Existence, Truth, and Method
In the Mathematical Theory of Sets

Curtis Franks
Professor Richard Grandy
Rice University 2000

Preface

In these three chapters I treat a variety of issues that surround the current state of set theory. Through an analysis of the relationships between set theory and other branches of mathematics, I try to paint a truer picture of the nature of the theory of sets than what seems to be the prevalent characterization of the subject. This same analysis directs my approach to questions about certain undecidable statements in set theory. My aim is to suggest that a proper understanding of set theory should provoke a realist attitude toward mathematical truth. One consequence of having such an attitude is that the present formal structure of set theory necessarily fails to capture all that there is to know about sets. With that in mind, I propose a change in the foundations of the theory which I hope is a beginning of a richer set theory.

This volume was written under the generous support of the Rice Undergraduate Scholars' Program in 1999 and 2000. I thank Professors James Kinsey, Scott Derrick, and Don Johnson for their enthusiasm and the opportunities offered in the program. Thanks are also due to Byeong Uk-Yi and Alasdair Urquhart, whose insights and encouragement greatly affected the direction of my research. Professor Richard Grandy of the Rice University Philosophy Department was my advisor during the course of this research. I am greatly indebted to his contributions in this project.

Curtis Franks
April 2000

Contents

I. Irreducibility in Mathematics

1. Introduction

Axioms are given, and theorems require proof. We have a comfortable familiarity with this much of mathematics. Still, several questions remain: What kinds of objects are mathematical objects? What is mathematical truth—if we can never know whether some seemingly mathematical statements are true, are we even clear on the meaning of truth and falsity in a mathematical context? How do we come to know about mathematical objects, if they are not physical objects, and how do we learn so much about the physical from them? Perhaps our comfortable familiarity is not characteristic enough of mathematics.

Perhaps it is not characteristic at all. The controversy surrounding Euclid's fifth postulate demonstrated that a system of axioms need not always be taken as a hypothetico-deductive system. Mathematicians traditionally regarded the axioms of geometry as intuitive truths which hardly required proof against our observations of the physical world. Yet the demonstrated independence of the parallel postulate inspired mathematicians to study seemingly deviant geometries before their applicability to physics was apparent. We were left not with one new geometry, but with several, whose axioms have since proven characteristic of many physical systems.

In topology classrooms, students are asked to prove the axioms of homology theory. These Eilenberg-Steenrod axioms from algebra are like theorems for the topologist. Students of differential geometry might also prove these axioms. Neither the topologist nor the geometer worry much about this. They are glad to use the algebraic results from homology theory to prove more theorems in their respective theories.

For mathematics is no single theory. It is a discipline of several theories, each with their own structures and objects. A mathematician can study a single theory independently, but often she is motivated by the fact that some physical or social system satisfies the axioms of her theory, and more often still she borrows results from other theories when the objects she studies satisfy the axioms of another part of mathematics.

In this essay I will argue that to understand mathematics properly we cannot ignore the variety of mathematical theories. I will critique a common trend of the century, reduction of mathematics to foundations, by examining ontological reduction in general and testing its application to mathematical theories. By the end I will have demonstrated how three distinct types of reduction are inadequate in characterizing mathematics. My hope is that a philosophical approach to mathematics tempered by these irreducibility arguments might provide a framework from which we may readdress questions in the philosophy of mathematics.

2. Inscrutability of reference

To what things, I might wonder, is a speaker of a foreign language referring? What objects do her words denote, that her speech might be meaningful? Luckily, I have with me a dictionary which glosses her language, term for term, into my own. Her sudarium, I know, is a kerchief; that it is album, I discover, means that it is white. For everything she speaks of and for every quality she attributes to her world, I need only time to translate her words into my own and I have grounded the mystery of her scattered staccatos and tones.

I can continue communication throughout the afternoon or develop a lifelong friendship. I can even commit the pages of my dictionary to memory and shift back and forth form my English thoughts to the foreign tongue. I can do this, in fact, and never mistake her words. We speak of the same things and only call them by different names, I know, because the authors of the dictionary were not in the business of deceiving me. The intertranslatability of our words matches perfectly our respective behaviors. But can I really be certain that everything survives the translation?

What if, when my new friend says of her kerchief (sudarium) that it is white (album) she has in mind that it is white and known to her before the turn of the new millennium or blue and as yet hidden from her sight. This circuitous definition certainly does apply to her white kerchief, and neither I nor any of the authors of my dictionary could ever know that *album* is so defined, and not simply our familiar term *white* (Goodman, 74). Furthermore, what if when she speaks of her sudarium she has in mind all but the edge stitching of the kerchief. Again, this subtlety would never come to light in our discussions, and neither would my dictionary have warned me of any discrepancy (Quine, OR, for similar examples).

We can, it would seem, never trust the reference of those who do not speak our language—after all, our translations could ‘mistake’ each foreign term for our own English words and we would never know the difference. But then, it is not even important that there be discrepancy in language. For what if I, when I say that my kerchief is white, mean what you would mean by ‘white and observed by me before the turn of the new millennium or blue and as yet hidden from my sight'? That our terms are recognizable as phonetic equivalents does nothing from preventing this situation. Are you, then, to distrust what I mean by my speech?

The answer, clearly, is that you should not, and moreover, that I should not doubt my understanding of foreign speech, which I know only through translation. For indeed, although we only know the extension of other's words through translation, between or within languages, the fact that we cannot point out discrepancies in our reference means that we understand one another after all. It is the correspondence between our behaviors and our sentences that establishes safe intertranslatability, not some fact of the matter about the objects and relations in our meanings (Quine, OR, 49).

The reason for the intertranslatibility among the thoughts and words of people who no more know the reference of others' terms than by what the translation allows is that the contents of our sentences, to whatever extent they are theory-bound, lie in the relationships between the objects of our reference and not in the objects themselves (Quine, 1981, 20). To illuminate this notion, let me introduce some notation. A theory, T(O,R) shall be a set of sentences on an ordered pair of names O of objects and relations R among those names. Given such a pair, all names and all relations can be freely combined to form sentences within the theory, so that a set of sentences which utilizes only a subset U of the names and a subset S of the relations in a pair (O,R) is not properly a theory on (O,R), but rather a theory on (U,S). The objects of the theory, by having different names within the theory, determine the theory's intension. The relations, by demonstrating which objects in the theory can be distinguished in any way other than by name, determine its extension. Sentences in a theory are assigned truth values according to the logic associated with the theory. Two theories are intertranslatable if each of the following conditions are met:

1. For any sentence in a theory T_1, its translation in a second theory T_2 has the same truth value.
2. From within either theory, one can determine the truth value of any sentence whose truth value is known in the other theory.

In other words, there exists a function $f:T_1(O_1,R_1)\rightarrow T_2(O_2,R_2)$ such that f assigns, to each name x of O_1 a name f(x) of O_2 and to each relation r of R_1 a relation f(r) of R_2 which preserves composition of relations (i.e. $f\,(r \circ s) = f\,(r) \circ f\,(s)$) and which preserves extension, i.e. If x and y in O_1 have the same extension in R_1, then f(x) and f(y) in O_2 have the same extension in R_2. This last condition is not, as Goodman appears to have thought, equivalent to f being a one to one, or injective, function, for the identity relation within a theory "=" is intensional identity, whereas the identity relation relevant to our condition is extensional identity. For instance, a theory with Mark Twain, Samuel Clemens, Kurt Vonnegut, and a few dozen novels and short stories as objects will translate into a theory with only Samuel Clemens, Kurt Vonnegut, and the literary works as objects by mapping Mark Twain to Samuel Clemens, and mapping the names of all other objects to themselves. This is because in the former theory, Mark Twain and Samuel Clemens, while intensionally distinct, are extensionally identical.

A translation scheme satisfying these conditions gives rise to what Quine has called the inscrutability of reference (Quine, 1981, 19). The mapping f, which makes explicit condition 2, is known as a proxy function (Quine, 1981, 20), for it designates which names and relations in T_2 are to stand for which names and relations in T_1. The modest restraints imposed on f are necessary and sufficient to preserve the structure of the theory in such a way that condition 2 is met. Without these constraints, a theory could model another theory with respect to the truth values of its corresponding sentences, but the theory would not be autonomous in that the truth values of its sentences could not be determined from within that theory alone. In the translation example that began this section, the dictionary, as a volume or in memory, serves as a proxy function.

The primary example of inscrutability of reference comes in the notion of a *reducing theory* in which all names of objects from a first theory which share an extension are identified. This is the formalization of the commonplace practice of eliminating distinctions without a difference from a theory. In such an instance, the relations are naturally identified, as is easy to check. There are two important ways in which this reducing theory can be constructed, depending on two different ways in which to make the aforementioned identities, which will be an important distinction throughout this essay.

The first way to design a reducing theory is to take an equivalence class of all the names of objects to be identified and to call that class an object in the reducing theory. The second way is to *choose* a single name from each equivalence class, and call it the object in the reducing theory. In either case, the reducing theory is autonomous and is preferable if our aim is to avoid a redundant ontology. The second method is commonly associated with cases of intensional confusion such as in my above example with Mark Twain and Samuel Clemens, and the first method is common in mathematics. Below is an example of the progression from one lower level mathematical theory to a higher, "more abstract" level.

Let T(O,R) be geometry. O is the set of the names of all geometrical objects—spheres, cubes, lines, discs, polyhedra, and the like, of all various sizes and dimensions, and R is the set of all geometrical relations: congruence, similarity, translations between the various objects, etc. In this theory, two named objects are the same if and only if the relation between them is an isometry (here the objects will be of the same size and rigid shape), in which case the theory identifies the objects. Suppose, however, that some mathematicians decide that they are only interested in certain properties of these objects, which are more evident when the translations are replaced with the more general continuous functions. In the new theory $T_1(O_1=O,R_1)$, entire classes of objects will behave exactly the same, because any statements about there being continuos functions from some object x_1 onto another y will also be true of all objects "homeomorphic to" x_1. For instance, if x_1 and x_2 are both polyhedra of the same dimension, then both x_1 and x_2 are homeomorphic to all spheres of that same dimension. Our theory $T_1(O_1, R_1)$ simply attempts too many distinctions, and effectively gives an infinite number of names to each object (topological space) whose extension is determined by the relations (continuous functions).

We can reduce this theory in either of the ways described above. Reduction in the first way, where equivalence classes become objects in the reducing theory $T_3(O_3,R_3)$, gives the mathematical theory "topology", when restricted to geometrical objects. Reduction in the second way (which is possible, even without the axiom of choice) gives a theory $T_4(O_4,R_4)$ which is truth-functionally equivalent to topology (again restricted) and autonomous. This is not the theory that topologists use, however, because any choice of objects from each equivalence class will necessarily result in O_4 objects with unnecessary information. O_3 objects are in some sense cleaner, in that only the properties which affect the relations of the theory are present in the objects.

If our theory were an economic one, say, one whose relations did not distinguish people with different incomes, this 'cleanliness' would be more apparent. We would not want our theory to have as objects an individual, named person for every possible income. While objects like these would in no way disrupt the structure of the theory, the names and lives of the people are inconsequential. Similarly, the objects O_4 above have relative sizes and rigidity which are inconsequential to the topologist. Just as it is preferable to consider only classes of people, or even incomes themselves, as objects in our hypothetical economic theory (Quine, OR, 55), mathematicians prefer to use the objects O_3 in topology.

A theory about geometrical objects whose relations only make use of the topological properties of those objects is reducible either to a theory about equivalence classes of objects or to a theory whose objects are single representative geometrical objects from each topological equivalence class. Moreover, each theory is autonomous. A genealogy of mathematics would suggest that the geometrical objects come first, and the very nature of topological objects—their being equivalence classes of geometrical objects—further suggests that the objects in geometry are somehow more real than those in topology. Yet we have seen that T_3 is the cleaner theory, despite the intuitive unnaturalness of its objects. In section 6 I will return to this question of naturalness and reality in mathematics and argue that each mathematical theory has equally real objects. But before taking up this argument, I want to say something more about intertheoretical mathematics.

3. Referencing theories

From what I have proposed, it is always preferable to identify all named objects of a theory which have the same extension with respect to the theory. We should neither pretend that our familiar names Mark Twain and Samuel Clemens distinguish separate writers; nor should we do topology with both spheres and cubes. The reason, in both cases, is that our two alleged objects behave identically in our theories, and that the respective identity relations, DNA testing and homeomorphisms, equate them. However, there is room for subtle errors in reasoning along these lines, particularly with respect to theories which explicitly reference different notions of equality which do not correspond with the identity relation in the theory itself.

Consider the following example from mathematical group theory:

> *Theorem: Suppose there exists a mathematical group G, which admits two subgroups H and K. Then the following statements are equivalent:*
>
> *(a) The product group HK is a subgroup of G.*
> *(b) $HK \subset KH$.*
> *(c) $HK = KH$.*

In light of the discussion in the last section, we may be tempted to eliminate talk of subsets and set equality and subgroups when the hypotheses of this theorem are met. After all, if these statements are equivalent, then any distinction among them seems like a redundant ontology. But this is a hasty conclusion. The objects of group theory are primitive elements and sets of primitive elements together with rules of multiplication, and the relations of group theory are homomorphic (product preserving) functions.

When there exists a homomorphism between two groups which also determines a one to one correspondence between them, the groups are in no way distinguishable *from within the theory* and are called isomorphic. Therefore, the only culprits for distinctions without a difference occur when elements are named twice or isomorphic groups are distinguished. There is theoretical content in the statement (b) ⊃ (c) because 'equivalence of statements', while an equivalence in the theory, is not an identity determined by group homomorphisms. It is, however, redundant to have the group of integers and the fundamental group of the topological circle both as objects in group theory. Because the integers, *qua group*, and the collection of closed paths in a circle, *qua group*, are identical: Both are infinite cyclic groups which cannot be further distinguished from within group theory.

Similarly, it is redundant to say that all bachelors are unmarried males. For instance, if I were to tell you that Isaac Newton was an unmarried male, you could sensibly reply "Why didn't you just say Newton was a bachelor." However, if it came to be known that at the turn of the sixteenth century, all and only European bachelors were scientists, and I told you that Isaac Newton was a bachelor who lived in Europe at the turn of the sixteenth century, you would sound rather odd replying, "Why didn't you just say he was a scientist?" The issue is the same as in the group theory example. Here, the identity relation is synonymy, whereas the statement that European bachelors at the turn of the sixteenth century were scientists, even if true, has theoretical content.

