From the Foundations of Mathematics to Mathematical Pluralism

Graham Priest

June 16, 2018

Departments of Philosophy, CUNY Graduate Center, and the University of Melbourne

1 Introduction

I suppose that an appropriate way to write an essay on the foundations of mathematics would be to start with a definition of the term, and then discuss the various theoretical enterprises that fall within its scope. However, I'm not sure that the term can be caught in a very illuminating definition. So I'm not going to do this.

What I want to do instead is to tell a story of how foundational studies developed in the last 150 years or so. I am not attempting to give an authoritative history of the area, however. To do justice to that enterprise would require a book (or two or three). What I want to do is put the development of the subject in a certain perspective—a perspective which shows how this development has now brought us to something of a rather different kind: mathematical pluralism. And what that is, I'll explain when we get there.

For the most part, what I have to say is well known. Where this is so, I shall just give some standard reference to the material at the end of each section.[1] When, towards the end of the essay, we move to material that is not so standard, I will give fuller references.

So, let me wind the clock back, and start to tell the story.

[1] A good general reference for the standard material is Hatcher (1982).

2 A Century of Mathematical Rigor

The 19th Century may be fairly thought of as, in a certain sense, the age of mathematical rigour. At the start of the century, many species in the genus of number were well known: natural numbers, rational numbers, real numbers, negative numbers, complex numbers, infinitesimals; but many aspects of them and their behaviour were not well understood. Equations could have imaginary roots; but what exactly is an imaginary number? Infinitesimals were essential to the computation of integrals and derivatives; but what were these 'ghosts of departed quantities', as Berkeley had put it?[2] The century was to clear up much of the obscurity.

Early in the century, the notion of a limit appeared in Cauchy's formulation of the calculus. Instead of considering what happens to a function when some infinitesimal change is made to an argument, one considers what happens when one makes a small finite change, and then sees what happens "in the limit", as that number approaches 0 (the limit being a number which may be approached as closely as one pleases, though never, perhaps, attained). Despite the fact that Cauchy possessed the notion of a limit, he mixed both infinitesimal and limit terminology. It was left to Weierstrass, later in the century, to replace all appeals to infinitesimals by appeals to limits. At this point, infinitesimals disappeared from the numerical menagerie.

Weierstrass also gave the first modern account of negative numbers, defining them as signed reals, that is, pairs whose first members are reals, and whose second members are "sign bits" ('+' or '-'), subject to suitably operations. A contemporary of Weierstrass, Tannery, gave the first modern account of rational numbers. He defined a rational number as an equivalence class of pairs of natural numbers, $\langle m, n\rangle$, where $n \neq 0$, under the equivalence relation, $\sim$, defined by:

$$\langle m, n\rangle \sim \langle r, s\rangle \text{ iff } m \cdot s = r \cdot n$$

Earlier in the century, Gauss and Argand had shown how to think of complex numbers of the form $x + iy$ as points on the two dimensional Euclidean plane—essentially as a pair of the form $\langle x, y\rangle$—with the arithmetic operations defined in an appropriate fashion.

A rigorous analysis of real numbers was provided in different ways by Dedekind, Weierstrass, and Cantor. Weierstrass' analysis was in terms of

[2] *The Analyst, or a Discourse Addressed to an Infidel Mathematician* (1734), §XXXV.

infinite decimal expansions; Cantor's was in terms of convergent infinite sequences of rationals. Dedekind's analysis was arguably the simplest. A *Dedekind section* is any partition of the rational numbers into two parts, $\langle L, R \rangle$, where for any $l \in L$ and $r \in R$, $l < r$. A real number can be thought of as a Dedekind section (or just its left-hand part).

So this is how things stood by late in the century. Every kind of number in extant mathematics (with the exception of infinitesimals, which had been abolished) had been reduced to simple set-theoretic constructions out of, in the last instance, natural numbers.

What, then of the natural numbers themselves? Dedekind gave the first axiomatisation of these—essentially the now familiar Peano Axioms. This certainly helped to frame the question; but it did not answer it.[3]

3 Frege

Which brings us to Frege. Frege was able to draw on the preceding developments, but he also defined the natural numbers in purely set-theoretic terms. The natural number n was essentially the set of all n-membered sets (so that 0 is the set whose only member is the empty set, 1 is the set of all singletons, etc). This might seem unacceptably circular, but Frege showed that circularity could be avoided, and indeed, how all the properties of numbers (as given by the Dedekind axioms) could be shown to follow from the appropriate definitions.

But 'follow from' how? The extant canons of logic—essentially a form of syllogistic—were not up to the job, as was pretty clear. Frege, then, had to develop a whole new canon of logic, his *Begriffsschrift*. Thus did Frege's work give birth to "classical logic".

Given Frege's constructions, all of the familiar numbers and their properties could now be shown to be sets of certain kinds. But what of sets themselves? Frege took these to be abstract (non-physical) objects satisfying what we would now think of as an unrestricted comprehension schema. Thus (in modern notation), any condition, $A(x)$, defines a set of objects $\{x : A(x)\}$. Because he was using second-order logic, Frege was able to define membership. Again in modern notation, $x \in y$ iff $\exists Z(y = \{z : Zz\} \wedge Zx)$.

Moreover, Frege took these set-theoretic principles themselves to be principles of pure logic. Hence all of arithmetic (that is, the theory of numbers)

[3]For the material in this section, see Priest (1998).

was a matter of pure logic—a view now called *logicism*. And this provided an answer to the question of how we may know the truths of arithmetic—or to be more precise, reduced it to the question of how we know the truths of logic. As to this, Frege assumed, in common with a well-worn tradition, that these were simply *a priori*.

