Parabola Volume 32, Issue 1 (1996)

ALMOST ALL REAL NUMBERS ARE TRANSCENDENTAL

C. Cox

(Reprinted from a Parabola article of 30 years ago!)

Doubtless the above heading will be to most of our readers quite cryptic and meaningless. It is the purpose of this article both to explain what it means, by defining the terms "almost all" and "transcendental", and also to outline how it may be proved.

Not only was the result received with astonishment by mathematicians in 1874, but the method used by the German mathematician G. Cantor in the proof sparked off a controversy regarding its validity, which was largely responsible for the investigations into the foundations of mathematics undertaken in the present century.

We begin with a discussion of some classes of real numbers. Our readers will certainly be familiar with the usual representation of real numbers as points on a number line.

$-2 \quad -1 \quad 0 \quad +1 \quad +2 \quad +3 \quad +4 \quad +5$

$\sqrt{2} \quad e \quad \pi$

Some of these points correspond to integers $(0,\ \pm 1,\ \pm 2,$ etc.) An apparently much more numerous class consists of points corresponding to numbers of the form $\frac{p}{q}$ where p and q are both integers $(q \neq 0)$. These are called rational numbers, and any interval of the line, however short, contains an infinite set of rational points. It is not immediately obvious that there are any points which do not correspond to rational numbers. The discovery that the number $\sqrt{2}$ could not be expressed as a vulgar fraction, $\frac{p}{q}$, came as something of a shock to early Greek mathematicians of the Pythagorean school.

That $\sqrt{2}$ is irrational can be proved by reductio ad absurdum:- Suppose $\sqrt{2} = \frac{p}{q}$, a rational number reduced to its lowest terms, p and q being integers with no common factor. Since $p^2 = 2q^2$, it follows that p^2, and therefore that p, is an even number. Substituting $p = 2P$, where P is an integer, we obtain $4P^2 = 2q^2,\ q^2 = 2P^2$. Thus q^2, and therefore q, are also even, so that both p and q have the factor 2, contradicting our assumption that they shared no common factor. This contradiction shows that $\sqrt{2}$ cannot be rational.

Having found one irrational number it is easy to see that the irrationals are, also, thickly distributed on the number axis; for example, the addition of any rational number r to $\sqrt{2}$ yields an irrational number $r + \sqrt{2}$, and this fact shows the assertion to be true. (Prove that $r + \sqrt{2}$ is irrational).

There is another important classification of the real numbers which we slall now discuss.

"Algebraic numbers" are numbers which satisfy an algebraic equation, i.e. one of the type

$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$ where $a_0,\ a_1, \cdots, a_n$ are all integers.

Every rational number $\frac{p}{q}$ is algebraic since it satisfies the equation $qx - p = 0$. The irrational number $\sqrt{2}$ is algebraic since it satisfies the quadratic equation $x^2 - 2 = 0$; and $r + \sqrt{2}$ is algebraic if $r = \frac{p}{q}$, since it satisfies $q^2x^2 - 2pqx + p^2 - 2q^2 = 0$.

Numbers which are not algebraic are called "transcendental numbers". Again it is not obvious that some real numbers are not algebraic. So far from obvious indeed, that it was not until 1844 that the first transcendental number was found; in that year, the French mathematician Liouville was able to show that, for example, the number $.110001000000000000000001 0\cdots$ (the nth "1" occurs in the $n!$'th place) was not algebraic.

Liouville's work remained for almost 30 years the only significant accomplishment in this field. Then Charles Hermite, in 1873, showed that the number $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \cdots$ was also transcendental, using methods whose originality and elegance won immediate admiration from his contemporaries. In view of its long history, with just two isolated successes, problems associated with transcendental numbers acquired a reputation (deservedly) for a level of difficulty demanding more than ordinary methematical ability (e.g. it was another 9 years before Lindemann was able to extend Hermite's method to show that π is transcendental). No wonder the mathematical world was taken by surprise when in 1874 Cantor asserted that "almost all real numbers are transcendental".

