Foundations of mathematics: an optimistic message*

Stephen G. Simpson
Department of Mathematics
Pennsylvania State University
State College, PA 16802, USA
http://www.math.psu.edu/simpson
simpson@math.psu.edu

First draft: November 18, 2015
This draft: December 31, 2015

Abstract

Historically, mathematics has often been regarded as a role model for all of science – a paragon of abstraction, logical precision, and objectivity. The 19th and early 20th centuries saw tremendous progress. The great mathematician David Hilbert proposed a sweeping program whereby the entire panorama of higher mathematical abstractions would be justified objectively and logically, in terms of finite processes. But then in 1931 the great logician Kurt Gödel published his famous incompleteness theorems, leading to an era of confusion and skepticism. In this talk I show how modern foundational research has opened a new path toward objectivity and optimism in mathematics.

SLIDE

SLIDE

Good evening to all. Basically I am a mathematics professor, and my task this evening is to explain something about my research area, mathematical logic and foundations of mathematics.

Throughout history, mathematics and logic have been regarded not only as indispensable components of most other sciences, but also as role models for all of science – the most objective and logically perfect of all the sciences. For these reasons, scientists throughout history have thought that it is particularly important to understand the logical structure of mathematics. And that is the need which my research area seeks to fulfill. *Foundations of mathematics* is the study of the most basic concepts and logical structure of mathematics.

*This document is the text of my Public Lecture at the National University of Singapore, to be delivered January 6, 2016.

Of course, mathematics is an ancient science which has evolved and continues to evolve over time. And it may be hard to believe, but throughout history there have been a number of *foundational crises* – times when there was *profound doubt* about the logical basis of mathematics. What I want to do here is to explore one of the issues which have dominated foundational research. The issue in question is centered on the concept of *infinity*.

SLIDE

I begin with a quotation from the greatest mathematician of the early 20th century, David Hilbert.

> "The infinite! No other question has ever moved so profoundly the spirit of man."

This ringing statement was the beginning of Hilbert's profound essay "On the infinite." Hilbert was a great mathematician, but why did he say this? Surely there are many *finite* things that move us profoundly. What is so moving about the infinite? Let me return to this question later, after giving you some background concerning the infinite and its role in mathematics.

Hilbert's essay "On the infinite" was a mature and definitive statement of his plan to place mathematics on a firm, solid, objective, logical foundation. But why was there even a need for such a foundation? After all, mathematics is normally regarded as the most basic, most certain, most logically perfect of all the sciences. Mathematics is the science of measurement and quantity, and *as* the science of measurement and quantity, it is an essential component of all of the hard sciences.

Furthermore, beyond its role in the hard sciences, mathematics has often been viewed as an excellent laboratory for intellectual training. Among the ancient Greek thinkers such as Plato, rigorous training in mathematics was regarded as an indispensable prerequisite for all serious intellectual activity. Above the gateway entrance to Plato's Academy was inscribed the phrase "Let none but geometers enter here." This widely accepted story is perhaps unverifiable, but at the same time it is certainly true in spirit. The point is that mathematics, perhaps moreso than other sciences, deals in powerful but precisely defined abstractions. By developing our ability to handle mathematical abstractions correctly, we also develop our ability to handle abstractions in other areas such as philosophy, law, politics, and even ethics.

So again, in an era when mathematics was flourishing and making great strides in many directions, why were great mathematicians such as Hilbert concerned about the foundations of mathematics, i.e., the most basic concepts and logical structure of mathematics? Why did they feel a need to shore up the foundations?

The answer is that, even as mathematics was flowering, some of the newer and more abstract mathematical abstractions became, well, a little bit hard to swallow. These new abstractions seemed to be more and more remote from observable reality, and this made mathematical practitioners uneasy. There was a real and growing concern that mathematics may have moved into some sort of "twilight zone."

