I think there are two parts to an effective response to this objection. The first part of the response is to point out that the model-theoretic objection, in some deep sense, begs the question as an argument against our foundational metaphysics. The objection *assumes* the model-theoretic definition of 'object of' and so *uses* mathematical language (e.g., the set theoretic notions of domain, model, satisfaction, etc.), thereby presupposing that the semantics of that language is clear. But the semantics of mathematical language is precisely what is in question. Our ontological project is to give an account of mathematical reference and truth in terms of a more basic, mathematics-free language and theory. So objections which presuppose an account of reference and truth in terms of mathematical language (the semantics of which, after all, is in question) lose their force. *Model theory just becomes another mathematical theory that is subject to a metaphysical reduction*. Our metaphysics tells us what the terms of model theory refer to and tells us the sense in which its claims are true.

The second part of an effective response is to reiterate the metaphilosophical claim that the two principal tasks of a philosophy of mathematics are to account for reference and truth. A philosophy of mathematics must not only identify the *referents* of the well-defined terms and predicates of mathematical theories but also precisely describe the conditions under which the theorems of mathematics turn out to be *true*. We now have accomplished both tasks—the analysis of mathematical reference is given by (47) and the analysis of mathematical truth is given by (56). The model-theoretic objection can be put to rest by the facts that we have stated truth conditions in $\mathcal{O}$ for *every* sentence of an arbitrary mathematical theory and that we can correlate *every* theorem of an arbitrary mathematical theory with a theorem of (extended) $\mathcal{O}$ that is unprefixed by the theory operator. It therefore becomes a mistake to suppose that in order to answer the ontological question about what the objects of a mathematical theory are, a foundational metaphysics has to be able to give a classical reduction, relative interpretation, or model of that theory.[34]

[34]In the case of the theory of dense linear orderings without endpoints (DLO), the

§7: Philosophical Observations

By showing that the individuals and relations of arbitrary mathematical theories are just abstract individuals and abstract relations found in the ontology of typed $\mathcal{O}$, we've produced a *prima facie* case for the main thesis of this paper. Of course, $\mathcal{O}$ has to be extended with new primitive symbols, with analytic truths of the form $\tau \models \varphi^*$, and with the analyses of statements which everyone assumes to be true, namely, $MathTheory(\tau)$, for recognizable mathematical theories τ. But the resulting system allows us to *prove* what many other philosophers stipulate, namely, that mathematical objects are abstract objects. Since $\mathcal{O}$ includes only primitive notions of logic and metaphysics,[35] we can conclude that mathematical objects fall into a more fundamental ontological category. This is an ontological reduction of one kind of entity to another.

Of course, there are numerous philosophical issues that arise in connection with our metaphysical reduction. Many of those issues were addressed in Linsky and Zalta [1995] and we shall not rehearse them in any detail here. We shall, however, consider the question of mathematical objectivity, but before we do so, it is important to consider the extent to which the present theory constitutes a kind of neo-logicism. Throughout this essay, we have presented $\mathcal{O}$ as a proper metaphysical theory. The comprehension principle for abstract objects appears, by most reasonable lights, to be a synthetic *a priori* truth and not an analytic truth of logic. However, there is a way to restate the comprehension principle so that it looks much more like a truth of logic, or at least more like an analytic truth. I shall not claim that this reformulated version of comprehension *is*

present theory analyzes the ordering relation $<_{\text{DLO}}$ by abstracting out the properties of relations that $<_{\text{DLO}}$ must encode in order to behave according the axioms of DLO. And the present theory tells us the sense in which the sentences of DLO are true. But then there are simply no further *ontological* questions to answer; in particular, there are no specifiable (type i) individuals which constitute objects of DLO that need to be identified. Sentences like "there are infinitely many points which are such and such" can be *true in* a mathematical theory even though there are no names for the points and no witnesses to the claim. Thus, everything that can actually be *said* in a mathematical theory gets an account.

