Set Theory and its Place in the Foundations of Mathematics- *a new look at an old question*

Mirna Džamonja *

(to appear in the Journal of Indian Council of Philosophical Research)

Abstract

This paper reviews the claims of several main-stream candidates to be the foundations of mathematics, including set theory. The review concludes that at this level of mathematical knowledge it would be very unreasonable to settle with any one of these foundations and that the only reasonable choice is a pluralist one. 

1 Introduction

Set theory is often cited as the foundations of mathematics. In this paper we review the history of that claim, argue that the other known foundations such as category theory and univalent foundations have the same but not larger right to exists, and conclude that the only choice which agrees with the mathematical practice of today is to take a pluralist view of the matter and accept all these as foundations of an interesting part of mathematics, each in its own right. We also lead the reader through the history of what the foundations of mathematics has actually meant in recent years, moving from the ontological to an epistemic point of view.

1

---

*School of Mathematics, University of East Anglia, Norwich, NR4 7TJ, UK, M.Dzamonja@uea.ac.uk, http://www.mth.uea.ac.uk/people/md.html

1The author thanks l’Institut d’Histoire et de Philosophie des Sciences et des Techniques, Université Paris 1 Panthéon-Sorbonne, where she is an Associate Member. She is grateful to Marco Panza, Bas Spitters and the anonymous referee for the improvements their comments brought to the paper, and to Michele Friend for acting as a native speaker overseer of the language matters.
2 A bit of history of set-theoretic foundations

Set theory started as a purely mathematical subject, brought into life by George Cantor. Usually the birth of set theory is traced back to his 1874 paper [6], in which Cantor shows that there are at least two sizes of infinity and derives Liouville’s theorem which states that there are infinitely many transcendental numbers. Although, initially, Cantor’s work was oriented more towards mathematics than logic, his view of mathematics was hugely influenced by that of Richard Dedekind who was working much more in an abstract logic approach. It is the meeting between Cantor and Dedekind in 1872 in Switzerland that should really be considered as the birth of the subject of set theory, that curious mixture of logic and mathematics that has formed it as a subject unto itself. As a result of the initial meeting with Dedekind, Cantor started looking into sets of uniqueness in the theory of trigonometric series, and this is where the idea of a set as a collection of objects satisfying some fixed property came from. In fact, there are even earlier mentions of a similar idea, since Bernard Bolzano considered in 1847 (published posthumously in [4]) the concept of a set as “an embodiment of the idea or concept which we conceive when we regard the arrangement of its parts as a matter of indifference” and worked on infinite sets, defending the idea of their existence and proving their characterization as sets that can be put in a bijective correspondence with a proper subset of themselves.

Cantor was dealing with what we today call naïve set theory, that is a theory without axiomatisation, where objects are taken as intuitively clear. For example, it was taken as evident that any mathematical object was a set, and indeed the ones that appeared in the practice of the time certainly could have been seen as sets. This is probably what mathematicians back then would have meant by the foundation of mathematics. However, paradoxes in naïve set theory started appearing as early as in 1897, when Cesare Burali-Forti [5] published work which, when some inaccuracies are corrected, shows that there is no set of all ordinal numbers. Some, not entirely verifiable claims, state that Cantor had discovered the same paradox in 1895. Yet, the inability to formulate arbitrary mathematical objects as sets was made fully transparent to mathematicians only in 1902 when Bertrand Russell proved (in a letter to Frege, [25]) that the set $U$ of all sets cannot exist, as the set \( \{ B \in U : B \notin B \} \) yields a contradiction. (It is less well known that Ernst Zermelo independently discovered this paradox, also in 1902 (see [23])). This was a bit of an embarrassing situation to say the least, since by 1902 the role of set theory in the foundations of mathematics has already been established and accepted. In spite of the resistance of mathematicians such as Leopold Kroenecker, now considered as the founder of constructivist mathematics,
set theory had become widespread, due to the utility of its concepts, its omnipresence in mathematical proofs, and the inviting concept of the “infinity of infinities” (the “Cantor’s paradise” of Hilbert) resulting from the iterative use of the power set operation. The contemporary appreciation of set theory can be seen by the article on set theory in Klein’s encyclopedia [26]. After Russell, it was immediately realised that paradoxes in set theory were coming from the lack of proper axioms. Many, including Zermelo, started working on the axiomatisation of set theory. In 1908, he published his results [30] with a resolution of the Russell paradox through the use of the restricted rather than full comprehension. In 1922, Adolf Fraenkel and in 1923 Thoralf Skolem independently improved Zermelo’s axiom system (see the book [10], pg. 22 for the history and further references). The resulting axioms system of axioms in the first order logic, called Zermelo-Fraenkel axioms (ZF), is now the most commonly used system for axiomatic set theory. It is usually used in conjunction with the Axiom of Choice, and known as ZFC. In spite of some historical controversy about the acceptance of the Axiom of Choice, since it is known from the work of Gödel in [13] that the consistency of ZF implies that of ZFC, it is common practice to accept ZFC as the axiomatisation of set theory. In this article we shall not discuss axiomatisations not related to ZFC, of which there are several. A celebrated part of the history of ZFC is that it does not resolve the continuum problem, as shown by Paul Cohen in 1963, see [7].

