

**Struggles of the Past: How
New Types of Numbers
Emerged in the History of
Mathematics**

Mike Boyle

University of Maryland

Maryland Mathematics Institute,

June 26, 2009

Outline of the talk

I. Introduction.

1. Counting.

2. Negative Numbers.

3. Zero.

4. Trying to define division by zero.

5. Why can't we divide by zero?

6. Harriot's Principle.

7. Beyond the rational numbers.

8. Proof that $\sqrt{2}$ is not rational.

9. The real numbers.

10. The complex numbers.
11. Complex numbers and the cubic.
12. Trouble with complex square roots.
13. Geometric interpretation of complex numbers.
14. Fundamental theorem of algebra.
15. Adding up infinitely many numbers.
16. Number and the foundations of probability.
17. Cardinal numbers.
18. Proof that $[0,1]$ is uncountable.
19. Conclusion.

I. Numbers

"What is a number?"

We take a lot of our knowledge of numbers for granted. So much seems easy – counting, zero, negative numbers, fractions, real numbers, etc. What's the problem!

But, these were hard-won advances over millennia. Even brilliant mathematicians (e.g. Brahmagupta, Leibniz, Euler) made mistakes in struggles with new numbers. [Perhaps an example could comfort a student?] Even after successful introduction, there were often long delays before the advances were fully accepted and used.

Also, in the history of numbers we can see some pro's and con's of the stereotypical "physicist's approach" and "mathematician's approach".

"The physicist" grabs a mathematical idea which seems to help and goes with it. Those bean counter mathematicians will get around to crossing the t's and dotting the i's.

"The mathematician" cares for Truth and struggles to produce definitions and arguments which are clear, sensible and reliable. So what if nonsense works – it is still nonsense!

In the end we need both approaches. Mathematics is like a tree, with roots and branches, and both keep growing.

1. COUNTING

There is a famous quote of Kronecker (1823-1891): "God created the integers, all else is the work of man."

I disagree!

Of all the great steps forward with numbers, I'm most impressed by the conception of the abstract counting numbers 1,2,3,4,... .

It was not automatic to think of two doves and two wolves as having something in common. A counting number makes a conceptual thing out of the process of being able to put two sets into one to one correspondence. Not all cultures have had a systematic way of naming arbitrarily large numbers. We have not observed this capability in a nonhuman species.

There is some history of the evolution of the idea of counting numbers the greatest was to conceive and reasoned speculation on how this capability emerged in human evolution, e.g.

- "Number, the language of science", Tobias Dantzig
- "Number words and Number Symbols", Karl Menninger
- "The Math Gene", Keith Devlin .

2. NEGATIVE NUMBERS

The Babylonians around 2000 BC were solving many problems which boiled down to solving a quadratic equation. But their problems were rooted in geometry and the material world, and they ignored negative number solutions.

The ancient Greeks likewise considered only positive number solutions. (The great Greek mathematician Diophantus, 3rd century AD, in "Arithmetica" referred to a negative solution of a problem as being "absurd".) For those Greeks, numbers were attached to explicit counting and ratios.

By 200 BC, in China, red and black counting rods were used for money held and owed.

By the Han dynasty (circa 200BC-200AD) there were written Chinese works with rules for manipulating negative quantities appearing in calculations (for problems which had positive solutions).

Circa 620 A.D. Brahmagupta in India essentially wrote down rules of arithmetic for negative and positive numbers (e.g. "the product or quotient of a fortune and a debt is a debt").

By the 1300's such rules had appeared in writing in Europe.

Cardano's *Ars Magnus* (1545) showed a clear acceptance of negative numbers as possible solutions to equations.

But even in the 18th century (!), there was still some resistance to negative numbers. The British mathematician Francis Maseres wrote (1759) that negative numbers *"darken the very whole doctrines of the equations and make dark of the things which are in their nature excessively obvious and simple"*.

Maseres was probably heretical among mathematicians for this view – but still, he was made a Fellow of the Royal Society in 1771, over 100 years after Newton invented calculus.

3. ZERO

It was hard to invent counting numbers. Much harder still to make a Thing which is nothing – i.e. to invent zero.

Zero appeared in numbers as a placeholder in place value systems. E.g. in our base 10 system, in 7105, the "0" means we add no multiple of 10: $7105 = 7(1000) + 1(100) + 5(1)$. But it is a big extra step to make zero a Thing and to integrate zero into arithmetic.

By the 7th century AD, the Hindus had begun to recognize "sunya", the absence of quantity, as a quantity in its own right. I.e., zero was treated as a number.

