EN ROUTE FOR INFINITY

Jan Coufal – Jiří Tobíšek

Abstract

This paper presents the birth and development of the term "infinity" in mathematics and in philosophy. Infinity is an abstract concept describing something without any bound or larger than any number. Ancient cultures had various ideas about the nature of infinity. The route for infinity goes from Sumer in the 4th millennium B. C. via Greece (Pythagoras, Aristotle, Euclid – the Ancient Greeks did not define infinity in precise formalism as does modern mathematics, and instead approached infinity as a philosophical concept) to the 19th century Prague (Bolzano), Braunschweig (Dedekind) and Halle (Georg Cantor formalized many ideas related to infinity and infinite sets during the late 19th and early 20th centuries). The article ends in 1900 Paris, at the Second International Congress of Mathematicians, where David Hilbert announced his famous list of 23 unsolved mathematical problems, now known as "Hilbert's problems" and in 1904 Heidelberg in the Third International Congress of Mathematicians, where Gyula Kőnig delivered a lecture where he claimed that Cantor's famous continuum hypothesis was false. An error in Kőnig's proof was discovered by Ernst Zermelo soon thereafter.

Key words: infinity, cardinality, ordinality, actual infinite, potential infinite

JEL Code: B16, C63

Introduction

Number theory is perhaps as old as civilization. Since the dawn of history, humanity has been fascinated by numbers and their properties. All this shows that the human mind asks itself questions so strange when infinity enters in, that one must not be surprised if there is difficulty in reaching an answer. The way to infinity has no end: and in order to follow it one has to make some choices. The starting point could be in the 19th century, the time when the mathematical idea of infinity could be said to have the right to exist. It could, alternatively, follow a route through various disciplines of mathematics to show how the idea of infinity has gradually infiltrated them. We have instead chosen a different path, leaving aside certain exceptional monuments on the way, such as the infinitesimal calculus. The decision looks at infinity in number, and tries to show how, throughout the course of time, men and women answered these two fundamental questions about cardinality and ordinality:

How does it count? How does it order?

Unfortunately, we will have to leave aside the contributions of some prestigious civilizations, such as those of the Arabs or of the Chinese. But think of the journey as a holiday journey: it takes motorways that lead to places of great reputation, but it has chosen an alternative route in order to visit a place of special interest. From small beginnings it will make the journey towards the infinite, enumerable or continuous, and reach right up to the foothills of the transfinite. The early Egyptian hieroglyphics and Chinese writing contain the first indications of “three” meaning “many” or “plural”. The step of Three is the decisive one, which introduces the infinite progression into the number sequence. Even though it cannot be for certain, it appears that language contains evidence of struggle to pass beyond the old barrier of two. “Three” is often associated with many as, for example, *three* and *through*, or the Latin *tres* and *trans*, or the French *trois* and *très*. After one and two, the ordinals were formed “third, fourth, fifth, ...” Formerly, first kept its meaning, which is before all the others, while we find second used for “the other” or “the one that comes after”, compare the Latin “secundus” (from *sequi*, *secutus* – to follow). It is difficult to draw a distinction between “second” and “twice”. Let us go on further, leaving the few for the many. They appear on the earliest documents that have come down to us and they remain unchanged throughout the three thousand years of Egyptian civilization. As for the number system that was used, it is similar to that used by most peoples who want to be able to deal with large numbers that is a base ten additive system. We find evidence of various ways of representing ordinals: two written forms for “twice” or “second” and one way of writing numbers from three to ten followed by a different way after ten. But guided by their care for purity and by their artistic taste, the Ancient Egyptians used different figures to refer to different powers of ten. They were obliged to use many ingenious methods for carrying out their calculations. These finite decompositions must have reinforced their taste for exactness and finiteness. History starts at Sumer, between the Tigris and Euphrates rivers. The proto-Sumerian scribes of the 4th millennium B. C. adopted a hybrid system of numbering deriving from a variety of units of measure and based upon 2, 3, 6 or 10. Sexagesimal (base 60) is a numeral system with sixty as its base. There is no concern here for arithmetic purity, the system being at the same time both additive and multiplicative, nor for aesthetics, the signs being simply the imprints of a reed or a piece of ivory in clay. The development of writing gradually led to cuneiform writing using a stylus. The Greeks and the Arabs were certainly aware of the value of such a system of numeration. Alongside their own additive systems they also used the Babylonian base sixty system. Both the Egyptian methods and the Babylonian sexagecimal methods had this in common: they used finite processes that fitted the needs of those who calculated. It is still used (in a modified form) for measuring time, angles, and geographic coordinates. The time for "theorizing" had yet to come. We shall leave this strictly numerical domain if we are to reflect upon the nature of its elements. The idea of number itself may help us better to understand what concerns us. In astronomical texts of Hellenic Greece, such as the writings of Ptolemy, sexagesimal numbers were written using the Greek alphabetic numerals, with each sexagesimal digit being treated as a distinct number. In medieval Latin texts, sexagesimal numbers were written using Hindu-Arabic numerals; the different levels of fractions were called "minuta" (i.e. fraction), "minuta secunda," "minuta tertia," etc. By the seventeenth century it became common to denote the integer part of sexagesimal numbers by a superscripted zero, and the various fractional parts by one or more accent marks.