3 Conceptions of set theory

To understand the place of set theory in the foundations at this moment, it is instructive to discuss the image that this subject presents both from the inside and the outside of its practice. It seems to be the case that set theory has quite a different image for those that practice it and for those who observe it. The observers of this subject consist of other logicians, mathematicians and philosophers, as well as rather frequent casual observers fascinated by the told and untold promises of the subject. Which other subject promises many versions of infinity, the basis for the whole of mathematics and then some unsolvable problems? Let us go into these conceptions a bit more seriously, to allow us to ask the question that we wish to address in the first place: what is the real role of set theory in the foundations of mathematics today? We also invite the reader to reflect on what one should mean by “the foundations of mathematics”. Should it be ontological or epistemic? We shall give our answer in Section §5.

For the outsider, set theory tends to be the subject axiomatised by ZFC.
Anecdotal evidence from working mathematicians suggests that this axiom system is viewed as more than sufficient for what mathematics needs. While the axiom of choice or the continuum hypothesis might still excite some occasional discussion, mathematicians in most areas seem to happily accept the “sufficiency” of ZFC and with somewhat less assurance, also its “necessity”. It might be worth mentioning that among the non-practioners, there are also proponents of ZFC2, the ZFC axiomatisation in full second order logic. Some philosophers, for example Stewart Shapiro (see [27]) are interested in this system as it offers certain categoricity. Namely, Zermelo proved in [31] that the only models of this system are of the form $V_\kappa$ for a strongly inaccessible cardinal $\kappa$. This result is interesting from various points of view, but not very relevant to the questions of independence in set theory since, for example, ZFC2 does not settle the continuum problem any more than ZFC does. (This is the case because by the results of Azriel Lévy and Robert Solovay[18] one can change the universe of set theory and change the values of the continuum function, while keeping the same cardinal $\kappa$ inaccessible.)

Given that there is no gain on the important question of independence, set theorists themselves do not like ZFC2 because it takes away a large chunk of the subject which is based on the soundeness, completeness (and consequently compactness) of the first order logic. Namely, taken together, these properties say that for any set $\Sigma$ of sentences of the first order logic, $|= \Sigma$ iff $\vdash \Sigma$, that is all sentences in $\Sigma$ are true in every model iff there is no proof of contradiction from $\Sigma$. Therefore to prove the consistency of a theory $T$, such as say ZFC+$\neg$CH, it suffices to find a model for it. A rather extensive discussion of these issues can be found in [16].

At any rate, in the first order or the second order axiomatisation, or even without any axiomatisation, set theory is considered important in foundations of mathematics because many of the classical notions are axiomatized by the theory and can be found in the cumulative hierarchy of sets. Let us also recall that Hilbert considered that all mathematics can be formulated in (naïve) set theory, which certainly hammered in the place of set theory in the foundations. Extensions of ZFC by axioms with strictly higher consistency strength are rarely considered outside of the subject itself, although we should mention the recent philosophical work by Giorgio Venturi [28] which argues for the acceptance of Forcing Axioms (see below for more discussion and [8] for a survey). Not to forget the work of Penelope Maddy, who has dedicated a large part of her work to explaining the methods that set theorists use in agreeing on axioms, especially those that go beyond ZFC. Or the

---

2 A theory $T$ is said to be of a strictly higher consistency strength than a theory $T'$ if from the consistency of $T$ we can conclude the consistency of $T'$, but not vice versa.
work of Peter Koellner. But all of these philosophers are so much specialists in set theory that for the purpose of this article and much wider than that, they certainly count as insiders of set theory rather than the outsiders.

