To appear in A. Caicedo et al, eds., *Foundations of Mathematics*, (Providence, RI: AMS).

Set-theoretic Foundations[1]

It's more or less standard orthodoxy these days that set theory -- ZFC, extended by large cardinals -- provides a foundation for classical mathematics. Oddly enough, it's less clear what 'providing a foundation' comes to. Still, there are those who argue strenuously that category theory would do this job better than set theory does, or even that set theory can't do it at all, and that category theory can. There are also those insist that set theory should be understood, not as the study of a single universe, V, purportedly described by ZFC + LCs, but as the study of a so-called 'multiverse' of set-theoretic universes -- while retaining its foundational role. I won't pretend to sort out all these complex and contentious matters, but I do hope to compile a few relevant observations that might help bring illumination somewhat closer to hand.

[1] It's an honor to be included in this 60th birthday tribute to Hugh Woodin, who's done so much to further, and often enough to re-orient, research on the fundamentals of contemporary set theory. I'm grateful to the organizers for this opportunity, and especially, to Professor Woodin for his many contributions.

I. Foundational uses of set theory

The most common characterization of set theory's foundational role, the characterization found in textbooks, is illustrated in the opening sentences of Kunen's classic book on forcing:

> Set theory is the foundation of mathematics. All mathematical concepts are defined in terms of the primitive notions of set and membership. In axiomatic set theory we formulate … axioms about these primitive notions … From such axioms, all known mathematics may be derived. (Kunen [1980], p. xi)

These days, familiarity has dulled our sense of just how astounding this fact really is. In his introductory text, Enderton makes sure that his students appreciate its scope and power:

> It is sometimes said that 'mathematics can be embedded in set theory'. This means that mathematical objects (such as numbers and differentiable functions) can be defined to be certain sets. And the theorems of mathematics (such as the fundamental theorem of calculus) then can be viewed as statements about sets. Furthermore, these theorems will be provable from our axioms. Hence our axioms provide a sufficient collection of assumptions for the development of the whole of mathematics -- a remarkable fact. (Enderton [1977], pp. 10-11)

The question for us is: what's the point of this exercise? What goal, properly thought of as 'foundational', is served by this 'embedding'?

A glance at the history delivers a hint. At the turn of the last century, Hilbert had just proved the consistency of geometry by relying on analysis, so he set the task of proving the consistency of analysis as the second on his famous list of problems in [1900]. In earlier correspondence with Cantor, Hilbert had come to understand that there could be no set of all alephs, and to feel the need for a consistent axiomatization of set theory, as well.[2] Zermelo arrived in Göttingen in 1897:

> When I was a *Privatdozent* in Göttingen, I began, under the influence of D. Hilbert, to whom I owe more than to anybody else with regard to my scientific development, to occupy myself with questions concerning the foundations of mathematics, especially with the basic problems of Cantorian *set theory*. (Quoted in Ebbinghaus [2007], p. 28)

By 1908, in his famous axiomatization, Zermelo drew the explicit connection between set theory and analysis:[3]

> Set theory is that branch of mathematics whose task is to investigate mathematically the fundamental notions 'number', 'order', and 'function', taking them in their pristine, simple form, and to develop thereby the logical foundations of all of arithmetic and analysis. (Zermelo [1908], p. 200)

Subsequent developments extended set theory's reach to the whole of classical mathematics, as indicated in the quotations from Kunen and Enderton.

Suppose, then, that your goal is to prove something or other about the vast variety of classical mathematics -- for the Hilbert school, its consistency. To do this, you first need to corral it all into some manageable package, and set theory turned out to be up to that task. Of course Gödel saw to it, with his second incompleteness theorem, that the project didn't turn out as Hilbert and his followers had hoped, but the fact remains that the most sweeping moral of Gödel's work -- that classical mathematics (if consistent) can't prove

---

[2] See Ebbinghaus [2007], pp. 42-43.

[3] Ebbinghaus ([2007], pp. 76-79), highlights the role of Hilbert's program among the motivations for Zermelo's axiomatization. Apparently Zermelo was keen to include a consistency proof in his [1908], but Hilbert encouraged him to publish it as it stood, in part, Ebbinghaus reports, because Hilbert knew 'that Zermelo needed publications to promote his career' (p. 78).

its own consistency -- is only possible given the set theory's codification of the entire subject into a neat set of simple axioms. And subsequently, the set-theoretic codification made it possible to settle other important questions of provability and unprovability, sparing the profession from sadly doomed efforts. So the 'embedding' of mathematics in set theory has this clear use. Presumably we'd all agree that the goal of proving something general about classical mathematics counts as 'foundational' in some sense, so here we do find set theory playing a foundational role.

Of course, this **Meta-mathematical Corral** is hardly the only foundational role set theory has been thought to play, but I think two of these can be dismissed as spurious. The first sees the 'embedding' of classical mathematics in set theory, often called a 'reduction' of classical mathematics to set theory, as the immediate descendant of Frege's Logicism. This line of thought takes Frege's project to be roughly epistemological: if mathematics can be reduced to logic, then knowing a mathematical fact is reduced to knowing a logical fact; assuming that we have an account of logical knowledge, or at least that finding such an account is a less daunting prospect than finding an account of mathematical knowledge had previously appeared to be, this reduction is a clear epistemological gain.[4] Of course Frege's logic turned out to be inconsistent and set theory has taken its place, but the epistemological analysis is supposed to carry over: we

---

[4] It isn't obvious that Frege himself saw the matter quite this way: much of his interest is in objective grounding relations between propositions, which are independent of our human ways of finding things out. See, e.g., Burge [1998].

know the theorems of mathematics because we know the axioms of set theory and prove those theorems from them. Thus the problem of mathematical knowledge reduces to the problem of knowing the set-theoretic axioms.

The trouble with this picture is that it's obviously false: our greatest mathematicians know (and knew!) many theorems without deriving them from the axioms. The observation that our knowledge of mathematics doesn't flow from the fundamental axioms to the theorems goes back at least to Russell -- who emphasized that the logical order isn't the same as the epistemological order, that the axioms might gain support from the familiar theorems they generate, not vice versa[5] -- and is prominent in Zermelo -- who defended the Axiom of Choice on the basis of its consequences.[6] As is often noted, a well-known fact of arithmetic or analysis or geometry may be considerably more certain than the axioms of set theory from which it derived. For that matter, to make a Wittgensteinian point,[7] in most cases we only believe that there is a proof of a certain mathematical theorem from ZFC because we believe the theorem on the basis of its ordinary proof and we believe that all theorems of mathematics are ultimately provable from ZFC! So this purported foundational use of set theory, as the **Epistemic Source** of all mathematical knowledge, is a failure. But this casts no doubt on the **Meta-mathematical Corral**.

---

[5] See Russell [1907].

[6] See Zermelo [1908a].

[7] Wittgenstein ([1978], Part III) is thinking of the relations between ordinary decimal calculations and the corresponding inferences in, say, the notation of *Principia Mathematica*, but the upshot is the same.

The other purported foundational role for set theory that seems to me spurious is what might be called the **Metaphysical Insight**. The thought here is that the set-theoretic reduction of a given mathematical object to a given set actually reveals the true metaphysical identity that object enjoyed all along. Benacerraf famously argued that this can't be right, because, for example, Zermelo took the natural numbers to be ø, {ø}, {{ø}}, … , von Neumann took them to be ø, {ø}, {ø, {ø}}, … , and there's no principled reason to choose one over the other.[8] There are practical reasons to prefer the von Neumann ordinals -- they generalize easily to the transfinite, for example -- but this sort of thing isn't an indicator of 'what the numbers really are'. Or so the argument goes.

Of course the practice of set theory is filled with even more arbitrary choices, like the conventional preference for the Kuratowski ordered pair. It's worth noticing that Kunen speaks of 'defining mathematical concepts', not identifying mathematical objects, and Enderton, who does refer to 'mathematical objects', only speaks of how they 'can be defined' and how theorems about them 'can be viewed'. In yet another textbook, Moschovakis makes the thought behind this circumspection explicit:

> A typical example of the method we will adopt is the 'identification' of [the geometric line] with the set … of real numbers. … What is the precise meaning of this 'identification'? *Certainly not that points are real numbers.* … What we mean by the 'identification' of [the line] with [the reals] is that the correspondence … gives a **faithful representation** … which allows us to give arithmetic definitions for all the useful geometric notions and to study the mathematical properties of [the line] **as if points were real**

[8] See Benacerraf [1965].

> **numbers**. … In the same way, we will discover within the universe of sets *faithful representations* of all the mathematical objects we need, and we will study set theory …[9] **as if all mathematical objects were sets***.* (Moschovakis [1994], pp. 33-34, emphasis in the original)

The trick, in each case, is to identify the conditions that a 'faithful representation' must satisfy. For the case of ordered pairs, this is easy: two of them should be equal iff their first elements are equal and their second elements are equal. The case of the natural numbers is more demanding: a set of sets with its operations should satisfy the (full second-order) Peano Postulates. For our purposes, the simple point is that these set-theoretic reductions don't give any sort of deep metaphysical information about the nature of the line or of ordered pairs or of natural numbers, nor are they so intended. **Metaphysical Insight**, like **Epistemic Source**, is spurious, leaving only **Meta-mathematical Corral**.

But this is hardly the end of the story. The impressive mathematical innovations present in the earliest invocations of sets -- Cantor's in his work on trigonometric series and Dedekind's on ideals in the early 1870s -- aren't actually foundational in character, but around the same time, Dedekind also undertook a project that was explicitly foundational: to 'find a purely arithmetical and perfectly rigorous foundation for the principles of infinitesimal calculus' (Dedekind [1872], p. 767). Charged to teach the subject, Dedekind laments 'the lack of a truly scientific foundation', finds himself forced to take 'refuge in geometric evidence' which 'can make

---

9 In the ellipsis, Moschovakis writes, 'on the basis of the lean axiom system of Zermelo', perhaps gesturing toward the **Meta-mathematical Corral**.

no claim to being scientific' (ibid.). Comparing the line with the rational numbers, we see in the rationals 'a gappiness, incompleteness, discontinuity' and in the line an 'absence of gaps, completeness, continuity', so the key question emerges:

> In what then does this continuity consist? Everything must depend on the answer to this question, and only through it shall we obtain a scientific basis for the investigation of *all* continuous domains. (Dedekind [1872], p. 771)

The then-current understanding rested only on those geometric intuitions, and on

> vague remarks upon the unbroken connection in the smallest parts [by which] obviously nothing is gained. (Ibid.)

So, the challenge was to replace these unscientific vagaries with

> a precise characteristic of continuity that can serve as the basis for valid deductions. (Ibid.)

And this, of course, is what Dedekind goes on to provide, in his theory of cuts in the rationals -- using set-theoretic machinery.