1 Crossing the Mediterranean Sea

Sailing across the waters of the Mediterranean, we may well follow the very route, as the legend has it, taken by Pythagoras. Going to Egyptian and Babylonian sources, he learned from them first, a certain mystique of numbers, and from the second their astronomical knowledge (Verdet, 1987, pp. 15–17): "*Order and disorder, this bipolarity can be found alike in the serious scientific study of the heavens as well as in popular works. If Aristotle, whose philosophy ruled western scientific thought for two thousand years, did not deal with comets in his* De Caelo*, but in his* Meteorologica*, it was because, by the disorder that they introduced, they could only belong to the lower levels of atmosphere, hardly higher than where storms were formed and winds were born, certainly not to the higher spheres where the stars moved according to immutable laws*." Well before Aristotle, the Pythagoreans had laid down the immutable laws of number. Their view that "all is number" is literally true in our modern digitalized world and led them to put representation, harmony and laws on the same footing. As order and disorder seemed to be mutually opposed in the heavens, might not number, whole and necessarily feint, be the means of better understanding the infinite?

Aristotle had been a teacher to the founder of this city and many famous mathematicians had connections to it, such as Euclid, Hero, Diophantus, Ptolemy, Pappus, Menelaus, and others. How did Greek mathematicians and philosophers understand the notion of infinity? The discovery of irrationality, the opposition between the finite and the infinitive, had proved a fatal blow for Pythagorean philosophy. Zeno, philosopher of the second half of the 5th century B.C., had set forward certain paradoxes to show the impossibility of deciding between a finite or atomistic view and a non-finite or continuistic view. The paradox of Achilles and the tortoise is the best known (Aristotle, 1910-1952, vol. II, V 239b): *[It] amounts to this, that in a race the quickest runner can never overtake the slowest, since the pursuer must first reach the point where the pursued started, so that the slower must always hold the lead […], but it proceeds along the same lines as the bisection-argument, for in both a division of the space in certain way leads to the result that the goal is not reached, […]. And the axiom that the one who holds the lead is never overtaken is false: it is not overtaken, it is true, while he holds the lead: but it is overtaken nevertheless if it is granted that he traverses the finite distance prescribed.* Aristotle, critical of these paradoxes, set up a theory of the infinite in distinguishing between the actual infinite and the potential infinite (Aristotle, 1910-1952, vol. III, 207b): *Our account does not rob the mathematicians of their science, by disproving the actual existence of the infinite in the direction of increase, in the sense of the non-traversable. In point of fact they do not need the infinite and do not use it. They postulate only that the finite straight line may be produced as far as they wish.*