Brahmagupta (7th century A.D.), was the first known mathematician to try to extend the rules of arithmetic to include zero. His rules for arithmetic involving zero were fine, apart from the problem of defining division by zero. There he and his Hindu successors had a heck of a time.

4. TRYING TO DEFINE DIVISION BY ZERO

- Brahmagupta (7th century A.D.) " *A zero divided by a zero is zero" and "a negative or a positive divided by zero has that zero as its divisor"* .

- Mahavira (9th century A.D.)

"A number remains unchanged when divided by zero."

- Bhaskara (II) (12th century A.D.)

"A quantity divided by zero becomes a fraction the denominator of which is zero. This fraction is termed an infinite quantity. In this quantity consisting of that which has zero for its divisor, there is no alteration, though many may be inserted or extracted; as no change takes place in the infinite and immutable God when worlds are created or destroyed, though numerous orders of beings are absorbed or put forth."

(He seems to be saying that $n/0 = \text{infinity}$. But that would mean that $n = 0 \times \text{infinity}$, for every number $n \dots$)

In his text, Bhaskara (II) posed problem of finding an unknown number "... whose multiplier is 0. Its own half is added. Its multiplier is 3; its divisor 0. The given number is 63."

We can translate this problem as: Find the number x such that

$$3\left(\frac{0x + (1/2)0x}{0}\right) = 63 .$$

It seems we are supposed to cancel 0's (!) to get $3(x + (1/2)x) = 63$ and then solve to get $x = 14$.

The correct resolution of these miseries is to give up. It's impossible to give a satisfactory definition of division by zero – so don't try.

5. WHY CAN'T WE DIVIDE BY ZERO?

This is a case where the modern view helps. At some point, to answer a question you realize that you need definitions and assumptions.

We write down some properties of numbers which we can't live without. These are assumptions. Then we show it's impossible to define division by zero and keep those essential properties. I'll write \times for multiplication.

One consequence of the properties is that zero times any number is zero, because

$$a \times 0 = a \times (0 + 0) = (a \times 0) + (a \times 0)$$

and then after subtracting $a \times 0$ from both sides we conclude $0 = a \times 0$.

Next, a huge step is to understand arithmetic as having ONLY TWO fundamental operations, addition and multiplication. Subtraction and division are not fundamental. They are defined in terms of addition and multiplication.

We say BY DEFINITION that $a/b = c$ is true if and only if two conditions hold:

- $a = b \times c$, AND
- c is the only number such that $a = b \times c$.

Now suppose we can find c such that $a/0 = c$. Then c is the unique number such that $a = 0 \times c$.

There are two cases.

I. If a is not zero, then $a = 0 \times c = 0$, which is a contradiction.

II. If a is zero, then $0 = a = 0 \times c$ for every c , and that contradicts that c is the only solution.

Score one for "the mathematicians"! It is really useful to know what is impossible (rather than continuing to try ...).

By the way: among those essential properties (like $a + b = b + a$ and so on) that give us all good arithmetical properties, the ONLY property connecting addition and multiplication is the distributive law: $a \times (b + c) = (a \times b) + (a \times c)$ and $(a + b) \times c = (a \times c) + (b \times c)$.

6. "HARRIOT'S PRINCIPLE"

Before moving on let's appreciate that "zero" allows a useful technique for solving polynomial equations. It was proposed by Thomas Harriot in the early 1600s and popularized later by Descartes. (The author Tobias Dantzig referred to this technique as "Harriot's Principle".)

The technique: subtract a polynomial from both sides to make one side zero. (This does not change the solution set.) Then find a root of the resulting polynomial.

Example. Solve $3x^2 + 2x - 5 = 2x^2 - 2$.

First subtract $2x^2 - 2$ from both side, to get $x^2 + 2x - 3 = 0$.

We can find the roots of $x^2 + 2x - 3$ with the quadratic formula, or by factoring. (What are they?) Those roots are the solutions of the original equation.

Problem: solve $3x^4 + 8x^3 + x^2 + 2x = 3x^4 + 8x^3 + 15$.

This method is a triviality today, but it was a great leap forward back in the day. It is another example of how we can forget the difficulty of progress (until we see our students struggle?).

7. BEYOND THE RATIONAL NUMBERS

The remarkable ancient Greeks thought very carefully about mathematics. They were the ones who introduced formal axioms and theorems.

A Greek would think of a magnitude as perhaps the length of a stick. Two sticks would have commensurable lengths if, for example, 3 copies of one stick would give the same total length as 5 copies of the other.