Frege's achievement was spectacular. Unfortunately, as is well known, there was one small, but devastating, fly in the ointment, discovered by Russell. The naive comprehension principle was inconsistent. Merely take for $A(x)$ the condition that $x \notin x$, and we have the familiar Russell paradox. If B is the sentence $\{x : x \notin x\} \in \{x : x \notin x\}$ then $B \wedge \neg B$. Given the properties of classical logic, everything followed. A disaster.

After the discovery of Russell's paradox, Frege tried valiantly to rescue his program, but unsucessfully. The next developments of the *Zeitgeist* were to come from elsewhere.[4]

4 Russell

Namely, Russell—and his partner in logical crime, Whitehead. Russell was also a logicist, but a more ambitious one than Frege. For him, *all* mathematics, and not just arithmetic, was to be logic. In the first instance, this required reducing the other traditional part of mathematics—geometry—to logic, as well. This was relegated to Volume IV of the mammoth *Principia Mathematica*, which was never published.

But by this time, things were more complex than this. The work of Cantor on the infinite had generated some new kinds of numbers: transfinite ones. These were of two kinds, cardinals, measuring size, and ordinals, measuring order. Russell generalised Frege's definition of number to all cardinals: a cardinal number was *any* set containing all those sets between which there is a one-to-one correspondence. He generalised it further again to ordinals. An ordered set is *well-ordered* if every subset has a least member. An ordinal is any set containing all those well-ordered sets between which there is an order-isomorphism.

Of course, Russell still had to worry about his paradox, and others of a similar kind which, by that time, had multiplied. His solution was *type theory*. The precise details were complex and need not concern us here. Essentially, sets were to be thought of as arranged in a hierarchy of types, such that

[4]For the material in this section, see Zalta (2016).

quantifiers could range over one type only. Given a condition with a variable of type i, $A(x_i)$, comprehension delivered a set $\{x_i : A(x_i)\}$; this set, however was not of type i, but of a higher type, and so it could not be substituted into the defining condition delivering Russell's paradox to produce contradiction.

Russell's construction faced a number of problems. For a start, it was hard to motivate the hierarchy of orders as *a priori*, and so as part of logic. Secondly, with his construction, Frege had been able to show that there were infinite sets (such as the set of natural numbers). The restrictions of type theory did not allow this proof. Russell therefore had to have an axiom to the effect that there was such a thing: the *Axiom of Infinity*. It was hard to see this as an *a priori* truth as well.[5]

On top of these, there were problems of a more technical nature. For a start, the hierarchy of types meant that the numbers were not unique: every type (at least, every type which was high enough) had its own set of numbers of each kind. This was, to say the least, ugly. Moreover, Cantor's work had delivered transfinite numbers of very large kinds. Type theory delivered only a small range of these. Specifically, if $\beth_0 = \aleph_0$, $\beth_{n+1} = 2^{\beth_n}$, and $\beth_\omega = \bigcup_{n<\omega} \beth_n$, then type theory delivered only those cardinals less than $\beth_\omega$. Of course, one could just deny that there were cardinals greater than these, but *prima facie*, they certainly seemed coherent.

Finally, to add insult to injury, one could not even explain type theory without quantifying over all sets, and so violating type restrictions.

Russell fought gallantly against these problems—unsuccessfully.[6]

5 Zermelo

New developments arrived at the hands of Zermelo. He proposed simply to axiomatize set theory. He would enunciate axioms that were strong enough to deliver the gains of the 19th Century foundational results, but not strong enough to run afoul of the paradoxes. His 1908 axiom system, strengthened a little by later thinkers, notably Fraenkel, appeared to do just this. The

[5] Earlier versions of type theory also required a somewhat problematic axiom called the *Axiom of Reducibility*. Subsequent simplifications of type theory showed how to avoid this.

[6] Starting around the 1990s, there was a logicist revival of sorts, neo-logicism; but it never deilivered the results hoped of it. For the material in this section, see Irvine (2015) and Tennant (2017).

axioms were something of a motley, and so all hope of logicism seemed lost;[7] but, on the other hand, the system did not have the technical inadequacies of type theory.

The key to avoiding the paradoxes of set theory was to replace the naive comprehension schema with the *Aussonderung* principle. A condition, $A(x)$ was not guaranteed to define a set; but given any set, y, it defined the subset of y comprising those things satisfying $A(x)$. An immediate consequence of this was that there could be no set of all sets—or Russell's paradox would reappear. Indeed, all "very large" sets of this kind had to be junked, but with a bit of fiddling, the mathematics of the day did not seem to need these.

In particular, the Frege/Russell cardinals and ordinals were just such large sets. So to reduce number theory to set theory, a different definition had to be found. Zermelo himself suggested one. Later orthodoxy was to prefer a somewhat more elegant definition proposed by von Neumann. 0 is the empty set. $\alpha + 1$ is $\alpha \cup \{\alpha\}$, and given a set, X, of ordinals closed under successors, the ordinal which is the limit of these is $\bigcup X$. A cardinal was an ordinal such that there was no smaller ordinal that could be put in one-to-one correspondence with it.

Logicism had died. The fruits of 19th Century reductionism had been preserved. The paradoxes had been avoided. The cost was eschewing all "large" sets; but this seemed to be a price worth paying.

The next developments came from a quite different direction.[8]

6 Brouwer

In the first 20 years of the 20th Century, Brouwer rejected the idea that mathematical objects were abstract objects of a certain kind: he held them to be mental objects. Such an object exists, then, only when there is some mental procedure for constructing it (at least in principle). In mathematics, then, existence is constructibility. Brouwer took his inspiration from Kant. Mental constructions occur in time. Time, according to Kant, is a mental faculty which enforms sensations—or intuitions, as Kant called them. Hence,

[7]About 20 years later later, in the work of von Neumann and Zermelo, a model of sorts was found: the cumulative hierarchy. This did provide more coherence for the axoms, but it did nothing to save logicism. On the contrary, it appeared to give set theory a distinctive non-logical subject.