[35]The implementation of $\mathcal{O}$ deployed in this paper has the following primitives: individual (type), relation (type), exemplification and encoding (i.e., modes of predication), the usual logical and modal primitives ($\neg$, $\rightarrow$, $\forall$, $\Box$, λ, $\imath$), the non-logical primitive $E!$, and the non-logical notions of '$Math(p)$' and 'Authorship'. None of these are mathematical notions—there are no mathematical constants like 0, $\emptyset$, etc., and no mathematical predicates such as membership, functions, maps, successor, etc., among our primitives.

a truth of logic, or analytic, but it will be recognized that some philosophers would conclude that it is. Our theory of abstract objects could have been presented by replacing the comprehension principle (3) by the theorem governing canonical descriptions (7):

$$(7)\quad \imath x^t(A!^{\langle t\rangle}x \,\&\, \forall F^{\langle t\rangle}(xF \equiv \varphi))G^{\langle t\rangle} \equiv \varphi^{G^{\langle t\rangle}}_{F^{\langle t\rangle}}$$

By elevating this theorem to the status of an axiom, with the understanding (stipulation) that all canonical descriptions of abstract objects denote, we have an equivalent formulation of $\mathcal{O}$. Moreover, if one considers what (7) *asserts*, then it clearly has 'the ring' of an analytic truth: the abstract object that encodes just the properties such that φ encodes property G iff G is such that φ. So is (7) an analytic truth? If so, does $\mathcal{O}$, when reformulated in this way, become a part of logic?[36] If the answers to these two questions are 'Yes', then our ontological reduction of mathematical objects might constitute a kind of neo-logicism.

As mentioned earlier, I do not claim that (7) is an analytic or a logical truth. At best, it is analogous to the following 'abstraction' principle (governing 'set abstracts') that might be employed as a substitute for (ST5):

$$z \in \{y \mid y \in x \,\&\, Fy\} \;\equiv\; z \in x \,\&\, Fz$$

Although this also has the ring of an analytic truth when introduced as a contextual definition, I doubt that it is analytic when introduced as a basic axiom that governs the *primitive notation* '$\{y \mid y \in x \,\&\, Fy\}$'. However, some philosophers have argued that axioms analogous to (7) and the above abstraction principle are analytic. Using Frege's [1884] Context Principle as a guide, Wright [1983] argues that Hume's Principle ($\#F = \#G \equiv F \approx G$) is an analytic truth.[37] If Wright considers the

[36]This question assumes that we can also justify the claim that the two other proper axioms of $\mathcal{O}$ can be understood as logical truths. These are (2) (abstract objects don't encode properties), and the axiom for the substitution of identicals. Since identity is *defined* in $\mathcal{O}$, and the definition involves our non-logical notion '$E!^t$', the principle for the substitution of identicals is correctly asserted as a proper axiom. Even though one could constrain interpretations of the theory so that these axioms turn out to be true in every interpretation, an appeal to such interpretations does not automatically constitute an argument for thinking that the principles in question are logical truths.

[37]For the uninitiated, Hume's Principle asserts that the number of Fs is identical to the number of Gs iff F and G are equinumerous (where 'equinumerous' has its usual definition in second-order logic). See the discussion of 'number-theoretic logicism' in Wright [1983], pp. 153-154.

result of adding Hume's Principle to second-order logic to be a logical system ('number-theoretic logicism'), then it would seem that he would have to regard the result of adding (7) to the logic of encoding as a logical system. Of course, there are, in the literature, trenchant criticisms of Wright's position concerning the analytic character of Hume's Principle.[38] Since I hope to discuss these issues at length on another occasion, I shall simply observe that the relationship between (7) and (3) is analogous to the relationship between Hume's Principle and the existence claim that Boolos calls 'Numbers' in his [1987].[39] The conclusion I wish to draw at this point is simply that the above treatment of mathematics constitutes a kind of neo-logicism if Wright's claim about the analyticity of Hume's Principle can be sustained.[40]

Consider, next, the question of mathematical objectivity. No doubt, it will be argued that if every mathematical theory is about a distinctive group of abstract objects, then there is no way to account for mathematical objectivity.[41] But, modulo our classical reduction of the natural numbers, if mathematical objectivity is correctly described in Linsky and Zalta [1995] and Field [1998b] as being limited to the objectivity of logical consequence, then there is no special problem of mathematical objectivity for the above theory.[42] These works deny that there is a single, objec-

[38]See, in particular, Field [1984] and Boolos [1997].

[39]Boolos formulates 'Frege Arithmetic' in terms of the axiom:

$$\text{Numbers: } \forall F \exists ! x \forall G (G \eta x \equiv G \approx F)$$

See [1987], p. 5 (or the reprint [1998], p. 186). Boolos discusses how Hume's Principle is grounded in Numbers. At some point, I hope to discuss the similarities between Boolos' η relation and the notion of encoding. To anticipate, compare the paradoxes of encoding described in Zalta [1983], Appendix A (pp. 158-159) with the paradoxes of η described in Boolos [1987], p. 17 (Boolos [1998], p. 198).