SLIDE
SLIDE

Actually, this kind of unease had a precedent in ancient Greece. More than 2500 years ago, the philosophical school of Pythagoras was obsessed with *pure numbers*. But consider for example the abstraction of a geometrical square. Clearly, the ratio of the diagonal of the square to any one of the four sides of the square is a pure number, independent of the physical lengths of the diagonal or of the sides. This number is now known as $\sqrt{2}$, the square root of 2. It is the unique positive number which, when multiplied by itself, gives exactly 2. For a long period of time, it was believed that each and every pure number could be described as a ratio of whole numbers, for example the ratio 4 divided by 3, or the ratio 29 divided by 19. However, the Pythagoreans were horrified to discover that this is not the case. In fact, they discovered that the square root of 2 *can never be described exactly* as the ratio of two whole numbers. It is *approximated* rather well as the ratio 99 divided by 70, and it is *approximated* even better as the ratio 665857 divided by 470832. But the Pythagoreans proved that there is no way to express $\sqrt{2}$ *exactly* as a ratio a divided by b where a and b are whole numbers, no matter how large a and b are. In other words, $\sqrt{2}$ was the first example of what is now called an *irrational number*. The Pythagoreans viewed this discovery as horrifying and subversive. For a significant amount of time they tried to keep the irrationality of $\sqrt{2}$ hidden as a deep, dark, esoteric secret, known only among their own exclusive group. Historically we may say that the discovery of the irrationality of $\sqrt{2}$ led to the first major crisis in the foundations of mathematics.

Moving forward through history, there has been a sequence of higher and higher mathematical abstractions, more and more remote from real-world measurements and real-world quantities. Even in the realm of *numbers*, the names chosen for these new classes of numbers give a hint of the kind of tension that has always existed in mathematics and science, between the abstract and the concrete. The very phrase *irrational number* may be viewed as an arrogant challenge to the power of reason and science. And far beyond the irrational numbers, 19th and 20th century mathematicians have introduced ever-more-remote abstractions such as the *transcendental* numbers, the *imaginary* numbers, the *complex* numbers, the *transfinite* ordinal numbers, the *uncountable* cardinal numbers, the *inaccessible* numbers, the *hyperinaccessible* numbers, the *ineffable* numbers, etc. Each of these advances met with considerable resistance when it was first introduced. Note also that each of these classes of numbers involves an appeal to the infinite. For instance, the decimal expansion of $\sqrt{2}$ is an infinite and nonrepeating sequence of digits,

$$1.4142135623730950488016887242096980785696718753769480731\ldots.$$

And the other kinds of abstract numbers which I just mentioned involve infinities which are even more extreme.

SLIDE
SLIDE

Throughout history, the tension between the abstract versus the concrete, or the infinite versus the finite, has been been pondered deeply. This tension is depicted graphically in a famous fresco, known as "The School of Athens," by the Italian Renaissance artist Raphael. In this 16th century masterpiece, Raphael depicts the two masters of ancient Greek philosophy, Plato and Aristotle, walking among their Athenian colleagues and debating the nature of the infinite and its relationship to the finite. Plato is pointing upward, toward the heavens, while Aristotle is spreading his hand powerfully and benevolently over the earth. This painting by Raphael is often regarded as the perfect embodiment of the classical spirit of the Renaissance. In the context of Raphael's fresco, Hilbert's statement about the spirit of man may be more readily understood.

Aristotle is arguably the greatest philosopher of all time. His major work on the philosophy of existence is known as *the Metaphysics*. The Metaphysics is a massive tome, hundreds of pages long. The purpose of the Metaphysics is to answer the most profound questions about *the nature of existence*. A typical question discussed in the Metaphysics is, do abstractions exist in reality, or only in the human mind? As part of the Metaphysics, Aristotle includes a highly nuanced and thorough discussion of the nature of *mathematical existence*. Books M and N of the Metaphysics are the first-ever treatise on the philosophy of mathematics.

As part of Books M and N of the Metaphysics, Aristotle fashions a *grand compromise* between the infinite and the finite. This grand compromise depends on a crucial distinction between two kinds of infinity: *potential infinity* versus *actual infinity*.