At first glance, this may look like just one more instance of the set-theoretic reduction that underlies the **Meta-mathematical Corral**, but in fact there's something more going on. It isn't that we have an explicit mathematical item -- the ordered pair, or the numbers as described by Peano -- which we then 'identify' with a set that can play the same role, do the same jobs. Instead, in this case, we have a vague picture of continuity that's served us well enough in many respects, well enough to generate and develop the calculus, but now isn't precise enough to do what it's being called upon to do: allow for rigorous proofs of the fundamental theorems. For that we need something more exact, more precise, which Dedekind supplies. This isn't just a set-theoretic surrogate, designed to reflect the features of the pre-theoretic item; it's a set-theoretic improvement, a set-theoretic replacement of an imprecise notion with a precise one. So here's another foundational use of set theory: **Elucidation**. The replacement of the imprecise notion of function with the set-theoretic version is another well-known example.[10]

To isolate another productive foundational use of set theory, we need to trace some of the profound shifts in the theory and practice of mathematics that took place over the course of the 19th century, starting with geometry. Since the ancients, geometry had been closely associated with ordinary diagrams, and late in the 1700s, Kant developed his elaborate theory of spatiotemporal intuition to undergird this approach. This comforting view of the matter was challenged early in the 19th century, as it became clear how much there was to be gained by viewing ordinary geometry from the richer, projective point of view -- with its 'imaginary points' (points with complex numbers as coordinates) and 'points at infinity' (points where parallel lines meet). Of course these new points can't be pictured, so the work was roundly resisted at first on the grounds that 'it keeps itself too much aloof from all intuition, which is the essential trait of mathematical knowledge'.[11] The very notion of invisible, unvisualizable 'points' where two disjoint circles or two parallel lines in fact 'intersect' was not only contrary to the intuitive ground of geometry, but to plain common sense. Still, as Nagel

---

[10] For a quick overview of the history, see [1997], pp. 118-126.

[11] The remark comes from Möbius, he of the strip. Quoted in Nagel [1979], p. 219.

remarks in his historical survey, on the closely related subject of negative and complex numbers:

> Scandal or not, there were few mathematicians who did not recognize their value or whose logical consciences were so tender than they would not use them. (Nagel [1979], p. 202)

The same could be said of the new projective geometry.

I bring up this episode because the ‘scandal’ was resolved in mid-century by von Staudt, using proto-set-theoretic techniques, in particular a precursor of the method of equivalence classes: for example, the point at infinity where two horizontal lines meet is identified with what we’d now see as the set of lines parallel to these two, and a given point at infinity is on a particular line if that line is in the set with which that point at infinity has been identified. In this and related ways, von Staudt managed to build surrogates for heretofore suspicious, possibly dangerous new items (like points at infinity) out of uncontroversial, unproblematic materials (ordinary lines), and to redefine the relevant relations so as to validate the existing, successful theory. His goal in all this is to remove any queasiness we might have about the legitimacy or coherence of the new, un-intuited items.

As time went by, it became clear that the construction tools needed for this ‘building’ process -- tools von Staudt regarded as ‘logical’ -- were actually set-theoretic in character. Speaking of the operations codified in Zermelo’s axioms, Burgess writes:

> A crucial fact … is that *these are essentially the existence assumptions needed to get new spaces or number systems or whatever from old ones* … in the manner of nineteenth-century introduction of auxiliaries for the study of traditional spaces or number systems. Indeed, the constructions of the auxiliaries in question can be, and now in retrospect are, viewed as

> essentially ‘set-theoretic’ constructions, though some of them actually antedate Cantor. (Burgess [2015], p. 76, emphasis in the original)

This striking fact -- that the methods of von Staudt and others all fall within the few closure principles used by the early set theorists and codified by Zermelo -- this fact is what eventually made possible what we now know as the set-theoretic reduction of classical mathematics.

But the story of queasiness-removal doesn’t end there. In the late 1800s, pure mathematics was on the rise, and with it, the axiomatic method; in place of von Staudt-like constructions, new fields were introduced instead by an explicit set of axioms. In his comprehensive history, Kline describes the situation this way:

> Mathematics, from a logical standpoint, was by the end of the nineteenth century a collection of structures each built on its own system of axioms. … As long as mathematics was regarded as the truth about nature, the possibility of contradictory theorems … would have been regarded as absurd. (Kline [1972], p. 1038)

-- but without that worldly backing, the question of which of these axiom systems could be trusted became acute. This new queasiness could best be removed by a proof of consistency, and set theory again presented itself, now as the source for such proofs. We return in a moment to the difference between von Staudt-like surrogates and axiomatic consistency proofs -- roughly, between proving from ZFC that there’s a structure of a certain sort, and proving from ZFC that there’s a model that thinks there’s a structure of a certain sort -- but for now what’s important is that neither of these effectively removes queasiness unless ZFC itself is known to be trustworthy.

Zermelo felt the force of these considerations, remarking that

> I have not yet … been able to prove rigorously that my axioms are consistent, though this is certainly very essential; instead I have had to confine myself to pointing out now and then that the antinomies discovered so far vanish one and all if the principles here proposed are adopted as a basis. But I hope to have done at least some useful spadework hereby for subsequent investigations of such deeper problems. (Zermelo [1908], pp. 200-201)

We now smile, perhaps a bit wistfully, at this optimism.[12] In the decades since Gödel dashed these fond hopes, the hierarchy of large cardinals has arisen as a measure of consistency strength, and the early foundational goal of conclusive queasiness-removal has given way to a more nuanced matter of **Risk Assessment**. So, for example, in the abstract for a recent ASL talk, Voevodsky speaks of the role of set theory in his program of ‘univalent foundations’:[13]

> Univalent foundations provide a new approach to the formal reasoning about mathematical objects. The languages which arise in this approach are much more convenient for doing serious mathematics than ZFC at the cost of being much more complex. In particular, the consistency issues for these languages are not intuitively clear. Thus ZFC retains its key role as the theory which is used [to] ensure that the more and more complex languages of the univalent approach are consistent. (Voevodsky [2014], p. 108)

Or as he puts it, more carefully, in the slides for this talk:

> Set theory will remain the most important benchmark of consistency. … each new addition to the … language will require formal ‘certification’ by showing, through formally constructed interpretation, that it is at least as consistent as ZFC. (Voevodsky [2013], slide 21)

Obviously this generalizes to ‘at least as consistent as ZFC + one or another large cardinal’.

---

[12] Recall footnote 3, above.

[13] For the foundational role ‘univalent foundations’ might itself be intended to play, see footnote 38.

Notice that **Risk Assessment** in either form isn't the same as **Meta-mathematical Corral**: the point isn't to round up all classical mathematical items into one simple package, so as to prove something about all of it all at once, but to assess a particular new, somehow dangerous or suspicious item to determine just how risky it is. And it differs from **Elucidation** as well: von Staudt had before him a perfectly functional geometric practice with ideal points; his task wasn't to make that practice more precise, and thus more functional, but to reproduce it chapter-and-verse in a way that was less worrisome; conversely, Dedekind's concern wasn't that the real numbers were somehow worrisome, but that they weren't sufficiently precise to support the practice, weren't sufficiently functional. So we have at this point these three distinct foundational uses for set theory, along with a pair of spurious ones -- **Metaphysical Insight**, and **Epistemic Source**.

There remains one more, quite familiar line of foundational thought, namely the idea that set theory provides decisive answers to questions of ontology[14] and proof: if you want to know whether or not a so-and-so exists, see whether one can be found in V; if you want to know whether or not such-and-such is provable, see whether it can be derived from the axioms of set theory. (In fact, both of these are provability conditions: a so-and-so 'can be found in V' iff the

---

[14] This term isn't intended in any philosophically loaded way: I just mean what the practice asserts to exist, leaving the semantic or metaphysical issues open. Mac Lane ([1981], p. 468) and Feferman ([1977], p. 151) both see set theory as inseparable from a kind of Platonistic metaphysics, but I've argued that employing a rich set-theoretic ontology is consistent with being an Arealist (as in [2011]).

existence of a so-and-so can be proved from the axioms.) This is sometimes expressed with the rhetorical flourish that set theory is the ‘final court of appeal’ on matters of proof and ontology. I should confess that I’ve indulged in this flourish myself ([1997], p. 26), making it sound as if classical mathematics must bow to the dictates of set theory, but in practice I’ve taken this foundational role to place methodological constraints on set theory, the founding theory, not on classical mathematics, the theory to be founded. In particular, I argued that set theory, if it was to play this role, should be as generous as possible -- so as not to curtail pure mathematics -- and should be given by one unified theory that’s as decisive as possible -- so as to provide unequivocal answers to questions of ontology and proof.[15]

Fearing, then, that the ‘final court of appeal’ is something of an exaggeration, let’s look a bit more closely to discern what foundational uses are actually in play. At the very least, there’s the plain sociological fact that derivation from ZFC is generally regarded as standard of proof in mathematics: in practice, the availability of the axioms of ZFC goes without saying; if stronger assumptions are in play, this is explicitly acknowledged;[16] if only weaker assumptions are needed, this is noted to give a more nuanced picture of the dependencies involved. Burgess observes:

---

[15] These are the methodological maxims MAXIMIZE and UNIFY from [1997].

[16] Burgess ([2015], p. 177, footnote 11) notes an exception: proofs appealing to Grothendieck’s work sometimes omit what comes to an appeal to inaccessibles. See §II below.

> There are … no official censors preventing a group of dissidents from founding a journal of their own, in which as a matter of editorial policy results must be proved according to the group's restrictive standards (or as the case may be, results may be proved making use of the group's preferred additional hypotheses), rather than presented as they would be in a journal enforcing the orthodox standard … No one dissident school of thought, however, produces work of sufficient volume at a sufficient pace to keep a journal of high standards following such a policy coming out regularly. (Burgess [2015], p. 118)

In this foundational role, then, formal derivation in set theory serves as a **Shared Standard** of what counts as a proof.

But what lies behind this sociological fact? Partly there's the recognition that formal derivation turns out to be a good mathematical model for the scope of human proving activity, but why from these particular axioms? To shed light on this question, recall the 'ontological' component of the 'final court of appeal': there is a so-and-so if one can be found in V, if the existence of a so-and-so can be proved from the axioms. A few pages back, in connection with **Risk Assessment**, we noted a contrast between queasiness-removal by outright existence proof and queasiness-removal by consistency proof. Now we find the purported standard of existence apparently promoting von-Staudt-like construction over the more lenient Hilbert-like idea that consistency of the appropriate set of axioms is enough. To illustrate with an example, suppose a mathematician wants to know: is there a definable (projective) well-ordering of the reals? In ZFC alone, the question can't be answered, but assuming, as many set theorists do, that ZFC + LCs is the appropriate measure, the answer is no. Still, the opposite answer can be had in L. In light of that fact, would we really want to shut the door on this mathematician? For that matter, why shouldn't we follow Hilbert and open that door to the existence claims of any consistent set of axioms? Why does 'final court' insist that we restrict ourselves to exactly what happens in V?