[40]See Rosen [1993] for an interesting discussion of this question.

[41]I think one way to defend the theory here would be to suggest that it simply offers a more well-developed account of Carnap's [1950] view that each 'linguistic framework' (substitute 'mathematical theory' for 'linguistic framework'), in some sense, presupposes its own group of objects. Carnap failed to explain how the language (predicates and constants) of each framework come to denote the right relations and objects, and our theory at least gives an account of this in the case of mathematical frameworks.

[42]It is important to remember the following. (1) In the present framework, logical consequence is a primitive notion that is axiomatized in the very specification of $\mathcal{O}$. (2) The intended understanding of the second order variables of $\mathcal{O}$ is that they range over *properties and relations*, where these are *not* construed as set-theoretic entities. The difference between sets and properties is vast—sets merely classify objects, whereas properties *characterize* objects. (3) Therefore, we are not presupposing a definition, based on standard models of second-order language, of the second-order logical consequence relation.

tively true set theory, that there is an objective fact of the matter as to whether the axiom of foundation is true (for there are perfectly good non-well-founded set theories), and that there is a fact of the matter as to the size of the continuum (there are perfectly good set theories which differ in their answer to the size to the continuum). Each set theory is simply about a different membership relation.

Though Field [1998b] concludes (p. 401) that an account of mathematical objectivity is more important than an account of mathematical objects, it may be more perspicuous to say that philosophers need a correct account of both if they are to have a comprehensive philosophy of mathematics: objects are to objectivity what reference is to truth. In addition to the ways mentioned in footnotes 17, 27, 31, and 32, the present analysis supplements Field's work as follows: (a) it gives a correct account of mathematical objects that is consistent with the view of mathematical objectivity he develops,[43] (b) it explains the indeterminacy in our mathematical concepts, discussed in Field [1994], without abandoning the idea that our mathematical predicates denote particular mathematical relations, and (c) it offers an account of the meaningfulness of the language of inconsistent mathematical theories.[44] This last fact deserves a brief discussion.

The analysis of mathematical language described above extends even to inconsistent mathematical theories. To take a classic example, consider Frege's *Grundgesetze der Arithmetik*. Recently, there has been a renaissance of interest in this work and it has become the subject of many philosophical and logical investigations. In the *Grundgesetze*, there are many hundreds of pages of formulas in Frege's special script, and despite the in-

[43]As mentioned above, Field defends the view that logical objectivity (suitably qualified) is all the objectivity that there is in mathematics. He clearly rejects the idea that there is one true set theory or one correct answer to such questions as the Continuum Hypothesis. Similar claims were defended in Linsky and Zalta [1995]. Moreover, the specific kind of mathematical objectivity inherent in number theory that he *would* accept is validated in $\mathcal{O}$ by the fact that the theory of natural numbers can be given a classical reduction, as described in Zalta [1999] and earlier in this paper! Here is where our work substantiates, to some extent, Kronecker's view that the natural numbers are made by God but that all the other numbers are man-made.

[44]It is interesting that in [1998b], Field (p. 398) seems to identify something like the present account with the structuralism of Resnik [1981] and Shapiro [1989]. (These works have been superceded by Resnik [1997] and Shapiro [1997], respectively.) I believe that the present account offers a more fine-grained account of the structuralist philosophy of mathematics than that found in these works, but I shall not argue for that here.

consistency of the system, these formulas are meaningful! How are we to describe the semantics of this language? The answer given by the present theory is that the terms and predicates of Frege's language denote abstract objects that encode properties that are inconsistent with one another.[45] Of course, the objects of an inconsistent theory τ will be uninteresting, for they will encode *all* properties (formulable in τ).[46] That explains why mathematicians try to avoid postulating inconsistent theories. But note that we now have a *unified* semantics of mathematical language.[47]