[8]For the material in this section, see Hallett (2013).

Brouwer's view came to be called *Intuitionism*. Intuitionism provides a quite different answer from that provided by logicism as to how we know the truths of mathematics: we know them in the way that we know the workings of our own mind (whatever that is).

Brouwer's metaphysical picture had immediate logical consequences. Given some condition, $A(x)$, we may have (at least at present) no construction of an object which can be shown to satisfy it; moreover, we may also have no way of showing that there is no such construction. In other words, both of $\exists x A(x)$ and $\neg\exists x A(x)$ may fail. Hence the Law of Excluded Middle fails. Nor is this the only standard principle of logic to fail. Suppose that we want to show that $\exists x A(x)$. We assume, for *reductio*, that $\neg\exists x A(x)$, and deduce a contradiction. This shows that $\neg\neg\exists x A(x)$; but this does not provide us with a way of constructing an object satisfying $A(x)$. Hence, it does now establish that $\exists x A(x)$. The Law of Double Negation (in one direction), then, also fails.

Brouwer did not believe in formalising logical inference: mental processes, he thought, could not be reduced to anything so algorithmic. But a decade or so later, intuitionist logic was formalised by Heyting and others. Unsurprisingly, it turned out to be a logic considerably weaker than "classical logic", rejecting, as it did, Excluded Middle, Double Negation, and other related principles.

Given the unacceptability of many classical forms of inference, Brouwer set about reworking the mathematics of his day. All proofs which did not meet intuitionistically acceptable standards had to be rejected. In some cases it was possible to find a proof of the same thing which was acceptable; but in many cases, not. Thus, for example, consider König's Lemma: every infinite tree with finite branching has at least one infinite branch. Such a branch may be thought of as a function, f, from the natural numbers to nodes of the branch, such that $f(0)$ is the root of the tree, and for all n, $f(n+1)$ is an immediate descendent of $f(n)$. We may construct f as follows. $f(0)$ has an infinite number of descendants by supposition. We then run down the branch preserving this property, thus defining an infinite branch. Suppose that $f(n)$ has an infinite number of descendants. Since it has only a finite number of immediate descendants, at least one of these must have an infinite number of descendants. Let $f(n+1)$ be one of these. The problem with this proof intuitionistically, is that we have, in general, no way of knowing which node or nodes these are, so we have no construction which defines the next value of the function. There is, then, no such function, since we have no way of constructing it.

Brouwer delivered ingenious constructions, which were intuitionistically valid, and which could do some of the things that classically valid constructions could do. Thus, in the case of König's Lemma, he established something called the *Fan Theorem*. However, it was impossible to prove everything which had a classical proof. Intuitionistic mathematics was therefore essentially revisionary. There are things which can be established classically which have no intuitionist proof.

This may make it sound as though intuitionist mathematics is a proper part of classical mathematics.[9] This, however, is not the case. True, not every proof that is classically valid is intuitionistically valid. But that means that there can be things which are inconsistent from a classical point of view, which are not so from an intuitionistic point of view. And this allows for the possibility that one may prove intuitionistically some things that are *not* valid in classical mathematics.

Take, for example, the theory of real numbers. Let U be the real numbers between 0 and 1; and think of these as functions from natural numbers to $\{1, 0\}$. Now consider a one-place function, F, from U to U. To construct F, we need to have a procedure which, given an input of F, f, defines its output, $F(f)$. And this means that for any n, we must have a way of defining $[F(f)](n)$. Since this must be an effective procedure, $[F(f)](n)$ must be determined by some "initial segment" of f—that is, $\{f(i) : i \leq m\}$ for some m. Hence, if f' agrees with f on this initial segment, $[F(f)](n)$ and $[F(f')](n)$ must be the same. It follows that $F(f)$ and $F(f')$ can be made as close as we please by making f and f' close enough. That is, all functions of the kind in question are continuous! This is simply false in classical real number theory: there are plenty of discontinuous functions.

Intuitionism is not, then, simply that sub-part of classical mathematics which can be obtained by constructive means: it is *sui generis*.

Ingenious though it was, though, intuitionist mathematics never really caught on in the general mathematical community. Mathematicians who did not accept Brouwer's philosophical leanings could see nothing wrong with the standard mathematics. Or perhaps more accurately, mathematicians were very much wedded to this mathematics, and so rejected Brouwer's philosophical leanings.[10]

---

[9] Though one might also simply consider simply the constructive part of classical mathematics. See Bridges (2013).

[10] Further on all these things, see Iemoff (2013).

7 Hilbert

The suspicion of classical reasoning was not restricted just to intuitionists, though. It was shared by Hilbert, who was as classical as they came. The discovery of Russell's paradox, and the apparently *a priori* principles that lead to it, was still something of a shock to the mathematical community; and Hilbert wanted a safeguard against things of this kind happening again. This inaugurated what was to become known as *Hilbert's Program*.

Hilbert wanted to *prove* that this could not happen. Of course, the proof had to be a mathematical one; and to prove anything mathematical about something, one has to have a precise fix on it. Hence, the first part of the program required such a fix on mathematics, or its various parts. This would be provided, Hilbert thought, by appropriate axiomatizations. Hilbert had already provided an axiom system for Euclidean geometry. So the next target for axiomatization was arithmetic. The axiomatization was to be based on classical logic—or at least the first-order part of Frege's logic, which Hilbert and his school cleaned up, giving the first really contemporary account of this.