The idea is that a theory is redundant in its ontology only when it distinguishes objects with the same extension. But it is important that the extension of a theory depends on the theory. In particular, objects of a theory are extensionally identical if and only if they are related to precisely the same objects by precisely the same relations. When this is the case, the objects are indistinguishable from within the theory in every respect other than their having different names. Whether the objects are distinguishable or not in some background theory is not important, because there is no fact of the matter about a theory's ontology and extension apart from what the theory's relations admit. And even in the case of theories which reference truths in background theories, it is important to understand the theoretic content in its various equivalence statements. Reduction of such referencing theories, like mathematical group theory, in an attempt to identify distinctions which are theoretically "equivalent", where "equivalent" does not denote the identity relation of the theory, is impossible. Such a reduction would be truth preserving, but would not yield an autonomous theory, because the relations would not be preserved appropriately.

4. Supervenience

Imagine that I propose as a class of objects, chesspieces. I insists, that is, that there is a property among things in the world that is the property of being a chesspiece. And suppose that my theoretical treatment of the world, my beliefs and statements about the relations among objects, is such that objects called chesspieces fall under the scope of quantifiers in my theory—I am, in other words, ontologically committed to their being chesspieces in the world.

I doubt that anyone would challenge my ontology at this point: we all speak of chesspieces—those knights, bishops, rooks, pawns, kings and queens which we use to play chess—in a most serious way. So far, so good. But what if I go further; what if I insist that there is something more to being a chesspiece than being any one of a knight, bishop, rook, pawn, king or queen? Imagine that my commitment is to a property, that of being a chesspiece, that nothing has without being one of our familiar chessboard characters, which I insist is a serious addition to our metaphysics. Without the independent notion of being a chesspiece, I might claim, our theory is somehow inadequate.

The common response to such a claim, and the response which I wish to uphold, is that such a claim is empty. After all, in any theoretical statement about chesspieces, the predicate for being a chesspiece can easily enough be replaced with a disjunction of predicates for being a knight, a bishop, a rook, a pawn, a king and a queen. Since there is no way to be a chesspiece other than being one of these objects, such a replacement would result in no discrepancy in truth-value, autonomy, or meaning. The predicate for being a chesspiece is mere redundancy. There is no lost matter in the world we can salvage with such a term.

The best arguments to the contrary—that a chesspiece could be anything one uses in the play of a game of chess, a stone, a twig, a well defined computer state, or even a mental abstraction—are naïve at best. To whatever extent any of these things are chesspieces, and my readers are free to disagree wildly on this matter, they must to that very same extent be one of the six familiar pieces. If our project is an efficient ontology, we must write off the property of being a chesspiece as a convenience, and not as a novel metaphysical class.

For some time philosophers have been skeptical of there being any real existence of so-called disjunctive properties, like being a chesspiece. Somewhat more recent has been the application of this principle to the alleged phenomenon known as supervenience. The apparent conclusion of this argument, which I attribute to Jaegwon Kim, is that properties which supervene on some properties at a lower level, are effectively reducible to those lower level properties (Kim, 107). Much remains to be said about this argument, its details, and its applicability to mathematical equivalence classes, but first I should ground the subject in some definitions.

Recent literature is filled with several notions of supervenience as well as discussion about the strengths and scope of each, but for our present discussion we need only consider the following:

1. A set of properties T supervenes on a set of properties S if and only if for every property E in T, everything which has property E has some property P in S (Kim, 9).

And

2. A set of properties T supervenes on a set of properties S just in case for every x and y, if precisely the same properties P_i in S are true of x as they are of y, then precisely the same properties E_i in T are true of x as they are of y (Kim, 10).

Now let T be a theory which proposes some such property E. Supposing that the above conditions of supervenience are met, there remains the question whether there is any theoretic role for E which could not exist without its inclusion in the theory. In section 4, I will argue that there can be and that mathematical theories with their ascending orders of equivalence classes are such a case. But first I will discuss the contrary argument—that supervening properties do in fact reduce to disjunctive properties and are therefore unwelcome—and conclude the present section with a discussion of the supervenient nature of mathematical equivalence classes.

Returning to our definitions of supervenience, particularly the first, we see that the alleged property E cannot be met of any object unless that object meets some property P in S. A typical, though rather deep example of such a case is the mental phenomenon of being in pain. With the common assumption that there must be some physical state that someone is in before she can be in pain, we see that no one can be in pain without it being true that she is in some “subvenient” physical state. Another trivial example comes from the above discussion about chess. Everything which is a chesspiece is necessarily either a knight, bishop, rook, pawn, king or queen. In this sense, the properties of “being in pain” and of “being a chesspiece” supervene on a set of physical, neural states and a set of types of chesspieces, respectively. So as in our earlier discussion, there is no way for someone to be in pain other than being in one of a set of physical states, and there is no way for something to be a chesspiece other than being one of our six subvenient object-types. Supervenient properties are disjunctive properties, for if there is any way to satisfy the supervenient property other than by being one of a set of properties in some (possibly infinite) disjunct, then neither definition of supervenience is met (Kim, 107). Therefore, there is no theoretical content in supervenient properties—they are only substitutions for disjunctions in theories on the subvenient level.

Now I wish to turn to the supervenient nature of mathematical equivalence classes. The proof that a mathematical theory of equivalence classes supervenes on the lower level theories whose objects form the equivalence classes can be presented generally, without reference to any specific mathematics. I think it is important to present the result in this form, to emphasize the threat of a possible reduction of mathematics to a single, low level theory. Later, however, I will return to our consideration of topology to make the ideas more tangible.

Let T(O,R) be a theory whose objects (named by elements of O) are equivalence classes of root objects from other mathematical theories. (Here the equivalence is due to the relations R, not the relations in the various theories which properly host the root objects.) Suppose E is the name an object of O. E can be rewritten with the names all root objects which belong in E, a trivial result from the definition of an equivalence class. In other words, there is no way to be in E other than by being one of the root objects equivalent under the relations R. So we could replace all instances of E in our theory with the disjunction of the root objects in E, without any loss of information. E is yet another unfriendly disjunctive property, which has no place in our theory.

For those of us who believe that there is something more to the varying levels of mathematics than can be said in the respective lower levels, this can be a bit disheartening. Some theories in mathematics are not subject to this argument, because their objects are not equivalence classes of objects in any other theory. But there remain the numerous theories where the above argument definitely applies, and something more needs to be said if those theories are to be spared from reduction. What I will suggest in the next section is that there really is nothing more in many mathematical theories than can be said in their lower levels, but that even this is not enough for reduction.

5. Two types of reduction

I have, up to this point, been using the term reduction in two distinct ways. In section 1 I mentioned reduction from one theory to another when the two theories have what I have called an inscrutability of reference. This type of reduction is well defined by a proxy function which preserves extension (defined by the relations in each theory) and transitivity of relations. This will hereafter be called "categorical reduction" because the process shifts from one category of objects and relations to another. In section 3 I spoke of the reduction of supervenient theories to their subvenient theories. This was also well defined by the procedure of replacing all objects of the supervenient theory with the disjunction of all objects of the subvenient theory over which the former objects supervene. I shall refer to this as "diminutive reduction" because the effect is to eliminate certain disjunctive properties from the ontology of a theory. By examining the differences between categorical reduction and diminutive reduction I hope to expose a problem with the alleged "reduction" of mathematical theories.

Recall that in section 1 I proved the equivalence between theories $T_1(O_1,R_1)$, $T_2(O_2,R_2)$, and $T_3(O_3,R_3)$. The objects in T_1, O_1, were all geometric objects, and the relations, R_1, were continuos functions between these objects. O_2 was the set of equivalence classes of objects in O_1 (where objects were equivalent if they behaved in the same manner with respect to the relations R_1). The categorical reduction from T_1 to T_2 is the natural one: all geometrical objects with the same extension determined by the relations are identified, and only the extension appears in the reducing theory. O_3 was the set consisting of one geometrical object in each equivalence class in T_1. This is also a fairly natural categorical reduction. Returning to our example of authors, the categorical reduction from T_1 to T_2 is equivalent to dropping all names of authors in our theory and speaking only of the people named. The categorical reduction from T_1 to T_3 is equivalent to dropping all but one name for every author.

	Intension	Extension
T1	Mark Twain, Samuel Clemens	The author of *Huckleberry Finn*
T2	The author of Huckleberry Fin	The author of *Huckleberry Finn*
T3	Samuel Clemens	The author of *Huckleberry Finn*

Also recall my claim that T_2 was the mathematical theory topology, restricted to geometrical objects. In other words, T_2 is the theory about the topological structure of geometrical objects alone. However, T_2 is not to be confused with the entire theory, topology, which deals with many objects which are not the equivalence classes of any geometrical objects. Similarly, T_3 is not equivalent to topology, because if we took one specific space from each equivalence class in topology, it would not be possible to take a geometrical space from each equivalence class—precisely because some topological spaces are not equivalent to any geometric space. Therefore, the categorical reduction from section 1 does not suffice to show that topology is reducible to geometry—only that some subtheory of topology is. This subtheory could be constructed from topology by eliminating all but those equivalence classes which have a geometric member and eliminating all but those continuous functions needed to have closure relative to the new set of objects.

Yet in section 4 I showed that those mathematical theories whose objects are equivalence classes of objects from other theories did in fact supervene on lower level theories, and that the supervenient theories were therefore diminutively reducible. In the case of topology, this means that we may replace all topological objects (equivalence classes of spaces) with the disjunction of all spaces in the respective classes. Each of these spaces belongs to some mathematical theory subvenient to topology (for example, an infinite number of these disjunctions will contain geometrical spaces), and so topology is diminutively reducible to the conjunction of all the subvenient theories needed to provide one space in each disjunct. At first glance it seems that the failure of topology to reduce to geometry via categorical reduction was a merely superficial result—topology reduces after all, diminutively, but more than one subvenient theory is required for the reduction.

But even this is insufficient for any serious form of reduction. All that diminutive reduction demonstrates is that topological classes can be replaced with disjunctions of the spaces from each class without disrupting the theoretical content. This, of course, is just our familiar trick of replacing extensions with intensions. Such a result is unsurprising if we are at all convinced by the arguments for inscrutability of reference. But this diminutive reduction fails to eliminate topology because topology is more than a set of equivalence classes—it is a theory on an ordered pair of equivalence classes and continuous functions. More precisely, diminutive reduction fails to reduce because the relations from the supervenient theory are untouched. Our claim was that topology was diminutively reducible to the conjunction of all the subvenient theories needed to provide one space in each disjunct—but can we really say, without the "diminutive" modifier, that topology is therefore reducible? I believe that we cannot, because each subvenient theory suffers a significant handicap, namely closure. Indeed, from within any single subvenient theory, it is impossible to apply the relations of that theory in such a way as to arrive at objects outside of the theory. The supervening theory, topology, is required to provide the generalized relations which interrelate the various subvenient theories.

Moreover, topology is in no way a special example. The discussion in this section demonstrates that one theory is categorically reducible to another just in case the subset of the first theory which admits a categorical reduction is the whole theory, and similarly, a theory is diminutively reducible to another just in case the second theory provides at least one disjunct in the disjunctive replacement of each element of the first theory. Equivalently, an *entire* theory is categorically reducible to a theory if and only if it is diminutively reducible to a *single* theory. When this occurs, the reducing theory of the categorical reduction and the subvenient theory of the diminutive reduction agree up to relativity of reference, and we may actually say that the first theory reduces, without a modifier. But this is not the case with mathematical theories of equivalence classes, precisely because these theories arise to relate multiple theories which are themselves closed and relatively disjoint. The role of the higher level mathematical theory is to provide relations which make precise the similarities between the lower level theories. If only one lower level theory exists, then by using the higher level theory we do nothing more than ignore the details in the lower level theory which are not relevant in a particular situation, and reduction is indeed possible, though inconvenient.

In the next section I will examine a third type of reduction in mathematics which threatens those theories not drawn from classes of objects in other theories. This type of reduction has inspired the various foundations projects of the last century. By refuting the ontological claim that this reduction to foundations tells us anything new about the number and types of objects in mathematics, I hope to excise the philosophical significance of set-theoretic reduction from the metaphysical reduction with which it is often confused.

6. Set-theoretic reduction

In distinguishing the mathematician from the philosophical logician, R. M. Martin rightly emphasizes ontology. “The attention of the mathematician,” he says, “focuses primarily on mathematical structure. . . [while] the philosophical logician . . . is more sensitive to matters of ontology and will be especially interested in the kind or kinds of entities there are actually” (Benacceraf, 272). This distinction is perhaps more true today—perhaps sadly so. An experienced student would be surprised to find a professional mathematician conscientious enough to worry about the actuality of any of the “objects” she invokes in her seminars. The philosopher, on the other hand, concerns herself with the nature of mathematical objects. Her interests are not so much in proving theorems within a mathematical theory, but in understanding how mathematics is able to describe the physical and social systems in our world, in questioning how we can come to know about mathematical objects (whatever they may be), and in deciding if anything might be true of mathematical objects other than what is decidably demonstrable of them. For her purposes, understanding the ontology of mathematical objects is crucial. For only then can we begin to approach such problems.

Martin continues, “[the philosophical logician] will wish to ask whether the entity dealt with is *sui generis* or whether it is in some sense reducible to (or constructible in terms of) other, perhaps more fundamental entities” (Benacerraf, 272). Here one wonders if such reduction to fundamental entities is all Martin has in mind when he speaks of “matters of ontology.” If so, I worry about his program.

True, it is old news now that “in some sense” there is mathematical reduction to fundamentals. The natural numbers can be fully characterized in terms of sets, and so we can, if we are so inclined, accomplish all the results of number theory with sets. But this phenomenon should not convince us that numbers, the objects of number theory, are ontologically reducible to sets, the objects of set theory. For one thing, there is no one choice way in which number theory reduces to set theory. Should we decide that the number 1 should be the unit set of the null set, we are still at liberty to construct the natural numbers in various ways. We could proceed as Von Neumann prescribed, and define natural number as follows: *The set n is a natural number if and only if n is a member of every set z such that the empty set is a member of z and all successors of members of z are members of z*. In this case, the natural numbers would form an ordered sequence $\{\varnothing\},\{\varnothing,\{\varnothing\}\},\{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$, ... Or we could use Zermelo’s construction, where *The set n is a natural number if and only if it is a unit set and its single element is a natural* *number*. And we would have, for the “objects” of number theory, the ordered sequence $\{\varnothing\},\{\{\varnothing\}\},\{\{\{\varnothing\}\}\}, \ldots$. On the surface, there remains the question which of these constructions is correct, for if numbers are objects, then both $\{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$ and $\{\{\{\varnothing\}\}\}$ cannot be the object 3. If we were to suppose that they both were, then we could ask the question, “What is the cardinality of the number 3?” and answer it both with “3” and with “1” (or strictly speaking, since we have allegedly reduced numbers to sets, both with “$\{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$” and with “$\{\varnothing\}$”, or both with “$\{\{\{\varnothing\}\}\}$” and with “$\{\varnothing\}$”), which would be absurd. Paul Benacerraf famously recognized this as yet another (particularly vivid) case of inscrutability of reference. So long as we restrict our language to the terms of number theory, I might have in mind Zermelo’s construction and my friend might have in mind Von Neumann’s, but we would never realize this discrepancy (Benacerraf, 278).