More formally, suppose x and y are the magnitudes (lengths of the sticks). They would be commensurable if $3x = 5y$. Or if $8x = 7y$. Or if there are any positive integers such that $px = qy$.

Another way to say $px = qy$ is to say $x = (p/q)y$. That is, y is a rational number multiple of x . In particular: x and 1 are commensurable if and only if x is a rational number.

These Greeks thought of numbers as commensurable magnitudes. The Pythagoreans circa 550 BC thought number was the basis for everything: “all is number”. So they naturally felt all magnitudes should be commensurable. But by 400 BC, they had learned this is false. (Equivalently, there exist numbers which are not rational numbers.) For them, this was a dramatic fact of great philosophical import.

8. A PROOF THAT THE SQUARE ROOT OF 2 IS NOT A RATIONAL NUMBER.

Suppose $\sqrt{2} = p/q$, with p and q positive integers. After cancelling 2's as much as possible, we can assume also that at least one of the numbers p and q is an odd integer.

Since $\sqrt{2} = p/q$, after squaring we get $2 = p^2/q^2$, and therefore $2q^2 = p^2$.

This tells us p^2 is divisible by 2. If p were odd, then p^2 would be odd; so, p is even.

Now, because p is even, we can divide the equation $2q^2 = p^2$ by 2 and get that $q^2 = p^2/2 = (p/2)p$. Since p is even, this tells us that q^2 is even, and again that implies q is even.

Therefore both p and q are even. That is a contradiction. So our original assumption that

we could write $\sqrt{2} = p/q$, with p and q positive integers, must be wrong.

Therefore $\sqrt{2}$ is irrational.

9. THE REAL NUMBERS

Today, we have an intuitive idea of a real number as indicating a point on a number line. Intuitively, a positive real number x indicates a point at a certain distance (or magnitude) to the right of zero, and a negative number indicates a point at a certain distance to the left of zero. We multiply, add, etc. these magnitudes with impunity. The existence of irrational numbers just means that some of these magnitudes aren't commensurable. No problema!

Euclid thought of numbers as magnitudes commensurable with some given unit. If we take

that unit to be the number 1, then those magnitudes correspond to the positive rational numbers. Euclid did not define an arbitrary magnitude as a number and he didn't multiply incommensurable magnitudes. (What would it mean?)

This careful distinction between magnitude and number turned out to be unnecessary. Eventually Islamic mathematicians (specifically, abu Kamil ibn Aslam, circa 900 AD) did useful arithmetic with irrational numbers given by expressions like $\sqrt{1/2 + \sqrt{5/4}}$ and simply disregarded the cautions of Euclid. No problema.

Only in the 1800s—long after calculus was invented—was there at last a satisfactory rigorous development of the real numbers and their arithmetic.

10. THE COMPLEX NUMBERS

The complex numbers involve taking square roots of negative numbers. They show up when one tries to solve polynomial equations, for example $x^2 = -1$, or equivalently $x^2 + 1 = 0$.

Already the Babylonians by 2000 BC were essentially solving quadratic equations. You might think that they would have discovered complex numbers. But remember, for them, only positive solutions were solutions. Square roots of negative numbers? No way, Jose.

The complex numbers first appeared in Cardano's "Ars Magnus" (The Great Work) (1545 AD). (At the end he has an inscription: "*Written in 5 years, may it last as many thousand*".)

Cardano's book contained a problem: divide 10 into parts such that the product is 40. If one of those parts is named x , then the other part is $10 - x$, and that number x must satisfy the equation

$$x(10 - x) = 40 .$$

Cardano found the solutions to this equation. (What are they?) They are complex numbers; they are not real numbers. He wrote

"So progresses arithmetic subtlety the end of which, as is said, is as refined as it is useless."

11. COMPLEX NUMBERS AND THE CUBIC

Cardano considered these numbers because his book gave a solution to the ancient problem of finding a root of a cubic polynomial with real coefficients. In Cardano's formula, square roots of possibly negative numbers appeared, but nevertheless one could produce from the formula roots which were real numbers.

For the equation $x^3 = 15x + 4$, Cardano's formula for a solution gave

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}} .$$

But how does that lead to the solution $x = 4$?

Cardano didn't understand these "fictitious" numbers very well; but he used them to good effect. Score one for "the physicists" .

12. TROUBLE WITH COMPLEX SQUARE ROOTS

Better understanding of complex numbers developed slowly (Bombelli, Viete, Euler, ...). Euler introduced the symbol i for $\sqrt{-1}$.

