# From here to infinity : Cantor and the transfinite numbers | By David Soquet

Accédez à la version française de cet article

"There are three kinds of persons: those who know how to count and those who don't." If the joke works, it is because very few of us get to have a close encounter with the third kind, those who know how to count to infinity. One of them, the German mathematician Georg Cantor (1845;1918) was the first to explain what it really means to count to infinity and also that there are many ways to do so.

**I. The origins of infinity.**

Before Cantor mathematicians were barely smarter than a 6th grader as regards infinity. The thing is that every child gets a fairly good appreciation of at least one aspect of infinity at a very young age. When he learns how to count, a child understands quickly that there is no limit to the sequence of natural numbers (positive integers). He knows that after 999,999 comes 1,000,000 as much as he knows that 1000 comes after 999, he just needs to learn how to name it. It is more a problem of vocabulary because some language is needed to speak the numbers.

I have been teaching to 6th graders for more than ten years and I know they have no problem imagining a straight line going to infinity, a very abstract concept. And when they are asked to place equally spaced dots on a straight line to represent numbers, they get it easily because they know that on an infinite line there will always be dots available for ever increasing numbers. One infinity for another. However, it takes a much greater effort for 6th graders to try and consider the infinite number of dots between any two dots forming a line segment. This infinity *within* seems harder to grasp than the infinity *without*. It seems that because the line segment is bounded while the line is limitless, some kids think it does not fit in, they believe there is not enough space. The problem with the segment line is that we can draw it completely and this makes us believe it is concrete when it actually has a purely abstract definition. In order to see it better, one can simply consider the amount of decimal numbers between 2 and 3. As one can always write a decimal number between any two decimal numbers, it is clearly infinite.

This fundamental distinction between the infinity of dots on a line (infinity at large) and the infinity of dots within a line segment, was already noticed by Aristotle. After some dubious considerations, Aristotle concluded that he could only accept one kind of infinity, the *potential* infinity of things limitless, as the straight line. More precisely, it looks as if Aristotle would accept that something goes to infinity but not that it would be made of infinity. The problem for Aristotle was the coexistence of an infinite number of things, what he called an *actual* infinite.

"*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 untraversable. 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.*"[1]

To make a long story short, the consequence of this was that it more or less forbid anyone to speak about or to probe infinity for about 2000 years. However it also allowed mathematicians to work with infinity without questioning the real nature of it. To illustrate this, let's see how easy it is to prove that the number 0.999... (with an infinite number of 9s) equals to 1.

(1) $X=0.999...$   (2) $10\times X=9.999...$   (3) $9\times X=9$ *subtracting (1) from (2)*   (4) $X=1$

[1] See Aristotle, *Physics*, Book III, 8

Line (3) assumes that there is the same infinite number of 9s after the dot in both numbers, thus explaining their disappearance. Using the same method, every number with a infinite but repeating sequence of decimals can be proved to be equal to a rational number (i.e. a fraction). This way, $5.42424242 = \frac{537}{99}$. This gave mathematicians the impression that infinity could be tamed.

Another example of mathematicians working with infinity is found in the development of Calculus at the end of the 17th century. Leibniz and Newton separately created their methods of differential and integral calculus by using *infinitely small quantities* but they never clearly defined them.

**II. Problems with infinity.**

One important problem in many fields of mathematics is that many rules go awry when one try to extend them from the finite to the infinite. To illustrate this, let's consider the first hundred natural numbers from 1 to 100. *Half of them are even numbers.* If one counts further, say to 10,000 then 5,000 of them are even. This easily brings the following statement :

(1) There are *twice as many* natural numbers as there are even numbers.

Since there is no limit to the sequence of numbers, these two additional statements are also true:

(2) The number of natural numbers is infinite.
(3) The number of <u>even</u> numbers is infinite.

Then, when one considers all the numbers, there is a problem: either statement (1) is not true anymore either one infinite is twice as big as the other. Both solutions go against common sense.

Another slightly different example is the following:

(1') The numbers of dots belonging to the line segment [AB] is infinite.
(2') The numbers of dots belonging to the straight line (AB) is infinite.
(3') The line segment [AB] is part of the straight line (AB).

Again because of (3') one may think that there are more dots on the straight line than on the segment but then that would mean that *the infinite number of (2') is larger than the infinite number of (1')*.

The "slight" difference between the two examples mentioned above is actually of great importance in our story. This is the difference between the *discrete* and the *continuous*. To explain this, it is enough again to consider the natural numbers on a dotted line. When placed on a straight line, natural numbers clearly leave holes as there are no integers between 1 and 2 or between any two consecutive integers. This is the discrete. On the other hand, the infinite number of dots completely fills up the straight line. This is the continuous. At the start of the 19th century, two related questions remained unanswered.