Benacceraf interprets this result as reason to believe that numbers are not objects at all (Benacceraf, 290). Putnam is only marginally more generous when he says, “Call [numbers] ‘objects’ if you like (they *are* objects in the sense of being things one can quantify over); but remember that these objects have the special property that each fact about them is, in an equivalent formulation, simply a fact about *any* ω-sequence” (Putnam, 49). So much for the hope that a reduction of mathematics to some fundamental theory in turn reduces all mathematical objects to the objects of that theory. There is, indeed, a world of difference in saying that the number 3 *is* the set $\{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$ or the set $\{\{\{\varnothing\}\}\}$ and merely *identifying* the number 3 with one of these sets (Benacerraf, 289). That these sets are identifiable with the number 3 simply shows that ω-sequences are constructible from the objects of set theory—it does not suggest that any such sequence of sets *are* the natural numbers.

However influential the variety of ways in which the natural numbers can be ‘reduced’ to set theory, it is worth further note that the reduction of other mathematical entities to set theory admit even more constructions. The real numbers, for instance, have been constructed from the natural numbers in several ways, independently of how the natural numbers themselves are constructed. One can likely reduce the objects of deeper, high level theories (however tedious the process) in astonishingly many ways, which appear wholly dissimilar. Yet even if there were some theory which reduced, in this sense of reduction, in only one way to the theory of sets, I see no reason to say that the relevant sets *are* the objects of that theory. If sets are objects, they are mathematical objects just like the objects of those theories which reduce to them. Reducing the ontology of mathematics to the objects of set theory is metaphysical parsimony for its own sake. It leaves us with the same question as before, namely "What types of objects are mathematical objects?" Never mind that there are suddenly fewer of them.

Recall that in section 5 I argued that entire mathematical theories of equivalence classes never categorically reduce to a lower level theory because they are never diminutively reducible to any single lower level theory. Yet I have spoken in this section as if the reduction of number theory—indeed the whole of mathematics—to the single theory of sets is not only possible, but something we can achieve in many ways. We can clarify this issue by returning to the issue of autonomy. Again, mathematics is not simply built from objects. Mathematics consists of theories on ordered pairs of names objects and relations. Because of this, the "reduction" I have discussed in the present section (properly the construction of mathematical objects from objects of another theory) cannot take place without defining operations in the reducing theory to correspond to the relations in the theory being reduced. In reducing number theory to set theory, at least the successor operation *s(x)* must be defined. Depending on your choice between Zermelo's construction ($s(x) = \{x\}$) and Von Nuemann's ($s(x) = x \cup \{x\}$), this definition and the usual arithmetic operations defined from it can either be straightforward or fairly awkward—but in any case the operation is novel in the sense that it was not originally part of set theory.

What these constructions really show is that the pattern determined by the objects of number theory occur in set theory. Many philosophers call this occurrence *reduction* because it seems to show that we no longer have need for both numbers *and* sets—that we can, in fact, get by with sets alone. This metaphysical reduction aside, there are several mathematical benefits to noticing this embedding of patterns. After constructing the natural numbers from Von Nuemann's sets, for instance, both finite and transfinite induction can be proven from the well-foundedness of set membership (Resnick 216). Once the functions from analysis were 'reduced' to sets of ordered pairs, the concept of function was generalized to its present form, and analysts were able to construct space filling curves, wildly discontinuous functions, and continuous but nowhere differentiable mappings. But I see no reason to confuse this with reason to believe in a metaphysical reduction to sets.

Again, to say that the natural number pattern occurs in set theory is only to say that the theory of sets is rich enough in its ontology to allow an ω-sequence to be defined in it. It is number theory, however, and not some aspect of sets, which draws attention to—and in some sense creates—this sequence. The belief that any number of objects has been eliminated by reducing a mathematical theory to set theory in this way is tied in with the misunderstanding of mathematical objects as existing independently of the theories they are objects of. By selecting an ω-sequence from the iterative hierarchy of sets and by defining the number theoretic relations among the elements of that sequence, we create number theory and begin studying mathematics beyond what we study in the theory of sets. But in creating number theory, we also create the objects of number theory, as they are determined by the relations we defined. These objects *are* not the sets in the particular ω-sequence we began with, they simply are identifiable with those sets. So we can *do* number theory, functional analysis, algebra, topology, and geometry with only sets, but we need not, and our ability to work through our various mathematical theories with sets does not make those sets the objects of our theories.

I want to point out the stakes that I am raising with this argument. In particular, I want to emphasize the ontological equality among mathematical theories that I first suggested in section 2. When the theories of mathematics are seen as essential for their various objects to exist, I see no place to say that the objects of one theory are somehow more real than those of another. I might want to say that a geometric square, for instance, is more real than a topological projective plane because I can find approximate squares in nature or draw them on paper, while I cannot do the same with projective planes. Yet the squares I locate or sketch are not objects in geometry any more than sets are objects of number theory. Similarly, the objects in subvenient theories which make up the equivalence classes in other theories are no more real than those classes. The objects of the theory of equivalence classes are wholly determined by the relations of that theory. It is only with respect to these relations that the objects in the subvenient theory are equivalent and can each act like the same object of another theory[*] .

As enlightening as his insights in this matter are, I still do not subscribe to Benacerraf's eliminative agenda. Inscrutability of reference alone does not entail the nonexistence of mathematical objects any more than it does the nonexistence of other objects—it simply makes precise exactly where our objects cease to answer for themselves. Again, this is when we ask questions of the objects with which we attempt to pry deeper than the relationships between the objects. For what we know from these relationships is all we can know about the objects when we speak in sentences from the theory that they are objects of. I maintain that mathematical objects, numbers, groups, topological spaces, sets, and the rest,

---

[*] Resnick (214) uses these arguments to suggest that the natural number 2, the real number 2, and the complex number 2 are all different objects. I want to reiterate this result as an aside so long as I am pointing out the stakes of my position.

exist. By "exist" I do not mean Putnam's watered down existence, where to say that mathematical objects exist is merely to say that they are mathematically possible and that there are necessary truths of the form *if some object* x *satisfies the postulates of this theory, then the following theorems are also true of* x (Putnam, 49).

Putnam's motivation to grant mathematical objects a hypothetical existence stems from the same phenomena that threatened ontological reduction above, viz. some objects can *act like* mathematical objects. I have reviewed how sets can act like natural numbers: It is possible to construct a sequence of sets that satisfy the Dedekind-Peano postualtes, and therefore the theorems of number theory are true of these sequences. This is the same phenomenon that accounts for the application of mathematics to social and physical systems. When these systems satisfy the axioms of some mathematical theory, we can apply the theorems from that theory to predict or understand the systems. Another rich example of one class of objects acting like mathematical objects appears in the application of homology theory which opens this essay. In this case, the objects of one mathematical theory act like the objects from another. The eight Eilenberg-Steenrod axioms of homology theory suffice to derive all of the theorems of homology. The results of the theory are relatively easy to prove, often times involving simple "diagram chasing", where a few easily proven rules are applied to exact sequences of Abelian groups. After proving that topological spaces satisfy the Eilenberg-Steenrod axioms, several questions in topology are readily solved, using only the algebraic techniques of homology theory[*] . Rather than suggesting that there are no mathematical objects or that mathematics has only a subjunctive ontology, these examples should demonstrate how and why the real objects of mathematical theories are applicable, both to the concrete world around us and to other mathematical theories.

7. Conclusion

I began this essay with a list of questions which I find central in philosophy of mathematics: What kinds of objects are mathematical objects? What is mathematical truth—if we can never know whether some seemingly mathematical statements are true, are we even clear on the meaning of truth and falsity in a mathematical context? How do we come to know about mathematical objects, if they are not physical objects, and how do we learn so much about the physical from them?

---

[*] A detailed discussion of this can be found in Munkres' text, chapters 3 and 4.

In this essay I hope to have made some progress toward answering this first question, if only to demonstrate the essential role of the relations of a theory in determining the objects of that theory, and the dependence of mathematical objects on the structure of the theory they are a part of. I have discussed three types of reducibility, each of which threatens the existence of ‘higher level’ mathematical objects. But I have shown that these reductions either do not apply to mathematical theories or do not suffice to eliminate mathematical objects. Our mathematical theories are ontologically independent of one another, though the objects of one theory may “act like” the objects of another. Understanding the objects of mathematics as real objects, inseparable from the relations among them, we can maintain the several irreducible theories that comprise the discipline. I hope that by establishing this groundwork we might shift attention from eliminating the deep questions of mathematical philosophy to answering them.

References

Benacerraf, Paul. 1983. “What Numbers Could Not Be” from *Philosophy of Mathematics: Selected Readings*. 2nd ed. Benacerraf and Putnam, editors. Cambridge University Press. New York.

Goodman, Nelson. 1979. “The New Riddle of Induction” from *Fact Fiction and Forecast*. 4th ed. Harvard University Press. Cambridge.

Kim, Jaegwon. 1998. *Mind in a Physical World*. MIT Press. Cambridge.

Munkres, James R. 1984. *Elements of Algebraic Topology*. Perseus Books. Reading, Massachusetts.

Putnam, Hilary. 1995. “Mathematics Without Foundations” from *Mathematics Matter and Method: Philosophical Papers vol. 1*. 2nd ed. Cambridge University Press. New York.

Quine, W. V. [OR]. “Ontological Relativity” from *Ontological Relativity and Other Essays*. Columbia University Press. New York.

Quine, W. V. 1981. “Things and Their Place in Theories” from *Theories and Things*. Harvard University Press. Cambridge.

Resnick, Michael. 1997. *Mathematics as a Science of Patterns*. Oxford University Press. New York.

II. Incompleteness and Mathematical Truth

1. Introduction

In this essay I am primarily concerned with the still unresolved conflict between mathematical epistemology and mathematical truth. I focus on the impact that incompleteness of mathematical theories has on a realist account of mathematical truth, where the truth conditions for quantified expressions of mathematical language are the same as those for quantified expressions in natural language and physical theories. Simultaneously I maintain that this realist account is the proper one and that whatever impact incompleteness has on such an account should be taken seriously. But from within a realist account for mathematical truth, taking seriously the difficulties of mathematical incompleteness means reassessing the epistemology with respect to which the truth conditions of mathematical statements are undecidable, not reconsidering which mathematical statements are properly meaningful. Throughout this essay I use the continuum hypothesis as the model undecidable mathematical statement. Toward the end I suggest that the continuum hypothesis is false, although our present formalization of set theory is incomplete with respect to its truth value. My hope is that my approach to the continuum hypothesis can serve as a paradigm for reshaping formalized mathematics in harmony with a realist account of mathematical truth.

My argument is of the following form: After examining the standard approach toward evaluating the truth values of quantified expressions within a theory, I demonstrate that certain mathematical expressions are analogous to those from theories of natural language and should therefore be subject to the same ‘standard approach’ used for natural language expressions. In particular, I show that not all undecidable mathematical statements are statements with respect to which we call a mathematical theory incomplete, and that among undecidable statements only those which do merit mathematical incompleteness should nevertheless have a truth value within a realist semantic account of truth.

One thing I do not suggest is that we should decide mathematical truths in any way other than by considering what we can and cannot derive in formal axiomatic systems. As I explain later, this does not mean that we require theoremhood in any one system before we can be sure of any mathematical truth. Rather, we should be attentive to the scope of individual systems and judge from their successes and failures how to form new systems, either to replace the old ones or to be used simultaneously. Since nearly all formal systems of interest to mathematicians are incomplete with respect to some mathematical statements, I argue that the task we face when constructing such formal systems is to establish axioms which can be used to evaluate mathematical propositions which are of interest prior to formalization. When a formal system whose theorems are both consistent and useful in determining the truth values of many mathematical propositions leaves other intuitively meaningful propositions undecidable, this alone should not cast doubt on the meaningfulness or importance of those propositions. When such a system somehow explains away the meaningfulness of such propositions, we are better off abandoning our intuitions and sticking with the formal system. But when, as is the present case with mathematical set-theory, a system is incomplete with respect to propositions which continue to be of interest (specifically the continuum hypothesis), our conclusion should be that the system is not characteristic of the theory which we intended it to formalize.

Mathematicians have provided us a history of examples of the cases I describe above. Prior to the modern formal definition of a function, it was common for mathematicians to question what sorts of functions could exist. At the time, mathematical functions were seen as abstractions from physical processes, and mathematicians not only thought that questions like “In how many locations can a continuous function fail to be differentiable?” and “What is the maximum amount of deviation that a function’s values can exhibit in a small interval?” seemed meaningful, they focused a great amount of attention on them. But with a precise definition of ‘function’, mathematicians shed these intuitions in favor of a more useful and general notion, despite the initial implausibility of discontinuous functions and functions which were continuous but nowhere differentiable.

On the other hand, strong intuitions have sometimes inspired the birth of new theories. From the axioms of Euclidean geometry, one could not prove that a line between two points on a plane—one inscribed in a circle on the same plane and the other lying outside that circle—must intersect that circle in at least one other point. Yet the pre-formalized notions of points, lines, and circles that we share all but demand that this be true, not only for lines but for continuous curves between points. Of course no mathematician ever cited the fact that Euclidean geometry was incomplete with respect to this question as reason to suppose that it was the wrong sort of question to ask. In fact, the lack of a formal proof of such a basic intuition motivated much of the foundation and practice in manifold topology, eventually leading to the Jordan Curve Theorem, a strong version of this very fact.

My purpose in the remainder of this essay is to explain the relationship between intuition and formal mathematics and to demonstrate how this provides us with a way of looking at mathematical incompleteness from a realist perspective. Later I suggest that the continuum hypothesis poses a question which is not explained away by its formal undecidability in set theory, but one which should motivate a reconsideration of mathematical foundations which will ultimately decide its truth value.

2. The realist account of truth

Consider the following sentence:

(i) There are exactly three rooms in Lawson's apartment.

The standard evaluation of the truth value of this sentence is that it is true if and only if there are exactly three things which are rooms and which are in Lawson's apartment. In other words, (i) is of the form,

(ii) There are exactly three A's that bear R to f.