Even Euler had some trouble. At one point Euler wrote that the general rule for square roots (he meant the rule $\sqrt{a}\sqrt{b} = \sqrt{ab}$) shows that

$$\sqrt{-1}\sqrt{-4} = \sqrt{(-1)(-4)} = \sqrt{4} = 2 .$$

But this reasoning leads to contradictions such as

$$-1 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1 .$$

Evidently that “general rule” cannot be extended from the positive real numbers to the complex numbers. Roots (and logarithms) are more complicated (and interesting) in the complex number system.

13. THE GEOMETRIC INTERPRETATION OF COMPLEX NUMBERS

A big step toward seeing and understanding the complex numbers was their geometric interpretation as elements in the plane, and the geometric interpretation of complex addition and multiplication. This only happened around 1800 (!), in publications of Wessel (1797) and Argand (1806). One considers the complex numbers as numbers of the form $a + ib$, where a and b are real numbers, and then pictures $a + ib$ as the point (a, b) in the plane.

For one of our expositions of the basic properties of complex numbers, with connection to calculus, see

www.math.umd.edu/undergraduate/courses/dcp/141/141cxnotes.pdf

14. THE FUNDAMENTAL THEOREM OF ALGEBRA

Today the complex numbers are indispensable to mathematics. One reason is the Fundamental Theorem of Algebra (FTA): every nonconstant polynomial has a root which is a complex (possibly real) number. (A fully correct proof was not given until the 1800's). An equivalent statement is that a nonconstant polynomial can be written as a product of factors of the form $(x-a)$, where a is a complex number (possibly real). For example,

$$\begin{aligned}x^5 + x^4 + x^3 + x^2 &= x^2(x + 1)(x^2 + 1) \\ &= (x - 0)(x - 0)(x - (-1))(x - i)(x + i) .\end{aligned}$$

(Notice, $(x - a)$ is a factor if and only if a is a root.)

Leibniz was smart enough to invent the calculus (independently of Newton). Even Leibniz had troubles here.

Leibniz claimed (for a little while) that for a positive real number a , the polynomial $x^4 + a^4$ has no complex number root. Let us use $a = 1$ as an example here, that is we consider the polynomial $x^4 + 1$. Leibniz noted that a solution would be $x = \sqrt{\sqrt{-1}}$, and he thought this could not be a number of the form $a + b\sqrt{-1}$ (with a and b real numbers).

Leibniz was wrong. Problem: find the four complex numbers x such that $x^4 + 1 = 0$.

15. ADDING UP INFINITELY MANY NUMBERS

Suppose a_1, a_2, \dots are nonnegative numbers. Then $a_1 + a_2 + \dots$ will either be infinity or a nonnegative real number (called the "sum" of the the infinite series $a_1 + a_2 + \dots$).

Many infinite series were "summed" long ago, for example

- ("Zeno")
 $1/2 + (1/2)^2 + (1/2)^3 + (1/2)^4 + \dots = 1.$
- (Oresme, ca. 1350)
 $(\frac{1}{2})1 + (\frac{1}{4})2 + (\frac{1}{8})3 + (\frac{1}{16})4 + \dots = 2 .$
- $1/3 + (1/3)^2 + (1/3)^3 + (1/3)^4 \dots = 1/2 .$

After a while summing series is as natural as adding up finitely many numbers.

16. NUMBER AND THE FOUNDATIONS OF PROBABILITY

In probability, if you have a (possibly infinite) list of disjoint (nonoverlapping) events A_1, A_2, A_3, \dots and if A is the union of all these, then it is a fundamental axiom that the probability of the union is the sum of the probabilities:

$$\text{Prob}(A) = \text{Prob}(A_1) + \text{Prob}(A_2) + \dots$$

Example. Suppose an immortal is working on a problem. Let A_n be the event that he succeeds on the n th day. Suppose $\text{Prob}(A_n) = (1/3)^n$. Now let A be the event that the immortal succeeds at all, i.e. on some day. Then A is the union of the events A_n , and

$$\text{Prob}(A) = 1/3 + (1/3)^2 + (1/3)^3 + \dots = 1/2$$

On the other hand, if in the example $\text{Prob}(A_n) = 0$ for every n , then $\text{Prob}(A) = 0 + 0 + \dots = 0$. If the immortal has probability zero of solving the problem on every n th day, well, he has probability zero of solving the problem at all.