Given the axiomatised arithmetic, this was then to be proved consistent. Of course given that the proof of consistency was to be a mathematical one, and the security of mathematical reasoning was exactly what was at issue in the project, there was an immediate issue. If our mathematical tools are themselves inconsistent, maybe they can prove their own consistency. Indeed, given that classical logic is being employed, if arithmetic is inconsistent, it can prove anything.

Hilbert's solution was to insist that the reasoning involved in a consistency proof be of a very simple and secure kind. He termed this *finitary*. Exactly what finitary reasoning was, was never defined exactly, as far as I am aware; but it certainly was even weaker than the constructive reasoning of intuitionists. The danger of contradiction, it seemed to Hilbert, arose only when the infinite reared essentially its enticing but dangerous head. Hence the reasoning of the consistency proof had to be something like simple finite combinatorial reasoning—most notably, symbol-manipulation.

This approach allowed a certain philosophical perspective. Take the standard language of arithmetic. Numerals are constituted by '0' followed by some number of occurrences of the successor symbol. Terms are composed from numerals recursively by applying the symbols for addition and multiplication. Equations are identities between terms. The Δ_0 fragment of the language is the closure of the equations under truth functions and bounded quantifiers (i.e., particular or universal quantifiers bounded by some particular number). The truth value of any Δ_0 statement can be determined in a finitary way. Terms can be reduced to numerals by the recursive definitions of addition and multiplication. Identities between numerals can be decided by counting occurrences of the successor symbol, and then truth functions do the rest, the bounded quantifiers being essentially finite conjunctions and disjunctions. So, according to Hilbert, we may take the Δ_0 statements to be the truly meaningful (contentful) part of arithmetic.

But what about the other statements? Given some axiom system for arithmetic, this will contain the finitary proofs of the true Δ_0 statements, but proofs will go well beyond this—notably, establishing statements with unbounded quantifiers. However, since the true Δ_0 statements are complete (that is, for any such statement, either it or its negation is true), the system is consistent iff it is a conservative extension of that fragment.[11] Thus, if the system is consistent, reasoning deploying statements not in the Δ_0 fragment can prove nothing new of this form. However, the statements might well have an instrumental value, in that using them can produce simpler and more expeditious proofs of Δ_0 statements. Hence, thought Hilbert, the non-Δ_0 sentences could be thought of as "ideal elements" of our reasoning—in much the same way that postulating an ideal "point at infinity" can do the same for proofs about finite points, or imaginary numbers can do the same for proofs about real numbers.[12]

The demise of Hilbert's Program is so well known that it hardly needs detailed telling. A young Gödel showed that any axiomatization of arithmetic of sufficient power—at least one that is consisent—must be incomplete. That is, there will be statements, A, such that neither A nor $\neg A$ is provable. Since (classically) one of these must be true, there was no complete axiomatization of arithmetic. Putting another nail in the coffin of the Program, Gödel also

[11]If the system is not a conservative extension it proves the negation of some true Δ_0 sentence, and so the system is inconsistent. Conversely, if it is inconsistent, since it can prove everything, it is not a conservative extension.

[12]In these cases, the ideal elements are not statements, but objects. Hilbert discovered that quantifiers could be eliminated by the use of his ε-symbol. Thus, $\exists x A(x)$ is equivalent to $A(\varepsilon x A(x))$. One might—though I don't think Hilbert ever suggested this—take ε-terms themselves to signify ideal objects. In this way non-Δ_0 statements might be thought of as statements of Δ_0 form, but which concern these ideal objects (as well as, possibly, real ones).

established that given such a system, there is a purely arithmetic statement which can naturally be thought of as expressing its consistency. However, this is not provable within the system—again if it is consistent. Since the system encodes all finitary reasoning (and much more), this seemed to show that a finitary proof of the consistency of even this incomplete system was impossible.[13]

8 Category Theory

So, by mid-century, this is how things stood: apart from some rearguard actions, the great foundational programs of the first part of the 20th Century were defunct. This, however, was not an end of the matter. New developments were to come from a quite new branch of mathematics: category theory.

It is common in mathematics to consider classes of structures of a certain kind: groups, topological spaces, etc. Important information about their common structure is delivered by the morphisms (structure preserving maps) between them. When the range of one morphism is the domain of another, such morphisms can be composed. If we write composition as $\bigcirc$, then the morphism $f \bigcirc g$ is a map which, when applied to an object x , delivers the object obtained by applying f to $g(x)$.

Starting in the late 1940s Eilenberg and MacLane generalised this way of looking at things, to deliver the notion of a category. The idea was taken up and developed substantially by later mathematicians, including Grothendieck and Lawvere.

A category is a bunch of objects, together with functions between them, thought of as morphisms, and often termed *arrows*, because of the way they are depicted diagrammatically. In fact, the objects may be dispensed with, since each may be identified with the identity function on it (which is a morphism). So the notion of a category may be be axiomatised with a number of axioms concerning functional composition. A category is, then, any model of this axiom system (in the same way that a group, e.g., is any model of the axioms of the theory of groups). Hence, there is a category of all groups, all topological spaces, all sets, etc. The category of a particular kind of structures (e.g., sets) may justify further axioms concerning functional

[13]For the material in this section, see Zach (2013).

composition, in the same way that a consideration of Abelian groups justifies axioms additional to those of groups in general.

But foundationally there is now a problem. Since the consolidation of set theory in the early part of the century, it had been assumed that all mathematics could be formulated within set theory. One can, indeed, think naturally of a category as a set of a certain kind. But the problem is that categories such as those of all groups, all topological spaces, all sets—large categories, as they are called—are of the very large kind that had been excised from set theory by Zermelo in order to avoid Russell's paradox. It would seem, then, that set theory cannot provide any kind of foundation for category theory.