The source of this more stringent 'final court' criterion is simple: the branches of modern mathematics are intricately and productively intertwined, from coordinate geometry, to analytic number theory, to algebraic geometry, to topology, to modern descriptive set theory (a confluence of point-set topology and recursion theory), to the kind of far-flung interconnections recently revealed in the proof of Fermat's Last Theorem. What's needed is a single arena where all the various structures studied in all the various branches can co-exist side-by-side, where their interrelations can be studied, shared fundamentals isolated and exploited, effective methods exported and imported from one to another, and so on. Burgess puts the point forcefully:

> Interconnectedness implies that *it will no longer be sufficient to put each individual branch of mathematics separately on a rigorous basis* … To guarantee that rigor is not compromised in the process of transferring material from one branch of mathematics to another, it is essential that the starting points of the branches being connected should … be compatible. … The only obvious way to ensure compatibility of the starting points … is ultimately to derive all branches from a common, unified starting point. (Burgess [2015], pp. 60-62)

Set theory's universe, V, provides the **Generous Arena** in which all this takes place, and that's why the 'final court' condition takes the form it does: to be a full participant in mathematical interaction, a so-and-so must appear along-side the full range of its fellows, with all the tools of construction and interaction fully available.[17]

[17] Of course **Shared Standard** and **Generous Arena** depend on the same facts of set-theoretic reduction as **Meta-mathematical Corral**: that formal proof is a good model of provability by humans and that the axioms of set theory codify

Viewed in this light, our flat answer to the mathematician's question deserves a bit of shading. A definable well-ordering of the reals occurs in L, a well-understood structure where all the axioms of ZFC are satisfied. This means that L itself is a fairly generous arena: all the usual constructions of ZFC are available; all the standard theorems from all the familiar branches of the subject are in place; so serious mathematical work can be carried out in the presence of a definable well-ordering. The drawback is that care has to be taken with export and import. But the pure Hilbert-style case is different: proving that there is a model for a set of axioms that implies the existence of a so-and-so can provide **Risk Assessment**, but it doesn't by itself install a so-and-so in the **Generous Arena** where classical mathematics takes place.

Stripped of its pretensions, then, the 'final court' condition comes down to this: a **Shared Standard** of proof designed to generate a **Generous Arena** for the pursuit and flourishing of pure mathematics. From this point of view, the requirement that assumptions beyond ZFC be noted explicitly makes perfect sense: our mathematician is welcome to work with his definable well-ordering in L as long as he recognizes that his conclusions can't be freely exported to the more standard arena with ZFC alone, and that work predicated on the popular assumption of large cardinals can't be imported. On the plus side, he

---

the fundamental assumptions of classical mathematics. What separates them are the uses to which these facts are being put: **in Meta-mathematical Corral**, 'derivable in ZFC' functions as model for 'provable in classical mathematics'; in **Shared Standard**, it's used as a benchmark for what counts as a legitimate informal proof; in **Generous Arena**, V brings all the objects and methods of classical mathematics together for fruitful interaction. As foundational uses, these are distinct.

gains all the benefits of our existing understanding of the details of life in L, and his work can be understood as further illuminating what goes on there. Work in ZFC + V=L has obvious value, but of course some extensions of ZFC are be more mathematically rewarding than others![18]

In sum, then, we've collected an array of important uses of set theory that ought to qualify as 'foundational' -- **Meta-mathematical Corral**, **Elucidation**, **Risk Assessment**, **Shared Standard** and **Generous Arena** -- as well as a pair that are spurious -- **Metaphysical Insight** and **Epistemic Source**. The famed set-theoretic reduction of classical mathematics lies in the background for most of this, fruitful and spurious alike. Details aside, we see that the remarkable fact of the reduction doesn't, by itself, dictate any particular foundational use. For this reason, it seems to me counter-productive to begin from the question: does a given theory provide a foundation for classical mathematics? Rather, we should be asking: what foundational purposes does the given theory serve, and how? With this in mind, let's turn to set theory's famous rival.

II. Foundational uses of category theory

Category theory was introduced in the 1940s, by which time the notion of set-theoretic foundations had become mainstream orthodoxy. Like Cantor's initial appeals to sets, Eilenberg and Mac Lane's categories emerged in the pursuit of straightforwardly mathematical

[18] For some purposes, it's useful to consider arenas not quite as generous as full ZFC, e.g., L(R), where Choice is false, but all sets of reals are Lebesgue measurable.

goals. Foundational concerns first came into the picture in the form of criticisms of set theory's foundational aspirations. Mac Lane grants the effectiveness to date of set-theoretic foundations:

> The prior situation in the foundations of Mathematics had in one respect a very simple structure. One could produce one formal system, say Zermelo-Fraenkel set theory, with the property that all ordinary operations of practising Mathematicians could be carried out within this one system and on objects of this system. … 'every' Mathematical object was or could be defined to be a set, and … all of the arguments about these objects could be reduced to the axioms of ZF set theory. … This one-formal-system 'monolithic' approach has … been convenient for specialists on foundations ever since Frege and Whitehead-Russell. On the one hand, all the classical nineteenth century problems of foundations (the construction of integers, real numbers, analysis …) could be stated in this one system. On the other hand, alternative formal system[s] could … be tested by comparison (as to strength or relative consistency) with this one system. (Mac Lane [1971], pp. 235-236)

Here we find gestures toward **Generous Arena**, **Risk Assessment**, **Elucidation** (for the case of the reals)[19], **Meta-mathematical Corral** (in the usefulness to 'specialists in foundations') and **Shared Standard**[20]. Mac Lane's only complaint is that 'this happy situation no longer applies to the practice of category theory' (Mac Lane [1971], p. 236).

What is it about category theory that purportedly ends the reign of set-theoretic foundations?

> Categorical algebra has developed in recent years as an effective method of organizing parts of mathematics. Typically, this sort of organization uses notions such as that of the category **G** of all groups. This category consists of two collections: The collection of all groups G and the collection of all

[19] Cf. Mac Lane [1986], p. 362: 'this approach to Mathematics has the advantage that every concept can be made absolutely clear and explicit'.

[20] Cf. Mac Lane [1986], p. 377: 'we have now stated an absolute standard of rigor: A Mathematical proof is rigorous when it is (or could be) written out in the first order predicate language [with membership as the only non-logical symbol] as a sequence of inferences from the axioms ZFC, each inference made according to one of the stated rules'.

> homomorphisms … of one group G into another one; the basic operation in this category is the composition of two such homomorphisms. To realize the intent of this construction it is vital that this collection **G** contain all groups; however, if 'collection' is to mean 'set' … this intent cannot be directly realized. (Mac Lane [1971], p. 231)

Because there is no set of all groups, set theory can't properly found category theory.[21] So the argument goes.

To overcome this problem, Grothendieck devised a system of 'universes' essentially equivalent to Zermelo's hierarchy of V_κ's for κ inaccessible,[22] and all parties agree that surrogates for any actual use of category theory can be found within some V_κ.[23] Burgess writes:

> For applications … one doesn't need a category of literally *all* groups … It is always enough to have a category of 'enough' groups, though how many is enough may vary from application to application. … Grothendieck's hypothesis is that every set, however large, belongs to some local universe [that is, some V_κ]: The 'global universe' … is simply the union of increasingly large local universes. (Burgess [2015], p. 174)

Why doesn't this settle the matter? Assuming that **Risk Assessment** is the foundational use of set theory in question here, as Mac Lane sometimes suggests,[24] then we have our answer: category theory is no

---

[21] A move to NGB would introduce the category of all groups as a proper class, but it would still leave out important functor categories. See MacLane [1971].

[22] See Zermelo [1930].

[23] See, e.g., Feferman [1977], p. 155. Grothendieck's accomplishment here runs parallel to von Staudt's: show that the job of a worrisome item (point at infinity, category of all groups) can be carried out with an uncontroversial item (a set of lines, a set of 'enough' groups).

[24] See Mac Lane [1986], p. 406: 'in one sense a foundation is a security blanket: If you meticulously follow the rules laid down, no paradoxes or contradictions will arise'. He goes on to point out that a risk of zero can't be achieved in this way, but the quotation in the text from Mac Lane [1971], p. 237, explicitly recognizes the worth of **Risk Assessment**: 'alternative formals system[s] could … be tested by comparison (as to strength or relative consistency) with this one system.'

worse off than 'ZFC + many Inaccessibles', which doesn't take us far beyond ZFC itself.[25] With this mild extension of ZFC, **Meta-mathematical Corral** would also be achieved, nor is there any obvious difficulty for any of the other foundational uses of set theory.

But Mac Lane knows all this and clearly isn't satisfied:

> Given any universe *U'*, one can always form the category of all categories within *U'*. This is still not that will-of-the-wisp, the category of all categories *überhaupt*. (Mac Lane [1971], p. 234)

What exactly is the imagined deficit? For any particular use of the category of 'all' groups or the category of 'all' categories, there's a set-theoretic surrogate that does the job. The objection here appears to be that the surrogate isn't the real thing. But why should it be? In this form, the objection begins to sound analogous to the complaint, in Benacerraf's context, that the von Neumann ordinals aren't acceptable set-theoretic surrogates for the natural numbers because the actual 2 isn't an element of the actual 3! In other words, Mac Lane appears to be drifting into the demand that a foundation provide **Metaphysical Insight** -- a demand that set theory never properly took on in the first place.

However that may be, the problem of 'founding' unlimited categories was taken up by Feferman in the late 1960s and revisited

---

[25] In fact, Burgess ([2015], p. 176) points out that the inaccessibles aren't strictly necessary, that they can be removed by careful use of reflection. But, he observes, 'in Grothendieck's kind of work the intellectual faculties are being strained to their uttermost limit, and one doesn't want the distraction of any sort of bookkeeping requirements'.

most recently in his [2013].[26] Feferman explicitly identifies the issue as one of **Risk Assessment**, and the theory whose risk he's interested in assessing is one that allows the formation of

> the category of all structures of a given kind, e.g., the category … of all groups, … of all topological space[s], … of all categories. (Feferman [2013], p. 9)

In other words, not just, for example, the category of 'enough groups' (as Burgess puts it) for some practical purpose, but the category of *all* groups. These

> unrestricted notions … are mathematically reasonable on the face of it and do not obviously lead to paradoxical conclusions.[27] (Ibid.)

The challenge, then, is to combine them with all the usual objects and operations that category theorists call upon in the course of their mathematical work, and to prove the resulting theory consistent relative to 'a currently accepted system of set theory' (ibid.).

Despite incremental partial successes from Feferman, McLarty and others over the years, this problem remained open until just recently. In his [2015], Ernst shows that any theory that allows the formation of the category of all graphs and that includes the required

[26] Feferman ([1977], p. 155) characterizes the deficit of the Grothendieck-style reduction as 'aesthetic', but it's hard to see why a foundational use has to be beautiful in some way or other.

[27] It might appear that a paradox is ready to hand: the category of categories is a category; the category of groups isn't a group; what about the category of all non-self-membered categories? The trouble with this tempting line of thought is that 'membership' isn't native to the world of naïve category theory. We could define what it is for the category of categories to be 'self-membered': it's for the category of categories to be a category. Likewise the category of groups is 'non-self-membered' because it's not a group. But categories in general aren't given in the form 'the category of all X's', so a general membership relation can't be defined in this way. As Feferman ([2013], p. 9) remarks, 'There is no sensible way … to form a category of all categories which do not belong to themselves'.

mathematical staples is in fact inconsistent.[28] So regardless of whether Mac Lane's demand that set theory provide for unlimited categories is based on **Metaphysical Insight** or on another motive, it's a demand that can't be met -- not because of any shortcoming on set theory's part, but because the objects he demands are themselves inconsistent, beyond the reach of any consistent foundation.