We have the potential of being able to write down all the natural numbers. But we can also see that the collection of all the natural numbers cannot be envisaged as a whole, it is impossible to write them all down. In other words, to use modern terminology, they do not form a set. This has been to be the fixed position for almost two thousand years. As for Euclid, his use of the infinite was at a number of levels. Not being able to use the idea of an actual infinite set, he had to set out his statements in the form of the potential infinite. Euclid's *Elements* IX, 20 (Heath, 1926, vol. 2, p. 412) reads: "*Prime numbers are more than the assigned multitude of prime numbers*." Today we would prefer to say: "*The set of primes is infinite*." This realisation or actualisation of the infinite makes no fundamental change to the Euclidean proof. If one takes the product of finite number of primes and adds 1, the result is either another prime, or has prime factors which are different from the original primes. In fact, the language here is illusory: there is no calculating procedure or theory that will allow us to understand this property. What, for example, is the (10^{1995})-th prime number? Are we forced to await the arrival of even more powerful computers before we can know? Euclid states this as a definition in Book V of the *Elements* (Heath, 1926, vol. 2, p. 114): *Magnitudes are said to have a ratio to one another and are capable, when multiplied, of exceeding one another*. Leave ratios aside: what this means is that given any two magnitudes, then adding a number a sufficient number of times to the smaller, we shall obtain a magnitude that is greater than the larger. The Archimedean character of Euclidean magnitudes has many advantages. Further, it can be used as the justification for the “potential limit” contained in Euclid’s *Elements* X (Heath, 1926, vol. 3, p. 14): *Two unequal magnitudes being set out, if from the greater there be subtracted a magnitude greater than its half, and from that which is left a magnitude greater than its half, and if this process be repeated continually, there will be left some magnitude which will be less than the lesser magnitude set out*. The potential of the infinite is no longer merely in the language, it has become operational. It involves a justification for the “exhaustion” of the method of exhaustion. It is sufficient to be able to find a process that leads to “subtracting a magnitude greater than its half.” However, recent readings of the Archimedes Palimpsest have hinted that Archimedes at least had an intuition about actual infinite quantities. The Aristotelian principle of the potential infinite is going to have a hard life. Stricter mathematical objections there were joined by religious ones. To quote Descartes (1973, p. 108), for example, “*...and because we do not know how to imagine how many more stars God may create, we assume that their number is indefinite. And we call these things indefinite rather than infinite in order that the word infinite be kept only for God.*”

2 From Prague to Braunschweig

Prague was once one of the great cities of cultural and scientific tradition. But by the end of the 18th century it had become rather a backwater, far removed from centres of political and intellectual decision making. This, to some extent, explains the lack of influence and the little that is known of one of its sons, Bernard Bolzano. He was ordained a priest in 1804 and was much occupied with religious, ethical, and social matters, the latter casting him as a radical and leading to his dismissal from the chair of philosophy. He was attracted to philosophy, methodology of science, and particularly to mathematics and logic. Indeed, it was the subject of one of his most celebrated writings, the posthumously published *Paradoxien des Unendlichen*. Bolzano was the first to “actualise” the infinite, considering collection as a whole (Bolzano, 1851, § 4, p. 77): “*I call a set a collection where the order of its parts is irrelevant and where nothing essential is changed if only the order is changed.*” But it is as if he drew back somewhat when he wished to give an example of an infinite set: what he offers is an example of the infinite that arises from continued increasing (Bolzano, 1851, § 13, p. 14): “*It can easily be seen that the set of all absolute and true propositions is an infinite set. [...] For any number, no matter how large, there exist the same number of distinct propositions, and beyond these we are able to construct new propositions or, rather, there are such propositions whether we construct them or not.*”