Now imagine you are telling a student that if you pick a number randomly from the unit interval $[0,1]$, and if $0 \leq a < b \leq 1$, then the probability of the number being in $[a,b]$ equals $b-a$. (For example, the probability of the number being between $1/3$ and $1/2$ is $1/6$.) In particular, the probability the number is between 0 and 1 is $1-0=1$ (of course!) and the probability of picking any particular number is zero.

But suppose your student is smart, and says,

This is ridiculous! Consider an infinite list of all the numbers in $[0, 1]$. Say A_n is the event that the n th number on the list is picked. Then the event (call it A) that a number gets picked is the union of the events A_n . And since the probability of picking any particular number is zero, that fundamental axiom would say

$$\begin{aligned} 1 &= \text{Prob}(A) \\ &= \text{Prob}(A_1) + \text{Prob}(A_2) + \dots \\ &= 0 + 0 + \dots \\ &= 0 . \end{aligned}$$

What's wrong with your student's argument?

To answer that we need cardinal numbers.

17. CARDINAL NUMBERS

The counting numbers describe the sizes of finite sets. They are a special case of "cardinal numbers", which describe the sizes of sets, which might be finite or infinite.

Two sets are defined to have the same cardinality (size) if the elements inside them can be put into one to one correspondence.

Example: the set of fingers on my left hand and the set of fingers on my right hand have the same cardinality (5).

Example of Galileo: the set of positive integers and the set of even positive integers have the same cardinality ! The rule $n \mapsto 2n$ gives the one-to-one correspondence:

1, 2, 3, 4, 5, 6, ...
2, 4, 6, 8, 10, 12,

A set is called countable if it is finite or has the same cardinality as the set of positive integers $1, 2, 3, \dots$.

A set of numbers is countable if and only if it is possible to make an infinite list of the numbers from the set such that every number is on the list.

The set of all integers is countable.

Here's the list: $0, 1, -1, 2, -2, 3, -3, \dots$

Clearly every integer appears on the list, exactly once.

The set of rational numbers is also countable. This is a little harder to show (next page)

Here's one way to make a list of all the rational numbers:

-1, 0, 1,

-2, 2,

-3/2, -1/2, 1/2, 3/2,

-3, 3,

-5/2, 5/2,

-8/3, -7/3, -5/3, -4/3, -2/3, -1/3, 1/3, 2/3,
4/3, 5/3, 7/3, 8/3,

and so on. At the n th stage, we are adding to the growing list all the rational numbers in $[-n, n]$ which can be written with denominators $1, 2, \dots, n$, and which are not yet on the list. Eventually every rational number gets on the list.

BUT! Cantor's genius was to conceive of and prove the following (late 1800s):

Not all infinite sets are countable. In particular:

The interval $[0,1]$ is not countable.

and this is the problem with your good student's argument. That fundamental axiom of probability tells you that the probability of the sum is the sum of the probabilities for COUNTABLY MANY disjoint events. There is no axiom for adding up probabilities of uncountably many events. We just have to stay away from that, just as we have to stay away from division by zero.

18. Proof that $[0,1]$ is uncountable.

We prove the unit interval $[0,1]$ is not countable by the method of contradiction, using "Cantor's Diagonal Argument".

For $[0,1]$ to be countable would mean that there is some (infinite) listing of numbers from $[0,1]$ such that every number in $[0,1]$ appears on the list.

So suppose we have such a listing of all the numbers in $[0,1]$. All we have to do is exhibit a number from $[0,1]$ which is not on the list. That contradiction proves no such list exists.

For example suppose we have (in decimal notation) an infinite list of numbers x_1, x_2, \dots from $[0, 1]$:

$$x_1 = .\boxed{6}24453$$

$$x_2 = .3\boxed{5}8711\dots$$

$$x_3 = .33\boxed{4}229\dots$$

$$x_4 = .458\boxed{7}28\dots$$

$$x_5 = .0015\boxed{2}2\dots$$

$$x_6 = .22756\boxed{6}\dots$$

and so on. We just pick a number y such that the n th digit in its decimal representation is not the n th digit of x_n . (I put boxes around the numbers x_n for easier viewing). For example, take $y = .448488\dots$. Now y can't be on the list, because it can't be any of the x_n . QED.

(This is called the “diagonal argument” because those numbers x_n lie on some diagonal line.)

19. CONCLUSION.

So what is a number? There are lots of numbers, including more we haven't considered. Quaternions, ordinal numbers, surreal numbers, algebraic numbers ... They have various different meanings.

The number on the clock means my time is up.