Potential infinity is the kind of infinity associated with a process that can be repeated or continued indefinitely. Many processes of this kind can be observed in reality. For instance, there is the earth turning on its axis, the planets revolving around the sun, the march of time, the motion of a body in outer space, the cycle of generations, the accumulation of wealth, the accumulation of knowledge, etc. Each of these processes is perfectly capable of repeating or continuing, without any forseeable limit. There are many examples of potential infinity in mathematics. The most important example is the potentially infinite sequence of whole numbers, 1, 2, 3, 4, 5, 6, 7, dot dot dot. Another example is a potentially infinite sequence of smaller and smaller rational numbers, for example 1, 1/2, 1/4, 1/8, 1/16, 1/32, dot dot dot. Yet another example is the potentially infinite sequence of decimal approximations to the square root of 2. The process of computing more and more decimal digits of the square root of 2 can be continued indefinitely. All of these potential infinities in mathematics play an important role in our understanding of real-world phenomena.

Actual infinity is another kind of infinity, standing in contrast to potential infinity. An *actual infinity* is a *completed infinite totality*, with emphasis on the word *completed*. As a mathematical example, imagine that we have just now finished enumerating *the entire sequence of all the whole numbers*, and then imagine that we pause to look backward over the entire sequence. This stopping point is sometimes denoted by the symbol ω, because ω (pronounced "omega") is the last letter of the Greek alphabet. In abstract mathematics, the symbol ω denotes an actual infinity, a completed infinite totality. The symbol ω may be viewed, and often *is* viewed, as standing for the collection of *all* whole numbers, all gathered together into one infinite basket.

The philosophical question at issue here is:

Which kinds of infinity have an objective existence?

Do potential infinities exist objectively? Do actual infinities exist objectively? Aristotle comes down on the side of a rather compelling philosophical position. The Aristotelean position may be called *finitism*, but it is really a compromise between the finite and the infinite. According to Aristotelean finitism, potential infinities are part of objective existence, but actual infinities are not.

Over the centuries, many thinkers have agreed with Aristotle's finitism, simply because in observable reality there are many obvious examples of potential infinity, but there are no convincing examples of actual infinity. On the other hand, there have also been many non-finitistic thinkers who profoundly disagreed with Aristotle. Some of them have endorsed a great variety of actual infinities, typically on the grounds that these actual infinities exist as an important component of theological, philosophical, or mathematical thought. As a possible response to these non-finitist thinkers, we hard-headed realists could note that existence-in-thought is different from the kind of objective existence which Aristotle was analyzing.

While we are at it, let me note that a great majority of contemporary mathematical practitioners are indifferent to philosophical questions such as potential infinity versus actual infinity. They wish only to carry on with their day-to-day mathematical work, extending current mathematical paradigms and unencumbered by philosophical issues.

We have been dwelling on ancient Greece, and especially Aristotle's finitism. And again I want to stress that Aristotle's finitism does not reject all infinities. Rather, Aristotle's finitism fully accepts the concept of potential infinity, while at the same time rejecting the concept of actual infinity.

SLIDE

But now let us fast-forward to the late 19th and early 20th centuries. During this relatively modern period of history, mathematics was undergoing a great expansion in all directions: differential equations, geometry, number theory, algebra, etc. And as part of this tremendous flowering of mathematics, great strides were being made in the direction of mathematical rigor and logical precision. The work of Cauchy, Weierstrass, Dedekind, Peano, Frege, Cantor, Russell, Whitehead, and Zermelo cleared the way toward a perfect logical framework for all of mathematics. A landmark was the treatise of Hilbert and his assistant Ackermann. Hilbert and Ackermann presented an elegant and completely symbolic formulation of the logical axioms and rules for what is now known as the predicate calculus. Even though digital computers did not yet exist, it was in some sense clear that mathematics and perhaps many other scientific disciplines could in principle be completely *formalized*, i.e., programmed for digital computers.