It is important to reflect next on the features of metaphysical reductions that contrast with other forms of reduction. Clearly, our metaphysical reductions are not classical reductions, for the theorems of arbitrary mathematical theories T do not constitute a subtheory of a definitional extension of $\mathcal{O}$. Moreover, our metaphysical reductions of the objects of arbitrary mathematical theories do not show that those mathematical theories are relatively interpretable in $\mathcal{O}$. Nor are we using $\mathcal{O}$ to build models for arbitrary mathematical theories. We are not claiming that mathematical notions can be *defined* in terms of the notions of pure logic and metaphysics. Nor are we suggesting that we can get along without the proper axioms of mathematics by being creative with the logical axioms, non-logical axioms, and definitions of $\mathcal{O}$. Instead, we've developed a new kind of reduction, which yields a precise philosophical account of mathematical objects. In a sense, our metaphysical reductions constitute a distinctive new kind of relative interpretation, for every theorem φ of an arbitrary mathematical theory τ can be correlated with a (specially

[45]Whereas it is a theorem of $\mathcal{O}$ that $\neg(xF \,\&\, \neg xF)$, $\mathcal{O}$ asserts the existence of all kinds of objects that encode inconsistent properties. There are abstract objects that encode a property P as well as its negation $\bar{P}$ (where $\bar{P} =_{df} [\lambda y\, \neg Py]$).

[46]Suppose that theory τ yields a contradiction and that κ^t is an object of τ. Then, for some φ, both $\tau \models \varphi$ and $\tau \models \neg\varphi$ will be true in $\mathcal{O}$. Not only does our Rule of Closure now allow us to infer $\tau \models \psi$ (for any ψ), but where ψ' is the result of substituting the new variable y^t for κ^t in ψ, we may infer $\tau \models [\lambda y^t\, \psi']\kappa^t$. So, by (48), we know that κ^t encodes every τ-formulable property $[\lambda y^t\, \psi']$.

[47]At this point, our semantics accounts for the *denotation* of mathematical terms and predicates. But the expressions of mathematical language also have a 'sense'. This Fregean sense can also be modeled in $\mathcal{O}$. See Zalta [1983], Chapter VI, and Zalta [1988], Chapters 9 – 12. Note that whereas the assignment of denotations of mathematical expressions, on the above analysis, is independent of the mental states of mathematicians, we might suppose that the sense of a mathematical expression *for* person x encodes the properties involved in x's conception of the object denoted by that expression. This is how we account for error and ignorance in mathematical beliefs.

identified) theorem of *extended* $\mathcal{O}$, namely, the definiens of (56). I think one conclusion we should draw from all of this is that no matter how mathematicians carry on with their work and no matter how the mathematics they produce might turn out, philosophers will have something metaphysically precise and circumspect to say about the subject matter of the resulting mathematics and about the proper semantic analysis of the language used to express it.[48]

There is one final observation to make before we conclude, namely, that it is fascinating (and possibly insightful) to consider that many of the ideas about the metaphysical reduction of mathematical objects expressed thus far presuppose a certain 'platonist' interpretation of the formalism of $\mathcal{O}$. In the present paper, we have employed the 'Quinean' understanding of the quantifiers of $\mathcal{O}$, in which the quantifier '$\exists$' is read 'there exists' and the predicate '$E!^t$' is read 'is concrete^t'. On this understanding, $\mathcal{O}$ asserts that there exist objects (individuals and relations) that couldn't possibly be concrete. This is just a consequence of the comprehension principle (3) and the definition (1) of 'abstract'. However, one can give the formalism of $\mathcal{O}$ a 'fictionalist' reading, by using the 'Meinongian' reading of the quantifier '$\exists$' as 'there is' (with no implication of existence) and by reading the predicate '$E!$' as 'exists'. On this reading, $\mathcal{O}$ asserts that there are objects that don't (and couldn't possibly) exist. On such a fictionalist reading of $\mathcal{O}$, one can say that abstract objects are *fictions*, since they don't exist. So mathematical objects become metaphysically reduced to the more general category of *fiction*.[49] The fact that $\mathcal{O}$ has these two fundamental readings is, in our opinion, what grounds the 'equivalence' of the platonist and fictionalist philosophies of mathematics described in Balaguer [1998].[50] The reader might find it worthwhile to consider just

[48]Our work may therefore also supplement the conclusion of Maddy [1997] with a precise philosophical account of the language of any theory that the *mathematicians* decide is the best way to extend ZF.

[49]Such a view seems to be consistent with the ontological views of Wagner [1982]. However, it does have the consequence that abstract objects couldn't possibly exist (since 'abstract' is defined as 'not the kind of thing that could exist' on this interpretation). So although this interpretation does preserve a large part of Field's fictionalism, it is inconsistent with his view that numbers are fictions that *contingently* fail to exist. See Field [1993].