"There are exactly three" is a numerical quantifier, which we can rewrite with existential quantifiers, A is to be replaced with a one place predicate, R with a two place predicate, and f with the name of an element from our universe of discourse. In the usual formal language QL (ii) is written as,

(ii') $(\exists x)(\exists y)(\exists z)(x\neq y \ \& \ x\neq z \ \& \ y\neq z \ \& \ Ax \ \& \ Ay \ \& \ Az \ \& \ Rxf \ \& \ Ryf \ \& \ Rzf$

$\& \ (\forall w)((w\neq x \ \& \ w\neq y \ \& \ w\neq z \ \& \ Aw) \supset {\sim}Rwf))$.

The standard evaluation, then, is that (i) is true if and only if when we replace f with 'Lawson's apartment' there are three distinct things (x, y, and z) which satisfy the predicate A (is a room) and the relation R (the first argument *is in* the second) to f (Lawson's apartment), and anything which is not identical to x, y, or z but nonetheless satisfies the predicate A (is a room) does not bear R to f (is not in Lawson's apartment).

There is no question as to whether (i) has a truth value so long as all our predicates are well-defined. We need only check to see whether it is true or false according to the properties of the things we are speaking about. The question I want to consider is whether mathematical statements which appear to have the same logical structure as (i) are subject to the same evaluation[1]. Consider the following sentence:

(iii) There are exactly three real solutions to von Neumann's function.

And define von Neumann's function as follows: Let $f : [-1,2] \rightarrow R$ be the function satisfying the following conditions.

---

[1] This is the central question Paul Benacerraf poses in his 1973 essay "Mathematical Truth". I have adopted the general form of his argument in the opening of this section, but I concentrate on only a subset of the issues he points out as relevant to mathematical truth.

$f(-1) = 0$;
$f(0) = -1$;
$f(1/3) = -\sum_{\infty} (e_{2n})/(2^{n})$;
$f(2/3) = \sum_{\infty} (e_{2n-1})/(2^{n-1})$;
$f(1) = 1$;
$f(2) = 0$;
f is linear between -1 and 0, between 0 and 1/3, between 1/3 and 2/3, between 2/3 and 1, and between 1 and 2;
$e_n = 0$ if 2n is the sum of two primes, and $e_n = 1$ otherwise[2].

Within the realist account of mathematical truth (iii) is of the form (ii) and an evaluation of the truth conditions of (iii) is analogous to the evaluation we gave to (i). We need only check whether there are three distinct things (x, y, and z) which are real numbers and which are solutions of von Neumann's function (i. e. $f(x) = f(y) = f(z) = 0$ and for every $r \in R$, if $r \notin \{x, y, z\}$, then $f(r) \neq 0$). This example presents two problems for the realist account, however, which I examine below.

Von Neumann's function is continuous and clearly has two solutions, -1 and 2, which are explicitly denoted as such in its definition. Furthermore, since $f(0) < 0$ and $f(1) > 0$, there must be some real number s between 0 and 1 for which $f(s) = 0$, by the intermediate value theorem. However, determining s depends on evaluating the expressions for f(1/3) and f(2/3), which depend on Goldbach's famous conjecture: every even number is the sum of two primes. Now if Goldbach's conjecture is true, then f(1/3) and f(2/3) each equal zero, so every real number between 1/3 and 2/3 is a solution to von Neumann's function. But if Goldbach's conjecture is false, $f(1/3) < 0$ and $f(2/3) > 0$, so there is a single solution s between 1/3 and 2/3. To date Goldbach's conjecture has neither been proved nor disproved, so we are unable to decide whether (iii) is true or false. The first problem this example poses for the realist account of mathematics, then, is that while (iii) appears to be of the form (ii), we can not presently evaluate its truth. The second problem is that while a proof of Goldbach's conjecture would suffice for us to say that (iii) is false—for there would then be uncountably many solutions to von Neumann's function—a

---

disproof would suggest to us that (iii) is true, and yet we would still be unable to determine what the third solution to von Neumann's function is.

Historically, these problems with statements like (iii) have posed more of a threat to the realist account of mathematics than they presently do. The first problem originates from the fact that the truth value of (iii) depends on the truth value of a presently undecided mathematical statement. So to maintain that we can evaluate (iii) according to its general logical form (ii) we necessarily must believe that Goldbach's conjecture, while today undecided in mathematics, nevertheless is either true or false. There is no way that Goldbach's conjecture could be undecidable, however, because a proof of its undecidability would have to show that there is no even number which is not the sum of two prime numbers (for otherwise Goldbach's conjecture would be false). This being the case, we might have proven the conjecture to be undecidable in that particular formal system, but we would know from a meta-theory that it was actually true. So Goldbach's conjecture must be either provably true or provably false, we simply have not discovered which. The first threat to the realist account disappears because there is no problem with leaving the truth value of a sentence dependent on a statement which has not been proven, so long as we know that the statement can be either proved or disproved. Goldbach's conjecture is *undecided*, but it is not *undecidable*, so there remains for us no worry about whether it is true or false other than that we do not yet know which. The only way to disagree with this is to say that mathematical statements have truth values only after they have been proved or disproved. But this objection rests on a confusion of truth and knowledge. To say that something is neither true nor false unless we know which is indeed a contradiction to the realist account of truth, but it is not one I take seriously.

The second problem in evaluating (iii) according to the standard account is that should Goldbach's conjecture be false, we would say that (iii) is true despite not knowing what all of the solutions to von Neumann's function are. Beginning with the Dutch mathematician Brouwer, the Intuitionists believe that the notion of mathematical existence without constructability is incoherent, and consequently, that sentences like (iii) are true only if we can say what the solutions are. I do not want to worry about the principles of Intuitionism in this essay any more than to say that the centrality of mathematical construction strikes me as largely uncharacteristic of mathematical existence. In our everyday discourses as

---

[2] The function von Neumann actually defines on page 62 of his essay "The Formalist Foundations of Mathematics" is the same as this one restricted to the interval [0, 1]. A more detailed discussion of the relationship between this function and intuitionism appears on page 62-3.

well as in our theories of physical science, we commonly arrive at the existence of things of which we have only indirect evidence, but which must exist in order for the rest of the theory they are a part of to make sense. The situation in mathematics is similar. If from an intermediate value theorem we know that some real number in a given range must be a solution to a function despite our inability to construct such a real number, I see no more reason to abandon the theorem and its consequences than I do some physical principle which allows us to know that a certain number of stars must exist in a galaxy without us having any direct evidence of those stars.

So far we have seen evidence that mathematical expressions can and should be treated according to the standard account of truth, and that an extension of realism as a general theory of truth is not threatened by the mathematical nature of such expressions. But the story is not over for two reasons. First there is the question of how the truth condition of a mathematical proposition could be defined in terms of its derivability in a formal system. A second problem is that some mathematical propositions are not simply *undecided* at the present time, like Goldbach's conjecture, but are actually *undecidable*—and provably so. As we will see, these two problems are related because Godel's incompleteness theorems show that every consistent formal system strong enough to contain arithmetic will contain undecidable statements. How we can take a realist approach to mathematical truth and simultaneously define the truth conditions of mathematical propositions in terms of derivability in formal systems is very problematic.

3. Incompleteness

On face value, the idea that the truth of a proposition could simply *be* proof of that proposition in some formal system is wrong. After all, formal systems could turn out to be inconsistent, and so the derivability of propositions in those systems would be trivial. Given any inconsistent formal system S, the fact that every proposition in the theory T that S is an attempt to formalize is true immediately demonstrates that 'derivability in S' could not be the same thing as 'truth in T'. If we take the standard route and understand 'derivability in S' to be 'truth in S', then there arises a similar disanalogy between 'truth in S' and 'truth in T'. If mathematical truth is therefore to be truth in some formal system, and truth in a formal system is simply derivability in that system, then something more is needed to distinguish which formal systems are allowable in this scheme.

Certainly, consistency is one such a criterion, but it alone cannot suffice for two reasons. First, as Paul Benacerraf points out, to think that consistency could be the sole distinction between those formal systems which generate mathematical truth from their derivations is to mistake "the significance of the fact that *in*consistency is *proof* that truth has not been attained" (Benacerraf 419). Suppose I propose a formal system for set-theory, say S_1, and later discover that it is inconsistent. I then propose a second formal system S_2, which also turns out to be inconsistent. After a series of thirty proposed systems, S_1, S_2, …, S_{30}, each turn out to be inconsistent, I arrive at S_{31}, which turns out to be consistent. Given the number of times that I thought I had produced a system which properly formalized set-theory, there seems to be little reason to suppose that with S_{31}, simply because it is consistent, I have finally arrived at a proper characterization of set theory so that derivation within S_{31} is the defining property of set-theoretic truth. Secondly, supposing that I have arrived at a consistent formal system in S_{31}, if it is sufficiently strong to contain all the statements of arithmetic, then it is necessarily incomplete. If we maintain realism so far as to say that all mathematical propositions are either true or false, neither derivability within S_{31} nor within any other consistent formalization of set theory could be the sole truth condition of set-theoretic propositions.

I belief that the above argument is in fact conclusive that derivability within a formal system cannot be the defining characteristic of mathematical truth. For to say that it is, one would necessarily have to give up any hope of extending the realist account of truth to mathematics, if for no other reason than because all undecidable statements within whatever formal system one is using would neither be true nor false. But while I believe that derivability within *any* formal system is not the defining characteristic of mathematical truth, I do not believe that any mathematical statement can be decided except by considering what can and cannot be derived within *some* formal system or systems. To make this claim more clear, I want first to be precise about what I mean by 'mathematical statement' and by 'incompleteness'.

Consider the following sentences:

(iv) All prime numbers are orange.
(v) Every even number is the sum of two prime numbers. (Goldbach's conjecture)
(vi) The greatest prime number is odd.
(vii) The elements of every infinite subset of real numbers can be put in one to one correspondence with either the elements of the set of natural numbers or with the elements of the whole set of real numbers. (continuum hypothesis)
(viii) The number of trees in Houston is odd.
(ix) The number of people on earth at this moment is odd.
(x) Hamlet's blood-type is A^-.

For now, the formal system under consideration will be the Zermelo-Fraenkel axioms of set theory (ZF), with the natural numbers, the real numbers, and arithmetic defined in ZF in some way or another[3]. To understand what incompleteness in mathematical theories is, I think it suffices to understand why only (ix) is a statement with respect to which we say that ZF is incomplete.

(iv) is certainly undecidable in ZF, in so far as from within ZF neither its truth or its falsity can ever be proven. But we do not say that ZF is incomplete with respect to (iv) because (iv) is not a statement expressible in the language of ZF, precisely because there is no predicate in ZF for 'orange'. Of course having a realist account of mathematical truth does not obligate anyone to be able to believe that (iv) is either true or false any more than having a realist account of truth in natural language requires us to believe that sentences like 'Three of the rooms in Lawson's house are brave' are either true or false. Realist accounts of truth only require that meaningful statements, whose predicates govern only names of objects over which the predicates are defined, have a determinate truth value.

Unlike (iv), (v) and (vi) seem to be mathematical statements, but still neither is a statement with respect to which ZF is incomplete. We saw above that while (v) is both meaningful and undecided, it cannot be undecidable. So not only does the mathematical realist believe that (v) is either true or false, it is provably either false in ZF or not false in ZF, we simply have not yet arrived at the proof. And if (v) is provably not false, then we can infer that it is in fact true. On the other hand (vi) does seem undecidable: we will never be able to determine whether the greatest prime number is odd or not because there is no greatest prime number. It is conceivable that prior to the proof that there was an infinite number of prime numbers, (vi) was believed to be true. Indeed, it was incoherent to believe that the number of primes was finite without also believing (vi), because of all the prime numbers, only 2 is even. But of course having a realist approach to mathematical truth does not commit anyone to believing that (vi) has a determinate truth value either. Whatever our prior conception of the truth of (vi) might have been, we know from a theorem provable in ZF that there is no greatest prime number. So depending on how we interpret (vi), it might not be a sentence that can be written in ZF after all[4].

---

[3] There are several known ways to identify the natural numbers and real numbers with sets in formalized set-theory. The identification I use later in the essay is von Neumann's.

[4] Alternately, we could interpret (vi) according to Russell's theory of descriptions. Rewritten as *There is a natural number x which is prime and which is greater than all other prime numbers and which is odd,*

(vii) is a form of the continuum hypothesis (CH) and is a sentence with respect to which ZF is incomplete. The undecidability of ZF is provable by constructing models which are consistent with ZF and in which the continuum hypothesis is either true or false. The first model of ZF in which CH was true was given by Kurt Godel, and the first one in which CH was false was given by Paul Cohen[5]. The two facts that (vii) is a coherent mathematical statement expressible in ZF and that (vii) is true in some models of ZF and false in others are what make CH a particularly vivid obstacle in bridging the gap between mathematical realism and the derivability criterion for mathematical truth.

Even before Cohen completed the proof that the continuum hypothesis was undecidable by the axioms of set theory, the result was expected and people had begun to try to make sense of its undecidability. At times Godel suggested that the continuum hypothesis was true (Maddy, 134), but he later believed that it was false. In any case, he always maintained that it was either true or false:

> It is to be noted … that … a proof of the undecidability of Cantor's conjecture from the accepted axioms of set theory … would by no means solve the problem. For if the meanings of the primitive terms of set theory … are accepted as sound it follows that the set-theoretical concepts and theorems describe some well-determined reality, in which Cantor's conjecture [the continuum hypothesis] must be either true or false. Hence its undecidability from the axioms being assumed today [ZF + the axiom of choice] can only mean that these axioms do not contain a complete description of that reality (Godel 1947, 476).

I agree with Godel on this point, except that I do not subscribe to the Platonism latent in Godel's belief in a well-determined mathematical reality. What I do believe is that it is our notion of set which makes the continuum hypothesis meaningful and which assigns to it a truth value. In my terminology, then, Godel's claim would be that the system ZF, or ZF+C, does not fully characterize our intended set theory.

I believe that Godel's realism places the continuum hypothesis in a similar position with sentence (ix), provided we first stipulate more rigid definitions for the natural language predicates in (ix). Once we circumvent the vagueness in human existence by defining precise stages of development and consciousness that someone must satisfy in order to be "a person living on earth at this moment", (ix) will have some definite truth value which we are unable to determine. I call the response to incompleteness that treats undecidable mathematical statements like (vii) as analogous to (ix) in this way 'the poverty of methods' response. As I mentioned above, it is possible to be a realist about mathematics and also to believe that

---

the sentence is false. In any case, we would not say that this sentence is one with respect to which number theory is incomplete.