There are certain remedial measures one might essay. A large category is not a set, but one can think of it as a proper class, in the sense of NBG set theory (a weak form of second-order ZF)—proper classes being, in effect, subcollections of sets which are not themselves members of anything. However, this is not generally good enough. For category theorists consider not only particular categories, but the category of functions between them. (Given two categories, the category of morphisms between them is called the *functor category.*) This is "too big" even to be a proper class.

One solution to these problems is to deploy the "Grothendieck hierarchy". This is the cumulative hierarchy, with levels, V_α, for every ordinal α, together with the assumption that there are arbitrarily large inaccessible cardinals. As is well known, if ϑ is an inaccessible cardinal, V_ϑ is a model of ZF set theory, so all the usual set theoretic operations can be performed within it. We may then think of the category of all sets, group, etc, as categories of objects in a V_ϑ. The categories themselves, their functor categories, etc, are not in V_ϑ, but are denizens of higher levels of the cumulative hierarchy. Category theory, then, must be thought of as "typically ambiguous", applying schematically to each V_ϑ.

The retrograde nature of this move is clear. The point of category theory is to chart commonalities of structure between *all* structures of a certain kind. The Grothendieck hierarchy explicitly reneges on this. One way to bring the point home is as follows. Suppose that we are considering a category of a certain kind, and we prove something of the form $\exists! x \forall y R(x, y)$. This might be some sort of representational theorem. Interpreting this at each level of the Grothendieck hierarchy, the uniqueness of the x in question is lost: all we have is one at every level.

An honest approach to category theory would seem, then, to take it to be a *sui generis* branch of mathematics. Some have even gone so far as to suggest that it should be taken as providing an adequate foundation for all mathematics, including set theory. The plausibility of this is delivered by the theory of *topoi*. Topoi are particularly powerful categories of a certain kind. (The category of all sets is one of them.) They can be characterised by adding further axioms concerning composition to the general axioms of category theory. All the standard constructions of set theory, at least all those which are involved in the reduction of the other normal parts of mathematics to set theory, can they be performed in a topos.

As a foundational strategy, the weakness of this move is evident. There are many topoi, and "standard mathematics" can be reconstructed in each one. We are, thus, back to the theoretical reduplication which plagued type theory.

I think it fair to say that what to make of all these matters is still *sub judice*. However, we are still not at the end of our story.[14]

9 Paraconsistency

We have so far met two formal logics in the foregoing: classical and intuitionistic. In both of these, the principle of *Explosion* is valid: $A, \neg A \vdash B$, for all A and B. The inference might be thought of as "vacuously valid" in these logics, since the premises can never hold in an interpretation. The principle is clearly counter-intuitive, though. Starting around the 1960s, the development of paraconsistent logic began, a *paraconsistent logic* being exactly one where Explosion is not valid. Using a paraconsistent logic we may therefore reason using inconsistent information in a perfectly sensible way. The information does not deliver triviality—that is, not everything can be established.

There are, in fact, many paraconsistent logics. Their key, semantically speaking, is to stop Explosion being vacuously valid, by including in the domain of reasoning, not only standard consistent situations, but also inconsistent ones. Thus, if p and q are distinct propositional parameters, there can be a situation where p and $\neg p$ hold, but q does not. This is not to suggest that these inconsistent situations may be actual. We reason, after all, about situations which are conjectural, hypothetical, etc. However, the view that

[14]For the material in this section, see Marquis (2014).

some of these inconsistent situations are actual (that is, that what holds in them is actually true), is called *dialetheism*.[15]

The possibility of employing a paraconsistent logic opens up new possibilities in a number of the foundational matters which we have met.[16]

Thus, one of the possibilities that has been of much interest to paraconsistentists is set theory—and for obvious reasons: using a paraconsistent logic allows us to endorse the unrestricted comprehension schema. Contradictions such as Russell's paradox can be proved in the theory, but these are quarantined by the failure of Explosion. Moreover, it was proved that, with an appropriate paraconsistent logic, naive set theory (that is, set theory with unrestricted comprehension) is non-trivial.[17]

This raises the prospect of regenerating Frege's foundational project. Of course, having an unrestricted comprehension schema does not guarantee that this project can be carried through. The set-theoretic principles are strong, but the logic is much weaker than classical logic. Things other than Explosion need to be given up. Notably, the principle of *Contraction*, $A \to (A \to B) \vdash A \to B$ cannot be endorsed, because of Curry's Paradox.[18]

It was only relatively recently that Weber was able to show that much of Frege's program *can* be carried out in such a theory.[19] He showed that virtually all of the main results of cardinal and ordinal arithmetic can be proved in this set theory. Moreover, the theory can be used to prove the Axiom of Choice, as well as results that go beyond ZF set theory, such as the negation of the Continuum Hypothesis, and several large-cardinal principles. (Of course, in the context, this does not show that one cannot prove the negations of these as well.)

Weber's proofs have a couple of very distinctive elements. First, they use comprehension in a very strong form, namely:

$$\exists x \forall y (y \in x \leftrightarrow A)$$

where A may contain x itself. This provides the potential for having a fixed

[15]On paraconsistency, see Priest, Tanaka, and Weber (2017). On dialetheism, see Priest and Berto (2013a).

[16]On inconsistent mathematics in general, see Mortensen (2017).

[17]See Brady (1989).

[18]With naive comprehension we can define a set, c, such that $x \in c \leftrightarrow (x \in x \to \bot)$. Contraction and *modus ponens* then quickly deliver a proof of $\bot$.