[50]Balaguer reaches this interesting 'equivalence' thesis by sketching what he takes to be the best version of platonism and fictionalism. Though his versions of platonism and fictionalism are not axiomatized, his version of platonism is, like ours, based on a plenitude principle. However, the plenitude principle of his 'full-blooded platonism'

how many of the remarks made in this last section apply, with minor readjustment, to the fictions described by this alternative reading of the formalism of $\mathcal{O}$.

Bibliography

Balaguer, M., 1998, *Platonism and Anti-Platonism in Mathematics*, Oxford: Oxford University Press

Balaguer, M., 1995, 'A Platonist Epistemology', *Synthese* **103**: 303-325

Benacerraf, P., 1965, 'What Numbers Could Not Be', *Philosophical Review*, **74**: 47-73

Boolos, G., 1987, 'The Consistency of Frege's *Foundations of Arithmetic*', in *On Being and Saying*, J. Thomson (ed.), Cambridge, MA: MIT Press; reprinted in Boolos [1998], pp. 183–201

Boolos, G., 1997, 'Is Hume's Principle Analytic?', in *Logic, Language, and Thought*, R. Heck (ed.), Oxford: Oxford University Press; reprinted in Boolos [1998], pp. 301-314

Boolos, G., 1998, , *Logic, Logic, and Logic*, Cambridge, MA: Harvard University Press

Carnap, R., 1967, *The Logical Structure of the World*, R. George (trans.), Berkeley: University of California Press

(FBP) cannot account for our *de re* mathematical beliefs. (Indeed, Balaguer rejects the claim that we have *de re* mathematical beliefs.) So FBP is not articulated in such a way that one can *prove*, for example, that $\emptyset_{\mathrm{ZF}}$ is a particular abstract individual. Nor is Balaguer's version of fictionalism developed in such a way that the truth conditions for theory-prefixed mathematical sentences can be precisely specified. But despite using 'naive' versions of platonism and fictionalism, Balaguer puts his finger on a deep kind of equivalence between these two philosophies of mathematics.

Indeed, in the last chapter of [1998], Balaguer concludes both (a) that the *only* point of disagreement between FBP and fictionalism is on the question of whether mathematical objects exist and (b) that there is no fact of the matter whether abstract objects exist. I suggest that the present work, in some sense, *validates* this idea by the fact that one (and possibly the only) way in which platonists and fictionalists could disagree on only one issue of substance is if they both adopted the *same* formalism to express their theory but with *different* readings or interpretations of the formalism's 'existential' quantifier. If the formalism of $\mathcal{O}$ is the most articulate formulation of both plenitudinous platonism and fictionalism, then if there is really no fact of the matter as to which reading, platonist or fictionalist, is better, then one might, by 'semantic descent', reach the conclusion that there is no fact of the matter as to whether abstract objects exist. I think this is the proper way to understand Balaguer's conclusions (a) and (b) in the last chapter of his [1998].

Carnap, R., 1950, 'Empiricism, Semantics, and Ontology', *Revue Internationale de Philosophie*, **4**: 20-40

Feferman, S., 1998, *In the Light of Logic*, Oxford: Oxford University Press

Feferman, S., 1998a, 'What Rests on What? The Proof-Theoretic Analysis of Mathematics', in Feferman [1998], pp. 187-208

Feferman, S., 1988, 'Hilbert's Program Relativized: Proof-Theoretical and Foundational Reductions', *Journal of Symbolic Logic* **53**: 364-384

Feferman, S., 1960, 'Arithmetization of Metamathematics in a General Setting', *Fundamenta Mathematicae* XLIX: 35-92

Field, H., 1998a, 'Which Undecidable Mathematical Sentences Have Determinate Truth Values?', in *Truth and Mathematics*, H. Dales and G. Olivieri (eds.), Oxford: Oxford University Press, pp. 291-310

Field, H., 1998b, 'Mathematical Objectivity and Mathematical Objects', in *Contemporary Readings in the Foundations of Metaphysics*, S. Laurence and C. Macdonald (eds.), Oxford: Blackwell, pp. 387-403

Field, H., 1994, 'Are Our Logical and Mathematical Concepts Highly Indeterminate?', *Midwest Studies in Philosophy*, P. French, T. Uehling, H. Wettstein (eds.), Volume 19, Notre Dame: Notre Dame University Press, pp. 391-429