Nevertheless, as a historical matter, many calls to replace set-theoretic foundations with category-theoretic foundations arose from the ill-fated hope that category theory, unlike set theory, could meet this impossible demand:

> One might hope for some … new foundational system (the category of categories?) within which all the desired objects live. (Mac Lane [1971], p. 236)

Though Lawvere (in [1966]) takes exactly this thought as his title -- 'The category of categories as a foundation for mathematics' -- his primary focus is actually on a foundational use distinct from those considered so far. But before we turn to this new sense of 'foundation', we need to ask how this theory of categories, or really an improved descendant due to McLarty [1991],[29] fares on the various, more familiar foundational uses that set theory has been seen to underwrite.

---

[28] The proof is non-trivial, structured roughly along the lines of a proof that there is no set of all sets via Cantor's theorem: Ernst assumes there is a category of all graphs ***R*** (actually reflexive graphs, but the result generalizes to all graphs), concocts a certain exponential, shows there can't be a map of ***R*** onto this exponential, then uses the fact that the exponential is a substructure of ***R*** to show that there is a map of ***R*** onto the exponential.

[29] The system of McLarty [1991] avoids the shortcomings identified in Isbell's [1967] review of Lawvere [1966].

The Category of Categories as a Foundation (CCAF) begins from a simple axiomatic theory that codifies the usual category-theoretic machinery to yield

> the usual results of general category theory … as hypotheticals of the form 'If **A** is a cartesian closed category, then … ' or 'If **A** is a non-trivial topos, then … '. … On the other hand, the axioms prove the existence of few specific cases of these general results. (McLarty [1991], pp. 1258-1259)

This minimal background theory is then supplemented 'with axioms on particular categories or functors' (ibid.), depending on the intended use. For foundational purposes, one adds the axioms of the Elementary Theory of the Category of Sets (ETCS) -- also introduced by Lawvere (in his [1964]) -- which characterize a weak set theory (equivalent to ZC with bounded separation). Mathias [2001] describes various shortfalls of ETCS, mostly the failure of iterative constructions, but these can be achieved by adding a category-theoretic version of Replacement to generate a system equivalent to ZFC, and large cardinal axioms to duplicate the full force of ZFC + LCs.[30]

If the foundational uses are to be recovered in this way, opponents might argue that ZFC + LCs is doing the true foundational work, and this Augmented-ETCS is just piggy-backing,[31] but both Mac Lane and McLarty argue that ETCS is itself an independent, thoroughly category-theoretic system:

---

[30] See McLarty [2004] for these extensions of ETCS. As Ernst [201?] points out, this augmentation is often resisted by advocates of category-theoretic foundations, because the mathematics it's designed to accommodate is regarded as expendable. This would be a new use of a foundation -- as a way of pruning mathematics! -- directly opposed to the thought behind MAXIMIZE (see footnote 15).

[31] Indeed they have so argued. See, e.g., Hellman [2003].

> The standard ‘foundation’ for Mathematics start[s] with sets and their elements. It is possible to start differently, by axiomatizing not elements of sets but functions between sets.[32] This can be done using the language of categories and universal constructions. (Mac Lane [1986], p. 398)

> ETCS is a set theory. It is not a membership-based set theory like ZF. It is a function-based set theory. Mac Lane generally uses the phrase ‘set theory’ to mean ZF, a habit of more than thirty years before ETCS was conceived. But we cannot let his terminology misdirect us. He is explicit that ETCS is his preferred account of sets. (McLarty [2004], p. 39)

The thought is that ZF-style set theory doesn’t enjoy exclusive rights to the pre-theoretic notion of ‘collection’. We’re to imagine ourselves starting from scratch, thinking purely category-theoretically, and devising a theory of collections in top-down function-based terms, rather than bottom-up element-based terms.

I should note that this particular objection to category-theoretic foundations is part of a broader concern about category theory’s ‘autonomy’, beginning in Feferman [1977].[33] Linnebo and Pettigrew [2011] summarize the complaints under three headings: logical autonomy -- can the theory be stated without appeal to set theory?, conceptual autonomy -- can the theory be understood without a prior understanding of set theory?, and justificatory autonomy -- can the theory be justified without appealing to set theory and/or its justifications? Though I make no claim to have grasped all the ins and outs of this debate, it seems clear that both ZFC + LCs and Augmented-CCAF (= CCAF + Augmented-ETCS) are straightforward first-order theories, each inspired by a range of intuitive, proto-

---

[32] See von Neumann [1925].

[33] The literature on this topic is copious and tangled. See Ernst [201?] for a recent overview.

mathematical notions: for set theory, these include collection, membership, iteration and a combinatorial idea of 'all possible subsets' as described, for example, by Bernays[34]; for category theory, perhaps collection, function, composition of functions, … .[35] As long as 'collection' isn't awarded exclusively to set theory by some kind of natural right, the two appear to be equally autonomous.

As for justificatory autonomy, Linnebo and Pettigrew propose that the iterative conception justifies ZFC, and argue that the category theorist is unable to come up with anything sufficiently comparable. My own view is that the iterative conception is a brilliant heuristic device, but that the justification for the axioms it suggests (and even for potential axioms it doesn't suggest!) rests on their power to further various mathematical goals of set theory, including its foundational goals.[36] This mode of justification (which Linnebo and Pettigrew don't consider) is open to category theory as well, and granting the dramatic success of category-theoretic methods in various areas of mathematics, there can be little doubt that its concepts and techniques are well-justified. So, for what it's worth, I see no threat to the autonomy of category theory in these justificatory pathways. What we've been asking here, in this exploration of

[34] Bernays [1934]. See also [1997], pp. 127-128.

[35] It might be of interest to investigate more carefully which intuitive notions each theory appeals to and which it explicates, and where these lists differ, to examine how successfully each explicates the other's primitives. For example: ZFC-style set theory explicates 'function' with the much-maligned set of ordered pairs; Mathias derides the category-theoretic treatment of 'iteration' as 'clumsy' (Mathias [2001], p. 227).

[36] See [1997], [2011].

category theory’s foundational aspirations, is whether the axioms of Augmented-CCAF are (partly) justified by their effectiveness toward the foundational goals, but our question is whether there is such a justification at all, not whether it’s parasitic on set theory.

Returning to our main theme, then: how are the foundational uses of set theory recovered in this category-theoretic context, in Augmented-CCAF? As an example, McLarty takes up the construction of reals as Dedekind cuts, concluding that

> once you get beyond axiomatic basics, to the level of set theory that mathematicians normally use, ZF and ETCS [ETCS plus Replacement?] are not merely intertranslatable. They work just alike. (McLarty [2004], p. 41)

So it seems **Elucidation** works much as before: a surrogate is found in ETCS rather than ZF, and the clarificatory benefits are pretty much the same. Presumably **Risk Assessment** makes use of the large cardinals of Augmented-ETCS in the familiar ways.[37]

I’m less sure how to think about **Meta-mathematical Corral**, **Generous Arena**, and **Shared Standard.** On the category theorist’s foundational scheme, what do we say to the mathematician who wants to know whether or not there’s a definable well-ordering of the reals? What theory do we turn to if we want to formulate questions of what can or can’t be proved in ‘classical mathematics’, or to determine conclusively whether or not a purported informal proof is legitimate? All the reduced items appear side-by-side, and theorems about them are provable, in the category satisfying Augmented-ETCS, which suggests Augmented-ETCS as a likely candidate for **Generous Arena** and the rest,

---

[37] One reason, perhaps, to resist the pruning in footnote 30.

but the fact that CCAF posits items outside that category casts doubt on this move. Perhaps Augmented-CCAF itself, then?

Surprisingly, this line of thought is undercut by passages where Mac Lane appears to deny the very desirability of **Generous Arena** in the first place, preferring a Hilbert-like approach to a von-Staudt-style construction in the case of the reals:

> This careful construction of the real numbers was long accepted as standard in graduate education in mathematics, even though many mathematicians did not much believe in it. … This viewpoint can be expressed … formally: Do not construct the reals, but describe them axiomatically as an ordered field, complete in the sense that every bounded set has a least upper bound. (Mac Lane [1981], p. 467)

Let's leave this unsettling thought aside for the moment and look first at the new foundational use that Lawvere and others see as deciding the case in favor of category theory over set theory.

It first appears as an objection to set-theoretic foundations:

> This *Grand Set Theoretic Foundation* … does not adequately describe which are the relevant mathematical structures to be built up from the starting point of set theory. *A priori* from set theory there could be very many such structures, but in fact there are a few which are dominant … natural numbers, rational numbers, real numbers … group, ring, order and partial order … . The 'Grand Foundation' does not provide any way in which to explain the choice of these concepts. (Mac Lane [1981], p. 468)

While set theory has the wherewithal to build all the mathematically important structures, its construction techniques are indiscriminate, generating a vast store of mathematically useless structures along the way and providing no guidance as to which are which. There's also some discomfort about the way in which those structures are built up:

> In the mathematical development of recent decades ones sees clearly the rise of the conviction that the relevant properties of mathematical objects are those which can be stated in terms of their abstract structure rather than in terms of the elements

> which the objects were thought to be made of. (Lawvere [1966], p. 1)

Set-theoretic constructions introduce a lot of irrelevant structure: for example, a Dedekind cut is a set of rationals, which are equivalence classes of sets of pairs of natural numbers, which are ordinals, and so on, but none of this detail has any direct connection to their intended behavior as surrogates for the reals, as the availability of alternatives like Cauchy sequences serves to demonstrate.

Complaints like these about set-theoretic foundations led to the suggestion that category theory might be better suited to the task. Lawvere hopes for a foundation that will bring 'abstract structure' to the forefront:

> The question thus naturally arises whether one can give a foundation for mathematics which expresses wholeheartedly this conviction concerning what mathematics is about, and in particular, in which classes and membership in classes do not play any role. (Lawvere [1966], p. 1)

Mac Lane stresses the replacement of static elements with dynamic functions:

> There are other possibilities. For example, the membership relation for sets can often be replaced by the composition operation for functions. This leads to an alternative foundation for Mathematics upon categories … much of Mathematics is dynamic, in that it deals with morphisms of an object *L* into another object of the same kind. Such morphisms … form categories, and so the approach via categories fits well with the objective of organizing and understanding Mathematics. (Mac Lane [1986], p. 359)

McLarty emphasizes that this particular foundational use should reflect what mathematicians actually do in their mathematical lives:

> Mac Lane had a different idea of foundations … He took a *foundation of mathematics* to be a body of truths which organize mathematics as do I here. More specifically, that is truths

> which actually serve in practice to define the concepts of mathematics and prove the theorems. I do not mean merely truths which could in some principled sense possibly organize the practice but truths actually used in textbooks and journal articles, and discussed in seminar rooms and over beer, so their notions do occur in practice. (McLarty [2013], p. 80)

The goal here is a foundation for mathematics that will capture the fundamental character of mathematics as it's actually done, that will guide mathematicians toward the truly important concepts and structures, without getting bogged down in irrelevant details. I propose **Essential Guidance** as an awkward label for this foundational goal, hoping to highlight its two aspects: such a foundation is to reveal the fundamental features -- the essence, in practice -- of the mathematics being founded, without irrelevant distractions; and it's to guide the progress of mathematics along the lines of those fundamental features and away from false alleyways.[38]

Of course, Lawvere, Mac Lane and McLarty are entirely correct when they point out that set theory does not provide **Essential Guidance**.[39] For the record, though, we should note that this fact in no way compromises its other foundations uses: the lack of guidance and the presence of extraneous details don't undercut **Meta-mathematical Corral**, **Elucidation**, **Risk Assessment**, **Generous Arena**, or

---

[38] It may be that this is the use Voevodsky has in mind for 'univalent foundations': 'The languages which arise in this approach are much more convenient for doing serious mathematics than ZFC' (Voevodsky [2014], p. 108).