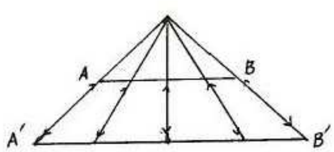
This statement will still seem meaningless to the reader even if we can show that the class of transcendental numbers is infinite. For we have seen that there are infinitely many algebraic numbers, and Cantor's result seems to imply that one infinite class contains, in some sense, many more elements than a second infinite class. Cantor's great contribution was indeed to give precise meanings to the phrases "set A contains the same number of elements as set B" and "set A contains a larger number of elements than set B" in such a way that they continue to make sense when the sets are infinite. We shall consider each of these phrases in turn.

Two finite sets contain the same number of elements (or, equivalently, "have the same cardinal number") if and only if their elements can be paired off in such a way that none are left over, unpaired, in either set. Thus if in a room every chair is occupied by one person, and no-one is left unseated, it is quite clear that the set of chairs and the set of people in the room have the same cardinal number. One says that there is a $(1-1)$ correspondence between the elements of the two sets. If it is possible to set up a $(1-1)$ correspondence between the elements of set A and set B, one says "set A is similar to set B". The process of counting the elements of a set consists in nothing more nor less than setting up a $(1-1)$ correspondence between the elements of the set, and the set of natural numbers $1, 2, 3 \cdots, n$.

Cantor now **defines**, for infinite sets, the phrase "set A contains the same number of elements as set B" to mean that it is possible to set up a $(1-1)$ correspondence between the elements of the two sets. This definition leads to some rather unexpected results. For example, consider

$$\begin{array}{ccccccc} 1, & 2, & 3, & 4, & \cdots, & n, & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ 2, & 4, & 6, & 8, & \cdots, & 2n, & \cdots \end{array}$$

which shows how a $(1-1)$ correspondence may be set up between the set of all positive integers, and its subset, the set of all even positive integers. Thus, by the definition, the set of even integers contains the same number of elements as the set of all integers. The cardinal number of the set of natural numbers was called by Cantor $\aleph_0$ (aleph null). You will notice that the statement "set A has cardinal number $\aleph_0$" is merely another way of saying that set A is similar to the set of positive integers.



The accompanying diagram shows how the points in a short line segment AB can be put $(1-1)$ correspondence with the points in a longer line segment. Thus any two line segments contain the same number of points.

Again consider

$$\begin{array}{ccccccccccccccccc} 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9, & 10, & 11, & 12, & 13, & 14, & 15, & 16, & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ \frac{0}{1}, & \frac{1}{1}, & -\frac{1}{1}, & \frac{1}{2}, & -\frac{1}{2}, & \frac{2}{1}, & -\frac{2}{1}, & \frac{1}{3}, & -\frac{1}{3}, & \frac{3}{1}, & -\frac{3}{1}, & \frac{1}{4}, & -\frac{1}{4}, & -\frac{1}{4}, & \frac{2}{3}, & -\frac{2}{3} & \cdots \end{array}$$

In the second row the elements are the rational numbers $\frac{p}{q}$ where p and q are relatively prime, $q > 0$. They are arranged as follows:- $\frac{p_1}{q_1}$ precedes $\frac{p_2}{q_2}$ if

$$|p_1| + q_1 < |p_2| + q_2; \quad \text{or if} \quad |p_1| + q_1 = |p_2| + q_2 \quad \text{but} \quad |p_1| < |p_2|;$$

$$\text{or if} \quad |p_1| + q_1 = |p_2| + q_2 \quad \text{and} \quad |p_1| = |p_2| \quad \text{but} \quad p_1 > 0.$$

A little thought will convince you that any rational number $\frac{p}{q}$ occurs sooner or later in the second row. According to our definition the set of positive integers has the same cardinal number as the set containing all rational numbers. A similar but slightly more complicated procedure which we will omit shows likewise that the set of algebraic numbers has cardinal $\aleph_0$, i.e. can be put into $(1-1)$ correspondence with the positive integers.