Later in the 20th century, levels of mathematical abstraction continued to increase rapidly, with a great variety of actual infinities. Consequently, due in part to the non-objective nature of these actual infinities, there were nagging and sometimes massive doubts about the validity of mathematics. The philosopher Russell famously characterized mathematics as "*the science in which we never know what we are talking about, nor whether what we say is true.*" The mathematical physicist Wigner declared that there is no rational explanation for the usefulness of mathematics in the physical sciences. The mathematician Morris Kline published an influential popular book entitled "*Mathematics: the loss of certainty.*" None of these scientists were viewed as enemies of mathematics. They were viewed as merely raising legitimate doubts about the validity of mathematics.

SLIDE

Highlighting a key step in this loss of certainty in mathematics, I want to say something about the ideas of great mathematician David Hilbert and the great logician Kurt Gödel.

Concerning our question about Aristotle's finitism, Hilbert had already made dramatic use of non-finitistic methods early in his career, in his work in a branch of algebra called invariant theory. Other finitistically-minded mathematicians criticized Hilbert's method. For instance, the mathematician Gordan said that Hilbert's work in this area was not mathematics but rather theology. Later, at the peak of his career, Hilbert argued forcefully for the acceptance of actual infinities in mathematics. "*Noone shall expel us from the paradise that Cantor has created for us!*" The paradise of which Hilbert was speaking was, of course, a massively non-finitistic paradise, chock full of actual infinities.

However, Hilbert also appreciated the finitistic viewpoint. Indeed, he saw finitistic mathematics as the most indispensable part, the *meaningful core* of mathematics. Attempting to reconcile these two viewpoints, Hilbert developed a sophisticated plan whereby non-finitistic mathematics would be "reduced" to finitistic mathematics, i.e., validated in a fully finitistic way. This would be achieved in three steps.

1. Formalize finitistic and non-finitistic mathematics as two distinct but logically perfect systems of axioms and rules of inference. The key technical tool here would be the predicate calculus of Hilbert and Ackermann.

2. Embed the finitistic system as a subsystem of the infinitistic system. In essence, these first two steps had already been accomplished by the late 1920s.

3. Argue finitistically that any finitistically meaningful conclusion which is deducible in the non-finistic system would also be deducible in the finitistic system.

In this way, all of the actual infinities would be seen as "ideal elements" or "harmless but convenient fictions." Because of this reduction, mathematicians would have full license to use the actual infinities freely. The actual infinities would be *guaranteed to be harmless*, once and for all, because we would know in advance that they could never contaminate the finitistically meaningful core of mathematics.

In short, when it came to the question of potential infinity versus actual infinity in mathematics, Hilbert was trying to ride to the rescue. He was proposing a bold, sophisticated, far-reaching plan to compromise between finitism and non-finitism. He wanted mathematicians to have it both ways, i.e., to have our cake and eat it too. This proposal is now known as *Hilbert's program* or, in a more technical vein, *finitistic reductionism.* The full statement of Hilbert's program is spelled out in his 1926 essay entitled "On the infinite."

SLIDE

But then, five years later in 1931, Gödel published his famous paper proving that Hilbert's program is impossible to carry out in its entirety. In fact, the work of Gödel revealed a hierarchy of stronger and stronger formal mathematical systems, relying more and more extensively on actual infinities, leading to more and more consequences which are finitistically meaningful but not finitistically verifiable. This hierarchy is known as *the Gödel hierarchy.* The Gödel hierarchy undercuts Hilbert's program, because it shows us that every time we introduce a new kind of actual infinity, it can result in more and more contamination of the finitistically meaningful core.

SLIDE

This discovery by Gödel was viewed as a complete and total defeat for mathematical finitism. The mathematical world seemed to be left with only two options. Either cling to objectivity by rejecting all non-finitistic methods, or abandon objectivity by embracing the panorama of actual infinities. Since mathematicians were by that time quite comfortable with actual infinities, they almost universally chose the second option. Finitism, finitistic reducibility, and objectivity were now apparently out the window. Even today, despite modern advances, the popular wisdom or "settled science" is that Gödel's theorems meant the end of Hilbert's program and the end of objectivity in mathematics.

In the title of this lecture, I promised an optimistic message. So let me now finish by reporting on some results of modern research, dating from the 1970s to the present. The big discovery here is that, despite Gödel, a large and significant part of Hilbert's program can in fact be carried out with brilliant success!