[39] Presumably, neither does ETCS. I take it ETCS and its extensions are to be called upon for some foundational jobs, like Elucidation and Risk Assessment, while something broader, like Augmented-CCAF, is to provide Essential Guidance.

Shared Standard.[40] They do, however, keep set theory from providing **Metaphysical Insight**; that's the moral of Benacerraf and the rest. As floated in passing earlier, could it be that something like this is part of what's at issue for some of these category-theoretic thinkers, some hope of uncovering the true nature of the mathematical structures? However that may be, it does seem likely that a version of **Epistemic Source** is implicit in the kind of guidance they have in mind: our 'understanding' of the mathematics is presumably based in our grasp of its fundamental concepts and techniques. In any case, I don't see that anything like **Essential Guidance** was among the ambitions of set-theoretic foundations in the first place, so to count them as 'failures' of set theory is to fault a cat for not being a dog. But if a category-theoretic foundation *does* deliver on this desideratum, it would enjoy a dramatic advantage over set theory.

Alas, it isn't clear that category theory does deliver on **Essential Guidance**. I think it's agreed on all sides that a category-theoretic conceptual framework is a remarkably effective way of thinking in fields like algebraic topology and algebraic geometry -- no one would suggest that specialists in these areas would do better to think more like set theorists. The contested claim -- if category-theoretic foundations are to capture what's most fundamental, to guide us to the mathematically significant concepts -- is that all mathematicians would do better to think like category theorists. But even Mac Lane, in his sober moments, doesn't believe this:

---

[40] Cf. Mathias [1992], p. 115: 'to reject a claim that set theory supplies a universal mode of mathematical thought … need not compel one to declare set theory is entirely valueless'.

> Categories and functors are everywhere in topology and in parts of algebra, but they do not yet relate very well to most of analysis. (Mac Lane [1986], p. 407)

Analysis is where set theory first arose, much as category theory arose in algebra, so it's no surprise that set-theoretic thinking is more suitable there.

Mathias, who has done much to bring out the difficulties of category theory in analysis,[41] gives this reading of Mac Lane's resistance to set theory:

> I would guess that his reason is not so much that he objects to the ontology of set theory but that he finds the set-theoretic cast of mind oppressive and feels that other models of thought are more appropriate to the mathematics he wishes to do. (Mathias [1992], p. 115)

But the analyst or set theorist might well feel the same way about the category-theoretic cast of mind:

> The CAT camp [who believe category theory has 'the one true view of pure mathematics'] may with justice claim that category theory brings out subtleties in geometry to which set theory is blind. … The SET camp [who believe set theory has 'the one true view of pure mathematics'] may with equal justice claim that set-theoretic analysis brings out subtleties to which the CAT camp is blind. (Mathias [2001], pp. 226-227)

It seems that both these camps are chasing a false goal: a foundation that delivers **Essential Guidance**, a single understanding of what mathematics is, a single recommendation on how mathematicians should think.[42]

---

[41] See Mathias [2000], [2001]. Beyond analysis, Ernst [2014] explores some potentially problematic examples from graph theory.

[42] Mathias himself advocates 'unity', akin to **Generous Arena** (see below), but not 'uniformity', or **Essential Guidance**: 'Is it desirable to press mathematicians all to think in the same way? I say not: if you take someone who wishes to become a set theorist and force him to do (say) algebraic topology, what you get is not a topologist but a neurotic' (Mathias [1992], p. 113).

Now the odd thing is that Mac Lane would apparently agree:

> We conclude that there is as yet no simple and adequate way of conceptually organizing all of Mathematics. (Mac Lane [1986], p. 407)

In other words, no foundation, not even a category-theoretic foundation, has a corner on **Essential Guidance**. Yet we've seen that he champions category-theoretic foundations over set-theoretic foundations because its approach 'fits well with the objective of organizing and understanding Mathematics' (Mac Lane [1986], p. 359). Perhaps this apparent conflict actually dovetails with the suggestion above that Mac Lane might reject some of the foundational desiderata that set theory successfully satisfies, for example, **Generous Arena**. If we deny the importance of bringing (surrogates for) all mathematical structures into a shared context where they can be compared side-by-side, their interrelations revealed, methods and results imported and exported, and so on, then the Hilbert-style approach of leaving the axiom system for each mathematical structure to stand separately, on its own, might seem preferable to the von-Staudt-style construction. The need for a single foundational scheme would be disappear, and the possibility of a range of schemes might seem attractive: 'the variety of proposals for organizations reflects the diversity and richness of Mathematics' (Mac Lane [1986], p. 407).

If this is right, then Mac Lane's ultimate point is more radical than the simple claim that category theory is a better foundation for mathematics than set theory. His point, rather, is that the most important function of a foundation is **Essential Guidance**, and if this conflicts with some of the traditional foundational uses set theory has been put to, like **Generous Arena**, then so much the worse for those traditional uses. I've tried to sketch how set theory came to be seen as the **Generous Arena,** why this was thought to be important, what work it did, and so on. Burgess [2015] gives a more extended review of the history; Mathias gives a mathematician's eye view:

> One of the remarkable things about mathematics is that I can formulate a problem, be unable to solve it, pass it to you; you solve it; and then I can make use of your solution. There is a unity here: we benefit from each other's efforts. … But if I pause to ask *why* you have succeeded where I have failed … I find myself faced with the baffling fact that you have thought of the problem in a very different way from me: and if I look around the whole spectrum of mathematical activity the huge variety of styles of thought becomes even more evident. … The purpose of foundational work in mathematics is to promote the unity of mathematics; the larger hope is to establish an ontology within which all can work in their different ways. (Mathias [1992], pp. 113-114)

I leave it to the reader's conscience to decide whether **Generous Arena,** and possibly other of the traditional foundational uses, should be jettisoned.

Perhaps the best course would to stop quibbling about the word 'foundation', leave set theory to the important functions it so ably performs, and turn serious philosophical and methodological attention to the matter of distinguishing and exploring the distinctive 'ways of thinking' that flourish in different areas of pure mathematics. Directly after his claim that category theory 'fits well with the objective of organizing and understanding Mathematics', Mac Lane continues, 'That, in truth, should be the goal of a proper *philosophy* of Mathematics' (Mac Lane [1986], p. 359, emphasis added). Perhaps the unremarked shift from 'foundation of Mathematics' to 'philosophy of Mathematics' is telling; perhaps what Mac Lane is really after isn't so much a replacement for set theory as a broader appreciation for, and more philosophical/methodological attention to, the organizational and expressive powers of category theory. I think we can all agree that this would be a good thing!

III. The multiverse

The most recent challenge to straightforward set-theoretic foundations in V comes, oddly enough, from among set theorists themselves. To quote yet another textbook:

> Should we suppose that the continuum hypothesis, for example, has a definite truth value in a well-defined canonical model? Or is there a range of models in which the truth value of the continuum hypothesis varies, none of which has any special ontological priority? Forcing tends to push us in the latter direction. (Weaver [2014], p. 118)

Hamkins, whose position Weaver endorses, describes the situation this way:

> Our most powerful set-theoretic tools, such as forcing, ultrapowers, and canonical inner models, are most naturally and directly understood as methods of constructing alternative set-theoretic universes. A large part of set theory over the past half-century has been about constructing as many different models of set theory as possible … As a result, the fundamental objects of study in set theory have become the models of set theory, and set theorists move with agility from one model to another. … Set theory appears to have discovered an entire cosmos of set-theoretic universes. (Hamkins [2012], p. 418)

This has come to be called a 'multiverse' view. Woodin observes of one such position that

> the refinements of Cohen's method of *forcing* in the decades since his initial discovery of the method and the resulting plethora of problems shown to be unsolvable, have in a practical sense almost compelled one to adopt the generic-multiverse position. (Woodin [2011], pp. 16-17)

If this is right, our fundamental theory shouldn't be a theory of sets, but a theory of set-theoretic universes.

There are actually several different multiverse proposals currently on offer.[43] The most generous is Hamkins version:

> The background idea of the multiverse … is that there should be a large collection of universes, each a model of (some kind of) set theory. There seems to be no reason to restrict inclusion only to ZFC models, as we can include models of weaker theories ZF, ZF^-, KP, and so on, and perhaps even down to second-order number theory, as this is set-theoretic in a sense. … We want to consider that the multiverse is as big as we can imagine. At any time, we are living inside one of the universes, referred to as V and thought of as the current universe, but by various means, sometimes metamathematical, we may be able to move around in the multiverse. (Hamkins [2012], pp. 436-437)

Hamkins is out to explain mathematical experience:

> This abundance of set-theoretic possibilities poses a serious difficulty for the universe view … one must explain or explain away as imaginary all of the alternative universes that set theorists seem to have constructed. This seems a difficult task, for we have a robust experience in those worlds … The multiverse view … explains this experience by embracing them as real. (Hamkins [2012], p. 418)

Steel and Woodin, in contrast, are concerned with the claim, suggested by Weaver above, that CH has no determinate truth value.[44]

Steel observes that

> we have good evidence that the consistency hierarchy is not a mirage, that the theories in it we have identified are indeed consistent. This is especially true at the lower levels, where we already have canonical inner models, and equiconsistencies with fragments of definable determinacy. This argues for

---

[43] Väänänen [2014] provides a framework that encompasses the positions of Hamkins, Steel and Woodin. S. Friedman sometimes uses multiverse language, but it isn't clear (at least to me) that his is a true multiverse view. See, e.g., Arrigoni and Friedman [2013].

[44] Cf. Woodin [2011], p. 16: 'The generic-multiverse position … declares that the Continuum Hypothesis is neither true nor false'. Steel ([2014], p. 168) takes his approach to open the possibility that CH and the like are 'meaningless … pseudo-questions'.