1. What type of numbers can fill up the straight line, leaving no holes?
2. Can there be a difference between two infinites?

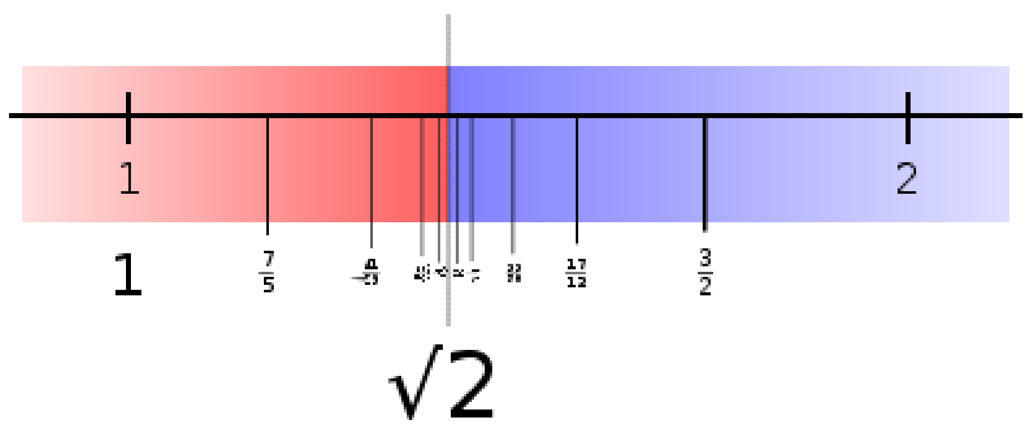
Question 1 suggests that other types of numbers can be placed on our straight line, e.g. rational numbers. Rational numbers are defined as the ratio of two integers. All of them can be placed on the "number line". But this doesn't take us very far. More interesting is the fact that -contrary to natural numbers- there is always a rational number between any two rational numbers. If $\frac{a}{b}$ and $\frac{c}{d}$ are rational numbers then $\frac{\frac{a}{b}+\frac{c}{d}}{2}$ is another rational number and is obviously between $\frac{a}{b}$ and $\frac{c}{d}$ because it represents the middle point. It follows easily that there is an infinite number of rationals between two numbers. Is that enough to fill up the number line? The answer is no. In fact, as was discovered by Pythagoras' clique, all numbers are not rational. The number measuring the diagonal of a square whose side is 1 can easily be "placed" on the number line (between 1 and 2) but it is not rational. This number $\sqrt{2}$ is now called an irrational number. Both rational and irrational numbers make up the group of *real numbers*[2].

**III. What is a real number ?**

Do real numbers leave holes on the number line? Before an answer could be given to this question, it was necessary to give a clear and precise definition of real numbers. In a concise article published in 1872 *On Continuity and Irrational Numbers*, **Richard Dedekind** (1831;1916) did exactly that. The true genius of Dedekind resides not only in the way he understood the necessity of a rigorous definition of real numbers but also in the method he used to define something that could seem obvious to others, especially to modern mathematicians who are used to work with the number line. Dedekind considered each dot on the number line as the intersection with another line falling on it. If this intersection represents a rational number Q then there is a clear division of the number line in two groups: numbers greater than Q and numbers smaller than Q. The main difficulty occurs when the intersection (*the cut* in Dedekind's own terms) is not a rational number. Because he wanted to define irrational numbers, he could not claim the same argument. However by considering the numbers whose square is smaller than 2 and the others, he could clearly define two subsets of numbers and the cut would serve as a definition of the number $\sqrt{2}$ . The details are more complicated and Dedekind gave a precise mathematical construction but for our concern, only the result matters.

"Whenever, then, we have to do with a cut produced by no rational number, we create a new, an *irrational* number, which we regard as completely defined by this cut. […] From now on, therefore, to every definite cut there corresponds a definite rational or irrational number.

—Richard Dedekind, *Continuity and Irrational Numbers*, Section IV




Born in Braunschweig, Germany in 1831, Dedekind entered the University of Gottingen in 1850 where he had the honor of being one of Gauss' last students. A contemporary of other great German mathematicians (such as **Riemann** or **Dirichlet**) his results on abstract algebra and number theory remain crucial today. In the early 1870's Dedekind also met Georg Cantor, the man at the center of our story, and became one of his early admirers and defenders.