SLIDE

To explain how this can be so despite Gödel, let me return to the Gödel hierarchy. There are two important points to make.

The first point is that Hilbert's program can be carried out for the so-called "weak" levels of the Gödel hierarchy. In other words, we can prove finitistically that for any proof of a finitistically meaningful statement in one of these "weak" non-finitistic systems, we can find an alternative proof which is purely finitistic. This is a consequence of the proof-theoretical methods pioneered by Hilbert and Gentzen, as outlined by Hilbert in his 1926 essay.

In order to explain the second point, I need to discuss an ongoing research direction in the foundations of mathematics. Beginning in the 1970s and continuing through the present, mathematical logicians have carried out a series of foundational case studies known as *reverse mathematics*. The goal of reverse mathematics is to determine which levels of the Gödel hierarchy are essential for the formalization of which parts of mathematics. This huge and ongoing research program, reverse mathematics, is the subject of an international conference or workshop being held this week and next (January 3–16, 2016) here at NUS, at the Institute for Mathematical Sciences.

As an outcome of reverse mathematics, we now know that the bulk of mathematics, including a great deal of non-finitistic mathematics, can in fact be formalized at the "weak" levels of the Gödel hierarchy. This is of course an ongoing project, but I estimate that it applies to something like 85 to 95 percent of mathematics as it exists today, including not only constructive or computable mathematics but also a great deal of highly nonconstructive mathematics.

In many cases, some care is needed. For instance, a central feature of the proof of Peano's existence theorem for solutions of ordinary differential equations is a sequential compactness argument involving the Ascoli lemma. In order to show that the Peano theorem is finitistically reducible, we need to rework this sequential compactness argument into a covering compactness argument. However, the reworking is purely technical and does not affect the mathematical significance.

So what is my optimistic message? My optimistic message is that, by focusing on the "weak" levels of the Gödel hierarchy, we can carry out Hilbert's program for at least 85 percent of mathematics. In other words, we can have our cake and eat at least 85 percent of it. Yes, some of the most blatant nonfinitistic moves can result in contamination, but not by any means all of them. Our message is that, if mathematicians exercise some care and work within the strictures of the known "weak" systems or other systems which are finitistically reducible, then all is well. The finitistic core is protected from contamination. Hilbert's program lives!

To summarize, what I am suggesting is that, even though Hilbert's program cannot be carried out in its entirety, there is a less sweeping version of Hilbert's program which can be carried out, and which represents the *proper balance* between the finite and the infinite in mathematics.

Now, if only we could find the proper balance between the finite and the infinite in theology!

References

[1] Hippocrates G. Apostle. *Aristotle's Philosophy of Mathematics*. University of Chicago Press, 1952. X + 228 pages.

[2] C.-T. Chong, Q. Feng, T. A. Slaman, and W. H. Woodin, editors. *Infinity and Truth*. Number 25 in IMS Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore. World Scientific, 2014. IX + 234 pages. 8

[3] S. Feferman, C. Parsons, and S. G. Simpson, editors. *Kurt Gödel: Essays for his Centennial*. Number 33 in Lecture Notes in Logic. Association for Symbolic Logic, Cambridge University Press, 2010. X + 373 pages. 8

[4] Stephen G. Simpson. Partial realizations of Hilbert's program. *Journal of Symbolic Logic*, 53(2):349–363, 1988.

[5] Stephen G. Simpson. *Subsystems of Second Order Arithmetic*. Perspectives in Logic. Association for Symbolic Logic and Cambridge University Press, 2nd edition, 2009. XVI + 444 pages.

[6] Stephen G. Simpson. The Gödel hierarchy and reverse mathematics. In [3], pages 109–127, 2010.

[7] Stephen G. Simpson. An objective justification for actual infinity? In [2], pages 225–228, 2014.

[8] Stephen G. Simpson. Toward objectivity in mathematics. In [2], pages 157–169, 2014.

[9] Richard Zach. Hilbert's Program. Stanford Encyclopedia of Philosophy, January 2015. http://plato.stanford.edu/entries/hilbert-program.