Not entering into more examples that form the outside image of set theory, probably the most important point to note here is that the outside view of set theory often takes it for granted that the set theorists believe in one universe of set theory, as much as mathematicians believe in one mathematics, and that set theorists are in search of the right axioms to describe that universe. So, in this view, the fact that for example CH is independent of ZFC would mean that the “right” axiomatisation of set theory would extend ZFC into something that would resolve CH. And this outsider is not entirely wrong, Gödel himself considered that his incompleteness theorems only show that the ZFC system is too weak to answer those questions, and therefore one should search for new axioms that once added to ZFC would answer them. The search for new axioms has been known as Gödel’s Program, see [1]. Yet, the search for the new axioms does not imply the belief in one universe.

While the insider also accepts the ZFC axioms, that is their “necessity”, it is clear to a set theorist that ZFC is insufficient. There is a universal belief in the consistency of large cardinal axioms and ZFC is viewed simply as a minimum of what one should consider as a reasonable system. Not only are the large cardinals needed for set theory but they are also known to be needed for some seemingly innocent statements about number theory. For example, Harvey Friedman [11] developed the Boolean relation theory, which demonstrates the necessity of large cardinals for deriving certain propositions considered “concrete”. Friedman and others view this as an obvious reason for a working mathematician to accept large cardinals. Moving on from the large cardinals and their widespread acceptance, we also claim that in set theory there is no belief in the single universe of the axioms. That is, even if the “right” axioms were to be found, there is no reason to hope that they would be categorical. Moreover, moving rather far from the foundational issues, as much as set theorists have understood the privileged position of the subject as regards to its ability to formulate mathematics, many in general consider that the subject of foundations is not worth pursuing and are much more interested in the mathematics of infinity. Most have the view that, certainly, the search for the unique perfect foundations of mathematics is closed since Gödel’s Incompleteness Theorems [12]. Since even the continuum problem is independent, most set theorists have long moved the emphasis of the subject somewhere else than the axioms, even if Gödel himself did not.

The philosophical views one can read or, more often, read between the lines, in the work of set theory today include the following: “Given the incompleteness, it is better to concentrate on what is possible to do in ZFC and
understand what is not by studying independence” (adopted by Saharon Shelah, some from his school and some others). The more persistent on the resolution of the continuum problem is the view of the California school that continues the Gödel program, now most prominently presented by Hugh Woodin. They do believe that we should find new axioms to complete ZFC, at least to the extent of resolving the continuum problem. Certainly one such axiom is that all sets are constructible (\(V = L\)). However, there is little enthusiasm among the practitioners to accept it as an axiom, since it gives a very restrictive view of the set-theoretic universe, excluding for example the notion of measurable cardinals and therefore most of the large cardinals, which are considered to be the backbone in modern set theory as discussed above. Another set of statements are the so called Forcing Axioms, such as MA, PFA and others. A set-theoretic universe in which such a statement holds is saturated under certain forcing constructions. This means that the universe itself already contains some of the objects that in other universes would have to be added by forcing. In most people’s mind this does not make for a natural axiom to add to ZFC. Certain among the forcing axioms, such as PFA, do settle the value of the continuum (to be \(\aleph_2\)), so some set-theorists, hold that such a statement could be an actual axiom. To mention some names, perhaps Menachem Magidor holds this view to a certain extent, as expressed in private conversations, but there is no written evidence of this. But really, to see how much set theorists do not take a position on any of these axioms, consider the Set Theory entry for [1], by Joan Bagaria. It says: “A central theme of set theory is thus the search and classification of new axioms. These fall currently into two main types: the axioms of large cardinals and the forcing axioms.” The mathematical properties of these axioms are then explained in two separate sections. That is the end. There is no philosophical conclusion or a choice made as to what and if any of these statements are indeed axioms. The author does not take sides.