[5] The original publication of Godel's proof is in his 1940, Cohen's appears in his 1966. A good discussion of both results appears in Tiles pages 175-191.

there is no way to determine a mathematical truth apart from deriving its corresponding sentence within some formal system. The reason for this is that the axioms of a formal system could be sound, so that they and their consequences are each true of the mathematical theories that they are meant to characterize, without being complete, so that not everything that is determined as true by our concepts is a consequence of the axioms of any one formal system.

The analogy with (ix) can be extended as follows: I could come up with verifiable methods for determining many things about the people living on earth at any given moment. I can, for instance, provide a reliable report of the number of people living in this room at this instance, and I can furthermore say whether that number is odd or not. The fact that my method (visual surveillance) cannot be used to determine whether the number of people living on Earth at this moment is odd or even does not make it any less reliable for what it *can* determine. It is my present belief that ZF, though incomplete, is sound, but that ZF+C is neither complete nor sound. In the next section I say more about the possibility of extending our formalization of set theory so as to preserve soundness and approach an answer to the continuum hypothesis, but first I will touch on some of the alternative responses to such an incompleteness proof.

Another possible response is to say that (vii) is more like (x). I call this the 'fictionalist response' because it is an attempt to expose our intuitions about the truth of mathematical statements as mistaken assumptions. Like Godel's Platonism, I believe that this response is inspired by belief in a well-determined reality, only one determined by the formal system at hand rather than by our pre-formalized mathematical theories and their primitive terms. If I made the claim, "Of course Hamlet has some blood-type. Everyone has a blood-type," I think it would be a reasonable response to say that Hamlet has no blood-type because Hamlet is only a character in a play. Since Shakespeare never mentioned a blood-type for Hamlet, and since Shakespeare could not possibly have had a blood-type in mind when he wrote *Hamlet*, it seems correct to say that Hamlet exists only as a character in the reality that Shakespeare created, and that in that reality no one has a blood-type. The fictionalist response seems to me more applicable to sentences like (iv). For while someone may suppose that the prime numbers either all are or are not all orange, because everything either is or is not orange, the reasonable response would be that numbers are not like other things in that respect. For while it is true that all medium sized physical objects are either orange or not, numbers are not objects of this type and the distinction does not apply to them (Maddy 129).

A third response is to say that (vii) should be treated more like (viii). I call this the conventionalist response. Despite the differences in his approach to mathematical philosophy and those approaches which are usually called conventionalism, I think that this is the type of response Hilary Putnam gives in his essay "Models and Reality". Putnam attends directly to the continuum hypothesis very rarely, but he is nonetheless concerned with the model for ZF that Godel used to show that CH is consistent with ZF. This model is the universe of constructible sets and is denoted V=L. According to Putnam the primary question which arises in instances of provable undecidability has been to ask which of the consistent models is the correct one. For if V=L is the intended model, then CH is true, and if Cohen's model which is rich in generic nonconstructible sets is the intended model, then CH is false. The analogy with (viii) arises because the question of whether the number of trees in Houston is odd or even depends on the total number of trees in Houston, which is not determined by the conventions established in the Houston 'city limits'. Given the conventions of the Houston city limits, they can be made more precise by arbitrarily selecting which trees near the edges of the city are within the limits and which are not. So we could construct precisifications of the Houston city limits in which the number of trees could be either even or odd.

There is no reason to say that either model—in the set theoretic case or in the case with trees in Houston—is the intended model if we admit that only "theoretical constraints" (that a model must be consistent with the axioms of the formal system) and "operational constraints" (that the extensions of all the predicates in the theory are preserved) determine which model is the intended one (Putnam 428-29). For both V=L and the model Cohen used to show that CH could be false without contradicting ZF, meet all these constraints. There is nothing new in saying that ZF does not determine its models. What Putnam brings into focus is the notion that no formal system can capture enough of our intuitive notion of set to determine that its models must all be uncountable (i. e. they must be of such a size that their members cannot be put into one to one correspondence with the set of natural numbers). This is a consequence of the Lowenheim-Skolem theorem, which says that any first order theory which is true on an uncountable model is also true on a countable model (a model whose elements can be put into one to one correspondence with the set of natural numbers). What Putnam leaves unattended are his reasons for believing that this is reason to suggest that no formal system could be sound with respect to our concept of set and satisfy the truth value of CH. If we take the Zermelo-Fraenkel axioms as fixed, as Putnam seems to do, then his argument does imply that settling the continuum hypothesis would be a matter of selecting one among many models which satisfy the axioms. His reasons for saying that this is inappropriate are convincing, but any reasons for thinking that the Zermelo-Fraenkel axioms are fixed are not. I examine the details of this claim below.

4. The concept of set and the continuum hypothesis

Central to any discourse on set-theory is a discussion of what we mean by a set. Cantor arrived at his distinctions between the different 'sizes' of infinite sets prior to the development of any formal system, and even today, the Zermelo-Fraenkel axioms are justified against our notion of set (Boolos 495-502, Tiles 124-134). It is this notion of set, rather than any particular formal system, which is primary. Consequently, it is this notion of set which we use to determine whether a formal system is sound or even worthy to be called a formalization of set-theory. Here I repeat Boolos' description of the notion:

> A set is any collection that is formed at some stage of the following process: Begin with individuals (if there are any). … At stage zero … form all possible collections of individuals. … [W]e assume that one of the collections formed at stage zero is the collection of all individuals, however many of them there may be.
>
> At stage one, form all possible collections of individuals and sets formed at stage zero. At stage two, form all possible collections of individuals, sets formed at stage zero, and sets formed at stage one. At stage three form all possible collections of individuals and sets formed at stages zero, one, and two. … Keep going in this way, at each stage forming all possible collections of individuals and sets formed at earlier stages.
>
> Immediately after all of stages zero, one, two, three, …, there is a stage; call it stage omega. At stage omega, form all possible collections of individuals [and sets] formed at stages zero, one, two, …. One of these collections will be the set of *all* sets formed at stages zero, one, two, ….
>
> After stage omega, there is a stage omega plus one … [of] all possible collections of individuals and sets formed at stages zero, one, two, …, and omega.

Similarly, construct sets at stages omega plus two, omega plus three, etc. After all of these, go to stage 2*omega, which contains all collections of individuals and sets at lower stages. Then 2*omega plus one, 2*omega plus two, …, 3*omega, 3*omega plus one, …, 4*omega, …, 5*omega, … omega*omega, etc (Boolos 491-2).

The Zermelo-Fraenkel axioms were developed to translate this notion of set into a formal system. Each of the axioms in ZF appear sound with respect to this notion, but since this cumulative hierarchy is strong enough to allow arithmetic to be defined within it, we know from Godel's incompleteness theorems that if ZF is consistent, then it cannot be complete. And of course, we have seen that it is incomplete with respect to CH, which I will restate now in more purely set-theoretic terms.

At each stage of set formation outlined above, there arise sets which appear at no earlier stage, any one of which is identifiable with the natural number corresponding with the stage at which it appears. The progression von Neumann gave is as follows: $0=\varnothing$, $1=\{\varnothing\}$, $2=\{\varnothing,\{\varnothing\}\}$, $3=\{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$, ... The set $\{0, 1, 2, 3, \ldots\}$ and every set which can be put in one to one correspondence with this set are said to be of size $\aleph_0$. This is the ‘smallest’ infinite set, when set size is defined in terms of the existence of one to one correspondences. In the same terms, the smallest sets which are larger than the sets of size $\aleph_0$ would be the same size as the set of all countable ordinals (an example of such a set would be a set with one element from each stage of the cumulative structure up to stage n*omega). These sets are said to be of size $\aleph_1$. One way to state the continuum hypothesis is to say that *the set of all subsets (or the power set) of a set of size* $\aleph_0$ *is of size* $\aleph_1$. This is equivalent to the statement given in (vii) above because it is provable that the real numbers, constructed from Dedekind cuts, can be put in a one to one correspondence with the power set of the natural numbers (Maddy 131). So to say that every infinite subset of the real numbers can be put in a one to one correspondence either with the set of natural numbers or with the set of real numbers implies that there is no set of a size in between $\aleph_0$ and the size of the power set of a set of size $\aleph_0$. Therefore, the size of the power set of a set of size $\aleph_0$ is $\aleph_1$.

Regardless of whether CH is true, I want to urge that the fact that it is provably undecidable in ZF should not suggest to us that it is somehow meaningless or lacking a truth value. To say that CH is neither true nor false is to say that the power set of a set of size $\aleph_0$ is not comparable with a set of size $\aleph_1$ with one to one correspondences. But I believe that either such a correspondence between elements of these sets exists or does not exist, and I do not think that we can abandon this intuition and still preserve the iterative concept of set that we are trying to formalize in the first place. What I do think the undecidability of CH suggests is that the notion of the power set of an infinite set is not precise in ZF. For in some models consistent with ZF we can prove that there is a one to one correspondence between the elements of the power set of a set of size $\aleph_0$ and the elements of a set of size $\aleph_1$ while in other models we can prove that no such correspondence exists.

Following Cohen, I believe that a proper precisification of the nature of power sets of infinite sets would likely show that CH is false because at no time in the ‘piecemeal’ progression through the cardinal numbers could we ever arrive at a set with the number of elements like what there must be in such a power set (Cohen 151). Penelope Maddy discusses this point in more detail, exposing how the power set axiom in ZF is intuitively stronger than the axioms used to proceed along the series $\aleph_0, \aleph_1, \aleph_2, \ldots, \aleph_\omega, \ldots$ (Maddy 130-131). F. R. Drake agrees with this notion, also with an appeal to intuitions of the cumulative hierarchy of sets. His results show that the continuum hypothesis depends on the relationship between the numbers of different types of orderings of infinite sets, which themselves seem intuitively comparable (Tiles 193-194).

Rather than dismissing an interesting and meaningful question in set theory in light of the relativity of models that arises from incompleteness, I think we should preserve realism about the continuum hypothesis and seek new axioms which correspond with our notions from pre-formalized set-theory. As I have suggested with the poverty of methods response to mathematical incompleteness, the resulting formal system will also be incomplete. But motivated by the specific task of establishing an axiom which is intuitively sound and from which we can derive a more precise notion of what subsets of infinite sets exist and what they are like, we might very well be able to settle CH by proving or disproving it within a formal theory[6].

If I have not provided a convincing argument for a realist approach to mathematical truth, I believe I have at least explained how one can be a realist about mathematics despite the threats posed by incompleteness *even while believing that mathematical truths can be known only by their derivation within some formal system whose foundations are sound with respect to our mathematical theories*. How to justify formal proof (opposed to reference and denotation) as a reasonable epistemological foundation for knowledge within a realist account of truth is yet another concern, but one which I believe can be satisfied. The incompleteness of formal systems alone does not put mathematical realism in serious jeopardy, however—it only shows that if the realist account of mathematical truth is correct, then no single formal system can be completely characteristic of a mathematical theory. I believe that a suitable project for philosophers of mathematics today is to take seriously their interests in mathematical statements, like the continuum hypothesis, and, rather than giving up all hopes of settling the truths of such statements in light of their

---

[6] One such axiom is the axiom of determinacy, first proposed by Mycielski. This axiom says a lot about what sorts of subsets exist. In particular, from it one can prove that all subsets of the real numbers are Lebesgue measurable (thereby contradicting the full axiom of choice and avoiding the very counterintuitive

undecidability within some particular formal system, to extend formal systems with new axioms which might be used to prove previously undecidable statements.

Banach-Tarski paradox), that the axiom of choice holds for countable collections, and several other results related to large cardinals.

References

Tiles, Mary. 1989. *The Philosophy of Set Theory: An Introduction to Cantor's Paradise*. Basil Blackwell. New York.

Maddy, Penelope. 1990. *Realsim in Mathematics*. Oxford University Press. New York.

Jech, Thomas. 1973. *The Axiom of Choice*. American Elsevier Publishing Company, Inc. New York.

von Neumann, John. "The Formalist Foundation of Mathematics" in *Philosophy of Mathematics: Selected Readings*, second edition. Ed: Paul Benacerraf and Hilary Putnam. Cambridge University Press. 1983.

Benacerraf, Paul. 1973. "Mathematical Truth" reprinted in *Philosophy of Mathematics: Selected Readings*, second edition. Ed: Paul Benacerraf and Hilary Putnam. Cambridge University Press. 1983.

Putnam, Hilary. 1977. "Models and Reality" reprinted in *Philosophy of Mathematics: Selected Readings*, second edition. Ed: Paul Benacerraf and Hilary Putnam. Cambridge University Press. 1983.

Boolos, George. 1971. "The Iterative Conception of Set" reprinted in *Philosophy of Mathematics: Selected Readings*, second edition. Ed: Paul Benacerraf and Hilary Putnam. Cambridge University Press. 1983.

Godel, Kurt. 1940. *The Consistency of the Axiom of Choice and the Generalized Continuum-hypothesis with the Axioms of Set Theory. Annals of Mathematics Studies*, no. 3. Princeton University Press. Princeton.

Cohen, Paul. 1966. *Set Theory and the Continuum Hypothesis*. W. A. Benjamin. New York.

III. Foundations of Set Theory

1. Introduction

This final installment in my examination of the relationship between truth and the mathematical theory of sets involves the rather ambitious beginnings of a reformalization. I want to remind the reader that I am ultimately more cautious about the matter than I might suggest at times. To that effect, I want first to re-examine the position from which I am beginning this essay.

I believe in a moderate form of set-theoretic realism: in particular, I believe that many undecidable statements in the theory of sets have determinate truth values. I think that I have presented some justification for this belief, but I do not think that I have come close to a final word on the matter. Compatible with my work are such 'anti-objectivist' positions as Hartry Field's, wherein our mathematical practice can drive us towards accepting particular truth values for statements like Cantor's continuum hypothesis for "aesthetic" reasons, but such reasons are not evidence that the continuum is of any particular size—only motivations for "refining our concepts so as to give the continuum" that size (Field 300). On the other hand, I *have* shown how it is *possible* to be a realist about such matters without adopting a bizarre mathematical epistemology. In the course of what follows I hope to bridge this gap. My broadest thesis will be that the aesthetic reasons which should inspire us to refine our set-theoretic concepts are also evidence that whatever more we can demonstrate with these refined concepts is true.

One reason why I have focused so closely on set theory is that I do not want to make realist claims about truth in formal systems in general. We can certainly create all sorts of formal systems without intending them to systematize some external matter important to us as humans. And the most interesting of these will be incomplete. We can also say interesting things about some of the incomplete statements in these purely conventional formal systems—we can often reason about their truth from a meta-theoretical point of view, and we can even come to conclusions about these matters from time to time. So long as these systems are purely conventional, however, we can only swallow the bulk of the incompleteness results and understand the corresponding statements as lacking a truth value.