> developing the theories in this hierarchy. (Steel [2014], p. 164)

He then defends the resulting ZFC + LCs on the grounds that it allows the set-theoretic reduction:

> The central role of the theories axiomatized by large cardinal hypotheses argues for adding such hypotheses to our framework. The goal of our framework theory is to *maximize interpretive power*, to provide a language and theory in which all mathematics, of today and of the future so far as we can anticipate it today, can be developed. (Ibid., p. 165)

The trouble, of course, is that this strong and attractive theory -- ZFC + LCs -- can't settle statements like CH, as forcing so dramatically demonstrates. Does it follow that CH has a no determinate truth value?[45]

Woodin addresses this question by formulating what he calls the generic multiverse, a collection of models of ZFC + LCs:[46]

> It is generated from each universe of the collection by closing under generic extensions (enlargements) and under generic refinements (inner models of a universe which the given universe is a generic extension of). (Woodin [2011], p. 14)

In a slogan: if N is a forcing extension of M, and one of them is in the generic multiverse, then so is the other. Obviously, he and Hamkins differ on the range of the universes in their multiverses --

---

[45] Notice the sharp contrast between Hamkins's goal -- embrace the widest possible range of universes -- and Woodin's and Steel's shared goal -- given what we take ourselves to know about sets (roughly ZFC+LCs), figure out whether there a fact of the matter about CH. Though both can reasonably be called 'multiverse' views, they are quite different undertakings, and can't be evaluated by the same standards. E.g., though Hamkins's all-inclusive project naturally prompts him to include universes in which V=L, it would be inappropriate to criticize Woodin and Steel for ruling them out.

[46] Actually, both Steel and Woodin begin with models of ZFC, with the addition of large cardinals as on option, but context (e.g., the above motivation from Steel) suggests they lean toward taking that option. This is irrelevant for the case of CH, which is their primary focus.

he and Steel both begin from our current best theory of sets[47] -- but they also disagree sharply on the relevant notion of truth: Woodin takes a claim to be true in the multiverse iff it holds in every universe of the multiverse; Hamkins rejects this notion[48] and holds that his many universes 'exhibit diverse set-theoretic truths' (Hamkins [2012], p. 416).

Woodin goes on to argue (assuming the Ω conjecture) that generic multiverse truth amounts to 'a brand of formalism that denies the transfinite', and hence that 'the generic-multiverse position … is not reasonable' (Woodin [2011], p. 17). Steel reports that

> Woodin's paper makes some arguments against the generic universe position, based on the logical complexity of certain truth predicates, but those arguments do not seem valid to me. (Steel ([2014], p. 170; for details, see his footnote 24)

I can only leave this call to the experts.

Steel's own multiverse approach is more syntactic than either Hamkins's or Woodin's, but before we get to that, let's pause to consider what becomes of set theory's foundational uses on a model-theoretic multiverse conception like these. As noted in §I, I did once believe there was a strong prima facia case against a multiverse of any kind, based on the assumptions that one of set theory's leading goals is to provide a foundation and that one function of a foundation is to serve as a 'final court of appeal'. The idea is simple: if set theory is to settle questions of proof and existence, then set

---

[47] See footnote 45.

[48] Cf. Hamkins [2012], p. 445: 'Woodin introduced [the generic multiverse] in order to criticize a certain multiverse view of truth, namely, truth as true in every model of the generic multiverse. … I do not hold such a view of truth'.

theorists should strive for a single preferred theory of sets that's as decisive as possible -- so as to produce unequivocal judgments when called upon. Of course this very same set-theoretic goal also counsels that set theory be as generous as possible, so as not to limit mathematics,[49] but at least in the most conspicuous cases of potential bifurcation -- like ZFC + LCs vs. ZFC + V=L -- it's possible to avoid an exclusive choice -- because ZFC + V=L can be viewed as the theory of L even in a universe with large cardinals. But this innocent observation actually paves the way for the less doctrinaire descendants of the 'final court' identified in §I, where **Shared Standard** and **Generous Arena** make room for stronger theories, as long as added hypotheses are flagged so that import/export restrictions can be observed -- and this begins to sound a bit like a multiverse view. So what I once took for an objection becomes a question: is the multiverse intended to continue to play the foundational roles that the universe once did, and if so, how does it do this?[50]

Hamkins addresses these questions directly:

> Set theorists commonly take their subject as constituting an ontological foundation for the rest of mathematics, in the sense that abstract mathematical objects can be construed … as sets. (Hamkins [2012], p. 416)

> The multiverse view does not abandon the goal of using set theory as an epistemological and ontological foundation for mathematics, for we expect to find all our familiar mathematical objects … inside any one of the universes of the multiverse. (Ibid., p. 419)

---

[49] See footnote 15.

[50] This is to entertain the possibility that UNIFY, in its original form, might not be required for set theory's foundational uses.

The idea is that when we do set theory, ‘we jump inside and explore the nature of set theory offered by that universe’ (ibid., p. 417), and whenever we do this, all the usual resources of classical mathematics will be available.[51]

So what happens when our mathematician asks about that definable well-ordering of the reals?

> When a mathematical issue is revealed to have a set-theoretic dependence … then the multiverse response is a careful explanation that the mathematical fact of the matter depends on which [universe] is used, and this is almost always a very interesting situation, in which one may weigh the desirability of various set-theoretic hypotheses with their mathematical consequences. (Hamkins [2012], p. 419)

This sounds very like what I said in the universe voice in §I: we explain to the mathematician that there is no such well-ordering in V, but that there is one in L; we lay out the mathematically attractive and unattractive features of L; and we caution that if he elects to take this route, he will have to be careful about exporting and importing resources to and from V.

Hamkins admits that avenues like this are available to the universe advocate, for outer models as well as inner:

> We do have a measure of access into the forcing extensions via names and the forcing relation, allowing us to understand the objects and truths of the forcing extension while remaining in the ground model. (Hamkins [2012], p. 419)

So what makes the multiverse approach preferable?

> The multiverse view explains our mathematical experience with these models by positing that, indeed, these alternative universes exist, just as they seem to exist, with a full mathematical existence, fully as real as the universe under the universe view. (Ibid., p. 419)

---

[51] Here, presumably, the multiverse doesn’t include models of theories too weak to do the job, a stipulation in apparent tension with the generosity cited a few pages back.

> The multiverse view … takes … forcing at face-value, as evidence that there actually are V-generic filters and the corresponding universes V[G] to which they give rise, existing outside the universe. This is a claim we cannot prove within set theory, of course, but the philosophical position makes sense of our experience -- in a way that the universe view does not -- simply by filling in the gaps, by positing as a philosophical claim the actual existence of the generic objects … With forcing, we seem to have discovered the existence of other mathematical universes, outside our own universe, and the multiverse view asserts that yes, indeed, this is the case. (Ibid., p. 425)

This is high metaphysics! Let's set it aside for the moment and return to the more tractable question of how Hamkins' multiverse carries out its foundational duties.

What we've been hearing about so far, over and above the reduction itself, is apparently the 'final court' descendants, **Shared Standard** and **Generous Arena**; in practice, these are treated much as the universe advocate would, with ZFC as the default **Shared Standard** and V the default **Generous Arena**, with added hypotheses noted and import/export carefully regulated. **Elucidation** seems unproblematic, as it, too, can be carried out within any universe of the multiverse.[52] For **Risk Assessment**, presumably we call on a universe with large cardinals. For purposes of **Meta-mathematical Corral**, given that there are many universes outside any given universe, perhaps we're to turn to the theory of the multiverse itself, much as we might turn to Augmented-CCAF in the case of category-theoretic foundations.[53] But this raises a new question: what is the theory of the multiverse?

---

[52] Again, assuming a more restricted range of models than Hamkins sometimes suggests (see previous footnote).

[53] For that matter, we might wonder whether, e.g., the **Generous Arena** should include the entire multiverse, just as we wondered in the case of Augmented-ETCS vs. Augmented CCAF.

What is the counterpart to Augmented-CCAF or ZFC + LCs? What fundamental theory is to tell us what universes there are, something we need to know before any of this can get off the ground?

Hamkins cautions that we shouldn't expect a first-order theory in the language of set theory,

> since the entire point of the multiverse perspective is that there may be other universes outside a give one. (Hamkins [2012], p. 436)

Still, he does offer a number of multiverse axioms, such as:

> For any universe V and any forcing notion P in V, there is a forcing extension V[G], where G is a V-generic subset of P. (Ibid., p. 437)

Obviously there is appeal here to quite technical set-theoretic notions, so we need to ask how principles like this one could be formulated without a prior, ordinary theory of sets in the background. We've met with this sort of concern before, in the 'logical autonomy' objection to category-theoretic foundations, but it seems, if anything, more apt here than it was there: can the theory of the multiverse be stated without presupposing a theory of the universe? Unless it can, no autonomous alternative is actually on offer.

Steel feels the force of this concern. As we've seen, his initial presentation (quoted above) puts the foundational goal front and center:

> The goal of our framework theory is to *maximize interpretive power*, to provide a language and theory in which all mathematics, of today and of the future so far as we can anticipate it today, can be developed. (Steel [2014], p. 165)

He particularly recognizes the importance of **Shared Standard** and **Generous Arena**, and of straightforward import/export:

> Why not simply develop all the natural theories … ? Let 1000 flowers bloom! … This problem with this … is that we do not want everyone to have his own private mathematics. We want one framework theory, to be used by all, so that we can use each other's work. It is better for all our flowers to bloom in the same garden. (Ibid., p. 164)

For this to work, Steel recognizes the need for an explicit, free-standing theory of the multiverse. In a 'historical note', he remarks:

> Neither Hamkins nor Woodin presented a language and a first-order theory in that language, both of which seem necessary for a true foundation. (Steel [2014], p. 170)

He provides this by introducing a multiverse language (ML) that speaks of both sets and universes, and a list of axioms in that language (MV) that includes assertions like these:

> For any axiom φ of ZFC + LCs, and every world W, φ^W.
>
> If W is a world and P in W is a poset, then there is a world of the form W[G], where G is P-generic over W.
>
> If U is a world, and U=W[G], where G is P-generic over W, then W is a world. (See Steel [2014], p. 165)[54]

With the help of some serious mathematics, the theory MV can be successfully formalized.[55]

Steel then constructs, from any transitive model M of ZFC + LCs, a set of worlds M^G that form a natural model of MV.[56] At this point, 'truth in the multiverse M^G' is a perfectly straightforward notion: a

[54] Woodin and Steel differ, e.g., over Steel's axiom of Amalgamation ([2014], p. 165): 'If U and W are worlds, then there are G, H sets generic over them such that W[G] = U[H]'.

[55] The key step for the last of these is a theorem of Laver and Woodin proved in an appendix to Woodin [2011].

[56] Let G be M-generic for $Col(\omega, <OR^M)$. Then M^G is the collection of all 'W such that $W[H] = M[G/\alpha]$, for some H set generic over W, and some α in OR^M' (Steel [2014], p. 166).

statement in ML is true or false in M^G in a standard model-theoretic sense. But this isn't the notion Steel is after. His leading question, like Woodin's, is whether CH, a statement in the pure language of set theory (LST), has a determinate truth value, and he approaches this question by asking whether CH is or isn't meaningful. (If it's not meaningful, it's not even a candidate for truth or falsity.) So the multiverse position Steel considers isn't the Woodin-like proposal that a set-theoretic claim is true in the multiverse iff it's true in every universe of the multiverse, but something he calls the Weak Relativist Thesis (WRT).