Could it be otherwise? The actual infinite presupposes a good acquaintance with the infinite and therefore with, on the one hand, the finite and on the other, with a set of natural numbers. In this way numbers become a free creation of human mind. In 1872 in Braunschweig, Dedekind followed these lines. He had just published his construction of the reals from the rationals. In his second publication, *Was sind und was sollen die Zahlen?*, he stated (Dedekind, 1888, p. 31): "*...numbers are free creations of the human mind, they serve as means of apprehending more easily and more sharply the difference of things*." But Kronecker took different view (Weber, 1991/92, p. 14): "*The integer numbers were made by God, everything else is the work of man.*" Here, in difference between the view that numbers are a creation of human mind and the view that numbers are God given, lies the essence of the two mathematical views of the potential and the actual infinite. Dedekind, for the first time, was bold enough to offer a definition of an infinite set: it is a set for which there is a bijection (one-one correspondence) between it and a proper subset (Cohen, 1964, § 64, p. 63): "*A System S is said to be infinite when it is similar to a proper part of itself; in the contrary case S is said to be a finite system*." In order to complete his construction of the set of natural numbers, he was led to use Bolzano's example of the set of thoughts. Zermelo (1908) set out a system of seven axioms for set theory: here is his axiom for the infinite (Zermelo, 1908, pp. 266–267): "*Axiom of infinity: There exists in the domain at least one set* Z *that contains the nullset as an element and is so continued that to each of its elements a there corresponds a further element of the form* $\{a\}$*, in other words, that which each of its elements a it also contains corresponding set* $\{a\}$ *as an element.*" In other words, the domain of objects considered to be sets (given that they satisfy Zermelo's axioms) contains a set Z which contains the set $\{a\}$ whenever it contains the element a. The smallest set Z_0 having the properties of Z allows us, purely symbolically, to define (Zermelo, 1908, p. 267) *the simplest example of countable infinite set*. This is the set whose first members are $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$ and, in general, each $\{a\}$ is the successor of a. In order, we have: $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\{\emptyset\}\}$, …, $n + 1 = \{n\} = \{\{\dots\{\emptyset\}\dots\}\}$. Other definitions were proposed for infinite sets or sets isomorphic to Z_0. A definition containing more ordinals is, e.g.: $0 = \emptyset$, $1 = \{0\} = \{\emptyset\}$, $2 = \{0, 1\} = \{\emptyset, \{\emptyset\}\}$, …, $n + 1 = n \cup \{n\} = \{0,1,2, \dots, n\}$ where the integer n contains n distinct elements. Now that the potential infinite has become actual, there remain other questions to face: How can infinite sets be counted? How can infinite sets be ordered?