Of course the fact that these formal systems are not meant to formalize some other aspect of our understanding is just what makes them different from mathematical systems. As humans, we are interested in such practical matters as efficient mail delivery, audio recording, and building patios as well as such impractical matters as classifying all the three dimensional objects we can visualize, seating dinner guests, and playing interesting games of *mu torere*[7] (Hersch 230). We can answer questions about these practical and impractical things by creating formal systems whose foundations are sound with respect to the corresponding activities. It is no coincidence that people who are talented at answering these mathematical questions are often also talented at deriving non-referring statements in purely conventional systems. This is because formal derivation is an essential part of mathematical epistemology, not because all formal systems are mathematical.

Earlier I have stressed the point that being a realist about mathematics is compatible with the standard notion that mathematical knowledge is dependent on formal proof. Formal derivation is necessary for mathematical knowledge, but it is sufficient neither in practice nor in theory. Mathematical practice is so rich with "physical intuition and [informal] experience" that many of the most difficult theorems to prove can be easily explained and defended (Jones 214). Derivation of a statement can never suffice for mathematical truth unless that derivation takes place within a formal system whose foundations are sound with respect to the relevant mathematics. We assume that all such formal systems will be consistent and ω-consistent, so no theorem of an inconsistent system can be a mathematical truth in virtue of its being a theorem of that system. But the relevant systems must also correspond to our mathematical theories by having foundations which are in some sense uncontroversially true of the relevant mathematics. The Peano Axioms, for instance, seem like a serious foundation for number theory despite their incompleteness. While we cannot prove from these axioms that they are consistent, it is so highly plausible that they are that the interest in their consistency is only tangential to their use. The individual axioms and axiom schema—that *every number has a unique successor*, for instance—are precisely the kinds of things that we believe about numbers and are simple enough to be axioms.

In this essay I will evaluate the theory of sets according to the general issues which I have mentioned here in brief. My first task will be a description of what set-theory is an should be—a point which I believe has been confused during the dynamic history of the subject. Only then will I return to the question of which formal system or systems have theorems, all of which are true about sets. I offer the details of my argument against one candidate system, the Zermelo-Fraenkel Axioms plus the Axiom of Choice (ZFC). This argument targets the choice axiom directly and the standard iterative hierarchy concept

---

[7] Mu torere is a game which originated in New Zealand. Several years before Euler announced his analytic solution to the famous Konigsberg bridge problem, an equivalent fact was known in New Zealand: The

of set-theory (with respect to which ZFC appears sound) indirectly. I will conclude with the presentation of an alternate system, an extension of ZF (without choice).

I do not believe that there is any convincing argument for accepting ZFC as the final formalization of set-theory. Many of the set-theoretic statements that are patently unprovable in ZFC continue to be of interest, and some additional work must be done in order to arrive at a system within which we will be able to decide these statements. However, I believe that any dogmatic extension of ZFC, such as choosing one from the plurality of available consistent models, would necessarily be an incautious one. Furthermore, systems which are not simply extensions of ZFC but, like the system I propose in this essay, are more radical alternatives which contradict ZFC in some way or another require even more justification. I cannot provide sufficient justification for the system I propose in this essay. Nevertheless, I will consider several essential criteria with which we should measure a formalization of set-theory. The advantages of the alternative system over ZFC will be clear throughout this consideration.

2. A history of the concept of set

"In [a] sense, set theory is the ultimate court of appeal on questions of what mathematical things there are" (Maddy, 1990, 4). But only in a very restricted sense. We can define the operations of a mathematical theory, for example the homomorphisms of group theory or the arithmetic operations of number theory, in terms of sets by identifying the groups and numbers from those mathematical theories with sets and by stipulating operations on sets in terms of operations allowable by the axioms of set theory[8]. We may then rephrase a higher level mathematical question such as whether there exists a group with some property P (expressible in group theory) purely in terms of set theory. This is often a useful procedure, because the questions from a higher level theory, once rewritten in the language of set theory, can often be answered immediately because of previous results in set theory.

I have argued that this sense in which mathematics is foundational cannot be the metaphysical one. For the purposes of this essay I will keep a fair distance from the ontology of mathematics. It will suffice for me to repeat my claim from chapter 1 that there is no reason to suppose that sets are in any sense more real than other mathematical objects because those objects can be identified with sets. I do not

---

most interesting games of mu torere are those in which each player has four markers and the board is an eight-pointed star.

[8] This is one of the main points I make in chapter I.

think that it is sensible to think of mathematical objects independently of the relationships between them, which, although definable in terms of operations on sets, are not properly part of set theory (Resnick's 222). This leaves us with the more modest sense of set theory as a foundation for mathematics: Set theory is a mathematical theory in which the objects and operations from all other mathematical theories can be defined.

My suggestion now is to take this modest foundational claim more seriously. Specifically I want to suggest that this is all that set theory is. The axioms of set theory are statements about what sorts of collections exist. There are collections of infinite size; for any two collections, there is another collection which is there union; etc. These are the most primitive notions we have of dealing with abstract collections, and it makes sense that they may be applied to all mathematical theories. For as much as these notions arose from our original understandings of collections of objects, they were also derived from an examination of how the rest of mathematics behaves.

There are two main histories of foundations projects from which the notion of set has evolved. The first is a route through Frege, Russell, and Whitehead. This was an attempt to found arithmentic. The second was an attempt to find a consistent foundation for the theory of real numbers and the calculus originating in the work of Cantor, Dedekind, and Zermelo. Each history had simultaneous development in two directions. The foundational work was shaped by the primitive, yet to be founded mathematics, and the mathematics was sharpened by the precisifications made in the foundations. Once functions were redefined in terms of sets, many old questions were answered and many others appeared meaningless. Number theory was founded in terms of logic, but only after sifting through several logical paradoxes and oddities which did not correspond to the facts about number theory that we sought to found. Both developments demanded that its sets justify those facts about mathematics which were uncontroversial, and both developments changed the nature of the mathematics as well. Ultimately, but only after a period of revision, the route through Cantor and Zermelo became standard. The set theory that emerged was of an iterative concept in a cumulative hierarchy.

Here is a first pass of what is meant by the iterative concept of set. The universe of sets is a sequence of stages. There exist sets at each stage of the sequence. These sets are formed from elements from previous stages "with complete freedom, without concern for any method of construction" (Maddy, 1990, 102). This universe of sets is free from the paradoxes of earlier set theories. There is no set of all sets, because it could not appear at any stage in the hierarchy. Likewise, there is no set which contains itself. Although this concept of set and the cumulative hierarchy were not the first approaches to set theory, they are very intuitive ones. While classes of objects are often very useful to speak about, the more combinatorial notion we need in order to found all of mathematics requires a sequential construction. Since I am arguing that the theory of sets is no more than the study of which collections exist and what we can say about them when we know nothing more than facts about membership and set size—those non-logical notions which are common to all branches of mathematics—this iterative concept seems to be the proper one.

A common opinion among set theorists today is that the axioms of ZFC are each true because they are sound with respect to the iterative concept that I outlined above. The choice axiom, for instance, cannot be controversial once we accept the iterative concept of set. This axiom is the statement that given any collection of non empty sets, there is a set which contains exactly one element from each of these sets. Since sets at higher levels can be formed with complete freedom, the choice set is necessarily one of the sets that exists at the very next stage beyond the highest stage of any set in the given collection. I agree that Zermelo's choice axiom is an obvious consequence from the iterative concept that I have discussed above. However, I do not believe that the choice axiom should be an axiom. In what follows I actually argue that the choice axiom is false. For this argument, then, I must explain that the iterative concept of set, as I have just presented it, is actually not the proper one.

The distinction I want to bring in to focus is an observation of Byeong-Uk Yi. The iterative concept that I presented above is actually a combination of a basic concept *and* a set of "alleged facts about that concept" (Martin 218). The concept is simply that sets are formed of elements from previous stages in a sequential hierarchy. This alone suffices to avoid the paradoxes of pre-axiomatic set theory. But this concept is usually presented along with the claim that absolutely all sets of all objects up to any stage in the hierarchy exist at the next stage, regardless of what sorts of constructions we can make from those sets. In the cumulative hierarchy that corresponds with the intended interpretation of this fact, the choice axiom and the other axioms of set theory are all uncontroversially true. What is not at all clear is that this fact, above and beyond the basic iterative concept, is true of the theory of sets which we abstract from the non-logical notions about collections and their relationships which are common to all branches of mathematics. My efforts will not be to reestablish controversy about whether the axiom of choice follows from the iterative concept as Maddy has characterized it. I want to shift the controversy to the question of whether the leap from the simple concept of sequential set formation to the notion of the existence of ‘absolutely all’ sets at each stage in the hierarchy is valid.

This question is precisely the subject matter of Charles Parsons’ 1975 essay, “What is the iterative conception of set?” Parsons is very attuned to the distinction between what is essential to the iterative concept and the alleged facts about that concept which usually accompany it. His statement of the iterative concept, “sets form a well-founded hierarchy in which the elements of a set precede the set itself”, leaves unanswered the question about exactly which sets actually exist at later stages (504). Of course, at all finite stages, the answer is ‘all of them’. But to prove this we do not need a separate axiom. When we consider infinite stages, things get more complicated. It is not even clear how we might interpret an answer like ‘all of them’ or whether any two people would mean the same thing by this.

I think that Hao Wang’s consideration of this issue is appropriate. “We can form a set from a multitude only in case the range of variability of this multitude is in some sense intuitive” (513). Wang suggests that the step from doing this in finite cases to doing it in infinite cases is a matter of extending the human faculty of overviewing “or look[ing] through or run[ning] through or collect[ing] together” to an idealized point of view, from which one can consider infinite multitudes (ibid). Parsons’ is critical of Wang’s discussion of this issue, because at times Wang seems to define a set so loosely that many proper classes—collections of objects that would appear nowhere in the cumulative hierarchy—would qualify as sets. This point aside, I think that Wang’s main idea is quite forceful. So long as we are looking for a measure of certainty in our set theory, we need some robust notion of what it takes for a set to exist at a stage. The vague claim that ‘absolutely all’ collections of objects from earlier stages appear at the next stage falls well short of this criterion. I also think that Wang is correct when he emphasizes the role of knowledge and intuition, albeit a hypothetical and ideal extension of our usual understanding of these terms, in this process.

I conclude this section with an example of a set which might or might not appear in the cumulative hierarchy, depending on how we might want to extend the iterative concept of set.

Let $\mathcal{R}$ be the set of real numbers, and let $\mathcal{Q}$ be the set of rational numbers. For every $r \in \mathcal{R}$ consider the translation $r + \mathcal{Q} = \{r + q : q \in \mathcal{Q}\}$ of the set $\mathcal{Q}$. We know that $r_1 \in r_2 + \mathcal{Q}$ if and only if $r_1 - r_2 \in \mathcal{Q}$. Furthermore, $r_1, r_2 \in \mathcal{R}$ implies either $(r_1 + \mathcal{Q} = r_2 + \mathcal{Q})$ or $((r_1 + \mathcal{Q}) \cap (r_2 + \mathcal{Q}) = \varnothing)$. So the translations of $\mathcal{Q}$ partition $\mathcal{R}$ into uncountably many disjoint sets. There is no way for us to determine in general, given any two real numbers, whether they belong in the same translation. The question I want to raise is whether there exists a set $\mathcal{S}$ which contains exactly one element from each translation.

The usual answer to this question is ‘yes’. If we extend the iterative concept of set with the additional statement that ‘absolutely all’ combinations of elements from all earlier stages are gathered into sets at the next stage in the hierarchy, then we are committed to this answer. This is precisely what the choice axiom allows. The partition of $\mathcal{R}$ into the translations r + $\mathcal{Q}$ provides us with a collection of nonempty sets, and the choice axiom is the statement that given any such collection, there will exist a set (here $\mathcal{S}$) which contains exactly one element from each set in the collection. This axiom is only useful in cases like the present one, when there would be no way to construct such a set, and no way even to prove the existence of such a set without the axiom.

But if we accept some criterion for set existence like Wang’s, and we say that a set exists at a stage only if there is some sense in which an idealized, infinite intuition could run through all the sets in previous stages and determine whether each set is or is not a member of the alleged set, then we should not want to say that 'absolutely all’ combinations of elements from earlier stages appear as sets at the next stage. Wang actually says that the meaning he intends to convey with his idealized intuition is one which secures a cumulative hierarchy from which the choice axiom follows. It is very unclear how he arrives at this point, however, because his reference to idealized intuitions is consistently obtuse. I am not in a position to say whether there is an interpretation of Wang’s discussion which in fact secures the obviousness of the choice axiom. The only coherent understanding of the notion of running through a infinite collection that I have, however, presupposes that every such collection can be well-ordered (so that when each set in the collection is situated in a linear ordering, every subset of that ordering will have a unique least element). And since this supposition is equivalent to the choice axiom (Stoll 116), it is hardly the type of assumption one should make when trying to justify the axiom.

Ultimately, the strength or circularity of Wang’s reasoning is irrelevant. Should it be the case that there is a coherent interpretation of his discussion which does not presuppose an equivalent to the axiom of choice and which secures the obvious existence of sets like $\mathcal{S}$ above, this will still rest on whether Wang’s criteria for set existence is the proper one. I am sympathetic to the idea that we should weigh heavily the social, human aspect of mathematics, either by searching for a criteria of set existence based on an idealization of human intuition or by some other route. But the only point I care to defend from Wang’s and Parsons' discussions as anything more than an attractive speculation is that the concept of set which we have come to understand through working with and abstracting from mathematical theories does not commit us to a free, combinatorial plethora of sets at all transfinite stages.

3. Evidence against the choice axiom

If the set ٯ whose existence I question above exists, it is not Lebesgue measurable. Lebesgue measure is a generalization of the idea of the relative size of a set which has a precise distance function defined on it. The Lebesgue measure of a line segment, for instance, is the length of the segment, the Lebesgue measure of a rectangular region of the plane is the region's area, discrete sets of points have Lebesgue measure of zero, etc. The details of Lebesgue measure are beyond the scope of this essay[9]. It is important to note, however, that the existence of a set which is not Lebesgue measurable is surprising and counter intuitive, but nothing to be too concerned about.