---

[2] Cf. Une histoire des équations

**IV. Cantor's first works.**

In the late 1860's, Cantor had already published several papers on number theory when he turned his interest on analysis. His work on trigonometric series (infinite sums of trigonometric functions, a subject beyond the scope of this article) lead him to a definition of irrational numbers and also forced him to consider infinite sets. Through this, he came to the invention of Set theory, a theory that would change the realm of mathematics forever. Put simply, set theory says everything consists of sets of objects and relations between these objects. For example, the number of girls in a class is a set as is the number of students wearing glasses. Today we are accustomed to speak about functions as a relation between numbers from one set (an interval of real numbers) with another. The function $f : x \mapsto x^2$ takes any number and returns its square, thus creating a relation between real numbers and positive real numbers. But speaking about intervals of real numbers required a definition of the latter, which Cantor also gave in a paper published in 1883. Furthermore, to explain his findings, Cantor devised a method to "count" a infinite number of elements.

First, Cantor defined the cardinal of a set as the number of elements contained in the set. For example, if one considers sets consisting of days, then the cardinal of a week is seven. This process of enumeration is quite obvious when it deals with a finite number of elements but, as we saw, can get more difficult when one starts counting elements in an infinite set. The cardinal of the set of positive integers is $\infty$ as is the cardinal of the set of real numbers. Does it mean that Card (N) = Card (R)? Not necessarily so.

To explain this, Cantor needed to find a method to compare the cardinals of two different sets. This method is now called a one to one correspondence. For example, the association Red Monday, Orange Tuesday, Yellow Wednesday, Green Thursday, Blue Friday, Indigo Saturday and Violet Sunday creates a relation between the days of the week and the colors of a rainbow. Each day is associated to a unique color and the same color is not associated to two different days. This is a one to one correspondence. Once again this example with finite sets doesn't help a lot because we already knew there were seven days and seven colors. This is also connected to *Dirichlet's drawer principle*: if there were less colors than days, then two different days would necessarily have the same color. This method can be used to compare numbers without counting. By pairing elements of each set, one can easily see which set has the greater cardinal. This is actually the method Cantor used to see which sets had the same cardinal as N, by trying to pair them with the sequence of positive integers.

A more interesting example is the function mentioned above which connects each real number with its square. Because two opposite numbers share the same square $(-2)^2 = 2^2 = 4$, the function is not a one to one correspondence between the set of real numbers and the set of positive real numbers. This doesn't mean these two sets don't have the same cardinal as there might be another function realizing the correspondence. This also doesn't mean the square function cannot realize a one to one correspondence, it does so between $[0;\sqrt{2}]$ and $[0;2]$. Cantor actually proved that every interval of R has the same cardinal as R.

With this invention of his, Cantor easily showed that the first paradox we encountered about even numbers is no more. The "doubling" function clearly realizes a one to one correspondence between all even integers and all positive integers : $1\to2$, $2\to4$, $3\to6$, $4\to8$, $5\to10$ etc… Following cantor's definition, this proves that the cardinal of the set of even numbers is equal to card(N).

Cantor also used geometrical correspondences to prove equality of cardinals. This is how he famously proved that there are as many dots on a square than there are on one side of the square. But to understand the method, we shall reflect on a simpler example :

Let's consider two concentric circles $C_1$ and $C_2$ (meaning they have the same center O) with respective radiuses of 1 and 10. Both circles are sets of an infinite number of dots and the length of $C_2$ is 10 times bigger than the length of $C_1$. Do both sets have the same cardinal ? The answer is yes and this is easily shown. Take any dot $A_1$ on the smaller circle $C_1$ and draw a half line from O and through A. It will intersect $C_2$ in one dot $A_2$. Of course, one gets the same result if one starts with a dot on the large circle. This way, each dot on the smaller circle can be associated with one and only one dot on the larger circle, thus creating a one to one correspondence. There are "as many" dots on both circles.

**V. Cantor and the transfinite numbers.**

Cantor answered our second question in a paper published in 1874 ("*On a Property of the Collection of All Real Algebraic Numbers*"). First he defined the first transfinite cardinal number as the cardinal of the natural numbers. He named it $\aleph_0$ (from aleph, the first letter of the Hebrew alphabet). All sets whose cardinal is $\aleph_0$ are called denumerable (or countable infinite) and it means there exists a one to one correspondence with the natural numbers. Examples of such sets include : the set of prime numbers, the set of even natural numbers (with the bijection already presented) but also, surprisingly, the set of all rational numbers. For the latter, Cantor found a clever way to order all fractions proving that they could be "enumerated".

The next step for Cantor was to prove that the set of real numbers is not denumerable. It was already known at the time that real numbers could be divided into two subsets : the set of algebraic numbers and the set of transcendental numbers. *(Footnote linking to article "Une histoire des equations".)*

First Cantor proved that "*The collection of real algebraic numbers can be written as an infinite sequence in which each number appears only once.*" which is to say that there exists a one to one correspondence between algebraic numbers and positive integers. In other words the set of algebraic numbers is countable.