Field, H. 1993, 'The Conceptual Contingency of Mathematical Objects', *Mind*, **102**/406 (April): 285-299

Field, H., 1989, *Realism, Mathematics, and Modality*, Oxford: Blackwell

Field, H., 1984, 'Critical Notice of Crispin Wright: *Frege's Conception of Numbers as Objects*', *Canadian Journal of Philosophy*, **14**: 637-662; reprinted in Field [1989], pp. 147-170 (with the new title 'Platonism for Cheap? Crispin Wright on Frege's Context Principle')

Field, H., 1980, *Science Without Numbers*, Oxford: Blackwell

Frege, Gottlob, 1884, *The Foundations of Arithmetic*, translated by J. L. Austin, Oxford: Blackwell, 1974 (second revised edition)

Frege, Gottlob, 1893/1903, *Grundgesetze der Arithmetik*, Band I/II, Jena: Verlag Hermann Pohle

Hempel, C., 1945, 'On the Nature of Mathematical Truth', *American Mathematical Monthly*, **52**: 543-556; reprinted in *The Philosophy of Mathematics: Selected Readings*, H. Putnam and P. Benacerraf, (eds.), second edition, Cambridge: Cambridge University Press, pp. 377-393.

Jubien, M., 1969, 'Two Kinds of Reduction', *The Journal of Philosophy*, **66**/17: 533-541

Kripke, S., 1959, 'A Completeness Theorem in Modal Logic,' *Journal of Symbolic Logic* **24**: 1-15

Lewis, D., 1986, *On the Plurality of Worlds*, Oxford: Blackwell

Linsky, B., and Zalta, E., 1995, 'Naturalized Platonism vs. Platonized Naturalism', *The Journal of Philosophy*, xcii/10 (October 1995): 525-555

Maddy, P., 1997, *Naturalism in Mathematics*, Oxford: Clarendon

Mundy, B., unpublished manuscript, 'Zalta's Logic of Encoding' (version #3, January 1996)

Niebergall, K. G., 2000, 'On the Logic of Reducibility: Axioms and Examples', *Erkenntnis*, **53**/1-2: 27–61.

Pelletier, J., and Zalta, E., 2000, 'How to Say Goodbye to the Third Man', *Noûs*, 34/2 (June 2000): 165–202.

Resnik, M., 1997, *Mathematics as a Science of Patterns*, Oxford: Clarendon

Resnik, M., 1981, 'Mathematics as a Science of Patterns: Ontology and Reference', *Noûs*, **15**: 529-550

Rosen, G., 1993, 'The Refutation of Nominalism(?)', *Philosophical Topics*, **21**/2: 149–186

Quine, W., 1976, 'Ontological Reduction and the World of Numbers', in *The Ways of Paradox and Other Essays*, rev. ed., (Harvard: Harvard University Press), pp. 212-220

Shapiro, S., 1997, *Philosophy of Mathematics: Structure and Ontology*, Oxford: Oxford University Press

Shapiro, S., 1989, 'Structure and Ontology', *Philosophical Topics*, **17**: 145-171

Tarski, A., Mostowski, A., and Robinson, R., 1953, *Undecidable Theories*, Amsterdam: North Holland

Visser, A., 1998, 'An Overview of Interpretability Logic', in *Advances in Modal Logic*, M. Kracht, M. de Rijke, H. Wansing, and M. Zakharyaschev (eds.), Volume 1, CSLI Lecture Notes No. 87, Stanford: Center for the Study of Language and Information Publications, pp. 307-359

Wagner, S., 1982, 'Arithmetical Fiction', *Pacific Philosophical Quarterly*, **63**: 255-269

Wright, C., 1983, *Frege's Conception of Numbers as Objects*, Aberdeen, Scotland: Aberdeen University Press

Zalta, E., 2000, 'A (Leibnizian) Theory of Concepts', *Philosophiegeschichte und logische Analyse / Logical Analysis and History of Philosophy*, 3: 137–183.

———, 1999, 'Natural Numbers and Natural Cardinals as Abstract Objects: A Partial Reconstruction of Frege's *Grundgesetze* in Object Theory', *Journal of Philosophical Logic*, **28**/6: 619-660

———, 1993, 'Twenty-Five Basic Theorems in Situation and World Theory', *Journal of Philosophical Logic*, **22**: 385-428

———, 1988, *Intensional Logic and the Metaphysics of Intentionality*, Cambridge, MA: MIT/Bradford

———, 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: D. Reidel