At this stage you may be beginning to believe that all infinite sets have the same cardinal number, that after all there is only one infinity. Such, however, is not the case. We return to the second statement:- "Set A contains a larger number of elements than set B". For finite sets this is true if and only if, when the elements of B are paired off with the elements of A, there is at least one element of A left over, unpaired. That is, a $(1-1)$ correspondence can be set up between the elements of B and the elements of a subset of A, but not between the elements of B and all the elements of A.

This last sentence is taken as the **definition** of the statement:- "The cardinal number of the set A is greater than that of the set B", when the sets are infinite. Note that the first part of the sentence is not sufficient by itself, and does not imply the second. We have indeed already seen several examples of $(1-1)$ correspondences between infinite sets and subsets of themselves. The existence of such a correspondence is a property which is always true of infinite sets and never of finite sets.

We can now produce an example of an infinite set whose cardinal number is larger than $\aleph_0$: the set of all real numbers lying between 0 and 1 is such a set.

To prove this we must, according to our definition, do two things. The first, which is very easy, is to show that there is a $(1-1)$ correspondence between the set of positive integers and a subset of our set of real numbers. For example,

$$\begin{array}{cccccc} 1, & 2, & 3, & \cdots & n, & \cdots \\ \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ \frac{1}{2}, & \frac{1}{3}, & \frac{1}{4}, & \cdots & \frac{1}{n+1}, & \cdots \end{array}$$

The second, much more difficult, problem is to show that no such $(1-1)$ correspondence between the positive integers and **all** the real numbers between 0 and 1 is possible. Suppose there is such a correspondence, and let c_n be the real number which corresponds to the integer n. i.e. we suppose it is possible to construct a list of real numbers $c_1, c_2, c_3, \cdots, c_n, \cdots$ which contains every real number between 0 and 1. Each of these numbers can be represented by its decimal expansion (if the expansion terminates we can pad it out by adding an unending string of zeros).

$$\begin{array}{ll} c_1 & = .c_{11}c_{12}c_{13}\cdots c_{1n}\cdots \\ c_2 & = .c_{21}c_{22}c_{23}\cdots c_{2n}\cdots \\ c_3 & = .c_{31}c_{32}c_{33}\cdots c_{3n}\cdots \\ & \cdots \quad \cdots \quad \cdots \\ & \cdots \quad \cdots \quad \cdots \\ c_n & = .c_{n1}c_{n2}c_{n3}\cdots c_{nn}\cdots \\ & \cdots \quad \cdots \quad \cdots \end{array}$$

We now show that contrary to our supposition there is at least one number d, between 0 and 1, missing from this list.
In fact, let $d = .d_1d_2d_3\cdots d_n\cdots$ where the digits $d_1, d_2, \cdots, d_n, \cdots$ are chosen as follows:

If $c_{11} = 2$, put $d_1 = 3$; otherwise put $d_1 = 2$.
If $c_{22} = 2$, put $d_2 = 3$; otherwise put $d_2 = 2$.
For every n, if $c_{nn} = 2$, put $d_n = 3$; otherwise put $d_n = 2$.

It is clear that d is a real number between 0 and 1. But it is not in the list. It is not c_1 because its decimal expansion differs from that of c_1 in the first decimal place. In fact, it is not c_n since its decimal expansion differs from that of c_n in the nth decimal place.

This contradiction shows the impossibility of including all real numbers between 0 and 1 in a single unending list: the cardinal number c of this set is greater than $\aleph_0$. It is easy to see that the set of all real numbers also has cardinal number c and it then follows that the set of transcendental numbers also has cardinal c.

In fact, it is clear from these results that not only is $c > \aleph_0$, but $\aleph_0$ is so negligible in comparison with c, that removal of a set of cardinal $\aleph_0$ from a set of cardinal c can never make any significant change at all in the number of elements in the set: the remaining set still has cardinal number c. A set of cardinal c cannot be built up by putting together any finite number (however large) of sets of cardinal $\aleph_0$: in fact, not even if $\aleph_0$ such sets are put together.
In view of these facts, the statement "almost all real numbers are transcendental" is quite justified.

(Reference: D. Pedoe, The Gentle Art of Mathematics, Ch. III).