Bagaria’s position is very representative of the modern set theorist. We do not feel that we know enough in order to take sides. Moreover, some of us do not believe that there ever will be a side to be taken. Whatever flavour of set theory we do, there is one common ground to all working set-theorists at the moment. We accept that there are many universes of the axioms of ZFC, which all have different properties and satisfy various additional statements or axioms. None of them is ultimate. This, evidently pluralist, view is certainly the only reasonable one at the moment, given the fact that mathematical evidence does not support the choice of any of these universes over any other ones. Therefore, not only the working set theorists of today do not see themselves as providing the unique ontological foundation of mathematics, but they do not believe in such a unique foundation and
do not claim it even for their own subject. I was asked if this means that practicing set theorists are no longer looking for the ultimate truth in the form of one universe or model of set theory. I think that this is the case. This is best expressed in the Multiverse View formulated by Joel Hamkins [14]. If the classical position, referred to by Hamkins as the Universe View, asserts that there is a unique background set-theoretic context or universe in which all our mathematical activity takes place, then the Multiverse View is exactly the pluralist opposite, asserting that there is no single universe but many and that only in understanding their multitude do we understand the actual nature of set theory. The Multiverse View accepts any universe consistent with ZFC.

4 Other Foundations

In addition to the set-theoretic view which we explained in the above, there have been several other approaches to foundations of mathematics. We shall briefly discuss two successful approaches: category theory and the, recently developed, univalent foundations.

ZFC set theory cannot formalise proper classes, such as the class of all sets, as seen by the Russell paradox. In other words, there are objects in mathematics, ‘collections’ which are not sets. Category theory proposes a formalization of such objects in terms of categories and their morphisms. Category theory is very appropriate for the study of mathematical objects by their abstract properties and has met with a large degree of success in subjects such as algebraic geometry.\(^3\)

There has been quite a bit of heated discussion on the subject of the preference for sets or for categories. For example, Saunders Mac Lane in [19] expresses the view that set-theoretical foundations are inappropriate for mathematics as practised and proposes category theory as an alternative. William Lawvere states in [17]: ‘In the mathematical development of recent decades one 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. 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 do not play any role”. This was part of Lawvere’s

\(^3\)There are other set theories that do formalise the notion of a proper class, but Russell’s paradox creeps in in other forms. This point does not change the conclusions here and discussing it is outside of the scope of the paper;
work in which he develops the foundations based on the elementary topoi, a concept that has grown out of his approach to the category of sets. More recent work on the theory of topoi is by Olivia Caramello who is building a methodology based on the theory of topoi to unify diverse parts of mathematics and logic. Gerhard Osius [20] gave an two extension ETS(ZF) of the elementary theory of topoi which is equivalent (by translation) to ZF. In the other direction, the straightforward set-theoretic foundations for category theory are inadequate as they, for example do not formalise the Grothendieck method of universes. Solomon Feferman then in [9] argues that even an unrestricted category needed for the method of universes can be given foundations within set theory. So perhaps this justifies the view expressed by MacLane in the introduction to [9], although Feferman argues against it, that set-theoretical foundations and categorical foundations are entirely equivalent, philosophically speaking. Discussion of relative merits of set theory versus category theory continues on the electronic forum FOM, initiated by H. Friedman. Yet in our opinion, the two foundations are complementary. For example, iterated forcing which in set theory we consider as an unquestionable part of the subject, is also used in topos theory and has even been used in arguments concerning computer sciences (see [3]). Yet, this does not mean that the open set-theoretic questions about it, such as for example what can be the cofinality of the first cardinal that fails $MA$ or if there is a forcing axiom at the successor of a singular cardinal or if forcing with the Talagrand algebra adds a random real, seem likely to be solved by translations to the topos theory, or that one can formulate the concrete questions such as these ones in a natural way using the topos. Vice versa, one would indeed be missing a mathematical opportunity if one would refuse to use categorical language to express the unifying results and conjectures such as the ones proposed by the Langlands programme.

A very recent, exciting, event in the foundations of mathematics is the discovery by Vladimir Voevodsky in [29] of univalent foundations. These are foundations for constructive mathematics and they have brought about a discovery of connection between the homology theory and the theory of types. Here the first order logic is replaced by the theory of dependent types. In this approach, it is the deduction system that changes, not just the axioms or the basic objects. This approach allows us, by using proof assistants such as Coq or Agda to formalise an important part of modern mathematics and to verify proofs of certain, quite involved, theorems, such as the Four Colour Theorem and the Feit-Thompson (Odd Order) Theorem. An important difference between the dependent type theory approach and the one offered by classical logic is the treatment of equality. In traditional foundations, equality carries no information beyond its truth-value: if two
things are equal, they are equal in exactly one way. So, the equality is just a two-place predicate on some set. In dependent type theory, equalities can carry information: two objects may be equal in multiple ways. The basic objects, which are called types, are equipped with an operation resembling a morphism and two are considered the same if there is a morphism between them. This type of equality conforms with the intuition that in mathematics we often consider objects that are equivalent as actually being equal. The Univalence Axiom formalises this practice, stating that two types are equal iff they are equal as the elements of the universe, which is a type of types. This gives for free many theorems stating that certain properties are preserved under equivalence, which otherwise need to be proved each time for each new construction. The same questions that one can ask in classical foundations can be asked in univalent foundations, starting from what we mean exactly by the foundations of mathematics. For example, Voevodsky’s initial view follows a platonist tradition, while Steve Awodey [2] argues for a structuralist point of view, based on the idea that the isomorphic objects are considered to be identical. Many developments in the univalent foundations are described in the collective HoTT book [22].