Of much greater importance is what one can prove from the existence of non measurable sets. An analogous proof of the existence of non measurable subsets of a sphere[10] (here we replace R with the set of all points within a fixed distance of the origin in three dimensional space, we replace Q with the subset of the sphere consisting of all points with three rational coordinates, and we define the translations of the rational subset as rotations about the origin) will allow us to prove the following:

Theorem (Banach-Tarski): Given two spheres S_1 and S_2 of radii r and 2r, respectively, there exist partitions of S_1 and S_2 into finite sets $A_1, A_2, \ldots, A_n$ and $B_1, B_2, \ldots, B_n$ ($A_i \subset S_1$ for all $1 \leq i \leq n$, $B_i \subset S_1$ for all $1 \leq i \leq n$) and there exist rigid motions $\varphi_1, \varphi_2, \ldots, \varphi_n$ such that $\varphi_i (A_i) = B_i$ for all $1 \leq i \leq n$.

A rigid motion is essentially a displacement of a set during which the size and shape of the set are completely unchanged. The only characteristics of a set which may change during a rigid motion are the set's location and its orientation.

In order to emphasize exactly why this result is so striking, I want to distinguish the result from a trivial claim. This theorem would have had no impact on the mathematical community if it were not for the fact that the partitions of the spheres were finite. One could easily prove the analogous result with

---

[9] An accessible treatment of the relevant notion of Lebesgue measure (Lebesgue measure on Euclidean space) can be found in Frank Jones' book.

[10] In my discussion of the Banach-Tarski theorem, I will deviate from the standard mathematical definition of sphere, which is the set of points in Euclidean 3-space all of which are a fixed distance from a specified

infinite partitions. Suppose that each set in the partition were a single point. Then only one rigid motion would be needed in order to satisfy the result: $\varphi(x) = 2x$. The amazing thing about this theorem is that the partitions are finite. You can get a general sense of the theorem by imagining two glass spheres, one with twice the volume as another. The theorem says that the two spheres can be taken apart into some finite number of pieces so that for every piece of one sphere, there is a piece of the other which is exactly identical to it. Of course, it is wrong to think about this in terms of shattering the glass spheres, because each shard of each sphere would be measurable. One characteristic of the pieces of the spheres one would need for this proof is that whatever the pieces are like, they are unlike anything that we could ever understand. Their existence is guaranteed only by an axiom which extends further than the human understanding of infinite collections.

This theorem is often called the Banach-Tarski paradox, instead of the Banach-Tarski theorem. Of course the result is not literally a paradox, like the paradoxes of pre-axiomatic set theory. The negation of the Banch-Tarski theorem is not provable, so the result does not demonstrate that the ZFC is inconsistent. On the other hand, the theorem is more than interesting or counter-intuitive. It is hard to believe that the statement of the theorem is true about the spheres. Of course the statement, rewritten in purely set theoretic terms, is a fact about the ZFC axioms, but if it is not also a fact about the geometrical notion of a sphere, then we are faced with an instance where geometry does not reduce (in the relevant sense) to ZFC. But if this is the case, given that my characterization of set theory is the proper one, then not all the theorems of ZFC can be true statements about sets.

I believe that this is indeed the situation that the Banch-Tarski theorem forces us into, although there seem to be two ways to argue otherwise. One could say that although the Banach-Tarski theorem is wildly counter-intuitive and threatens our understanding of rigidity in geometry, the situation is merely a case where we are forced to abandon our intuitions in favor of the foundations project and the more precise notion of our mathematical concepts that arise from it. This is similar to what happened with the mathematical notion of function. Once the term function was redefined in terms of sets, so many counterintuitive results became theorems about functions that many people resisted the reduction of functional analysis to set theory for years. But ultimately, the benefits of the set theoretic definition became clearer, and the opposition disappeared.

---

center. By sphere, I will mean a "solid sphere": all such points which are at most that distance from the

The present situation is different for several reasons. Unlike the case with functions in the late nineteenth century, there seem to be no ambiguities in the notions of rigidity or finite partition, and so there is no precision to be gained from an appeal to set theory. Furthermore, this line of argument rests on the assumption that the choice axiom is a fact about sets. But if our theory of sets is an formalization of the basic principles of relationships between collections which we abstract from our mathematical theories, it makes no sense that any axiom of set theory could contradict so strongly our intuitions about any one mathematical theory. I have argued above that there should be some independent reason for expanding on the iterative concept of set before we commit to a model of the set theoretic hierarchy from which the choice axiom is immediately true or false. If anything, the fact that the choice axiom leads to such a striking geometrical result should incline us toward reasons why the choice axiom might be false.

Another way to defend the choice axiom is to say that while the Banach-Tarski theorem is a very unattractive result, the benefits of the axiom outweigh its less fortunate results. The benefits seem to come in two varieties. The first is that without at least a weaker form of the choice axiom such as the axiom of countable choice, the axiom of dependent choice, or the Stone representation theorem, there are far too many useful and uncontroversial results in numerous branches of mathematics that would not be provable. The second type of benefit of the choice axiom is the idea that the set theory that results from its inclusion is a natural extension of undisputed facts about finitary combinatorics into the infinite.

I think that this latter defense of the choice axiom is very powerful. However, I think that the very arguments used in the defense are actually more forceful when turned against the axiom. I will close this section with the observation that while the Banach-Tarski theorem is a very serious obstacle for any defense of the choice axiom, the benefits of the axiom of countable choice, the axiom of dependent choice, and the Stone representation theorem are so great that accepting a pathological deconstruction of a sphere might very well be a reasonable price for them. In fact, if there were no way to salvage the weaker forms of choice in the process and to replace it with an at least equally natural principle, it would not be clear that there could be any conceivable way to reject the truth of the Banach-Tarski theorem. In such a case, the same reasons for challenging the foundations of set theory would remain, because the undecidability of the continuum hypothesis would still suggest that ZFC was not fully characteristic of set theory. But the

---

center.

direction which I propose in the next section, namely to begin reaxiomatizing set theory by eliminating the axiom of choice, would not be open.

But supposing that there is some way to preserve the useful results of the choice axiom while avoiding those results which are more pathological, the most plausible interpretation of the Banach-Tarski theorem is that it is evidence against the axiom of choice. This turns out to be the case.

4. An approach to set theory

One common criticism of the choice axiom is that it is too complex a statement to be an axiom. Rather than deciding whether or not the statement is obvious from our understanding of sets, perhaps we should find some other more intuitive principles which make either the statement of the choice axiom or its negation a theorem. If it is true that by restricting ourselves to the understanding of sets which we can have just from abstracting basic principles about the relations between collections from our various mathematical theories we must necessarily remain agnostic about what sorts of sets exist at each transfinite stage of the cumulative hierarchy, then this seems like a valid criticism. While the choice axiom is obviously true with respect to some alleged facts about the iterative concept of set, the concept itself does not commit us to the axiom being true or false. From that point of view, neither the axiom nor its negation could be obvious enough to be an axiom. For it to be true or false, it should be provably so from some more primitive statement.

The same argument will also prevent statements like the principle of countable choice, which seem indispensable, from being axioms. In this section I will discuss a logical principle which appears very intuitive, and with which we can demonstrate that the choice axiom is inconsistent with the other ZF axioms. This principle has the additional remarkable feature that with it the principle of countable choice and the principle of dependent choice are theorems in ZF. A final, rather surprising feature of this principle is that from it we can prove that every model of set theory is uncountable and contains non-constructive sets. This last fact suggests that the principle might be a step towards disproving the continuum hypothesis.

Here is the principle. Retain all the rules from classical quantificational logic, and add the following two rules: (1) Expressions with a countable number of quantifiers are valid sentences iff the expression is equivalent to an expression in Kleene Normal Form with a countable string of quantifiers and a leftmost quantifier. (2) The principle of quantifier negation applies to infinite strings of quantifiers. Call the resulting infinitary logic $L_{\omega 1,\omega}$. Let AC be the choice axiom, let CC be the principle of countable choice, and let DC be the principle of dependent choice.

Theorem: $L_{\omega 1,\omega}$ (ZF) |= ~AC.

Assume the well ordering property. Let $\{\sigma_\alpha , \langle\}$ be a well ordering on the set of infinite sequences of natural numbers. There are $2^{\aleph_o}$ such sequences.

Let $\aleph_\gamma$ be an aleph such that $2^{\aleph_o} = \aleph_\gamma$ (Since we are assuming the well ordering property this aleph is unique, but this need not be true in general, when cardinal numbers are only partially ordered.) We will construct two transfinite sequences of sets $\{X_\alpha\}$ and $\{Y_\alpha\}$, $\alpha < \aleph_\gamma$, $|X_\alpha| = |Y_\alpha| \leq \alpha$, ordered by inclusion as follows.

Base case: $X_0 = Y_0 = \varnothing$.

Induction: We will define X_α and Y_α for arbitrary $0 < \alpha < \aleph_\gamma$. Either $\alpha = \beta+1$ for some β or α is a limit ordinal (e. g. ω is a limit ordinal). If $\alpha = \beta+1$, then consider the set of sequences of the form $\{x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3, . . .\}$, where the subsequence $\{x_0, x_1, x_2, x_3, . . .\} = \sigma_\beta$. There are $2^{\aleph_o}$ such sequences because there are $2^{\aleph_o}$ possibilities for the subsequence $\{y_0, y_1, y_2, . . .\}$. Since $|X_\alpha| \leq \alpha < \aleph_\gamma = 2^{\aleph_o}$, there exist sequences of the form $\{y_0, y_1, y_2, . . .\}$ such that $\{x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3, . . .\}$ is not in X_α. Let σ_y be the least such sequence, in the ordering $\langle$, and let $\sigma_{\alpha'}$ be the alternating sequence of elements from σ_β and σ_y. Then $Y_\alpha = Y_{\beta+1} =^{df} Y_\beta \cup \sigma_{\alpha'}$. Similarly, $X_\alpha = X_{\beta+1} =^{df} X_\beta \cup \sigma_\alpha$, where σ_α is the alternating sequence formed from elements of least sequence in $\{\sigma_\alpha\}$ not in Y_α and σ_β. Alternatively, if α is a limit ordinal define $X_\alpha = \cup_{\beta<\alpha} (X_\beta)$ and $Y_\alpha = \cup_{\beta<\alpha} (Y_\beta)$.

Let $A = \cup_{\alpha<\aleph\gamma} (X_\alpha)$. Clearly $A \subseteq \{\sigma_\alpha\}$. Then there does not exist a sequence $\sigma_x \in \{\sigma_\alpha\}$, such that for any sequence $\sigma_y \in \{\sigma_\alpha\}$, the alternating sequence $\{x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3, . . .\}$ is in A. And there is no sequence $\sigma_y \in \{\sigma_\alpha\}$ such that for any sequence $\sigma_x \in \{\sigma_\alpha\}$, the alternating sequence $\{x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3, . . .\}$ is not in A.

But let $S \subseteq \{\sigma_\alpha\}$ be given, and suppose that $\sim(\exists x_0)(\forall y_0)(\exists x_1)(\forall y_1)(\exists x_2)(\forall y_2) . . . (\{x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3, . . .\} \in S)$. Then by quantifier negation we have $(\forall_{x0})(\exists y_0)(\forall x_1)(\exists y_1)(\forall x_2)(\exists y_2) . . . \sim(\{x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3, . . .\} \in S)$. Or $(\forall_{x0})(\exists y_0)(\forall x_1)(\exists y_1)(\forall x_2)(\exists y_2) . . . \sim(\{x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3, . . .\} \notin S)$. But this contradicts the construction of A.

In addition to this result, there are theorems which insure that the benefits of the choice axiom, which might be understood as outweighing the negative impact of the Banach-Tarski theorem, are preserved in the logical extension $L_{\omega 1,\omega}$. In particular, a trivial theorem of $L_{\omega 1,\omega}$ (ZF) is that a class of infinite games are all determined. I will postpone the discussion of determinacy and infinite games until the next section. At present, I will discuss some of the results that follow from this theorem.

$L_{\omega 1,\omega}$ (ZF) |= CC; $L_{\omega 1,\omega}$ (ZF) |= DC[11]

Both the principles of countable choice and dependent choice follow from the fact about the determinacy of certain infinite games mentioned above, and so these principles are theorems of $L_{\omega 1,\omega}$(ZF). This means that using this infinitary logic with set theory insures the uncontroversial results that every infinite set contains a countable subset, that a linear ordering with no infinitely descending subsequence is a well-ordering, that every subspace of a separable metric space is separable, etc. Not preserved in $L_{\omega 1,\omega}$(ZF) are the well-ordering principle (the statement that all sets can be well-ordered), Zorn's Lemma, and other results that are far more controversial.

In addition, we have $L_{\omega 1,\omega}$(ZF) |= every set of real numbers is Lebesgue measurable. This insures that the Banach-Tarski theorem fails, because the pathological decompositions require unmeasurable subsets of the sphere. A final result, $L_{\omega 1,\omega}$(ZF) |= $\aleph_1$ is a measurable cardinal, puts heavy restrictions on the number of models for set theory. In particular, no models in which all sets are constructible (V=L), including the model Godel used in 1938 to show that CH was relatively consistent with ZF, are allowed models for $L_{\omega 1,\omega}$(ZF). Any models for set theory with infinitary logic necessarily contain generic, nonconstructible sets at infinite stages. Such models converge towards models in which CH is false, as we add more nonconstructible sets.

5. Evidence for $L_{\omega 1,\omega}$(ZF)

The undecidable statements of set theory, like the continuum hypothesis, are equivalent to propositions about the existence and behavior of infinite collections, so far removed from human experience that all our attempts at setting a hypothetico-deductive system which would contain either them or their negations as theorems have been consistent with models in which they are true and with models in which they are false. It seems like too great a demand on human intuition to be able to secure uncontroversial axioms which will settle the truth values of these statements.

In his 1925 address to the Westphalian Mathematical Society, David Hilbert gave an interesting characterization of these questions. Speaking specifically about the continuum hypothesis, he said, "On the one hand, new methods are required for its solution since the old methods fail to solve it; on the other hand, its solution itself is of the greatest importance because of the results to be gained" (201). In my earlier discussions of the continuum problem, I have stayed close to Hilbert's sentiments on the matter. I

---

[11] For clear presentations of how CC and DC follow from the supposition of determinacy that I mention,

am also indebted to Hilbert's proposed method of establishing new methods: ". . . [T]he infinite is nowhere to be found in reality," he continues. "It neither exists in nature nor provides a legitimate basis for rational thought . . . operating with the infinite can be made certain only by the finitary" (ibid). Hilbert is certainly attuned to the serious demands we place on human intuition, when we expect to find axioms for set theory which will determine the truth-value of statements like the continuum hypothesis. The best intuitive conjectures we can make about the infinite are probably all generalizations from what we know about finite collections. This seems to be the case both with the choice axiom and with the logical extension which I am proposing. There is no controversy about the existence of choice sets for finite collections—we do not even need a separate axiom to insure their existence. A reasonable supposition is that similar sets exist for all collections of sets. Hence, the choice axiom. But this supposition is no more reasonable than many other generalizations we might make from facts about finite structures. One example is the quantification negation rule for infinite strings of quantifiers. But as we have seen, generalized quantification negation and the choice axiom cannot both be true. There must be some criteria which should lead us to prefer some extensions over others.