3 Correspondence between Braunschweig and Halle

Hardly hundred miles lies between Braunschweig and Halle where Cantor spent his whole life as a teacher, and correspondence between Cantor and Dedekind was rapid. In 1872 both published their constructions for the real numbers and consequently came to introduce the actual infinite in their work. Cantor wanted to be able to count these new infinite sets, both the rationals and the reals. To do this he used the idea of a one-one correspondence. This was a natural extension of countability from finite sets to infinite ones. But the task was not without difficulties and came up against many firm convictions, not least the Euclidian adage (Heath, 1926, vol. 1, p. 155): "*Whole is greater than the part*." On the 29th November 1873, Cantor wrote to Dedekind (Noether, Cavaillès, 1937, pp. 187-188): "*Let me to present to you a question which has a certain theoretical interest for me, but to which I can find no answer. [...] Suppose we take the set of all the individual positive integers* n *and represent it by* (n): *then we take the set of all the real positive numbers* x *and represent it by* (x): *the question is simply to know whether* (n) *can be put into a correspondece with* (x) *in such a way that to each element of one of the sets there correspondens an element and one only of the other*." Cantor was in fact perplexed, as this at first sight appears impossible since (n) is discrete and (x) is continuous. On the other hand, it is possible to set up one-one correspondence between (n) and the set of rationals and even with the set of integer n-tuples. On the 2nd December, Dedekind added to Cantor's confusion when he showed that the set of algebraic numbers, real or complex, is also countable. This result seems rather less surprising when one considers that all algebraic numbers are roots of polynomials with integer coefficients and for any given degree they are enumerable. Consequently, we obtain a countable union of countable sets. It remains to show that this union is also a countable set. Now this is not at all evident unless we introduce a further axiom for the infinite, the axiom of choice, which was not formulated until 1904 by Zermelo! Dedekind did not give much encouragement to Cantor to continue to pursue the question of countability of the reals (Noether, Cavaillès, 1937, p. 194): "*[It] does not justify spending too much time on it, since it is of no practical importance.*" In no way discouraged, Cantor sent Dedekind his first proof that the reals were uncountable, in a letter of 7th December. This was a scornful rejection of practical importance! There are many more transcendental numbers than algebraic numbers yet only a few are known to us. In other words, Cantor had just discovered that there are two sorts of infinities: countable and continuous. Stung by curiosity, Dedekind sent a simplified proof in replying letter (Cantor, 1878): "*Suppose that an interval* $\langle a, b\rangle$ *is countable: then it is possible to construct a sequence of enclosed intervals* $(\langle a_n, b_n\rangle)$: a_0 *and* b_0 *are respectively equal to* a *and* b *and for each integer* n, a_{n+1} *and* b_{n+1} *are equal to the* *first two numbers of* $\langle a, b \rangle$ *which belong to an open interval* (a_n, b_n)*. The convergence of the sequences* (a_n) *and* (b_n) *leads to a contradiction*." Having established two sorts of infinities, countable infinities and the infinity of the continuum, Cantor went on to develop the idea of cardinality for infinite sets using the idea of *power* (Cantor, 1878, p. 242): "*If it is to set up a one-way correspondence, element by element, between two well defined sets* M *and* N*[...] it will be covenient to express this by saying that these two sets have the same power.*" He already was aware of the powers of countable sets and of the continuum. Were there others? Cantor was able to give a positive reply, in that the continuum has the same power as the set of subsets of N and a power strictly greater than that of N. More generally, the set $P(X)$ of subsets of X has a power strictly greater than that of X. In other words, X and $P(X)$ do not have the same power and there exists a mapping of X to $P(X)$ which is *into* and not *onto*. But hardly had this property been discovered when there arose one of the first paradoxes of set theory which Cantor identified. If the collection of all sets is some set A, then the power set of A, $P(A)$, must also be an element of A, yet it has power strictly greater than that of A. In 1908 Zermelo found a way round this paradox by introducing the axiom of selection. This said that a new set could only be made up from a section of elements of an existing given set, in other words, the collection of sets could not be set. This paradox being resolved, other question arose. Did another power exist between the countable infinite and the continuum? More generally, was there a power between that of a set and that of its power set? Can any two arbitrary sets be compared in terms of their powers? Since Cohen (1964) we have known that we cannot give an answer to all these questions. Cantor (1878, p. 242) asserted that *if two sets* M *and* N *are not of the same power, then* M *has the same power as an integral part of* N*, or* N *has the same power as an integral part of* M. Cohen (1964) showed precisely that the axiom of choice is unprovable. It is just the same with the hypothesis of the continuum and with the generalized hypothesis of the continuum which state, respectively, that there does not exist any power between that of the countable and the continuous and between that of a set and its set of subsets. In other words, we can claim that all sets have powers that can be compared. But we could just as easily claim the opposite. Mathematics is not one and indivisible!

We are using the fact that the set of non-zero natural numbers has the same power as the set of naturals including zero. We create a new number ω which comes after all the natural numbers. This is the first transfinite number. This transfinite labelling or indexation can be extended. Knowing that there are as many integers as there are even integers or odd integers, we can assign the rank p to the even integer $2p$. We can dispose of the integers by classifying the odd numbers, following Cantor, by creating new transfinite numbers $\omega + p$ which can be used to to assign the rank $\omega + p$ to the odd number $2p + 1$. After this, a new arrival could be given the number $\omega + \omega$, labeled 2ω. Cantor went on, step by step, to define a countable infinity of transfinite numbers: $\omega,\ \omega + 1,\ \omega + 2, \dots \omega + p, \dots, 2\omega, 2\omega + 1,\ 2\omega + 2, \dots, n\omega, n\omega + 1,$ $n\omega + 2, \dots,\ \omega^n,\ \omega^n + 1,\ \omega^n + 2, \dots,\ \omega^\omega,\ \omega^\omega + 1,\ \omega^\omega + 2, \dots$ The set of these transfinite numbers is countable: writing out all the real numbers is not enough. Cantor did not stop here, he extended the scheme by which he had defined ω. Following the class of all transfinite numbers as defined here, there exists a new transfinite number, as there did after all the integers.