[19]See Weber (2010) and (2012).

point, and so self-reference, built into the very characterisation of a set.[20] Next, Weber not only accommodates inconsistencies, but makes constructive use of them. Thus, a number of the results concerning cardinality, such as Cantor's Theorem, make use of sets of the form $\{x \in X : r \in r\}$, where r is the set of all sets which are not members of themselves, and so inconsistent.

Whether Weber's proof methods, and the various distinctions they require one to draw, are entirely unproblematic; and whether other aspects of set-theoretic reasonings (such as those required in model theory) can be obtained in naive set theory, are still questions for investigation.

Another of the foundational matters we have met, and with which paraconsistency engages, is that concerning Gödel's theorems. Gödel's first theorem is often glossed as saying that any "sufficiently strong" axiomatic theory of arithmetic is incomplete. In fact, what it shows is that it is either incomplete *or* inconsistent. Of course, if inconsistency implies triviality, it is natural to ignore the second alternative. However, paraconsistency changes all that, since the theory may be complete, inconsistent, but non-trivial. (I note that there is nothing in the use of paraconsisent logic, as such, which problematises the proof of Gödel's theorem. The logic required of an arithmetic theory in order for it to hold is exceptionally minimal.)

Indeed, it is now known that there are paraconsistent axiomatic theories of arithmetic which contain all the setences true in the standard model, and so which are complete (that is, for any sentence of the language, A, either A or $\neg A$ is in the theory). These theories are inconsistent, but non-trivial.[21]

What foundational significance this has depends, of course, on the plausibility of the claim that arithmetic might be inconsistent.

Implausible as this may seem, Gödel's theorem itself might be thought to lead in this direction. At the heart of Gödel's proof of his theorem, there is a paradox. Consider the sentence 'this sentence is not provable'. If it is provable, it is true; so it is not provable. Hence it is not provable. But we have just proved this; so it is.

Of course, this argument cannot be carried through in a consistent arithmetic, such as classical Peano Arithmetic, when proof is understood as proof in that system. This may be a matter of relief; or it may just show the inadequacy of the system to encode perfectly natural forms of reasoning.

[20] Perhaps surprisingly, Brady's proof shows this strong form of comprehension to be non-trivial.

[21] See Priest (2006), ch. 17.

Indeed, if one takes one of the axiomatic arithmetics containing all the truths of the standard model, there will be a formula of the language $Pr(x)$ which represents provability in the theory in the theory itself. It is then a simple matter to construct a sentence, G, in effect of the form $\neg Pr(\langle G \rangle)$ (where angle brackets indicate gödel coding), and establish that both $\neg Pr(\langle G \rangle)$ and $Pr(\langle G \rangle)$ hold in the theory.[22]

What of Gödel's second theorem? Since a paraconsistent theory of the kind we have just been talking about is inconsistent, one should not expect to prove consistency. But it is also non-trivial, i.e., some statements are not provable. This can be expressed by the sentence $\exists x \neg Pr(x)$, and this sentence can be proved in theories of the kind in question. Of course, this does not rule out the possibility that one may be able to prove $\neg \exists x \neg Pr(x)$ as well; and in what sense the proof of non-triviality is finitary may depend on other features of the arithmetic. To what extent these matters may be thought to help Hilbert's program is, then, still a moot point.[23]

10 Intuitionist and Paraconsistent Mathematics

We have now been rather swiftly through a story of the development of studies in the foundations of mathematics in the last 150 years. As we have seen, none of the foundational ideas we have looked at can claim to have met with uncontestable success. But looking back over developments, we can see that something else has emerged.

As we saw in Section 6, there are fields of intuitionist mathematics that are quite different from their classical counterparts.[24] Indeed, there are fields of intuitionist mathematics that have no natural classical counterpart.

Let me give just one example of this. This is the Kock-Lawvere the-

[22]Additionally, one would expect that the schema $Pr(\langle A \rangle) \supset A$ would be provable in a theory of arithmetic in which $Pr(x)$ really did represent provability. In a consistent theory, it is not, as Löb's Theorem shows. However, the schema is provable in the above theories.

[23]A third foundational issue opened up by paraconsistency concerns category theory. Given that one can operate in a set theory with a universal set, it is possible to have a category of all sets, all groups, etc., where 'all' means *all*. The implications of this for the relationship between set theory and category theory are yet to be investigated.

[24]For a further account of some of these enterprises, see Dummett (2000), chs. 2, 3.

ory of smooth infinitesimal analysis.[25] To motivate this, consider how one would compute the derivative of a function, $f(x)$, using infinitesimals. The derivative, $f'(x)$, is the slope of the function at x, given an infinitesimal displacement, i; so $f(x+i) - f(x) = if'(x)$. Now, as an example, take $f(x)$ to be x^3. Then $if'(x) = (x+i)^3 - x^3 = 3x^2i + 3xi^2 + i^3$. If we could divide by i, we would have $3x^2 + 3xi + i^2$. Setting i to 0 delivers the result—though how, then, did we divide by i?[26] If $i^2 = 0$, we have another route to the answer. For then it follows that for any infinitesimal, i, $if'(x) = 3x^2i$. We may not be able to divide by i, but suppose that $ai = bi$, for all i, implies that $a = b$ (this is the *Principle of Microcancellation*). $f'(x) = 3x^2$ then follows. This is exactly how the theory of smooth infinitesimal analysis proceeds.