One plausible reaction to the mutual incompatibility among various reasonable extensions of our finitary knowledge into general rules is to say that there just is no reason to prefer any of these extensions. Perhaps the fact that our intuitions lead us astray of certainty once we begin to conjecture about the infinite is reason to limit our treatment of mathematics to finite domains. Hilbert was probably the leading skeptic about human endeavors towards understanding the infinite, but even he did not make this move. An amazing thing about statements like the continuum hypothesis is that its solution would reflect into other domains of mathematics. I will suggest, as Hilbert did, that there can be evidence for axioms of set theory about which we are intuitively uncertain when the set theory we get from these axioms is successful in solving "problems which, though known for a long time, the theory was not expressly designed to solve" (200).

I have already discussed one application of this "acid test" for axioms, by citing the Banach-Tarski theorem as evidence against the choice axiom. While our intuitions about infinite collections several levels deep in the cumulative hierarchy of sets is all but certain, our intuitions about finite spheres is fairly concrete. One reason for doubting the generalization of the existence of choice sets is that the existence of

---

see Jech 177-181. Here are also proofs that determinacy entails that there exists a measurable cardinal, etc.

such sets for uncountable collections of uncountable sets guarantees decompositions of spheres which challenge our more certain intuitions. By the same token, the fact that the general principle of quantification negation insures that such pathological results are impossible should be seen as evidence—though not conclusive evidence—for accepting that principle.

As more questions from different branches of mathematics are shown to be independent of ZF and ZFC but dependent on the size of the continuum, more possibilities for evidence for new axioms arises. The question of which homomorphisms from the Banach algebra C(I) of all continuous functions on the closed unit interval are themselves continuos is one such question. Other questions about the measurability of functions on identification spaces in real analysis are presently undecidable, but would all be solved if there were axioms about the existence of large cardinals (or if the general quantification negation principle were accepted, then the size of the continuum would itself be a very large cardinal—a *measurable* cardinal—and these same questions from analysis would be solved) (Dales 195). At the present, however, I doubt that many people other than few working analysts have any intuitions about these questions which are stronger than alleged intuitions about the set-theoretic axioms themselves. So there are no preferred solutions to these analytical questions which could serve as teleological justification for large cardinal axioms or for stipulations about the size of the continuum. In the remainder of this section, I will discuss some problems which do have intuitive solutions, and I will explain how this fact is further evidence for $L_{\omega_1,\omega}$(ZF).

One reason why Hilbert, Godel, and others' beliefs that axioms could ever be justified by the theorems which we can derive from them strike many modern readers as bizarre is that the history of set theory has obscured the role of its axioms. In the first half of the century, Kurt Godel spoke encouragingly about the prospect of reaching "a probable decision about" the truth of a new axiom, "even in case[s] [where] it has no intrinsic necessity at all," by studying the "verifiable consequences" of adopting the new axiom (477). At the time when this enthusiasm existed most prominently among philosophers, however, there were pragmatic obstacles in the way of their progress: "This criterion [of justifying a new axiom in light of its theorems], however, though it may become decisive in the future," Godel wrote, "cannot yet be applied to the specifically set-theoretical axioms . . . because very little is known about their consequences in other fields" (485). Today, of course, this is no longer true. I have mentioned only a few consequences of new set theoretic axioms in real analysis—there are several others. The consequences of such axioms in abstract algebra and topology are even more numerous. Yet in the time that has passed since Godel and Hilbert wrote about these matters, the axioms of Zermelo-Fraenkel set theory have so ingrained themselves in the minds of working mathematicians that few question either the validity or exhaustiveness of these axioms.

I am seriously inclined to question this attitude. If Hilbert was right, and our intuitions about the infinite are mere speculations—or, at best, plausible generalizations from finitary facts—it cannot be the case that an axiom like choice is self-evident. Maybe some would say that the present axioms are true by convention. But this, too, seems mistaken. I do not doubt that modern mathematicians' unquestioning behavior towards the axioms of set theory is the product of habit or convention, but I see no explanation of how the historical use of any axiom could establish its truth. Hibert's arguments alone should suffice "to counteract the impression that [the presently accepted axioms of set theory] enjoy a preferred epistemological status not shared by new axiom candidates" (Maddy, 1988). There simply is no hope that there could ever be a set of self-evident axioms at the foundation of a robust theory about infinite collections. Below, then, I will be rejuvenating Godel's hopes at a time when consequences of new foundations in set theory do appear in all branches of mathematics. These consequences can mold our preferences for foundational changes when there are no intrinsic reasons to adopt one principle over another among the wealth of available changes.

I mentioned in the previous section that one can prove in $L_{\omega 1,\omega}$ (ZF) that a certain class of infinite games can are all determined. I will say more about this now. Infinite games are best understood in the following dramatization: Two players take turns making moves. Each player is aware of all the moves that the other player has made, and each player is aware of what it would take to win the game even though a win is insured only after a completed sequence of infinite (countably infinite) moves. A winning strategy for either player (players are distinguished according to who makes the first move) is a sequence of moves available to her which will insure that she will win the infinite game. An infinite game is determined if the rules and the criterion for a win are such that one player has a winning strategy.

In 1962 Mycielski proved that there were well-defined infinite games which were not determined, but he proposed a class of infinite games which he thought might all be determined. This class of games is characterized by two requirements: each player is required to choose exactly one natural number for each of her moves, and the sequence of natural numbers formed from combining both players' moves in turn—ad infinitum—determines which player wins according to whether or not it is a member of a set $\mathcal{S}$ of sequences of natural numbers. We say that player 1 (the player who moves first) wins if and only if the sequence constructed from both players' moves is in $\mathcal{S}$, and player 2 wins if and only if the sequence is not a member of $\mathcal{S}$. Player 1, then, is said to have a winning strategy if there is a sequence of natural numbers which she can play which will guarantee that the sequence that results from combining her moves with player 2's moves will be in $\mathcal{S}$. Player 2 has a winning strategy is she has a series of moves which will ensure that the sequence will not be in $\mathcal{S}$. Mycielski's Axiom of Determinacy (AD) is the proposition that either player 1 or player 2 will have a winning strategy for any set $\mathcal{S}$.

AD is a cumbersome statement, and its chances of every becoming an axiom of set theory are consequently doubtful. However, AD is a theorem of $L_{\omega 1,\omega}$(ZF). (This fact is immediate from the proof of $L_{\omega 1,\omega}$(ZF) |= ~AC above.) For this reason, two important and useful theorems in recursion theory, which follow from AD, are further evidence for $L_{\omega 1,\omega}$(ZF), given Godel's criteria for extrinsic evidence for new axioms. Both results concern degrees of complexity of sets. The first tells us that if AD is true, then the arrangement of all functions according to how far they are from being recursive functions has a very intuitive order. The second result deals with a different notion of complexity. It says that the ordering of all sets of sequences of natural numbers by their relative complexity (i. e. a set "comes before" a second set in this ordering if it can be demonstrated to be less complex than the second.) is "almost" a linear ordering—another intuitive result. I will explain just how each of these results are intuitive after stating the respective theorems.

Our first theorem deals with the Turing degrees of unsolvability. A well-known result from computation theory is that some functions defined on natural numbers are recursive, although most are not. The Turing degree of any recursive function is 0, because all of these functions are computable with Turing machines. Beyond that, the Turing degree of a function is the measure of how incomputable that function is. More precisely, we order the Turing degrees of functions of natural numbers with $\langle$, so that degree(f) $\langle$ degree(g) means that the function g "has enough information to compute f mechanically". We then define a Turing cone as follows: Let δ be the Turing degree of a function on the natural numbers. The Turing cone of δ is the set of all δ' such that $\delta \langle \delta$'. It follows from AD that given any set of Turing degrees, either that set or its compliment contains a Turing cone.

Our second theorem concerns the complexity of sets of sequences of natural numbers. We define the following ordering of all such sets: Given any two sets, A and B, of sequences of natural numbers, we say that $\omega(A) < \omega(B)$ if there is a continuos function F on the set of all sets of sequences of natural numbers ($F: N^N \rightarrow N^N$) such that $F^{-1}(B)=A$. This ordering corresponds to a robust sense of complexity in the following sense: Given that $\omega(A) < \omega(B)$, if B is an open set, then A is an open set; if B is a projective set, then A is a projective set, etc. In short, if B is definable at any level of the analytic hierarchy, then its continuous pre-image, A, is also definable at that level. It follows from AD that given any two sets, A and B, of sequences of natural numbers, either $\omega(A) < \omega(B)$ or $\omega(N^N\text{-}B) < \omega(A)$. In other words, the ordering of the complexity of all such sets is linear, with the one exception that a set and its compliment might have incomparable degrees of complexity (both theorems in Martin 223-226).

There is an ample amount of both weak and strong evidence for AD to be gleamed from these theorems. A first point of interest is that whatever the ordering of complexities of functions on natural numbers and sets of sequences of natural numbers are like, it is almost inconceivable that there is no fact of the matter about their structure. Yet ZF (as well as ZFC) has a multitude of models, in which the orderings of these complexities take on a variety of structures. So whether or not the relatively linear ordering of set complexity and the cone structure of functional complexity that these theorems secure are correct, they are at least worth exploring, because they are possible answers to important questions which are otherwise unanswerable.

Of greater importance in this essay is what we have learned from exploring the consequences of these two theorems. It turns out that the orderings which follow from AD are precisely the types of orderings that complexity theorists expected (ibid). Almost inherent in the very notion of complexity are the ideas that a) the progression of set complexity is linear and b) any partitioning of the set of partitions of Turing degrees into two classes must have at least one class containing a Turing cone. In his discussion of these two theorems, Donald Martin concludes that their consequences are so acceptable, either because of how they measure up with our intuitions about complexity or because many of their corollaries provide simple proofs of facts which have since been proven directly from ZF, that in just about any sense in which theoretical consequences of a principle can justify accepting that principle as a mathematical truth, we are justified in accepting AD (227). But Martin is not willing to accept AD as true because it contradicts the choice axiom. In a sense, Martin is trying to balance two incompatible impulses: He sees as well as anyone that certain determinacy hypotheses have been considerably confirmed by theorems in complexity theory, but he is willing to endorse only that amount of determinacy which is consistent with the choice axiom, presumably because "few mathematicians [are] anxious to give up choice, which [has] become a well-entrenched axiom" (221). The determinacy hypotheses to which Martin thinks the above theorems have given evidence are weaker statements about projective sets and constructable sets—not "AD itself, which was never taken seriously as an axiom" (223).

It is not clear what to make of Martin's claim that AD has never been seriously considered as an axiom. Mycielski and Steinhaus certainly took it seriously when they first proposed it and championed the fact that it contradicted the more problematic consequences of choice. Moreover, whether or not AD ever appeared true in the past, if Martin is sincere about his claim that new axiom candidates are no less certain, prima facia, than the well-entrenched axioms of ZFC, I see no reason why the evidence for determinacy should be accepted as such whenever the consequences are consistent with ZFC but curtailed whenever they contradict the choice axiom. I am more inclined to take the consequences of the two theorems of this section as evidence, not only for the truth of an innocent amount of determinacy which is consistent with ZFC, but for further evidence (in conjunction with the problems from the Banach-Tarski theorem) against choice.

The virtues of using an extended logic in set-theory should now be clear. We have seen that the choice axiom is false in $L_{\omega_1,\omega}$(ZF), but that weaker principles of choice, which are often cited as grounds for accepting the full principle of choice as an axiom, are theorems of $L_{\omega_1,\omega}$(ZF). Another consequence of $L_{\omega_1,\omega}$(ZF) is that all infinite games with natural numbers as moves are determined. Given that the pathological consequences of the Banach-Tarski theorem shed some doubt on the truth of the choice axiom and that the consequences of determinacy in complexity theory are evidence for accepting principles of determinacy as true, extending the logic of set theory to include infinite strings of quantifiers is justified by these principles. Intuitively, this logical extension is roughly as plausible as the choice axiom itself. Both principles are generalizations from concrete facts about the finite. But when we measure the principles against one another by weighing their consequences, especially where their consequences correspond or fail to correspond with our intuitions in other branches of mathematics, the evidence mounts on the side of $L_{\omega_1,\omega}$—at least enough to justify $L_{\omega_1,\omega}$(ZF) as a potential foundation for set theory worth exploring.

References

Field, Hartry. "Which Undecidable Mathematical Sentences Have Determinate Truth Values?" in *Truth in Mathematics*. Edited by H. G. Dales and G. Oliveri. Oxford University Press. 1998.

Hersch, Reuben. *What is Mathematics, Really?* Oxford University Press. 1997.

Hilbert, David. "On the Infinite" in *Philosophy of Mathematics: Selected Readings*. Edited by Paul Benacerraf and Hilary Putnam. Second Edition. Cambridge University Press. 1983.

Jech, Thomas J. *The Axiom of Choice*. North-Holland Publishing Company. 1973.

Jones, V. F. R. "A Credo of Sorts" in *Truth in Mathematics*. Edited by H. G. Dales and G. Oliveri. Oxford University Press. 1998.

Jones, Frank. *Lebesgue Integration on Euclidean Space*. Jones and Bartlett Publishers. 1993.

Maddy, Penelope. *Realism in Mathematics*. Oxford University Press. 1990.

Maddy, Penelope. "Believing the Axioms part I" in *Journal of Symbolic Logic*, **53**, pp 481-511. 1988.

Martin, Donald. "Mathematical Evidence" in *Truth in Mathematics*. Edited by H. G. Dales and G. Oliveri. Oxford University Press. 1998.

Mycielski, J. and Steinhaus, H. "A Mathematical Axiom Contradicting the Axiom of Choice" in *Bulletin de l'Academie Polonaise des Sciences, Serie des Sciences Mathematiques, Astronomiques et Physiques*, **10**, pp 1-3.

Parsons, Charles. "What is the Iterative Concept of Set?" in *Philosophy of Mathematics: Selected Readings*. Edited by Paul Benacerraf and Hilary Putnam. Second Edition. Cambridge University Press. 1983.

Resnick, Michael. *Mathematics as a Science of Patterns*. Oxford University Press. 1997.

Stoll, Robert R. *Set Theory and Logic*. Dover Publications. 1961.

Wang, Hao. "The Concept of Set" in *Philosophy of Mathematics: Selected Readings*. Edited by Paul Benacerraf and Hilary Putnam. Second Edition. Cambridge University Press. 1983.