To sneak up on the WRT, first notice that since the statement of concern, CH, is a statement of LST, and Steel's (perfectly ordinary) multiverse truths are statements of ML, a link between the two languages is crucial. Fortunately, a theorem of Laver and Woodin implies that there's a recursive translation function t from ML to LST such that for all φ in ML, M^G thinks φ iff M thinks $t(\varphi)$.[57] The multiverse idea behind the WRT is that the only meaningful statements of set theory are those expressible in ML, so the Thesis says that a statement ψ of LST is meaningful iff there's a φ in ML such that ψ is $t(\varphi)$ -- or more succinctly, iff ψ is in the range of t.

[57] In fact, Steel suggests that we only understand ML via the translation t. If so, this raises a question of 'psychological autonomy' for MV that runs parallel to the case of category theory: can we understand ML without first understanding LST? Steel recognizes the problem and offers a rebuttal: 'This is not a very strong objection, as one could think of what we are doing as isolating the meaningful part of the standard language, the range of t, while trimming away the meaningless, in order to avoid pseudo-questions. After climbing our ladder, we throw it away, and from now on, MV can serve as our foundation' (Steel [2014], p. 168). This move deserves some scrutiny, but that's a topic for another occasion.

So, is CH meaningful? As a warm-up, consider another statement in LST: ‘there is a measurable cardinal’. It’s an axiom of MV that this holds in every world of the multiverse, but that’s not enough to show that it appears in the range of t. In fact, it doesn’t. But Steel suggests that it’s reasonable to assume that t(for all W, ‘there is a measurable cardinal’^{W}) and ‘there is a measurable cardinal’ are synonymous.[58] And since ‘for all W, ‘there is a measurable cardinal’^{W}’ is true in M^G, the synonymous set-theoretic statement ‘there is a measurable cardinal’ is both meaningful and true. So far so good, but the same trick won’t work for CH, because ‘for all W, CH^W’ won’t translate to anything like the LST statement that CH. If this were the end of the story, CH would be meaningless and WRT would yield the same conclusion as in Woodin’s multiverse: CH has no determinate truth value; it would be, as Steel puts it, a ‘pseudo-question’ (Steel [2014], p. 154). But this isn’t the end of the story, because, Steel notes, there may be ‘traces of CH’ elsewhere in ML, that is, there may be statements φ of ML other than ‘for all W, CH^W’ such that t(φ) *is* synonymous with CH.

How could this happen? Suppose that the multiverse language is expressive enough to single out one among its worlds with an explicit definition. Woodin has observed

> that if the multiverse has a definable world, then it has a unique definable world, and this world is included in all the others. (Steel [2014], p. 168)

[58] This isn’t explicit in the relevant passage on p. 167 of Steel [2014], but Steel has confirmed it in correspondence (cited with permission). He also notes a residual concern that the MV statement may involve a refinement of the original meaning, but I leave this aside (except for the next footnote).

This minimal world, if there is one, is called the core of the multiverse. If the multiverse has a core (*C*) -- the Weak Absolutist Thesis (WAT) -- then it might be reasonable to regard $t(\mathrm{CH}^C)$ as synonymous with CH.[59] In that way, CH could turn out to be meaningful even on the assumption of WRT, and its truth value would depend on what happens in *C*. But this is speculative for now; much remains to be explored.

In sum, then, the multiverse position Steel considers (WRT) would take the axiom system MV as fundamental, in the sense that it serves to circumscribe our official theory of sets (the range of *t*). From this perspective, ZFC + LCs (at least) would be both meaningful and true, and thus available to play the usual foundational roles in the usual ways. Whether CH has a determinate truth value would remain, for now, an open question.

So, should we prefer a multiverse foundation to the familiar universe foundation? I think we have to allow that the study of multiverse conceptions is in its infancy, so firm conclusions aren’t possible at this stage (even if I had the wit to draw them), but I would like to register some discomfort over the terms in which the debate is often couched. We witnessed Hamkins’ ontological flight above; he proposes that

> each … universe exists independently in the same Platonic sense that proponents of the universe view regard their universe to exist. (Hamkins [2014], pp. 416-417)

---

[59] The residual concern in the previous footnote would arise here as well, perhaps even more acutely: $t(\mathrm{CH}^C)$ involves a great deal of mathematical machinery unknown to many people who presumably do understand CH (like Cantor, for instance).

Of course, a ‘proponent of the universe view’ might just hold that we should pursue one preferred theory of sets, without metaphysical addenda,[60] but Hamkins, qua proponent of the multiverse view, takes a stronger line: ‘The multiverse view is one of higher-order realism -- Platonism about universes’ (ibid.).

In contrast, Steel’s multiverse position involves no such ontology. Instead, he appeals to meaning: the multiverse language is used (via WRT) as an indicator of which ordinary set-theoretic claims are meaningful, capable of truth or falsity, and which are not. As we’ve seen, this raises the possibility that

> the truth value of CH is not determined by the meaning we currently assign to the syntax of LST [the language of set theory]. (Steel [2014], p. 154)

… and the suggestion (quoted above) that we should

> trim back the current syntax, so that we can stop asking pseudo-questions. (Ibid.)

Talk of indeterminacy in ‘the concept’, of diverse ‘concepts of set’, is also common in multiverse thinking. Judging from these discussions, it appears that the overarching goal of set-theoretic practice is to get these things right, to determine the true Platonic ontology, the true contours of the meaning of the word ‘set’, or the true nature of ‘the concept of set’.

But beneath the rhetoric, it emerges that this way of framing the question can’t be quite right. For example, Hamkins admits that

> we may prefer some of the universes in the multiverse to others, and there is no obligation to consider them all as somehow equal … we may simply be more interested in parts of the multiverse

[60] See footnote 14.

> consisting of universes satisfying very strong theories, such as ZFC plus large cardinals. (Hamkins [2012], pp. 417, 436)

For that matter, it's even possible that we might have good mathematical reasons to seek out just one of the universes, just one extension of ZFC, as our unique preferred theory. As far as meanings and concepts go, I personally doubt there's a fact of the matter about what is or isn't part of the current meaning/concept of set, but even if there were, and even if it didn't settle CH, we'd be perfectly free to decide that there's good reason to move on to an enhanced meaning/concept that does. Much like Hamkins, Steel admits this possibility:

> Certainly we do not want to employ a syntax which encourages us to ask pseudo-questions, and the problem then becomes *how to flesh out our current meaning*, or trim back the current syntax, so that we can stop asking pseudo-questions. (Steel [2014], p. 154, emphasis added)

So the metaphysics of abstracta or meanings or concepts are all really beside the point. The fundamental challenge these multiverse positions raise for the universe advocate is this: are there good reasons to pursue a single, preferred theory of sets that's as decisive as possible, or are there not?

Now Steel laments that with matters of justification or 'good reasons', the 'general philosophical questions concerning the nature of … evidence … rear their ugly heads' (Steel [2014], p. 154), but I don't think matters are so dire. I would argue[61] that the relevant reasons are all of a type Steel knows well: straightforwardly mathematical reasons. What mathematical jobs do we want our theory of

[61] Have argued, in [1997] and [2011].

sets to do? One answer is that we want it to serve the various foundational roles of §I, but there are many others: Cantor was after a theory of trigonometric series; Dedekind sought representation-free definitions; contemporary set theorists hope for a rich theory of sets of reals; and so on. The choice between a universe approach and a multiverse approach is justified to the extent that it facilitates our set-theoretic goals. The universe advocate finds good reasons for his view in the many jobs it does so well, at which point the challenge is turned back to the multiverse advocate: given that we could work with inner models and forcing extensions from within the simple confines of V, as described by our best universe theory, what mathematical motivation is there to move to a more complex multiverse theory?

Hamkins gestures toward this perspective on the question in his appendix:

> The mathematician's measure of a philosophical position may be the value of the mathematics to which it leads. (Hamkins [2012], p. 440)

He goes on to describe a pair of research projects inspired by the multiverse perspective. I'm in no position to evaluate the mathematics; my question is whether multiverse thinking is playing more than a heuristic role, whether there's anything here that couldn't be carried out in our single official theory of sets. If not, then it's not clear these examples give us good reason to incur the added burden of devising and adopting an official multiverse theory as our preferred foundational framework. Presumably this same measure could be applied to Woodin's generic multiverse: unless there's at least some hint that it enjoys mathematical advantages over the universe approach, we needn't even concern ourselves about the complexity of its truth predicate.

For the supporter of Steel's WRT, the purported advantage of a multiverse foundation appears to be that it saves us from the misguided pursuit of pseudo-questions, from doomed efforts to settle matters that are indeterminate.[62] If the central goal of set theory is to identify the features of 'the meaning we currently assign to the syntax' of set-theoretic language, and if the range of *t* marks the outer limit of that meaning, and if no 'trace of CH' is hidden in the ML, then efforts to settle CH would be hopeless, inevitably sterile, and the multiverse rejection of CH as a legitimate question might be a valuable outcome. Of course a universe advocate, convinced, say, that ZFC + LCs marks the outer limit of what we'll ever know about V, could reach the same goal of deterring work on CH by epistemological means, without recourse to multiverse thinking. But more to the point, there remains the live possibility that the pursuit of CH isn't in fact doomed. There might even be what Hamkins calls a 'dream solution': a single new axiom that settles CH.[63] Of an even more ambitious dream, Woodin writes:

> I am an optimist … There is in my view no reason at all, beyond a lack of faith, for believing that there is no extension of the axioms of ZFC, by one axiom, a posteriori true, which settles all

---

[62] A defender of Woodin's multiverse might also argue that it saves us from doomed efforts, though Woodin himself -- certainly no defender! -- doesn't say this.

[63] Hamkins ([2012], pp. 429-230) argues that a dream solution to CH is impossible, but he requires that the axiom in question be 'obviously true' in place of Woodin's 'a posteriori true' in the following quotation. (I take 'a posteriori' here to mean that the justification is extrinsic rather than intrinsic.)

> the instances of the Generalized Continuum Hypothesis. (Woodin [2011], p. 30)

Indeed, Steel himself sketches a scenario that could lead even a tentative proponent of WRT back to a strengthened universe theory that 'fleshes out' the current meaning to something that settles CH in the positive.[64] So I don't see that there's call to throw over our universe foundation in favor of a multiverse quite yet!

IV. Inconclusive conclusion

We've seen that set theory, largely via the well-known set-theoretic reduction, serves a number of valuable mathematical ends that ought to qualify as 'foundational': in the current form of ZFC + LCs, it provides a simple first-order theory that interprets all of classical mathematics, so as to allow for meta-mathematical consideration of the whole expanse of that vast subject at once (**Meta-mathematical Corral**); it provides the conceptual resources and construction techniques to clarify old mathematical notions in order to take on new demands (**Elucidation**); in the hierarchy of large cardinals, it provides a flexible scale of consistency strength (**Risk Assessment**); it serves as a benchmark of mathematical proof (**Shared Standard**) and a framework in which the various branches of mathematics appear side-by-side, so that results, methods and resources can be pooled (**Generous Arena**). On the other hand, it doesn't tell us anything about the underlying nature of mathematical objects

[64] See Steel [2014], section 7. In this way, WRT -- the assumption that our current meaning is limited to the ML -- could ultimately point the way to a mathematically attractive extension of the current meaning.