Call a real number, i, a *nilsquare* if $i^2 = 0$. Of course, 0 itself is a nilsquare, but it may not be the only one! We may think of the nilsquares as infinitesimals. The theory of smooth infinitesimals takes functions to be linear on these. Given a function, f, there is a unique r such that, for every nilsquare, i $f(x+i) - f(x) = ri$. (In effect, r is the derivative of f at x.) This is the *Principle of Microaffineness*.[27]

Microaffineness implies that 0 is not the only nilsquare. For suppose that it is, then all we have is that $f(x+0) - f(x) = r0$, and clearly this does not define a unique r. So:

[1] $\neg\forall i(i^2 = 0 \rightarrow i = 0)$

But now, why do we need intuitionist logic? Well, one might argue that 0 *is* the only nilsquare, which would make a mess of things. A typical piece of reasoning for this goes as follows. Suppose that i is a nilsquare and that $\neg i = 0$. Then i has an inverse, i^{-1}, such that $i \cdot i^{-1} = 1$. But then $i^2 \cdot i^{-1} = i$. Since $i^2 = 0$, it follows that $i = 0$. Hence, by *reductio*, we have shown that $\neg\neg i = 0$. If we were allowed to apply Double Negation, we could infer that $i = 0$—and so we would have a contradiction on our hands. But this move is not legitimate in intuitionist logic. We have just:

[2] $\forall i(i^2 = 0 \rightarrow \neg\neg i = 0)$

[25] On which, see Bell (2008).

[26] One answer to the conundrum is provided by non-standard analysis, an account of infinitesimals developed in the 1960s by Robinson. This deploys non-standard (classical) models of the theory of real numbers.

[27] Microcancellation follows. Take $f(x)$ to be xa. Then, taking x to be 0, Microaffineness implies that there is a unique r such that, for all i, $ai = ri$. So if $ai = bi$ for all i, $a = r = b$.

And we may hold both [1] and [2] together.

What we see, then, is that there are very distinctive fields of intuitionistic mathematics, quite different from the fields of classical mathematics. Moreover, one does not have to think that intuitionism is *philosophically* correct to recognise that these are interesting mathematical enterprises with their own integrity. They are perfectly good parts of pure mathematics. (Whether they have applications to areas outside of mathematics is a quite separate matter, and irrelevant to the present point.)

A similar point can be made with respect to paraconsistent mathematics. In Section 9 we saw that there were inconsistent mathematical theories relevant to various foundational enterprises: set theory and arithmetic. There are, however, many interesting inconsistent mathematical theories based on paraconsistent logics, which have no immediate application to foundational matters.[28] These include theories in linear algebra, geometry, and topology.

Let me give an example of one of these. This concerns boundaries. Take a simple topological space, say the one-dimensional real line. Divide it into two disjoint parts, left, L, and right, R. Now consider the point of division, p. Is p in L or R? Of course, the description under-determines an answer to the question. But when the example is fleshed out, considerations of symmetry might suggest that it is in both. Then, $p \in L$, and $p \in R$ so $p \notin L$—and symmetrically for R. So a description of the space might be that if $x < p$, x is (consistently) in L; if $p < x$, p is (consistently) in R; and p is both in and not in L, and in and not in R. Given an appropriate paraconsistent logic, the description is quite coherent.[29]

This might not seem particularly interesting, but the idea of inconsistent boundaries has interesting applications. One of these is to describe the geometry of "impossible pictures".[30] Consider the following picture:

[28] See Mortensen (1995).

[29] Further on inconsistent boundaries, see Cotnoir and Weber (2015).

[30] For more on the following, see Mortensen (2010).

The three-dimensional content of the picture is impossible. How should one describe it mathematically? Any mathematical characterisation will specify, amongst other things, the orientations of the various faces. Now, consider the left-hand face, and in particular its lighter shaded part. This is 90° to the horizontal. Next, consider the top of the lower step on the right-hand side of the picture. This is 0° to the horizontal. Finally, consider the boundary between them (a vertical line on the diagram). This is on both planes. Hence it is at both 90° and 0° to the horizontal.[31] That's a contradiction, since it cannot be both; but that's exactly what makes the content of the picture impossible. Note that the characterisation of the content must deploy a paraconsistent logic, since it should not imply, e.g., that the top of the higher step is at 90° to the horizontal.

Hence, we see, again, that whatever one thinks about the truth of naive set theory, or inconsistent arithmetic, there are perfectly good mathematical structures based on a paraconsistent logic.

Both intuitionist logic and paraconsistent logic, then, provide perfectly coherent and interesting areas of mathematical investigation to which classical logic can be applied only with disaster.[32]

[31] And one can set things up in such a way that this does not imply that $90 = 0$.

[32] It seems to me that given any formal logic there could, at least in principle, be interesting mathematical theories based on this. However, intuitionistic logic and paraconsistent logic (and perhaps fuzzy logic; see Mordeson and Nair (2001)) are the only logics for which this has so far really been shown.

11 Mathematical Pluralism

What we have seen is that there are areas of relatively autonomous mathematical research. That is, there is a plurality of areas, such that there is no one of them to which all others can be reduced. This is mathematical pluralism. In truth, this pluralism was already clear in the case of category theory: attempts to reduce it to set theory or vice versa, were always straining at the seams. But intuitionist and paraconsistent mathematics have put the matter beyond doubt.[33] We have, in these cases, mathematical enterprises which are completely *sui generis*. (Note that I am not maintaining that all these branches of mathematics are equally deep, rich, elegant, applicable, etc. That is a quite different matter. All I am claiming is that they are all equally legitimate pure mathematical structures.) What to make of the situation certainly raises interesting issues,[34] but that it now obtains cannot be gainsaid.[35]

[33]In fact, the matter is arguably the case even in classical set theory. We can investigate set theory in which the Axiom of Choice holds, and set theory in which the Axiom of Determinacy, which contradicts it, holds. In that case, however, a monist might claim that we are simply doing model theory, and investigating what holds in models of certain set-theoretic axiom systems. One might consider a similar claim in the cases mentioned in the text: we are just doing model-theory using classical logic, albeit of models of non-classical logics. But this suggestion seems lame. First, intuitionist real number theory and paraconsistent set theory are *not* done in this way. So the suggestion gets the mathematical phenomenology all wrong. Secondly, the insistence that the model theory be classical seems dogmatic. Investigations could proceed with intuitionist model theory, which would give quite different results. (Note that intuitionist model theory is well established, but there is as yet no such thing as paraconsistent model theory.) Thirdly, it is entirely unclear how to pursue this strategy in the case of category theory, simply because the foundational problem with category theory was precisely that its ambit appears to outstrip the classical models.