Cantor's second theorem is the following : "*Given any sequence of real numbers $x_1$, $x_2$, $x_3$, … and any interval [a, b], one can determine numbers in [a, b] that are not contained in the given sequence.*"

And here is Cantor's conclusion :

*"The reason why collections of real numbers forming a so-called continuum (such as, all real numbers which are ≥ 0 and ≤ 1) cannot correspond one-to-one with the collection (v) [the collection of all positive integers]; thus I have found the clear difference between a so-called continuum and a collection like the totality of real algebraic numbers."*

**VI. The diagonal method.**

Some years later, in 1891, Cantor gave another proof of the uncountability of real numbers using the famous diagonal method. To understand how this method works, let's consider infinite sequences of 0s and 1s as for example: 0 1 0 1 0 1 0 1… etc.

We know define the set S which consists of an infinite number of such sequences which we shall name $s_1$, $s_2$, $s_3$, ... The set S is obviously countable because there is a one to one correspondence between the elements $s_1$, $s_2$, $s_3$, ... and the positive integers.

Now we write these elements as follows.

$s_1$ = (**0**, 0, 0, 0, 0, 0, 0, ...)
$s_2$ = (1, **1**, 1, 1, 1, 1, 1, ...)
$s_3$ = (0, 1, **0**, 1, 0, 1, 0, ...)
$s_4$ = (1, 0, 1, **0**, 1, 0, 1, ...)
$s_5$ = (1, 1, 0, 1, **0**, 1, 1, ...)
$s_6$ = (0, 0, 1, 1, 0, **1**, 1, ...)
$s_7$ = (1, 0, 0, 0, 1, 0, **0**, ...)

We then create a new sequence $s_0$ by taking the diagonal sequence and switching 0s and 1s. We get the sequence $s_0$ = (**1**, **0**, **1**, **1**, **1**, **0**, **1**, ...). Because of its construction, the sequence $s_0$ has at least one different digit with each of the sequences of S (the 1st digit in $s_1$, the 2nd in $s_2$ etc... the nth in $s_n$). Consequently, $s_0$ cannot belong to S. It follows that the set T of all infinite sequences of 0s and 1s has a cardinal greater than the cardinal of S, which is $\aleph_0$.

Cantor exhibited a one to one correspondence between the set of real numbers and T, a proof that the set of real numbers is not countable.

**VII. Cantor's theorem and the continuum hypothesis.**

This is in the same paper that Cantor proved what is now referred to as Cantor's theorem.

*"For any set A, the set of all subsets of A has a strictly greater cardinality than A itself."*

The set of all subsets of A is called the power set of A or P(A) and its cardinal is $2^n$ where n is the cardinal of A. Again, this statement can be easily verified for finite sets. Let's A be the set {a,b,c}. Then the subsets of A are {a}, {b}, {c}, {a,b}, {a,c}, {b,c}, {a,b,c} (which is A itself) and { }, the empty set. We can easily check than card(A) = 3, card(P(A))= 8 and $2^3$=8.

As for infinite sets, Cantor proved that the power set of a countable infinite set is uncountable infinite. As an example, the power set of the set of natural numbers (the set of all subsets of natural numbers) has the same cardinal as the set of real numbers.

It follows that card(R)=$2^{\aleph_0}$.

A simple and last question remains to be asked:

*Is there a set whose cardinality is strictly between that of the integers and that of the real numbers ?*

This problem is now called the *continuum hypothesis* and it was considered so important that **David Hilbert** (1862;1943), one of the greatest german mathematicians of the time, made it the first of his list of 23 problems to be solved for the century. *(Footnote linking to Marie's article on Godel.)*

Cantor spent many years vainly trying to prove the hypothesis which he believed to be true. According to many authors, this may be one of the reasons for his repeated depressions after 1884.

Sadly, and following two results by mathematicians Kurt Godel (in 1940) and Paul Cohen (in 1963), mathematicians know now that the continuum hypothesis can neither be proved nor disproved within the usual axioms of set theory. It is an undecidable proposition. *(Footnote linking to Marie's article on Godel.)*

Cantor's creations were not welcomed by everyone. **Leopold Kronecker** (1823;1891) another german and later the french polymath **Henri Poincaré** (1854;1912) were firmly opposed to tranfinite numbers and participated in the *intuinionist* movement which was a reaction againts Hilbert's *formalism*. However, whatever one may think about Cantor's achievements, one cannot deny his influence on the mathematics of the 20th century.

SOURCES :

*Everything and more*… by David Foster Wallace.
http://descmath.com/diag/ancients.html

Wikipedia

http://en.wikipedia.org/wiki/Actual_infinity
http://en.wikipedia.org/wiki/Georg_Cantor
http://en.wikipedia.org/wiki/Richard_Dedekind
http://en.wikipedia.org/wiki/Cantor%27s_first_uncountability_proof
http://en.wikipedia.org/wiki/Cantor%27s_diagonal_argument