(**Metaphysical Insight**) or of mathematical knowledge (**Epistemic Source**).

Criticisms of this set-theoretic foundation from category theorists may falsely assume that it aspires to **Metaphysical Insight**, but however that may be, they explicitly demand that it found a theory that can't in fact be founded (unlimited categories). Advocates of a category-theoretic replacement for set-theoretic foundations appeal to the category of sets for **Elucidation** and presumably for **Risk Assessment**, but the intended execution of the other traditional foundational roles is less clear, and the attractions of **Generous Arena** may even been rejected. Their central hope is for a foundation that provides a range of concepts and methods that capture and guide the productive ways of thinking that mathematicians actually do and should employ (**Essential Guidance**) -- a role set theory was never designed to play. Proponents are surprisingly unconcerned that category theory doesn't appear equipped to play this role for all areas of mathematics, another indication that **Generous Arena** may have fallen away; different foundations might serve for different branches of the subject. My suggestion is that we do best to retain set theory in the foundational roles is plays so well, retain **Generous Arena** in particular, but also pursue a serious philosophical/methodological investigation of the various 'ways of thinking' in mathematics, beginning with the contrast between those whose 'essence' is well-captured by category theory or by set theory.

Another challenge to ZFC-style set theory in its familiar foundational role comes from the advocates of a multiverse conception of the subject. Steel in particular offers an explicit, free-standing multiverse theory, MV, as fundamental -- in place of a theory in the language of set theory. With the multiverse assumption WRT, MV ratifies a sublanguage of LST that includes ZFC + LCs (more or less by fiat), though perhaps not the meaningfulness of CH. On this picture, the theory ZFC + LCs continues in its usual foundational uses, but only at the behest of MV, and the prospects for extending it are limited. Faced with the charge that this approach shuts down free inquiry into extensions of ZFC + LCs prematurely, an advocate might respond that the multiverse investigation has revealed this inquiry to be misguided, because the meaning or concept of 'set' is simply indeterminate beyond certain limits. Still, even if it's true that our current meaning or concept is indeterminate in this way, there remains the possibility that it might be more mathematically productive, not to give up the quest for an answer to CH, but to seek out a fruitful successor to our current meaning or concept -- a possibility that Steel himself clearly acknowledges and pursues. So far at least, the grounds for replacing the universe with a multiverse are inconclusive.

In sum, then, it seems to me that the familiar set-theoretic foundations, rough and ready as they are, remain the best tool we have for the various important foundational jobs we want done.[65]

Penelope Maddy

[65] I'm grateful to Michael Ernst, for his [2014], [2015], [201?], and many enjoyable conversations on the aspirations of category-theoretic foundations, and to John Steel, for his patience with my questions and confusions about his views. My admiration for John Burgess's [2015] should be obvious. Thanks also to Ernst, Steel, David Malament, Colin McLarty, and an anonymous referee for helpful comments on earlier drafts.

References

Arrigoni, Tatiana, and Sy-David Friedman

[2013] 'The hyperuniverse program', *Bulletin of Symbolic Logic* 19, pp. 77-96.

Benacerraf, Paul

[1965] 'What numbers could not be', reprinted in P. Benacerraf and H. Putnam, eds., *Philosophy of Mathematics*, second edition, (Cambridge: Cambridge University Press, 1983), pp. 272-294.

Bernays, Paul

[1934] 'On platonism in mathematics', C. Parsons, trans., reprinted in P. Benacerraf and H. Putnam, eds., *Philosophy of Mathematics*, second edition, (Cambridge: Cambridge University Press), pp. 258-271.

Burge, Tyler

[1998] 'Frege on knowing the foundation', reprinted in his *Truth, Thought, Reason: Essays on Frege*, (Oxford: Oxford University Press, 2005], pp. 317-355.

Burgess, John P.

[2015] *Rigor and Structure*, (Oxford: Oxford University Press).

Dedekind, Richard

[1872] 'Continuity and irrational numbers', W. Beman and W. Ewald, trans., reprinted in W. Ewald [1996], pp. 766-779.

Ebbinghaus, Heinz-Dieter

[2007] *Ernst Zermelo: an Approach to his Life and Work*, (Berlin: Springer).

Enderton, Herbert

[1977] *Elements of Set Theory*, (New York: Academic Press).

Ernst, Michael

[2014] *The Category-theoretic Imperative*, PhD dissertation, University of California, Irvine.

[2015] 'The prospects of unlimited category theory: doing what remains to be done', *Review of Symbolic Logic* 8, pp. 306-327.

[201?] 'Category theory and foundations', to appear in E. Landry, ed., *Categories for the Working Philosopher*.

Ewald, William, ed.

[1996] *From Kant to Hilbert*, volume II, (Oxford: Oxford University Press, 1996).

Feferman, Solomon

[1977] 'Categorical foundations and foundations of category theory', in R. Butts and J. Hintikka,, eds., *Logic, Foundations of Mathematics and Computability Theory, Proceedings of the Fifth International Congress of LMPS, University of Western Ontario, 1975,* (Dordrecht: Reidel), pp. 149-169.

[2013] 'Foundations of unlimited category theory: what remains to be done', *Review of Symbolic Logic* 6, pp. 6-15.

Hamkins, Joel

[2012] 'The set-theoretic multiverse', *Review of Symbolic Logic* 5, pp. 416-449.

Hellman, Geoffrey

[2003] 'Does category theory provide a framework for mathematical structuralism?', *Philosophia Mathematica* 11, pp. 129-157.

Hilbert, David

[1900] 'Mathematical problems', M. Winston, trans., reprinted (in part) in W. Ewald [1996], pp. 1096-1105.

Isbell, John

[1967] 'Review of Lawvere [1966]', *Mathematical Reviews* 34, pp. 1354-1355.

Kennedy, Juliette, ed.

[2014] *Interpreting Gödel*, (Cambridge: Cambridge University Press).

Kline, Morris

[1972] *Mathematical Thought from Ancient to Modern Times*, (New York: Oxford University Press).

Kunen, Kenneth

[1980] *Set Theory: an Introduction to Independence Proofs*, (Amsterdam: Elsevier).

Lawvere, F. William

[1964] 'An elementary theory of the category of sets', *Proceedings of the National Academy of Science of the USA* 52, pp. 1506-1511, reprinted in expanded form with commentary by C. McLarty in *Reprints in Theory and Application of Category Theory* 12 (2005), pp. 1-35.

[1966] 'The category of categories as a foundation for mathematics', in S. Eilenberg, D. Harrison, S. Mac Lane, and H. Röhrl, eds., *Proceedings of the Conference on Categorical Algebra, La Jolla, CA, 1965*, (Berlin: Springer), pp. 1-20.

Linnebo, Øystein, and Richard Pettigrew

[2011] 'Category theory as an autonomous foundation', *Philosophia Mathematica* 19, pp. 227-254.

Mac Lane, Saunders

[1971] 'Categorical algebra and set-theoretic foundations', in D. Scott and T. Jech, eds., *Axiomatic Set Theory*, *Proceedings of the Symposium in Pure Mathematics of the AMS, UCLA 1967*, (Providence, RI: AMS), pp. 231-240.

[1981] 'Mathematical models: a sketch for the philosophy of mathematics', *American Mathematical Monthly* 88, pp. 462-472.

[1986] *Mathematics Form and Function*, (New York: Springer-Verlag).

Maddy, Penelope

[1997] *Naturalism in Mathematics*, (Oxford: Oxford University Press).

[2011] *Defending the Axioms*, (Oxford: Oxford University Press).

Mathias, A. D. R.

[1992] 'What is Mac Lane missing?', in H. Judah, W. Just, and H. Woodin, eds., *Set Theory of the Continuum*, (Berlin: Springer), pp. 113-118.

[2000] 'Strong statements of analysis', *Bulletin of the London Mathematical Society* 32, pp. 513-526.

[2001] 'The strength of Mac Lane set theory', *Annals of Pure and Applied Logic* 110, pp. 107-234.

McLarty, Colin

[1991] 'Axiomatizing a category of categories', *Journal of Symbolic Logic* 56, pp. 1243-1260.

[2004] 'Exploring categorical structuralism', *Philosophia Mathematica* 12, pp. 37-53.

[2013] 'Foundations as truths which organize mathematics', *Review of Symbolic Logic* 6, pp. 76-86.

Moschovakis, Yiannis

[1994] *Notes on Set Theory*, (New York: Springer).

Nagel, Ernest

[1979] 'The formation of modern conceptions of formal logic in the development of geometry', in his *Teleology Revisited*, (New York: Columbia University Press), pp. 195-259.

Russell, Bertrand

[1907] 'The regressive method of discovering the premises of mathematics', reprinted in his *Essays in Analysis*, D. Lackey, ed., (New York: George Braziller), pp. 272-283.

Steel, John

[2014] 'Gödel's program', in Kennedy [2014], pp. 153-179.

Väänänen, Jouko

[2014] ‘Multiverse set theory and absolutely undecidable propositions’, in Kennedy [2014], pp. 180-208.

Van Heijenoort, Jean, ed.

[1967] *From Frege to Gödel*, (Cambridge, MA: Harvard University Press).

Voevodsky, Vladimir

[2013] Slides for his [2014], available at https://www.math.ias.edu/vladimir/lectures (#10).

[2014] ‘Univalent foundations and set theory’ (abstract), *Bulletin of Symbolic Logic* 20, p. 108.

Von Neumann, John

[1925] ‘An axiomatization of set theory’, reprinted in van Heijenoort [1967], pp. 394-413.

Weaver, Nic

[2014] *Forcing for Mathematicians*, (Hackensack, NJ: World Scientific Publishing).

Wittgenstein, Ludwig

[1978] *Remarks on the Foundations of Mathematics*, revised edition, G. E. M. Anscombe, trans., G. H. von Wright, R. Rhees, and G. E. M. Anscombe, eds., (Cambridge, MA: MIT Press).

Woodin, W. Hugh

[2011] ‘The continuum hypothesis, the generic-multiverse of sets, and the Ω conjecture’, in J. Kennedy and R. Kossak, eds., *Set theory, Arithmetic, and Foundations of Mathematics*, (Cambridge: Cambridge University Press), pp. 13-42.

Zermelo, Ernst

[1908] ‘Investigations in the foundations of set theory’, S. Bauer-Mengelberg, trans., reprinted in J. van Heijenoort [1967], pp. 200-215.

[1908a] ‘A new proof of the possibility of a well-ordering’, S. Bauer-Mengelberg, trans., reprinted in van Heijenoort [1967], pp. 183-198.

[1930] ‘On boundary numbers and domains of sets: new investigations in the foundations of set theory’, M. Hallett, trans., in Ewald [1996], pp. 1219-1233. Also in his [2010], E. de Pellegrin, trans., pp. 400-431.

[2010] *Collected Works,* volume I, H.-D. Ebbinghaus and A. Kanamori, eds., (Berlin: Springer).