[34]Some discussion of it can be found in Priest (2013b) and Shapiro (2014). Shapiro takes mathematical pluralism to entail logical pluralism. I am not inclined to follow him down that path. Given a mathematical structure based on a logic, L, reasoning in accord with L preserves truth-in-that-structures. This may not be truth *simpliciter* preservation.

[35]There are, as far as I can see, only two strategies for maintaining mathematial monism. One is the production of some kind of *ur*-mathematics to which all the kinds of mathematics we have met can be reduced. Maybe this could be some kind of foundational project for the 21st century; but nothing like this is even remotely on the horizon. The other strategy is simply to deny that the non-favoured kinds of mathematics really *are* mathematics. Now whether these theories are as deep, elegant, applicable, or whatever, as the favoured mathematics, might certainly be an issue. But it seems to me that denying that they are mathematics is the equivalent of the proverbial ostrich burying its head in

Mathematical pluralism was certainly not an aim of work in the foundations of mathematics, but it has emerged none the less. In the 18th century, some mathematicians investigated Euclidean geometry, trying to prove that the Parallel Postulate was deducible from the other axioms. They did this by assuming its negation, and aiming for a contradiction, which, it turned out, was not coming. Their aim was not to produce non-Euclidean geometries; but by the 19th century, it became clear that this is exactly what they had done. In a similar way, the aim of research in the foundations of mathematics was not to establish mathematical pluralism—indeed, most researchers in the area took themselves to be mathematical monists. But that, it seems, is what, collectively, they have done. If one were Hegel, one would surely diagnose here a fasinating episode in the cunning of reason.[36]

References

[1] Bell, J. L. (2008), *A Primer of Infinitesimal Analysis*, 2nd edn, Cambridge: Cambridge University Press.

[2] Brady, R. (1989), 'The Non-Triviality of Dialectical Set Theory', pp. 437–71 of G. Priest, R. Routley, and J. Norman (eds.), *Paraconsistent Logic: Essays on the Inconsistent*, München: Philosophia Verlag.

[3] Bridges, D. (2013), 'Constructive Mathematics', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/mathematics-constructive/.

[4] Cotnoir, A., and Weber, Z. (2015), 'Inconsistent Boundaries', *Synthese* 192: 1267-94.

[5] Dummett, M. (2000), *Elements of Intuitionism*, 2nd edn, Oxford: Oxford University Press.

[6] Hallett, M. (2013), 'Zermelo's Axiomatization of Set Theory', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/zermelo-set-theory/.

the sand: these theories have clear mathematical interest.

[36] Many thanks go to Hartry Field, Arnie Koslow, and Zach Weber for very helpful comments on an earlier draft of this essay.

[7] Hatcher, W. S. (1982), *The Logical Foundations of Mathematics*, Oxford: Pergamon Press.

[8] Iemhoff, R. (2013), 'Intuitionism in the Philosophy of Mathematics', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/intuitionism/.

[9] Irvine, A. (2015), '*Principia Mathematica*', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/principia-mathematica/.

[10] Marquis, J.-P. (2014), 'Category Theory', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/category-theory/.

[11] Mordeson, J., and Nair, P. (2001), *Fuzzy Mathematics: and Introduction for Engineers and Scientists*, Heidelberg: Physica Verlag.

[12] Mortensen, C. (1995), *Inconsistent Mathematics*, Dordrecht: Kluwer.

[13] Mortensen, C. (2010), *Inconsistent Geometry*, London: College Publications.

[14] Mortensen, C. (2017), 'Inconsistent Mathematics', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/mathematics-inconsistent/.

[15] Priest, G. (1998), 'Number', pp. 47-54, Vol. 7, of E. Craig (ed.), *Encyclopedia of Philosophy*, London: Routledge.

[16] Priest, G. (2006), *In Contradiction*, 2nd edn, Oxford: Oxford University Press.

[17] Priest, G., and Berto, F. (2013a), 'Diatheism', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/dialetheism/.

[18] Priest, G. (2013b), 'Mathematical Pluralism', *Logic Journal of IGPL* 21: 4-14.

[19] Priest, G., Tanaka, K., and Weber, Z. (2017), 'Paraconsistent Logic', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/logic-paraconsistent/.

[20] Shapiro, S. (2014), *Varieties of Logic*, Oxford: Oxford University Press.

[21] Tennant, N. (2017), 'Logicism and Neologicism', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/logicism/.

[22] Weber, Z. (2010), 'Transfinite Numbers in Paraconsistent Set Theory, *Review of Symbolic Logic* 3: 71-92.

[23] Weber, Z. (2012), 'Transfinite Cardinals in Paraconsistent Set Theory, *Review of Symbolic Logic* 5: 269-93.

[24] Zach, R. (2013), 'Hilbert's Program', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/hilbert-program/.

[25] Zalta, E. (2016), 'Frege', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/frege/.