Conclusion

Hilbert's famous speech on "The Problems of Mathematics" was delivered to the Second International Congress of Mathematicians in Paris on 8 August 1900 in which he described 10 from a list of 23 problems (Hilbert, 1900). The first problem related to the continuum (Hilbert, 1900, p. 263): "*The question now arises whether the totality of all numbers may not be arranged in another manner so that every partial assemblage may have a first element, i. e., whether the continuum cannot be considered as a well ordered assemblage – a question which Cantor thinks must be answered in the affirmative. It appears to me most desirable to obtain a direct proof of this remarkable statement of Cantor's, perhaps by actually giving an arrangement of numbers such that in every partial system a first number can be pointed out*.

At the Third International Congress of Mathematicians held at Heidelberg in 1904, Kőnig (1905) claimed to be able to give a negative reply, i.e. Kőnig gave a speech to disprove Cantor's continuum hypothesis. The announcement was a sensation and was widely reported by the press. Ernst Zermelo, the later editor of Cantor's collected works, found the error already the next day. This bolt from the blue shook the Cantorian community and on 24th September 1904, Zermelo (1908) showed that every set could be well ordered. But he used what come later to be called the axiom of choice. We are today aware of more than two hundred statements that are equivalent to the axiom of choice. Astounding! And yet it is not just the miracle of the multiplication of peas, but also that of our questions about the infinite. We have certainly not come to an end of questions, nor of possible answers to them.

Acknowledgment

This contribution is a follow-up to the project of Centre of Economic Studies of University of Economics and Management Prague.

References

Aristotle. (1910-1952). *Works of Aristotle translated into English*. Ross, W. D. (ed.), Oxford Translation (12 vols.), Oxford.

Bolzano, B. (1851). *Paradoxien des Unendlichen*. Příhonský, Leipzig

Cantor, G. (1878). Ein Beitrag zur Mannigflatigkeitslehre. Journal für die reine und angewandte Mathematik, 84, pp. 242-258

Cohen, P. J. (1964). The independence of the continuum hypothesis I. *Proceedings of the National Academy of Sciences of the United States of America*. 51 (1), January, pp. 105–110

Dedekind, R. (1888). *Was sind und was sollen die Zahlen*? Vieweg, Braunschweig

Descartes, R. (1973). *Œuvres philosophiques*. Ed. Alquier Garnier, Paris

Heath, T. L. (1926). *The Thirteen Books of Euclid's Elements*. 2nd edition, 3 vols., Cambridge University Press, Cambridge

Hilbert, D. (1900). Mathematische Probleme. *Nachrichten von der Königl. Gesellschaft der Wiss. zu Göttingen 1900*, pp. 253–297

Kőnig, G. (1905). Zum Kontinuum-Problem. *Mathematische Annalen*, 60, pp. 177-180

Noether, E., Cavaillès, J. (1937). *Briefwechsel Cantor-Dedekind*. Hermann, Paris

Verdet, J.-P. (1987). *Le ciel, ordre et désordre*. Gallimard, Paris.

Weber, H. (1991/92). Leopold Kronecker. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 2, pp. 5–31

Zermelo, E . (1908). Untersuchungen über die Grundlagen der Mengenlehre. *Mathematische Annalen*, 65, pp. 261–281

Contact

Jan Coufal
University of Economics and Management, Nárožní 2600/9a, 158 00 Praha 5, Czech Republic
jan.coufal@vsem.cz
Jiří Tobíšek
University of Economics and Management, Nárožní 2600/9a, 158 00 Praha 5, Czech Republic
jiri.tobisek@vsem.cz