The univalent foundations with the axiom of univalence are consistent modulo the consistency of ZFC and the existence of two inaccessible cardinals (due to Voevodsky), as reported in a paper by Chris Kapulkin and Peter Lefanu Lumsdaine [15]). Other models of univalence are now known, using less power of the axioms, notably the cubical sets model which gives a computational interpretation of the univalence axiom. In fact a very fine analysis the correspondence between the weak subsystems of univalence and weak subsystems of ZFC is carried out in a recent, not yet published work, by Ulrik Buchholtz.

Some discussion has been going on as to the relationship between the set-theoretic foundations and the univalent foundations. Egbert Rijke and Bas Spitters in [24] show that sets in homotopy type theory form a $\Pi W$-pretopos, and under certain additional assumptions, even a topos of sets. This shows in some sense that set theory is a special case of the homotopy type theory. This is a step in the programme of rendering constructive and verifiable the arguments that are obtained by classical set-theoretic foundations, or even all of mathematics. The programme is now known as the formalization of mathematics. Yet, without entering into details of how far exactly this programme has reached, as the authors of [24] state in the conclusion of the paper, this is still a rather far away goal, but, as seen by further papers, also a continuing area of research.
5 What is the Foundation, a pluralist answer

The above short excursion in to the present state of knowledge about the main-stream candidates to the foundations of mathematics leaves little doubt as to what foundations to take for mathematics. It would be very unreasonable at this point of our knowledge to take any particular one of these candidates as foundations of all of mathematics. They each, including set theory, have their limits, and have their points of breakthrough. They are all basically equiconsistent, differing from each other over a few large cardinals, and moreover, they can be studied more locally to obtain a finer distinction between their subtheories, with respect to consistency.

There is even a question to be asked, is it really that obvious that mathematics should have an ontological foundation? The fact that we can axiomatise certain parts of mathematics, starting with Euclidean geometry, has traditionally been taken as a proof-of-concept idea for axiomatisation of all of mathematics and a guideline for an ontological foundation. But why should this really be so? Perhaps the time has come in the foundations of mathematics to speak differently than the Greek ‘gods of mathematics’, in the sense of moving away of the Euclidean view of axiomatisation (or to be more precise, the view of axiomatisation traditionally ascribed to Euclid, as there is some question as to whether he took this view himself, see [21]). Just as Hamkins convincingly argues in [14] that one should take an ultimately pluralist view of set theoretic universes, we feel that mathematical practice and our state of knowledge do not justify a choice of any particular foundation as the unique one. Such a choice in each instance leads to a loss of a considerable part of mathematics that is invisible to the foundations chosen. Foundations of mathematics the way envisioned since the time of the ancient Greeks leave us with the image of mathematics as a building based on these foundations. However, modern practice suggests that mathematics is more like a complex interconnected city with many buildings and many foundations.

This represents a shift in the use of the word ‘foundation’ from the ontological view of the 20th century philosophers of mathematics who were looking for the truth or the essence of mathematics. The more recent view is more epistemic, since the practicing mathematician is more after understanding than after ‘knowledge’. 4 This shift is probably what distinguishes the philosophy of mathematics from that more recent subject which some would call the philosophy of mathematical practice.

4We thank the referee for formulating the last two sentences, which describe exactly the point discussed.
We should also like to point out that this pluralist view of foundations of mathematics does not contradict non-pluralists views of what mathematics is itself. In particular, it does not contradict a platonist view of mathematics as a unique and deterministically defined object. The point is not if there is a uniquely defined mathematics, but whether our ability to describe it is uniquely defined. The former question is a question of belief, intuition and to some extent philosophy. The latter one is a question of mathematical practice and its philosophy, which must by definition be based on what we know rather than on what we believe to be true.

References


