THE NUMBER CONCEPT AND NUMBER SYSTEMS

John Stillwell
Department of Mathematics, University of San Francisco, U.S.A.
School of Mathematical Sciences, Monash University Australia.

Keywords: History of mathematics, number theory, foundations of mathematics, complex numbers, quaternions, octonions, geometry.

Contents

1. Introduction
2. Arithmetic
3. Length and area
4. Algebra and geometry
5. Real numbers
6. Imaginary numbers
7. Geometry of complex numbers
8. Algebra of complex numbers
9. Quaternions
10. Geometry of quaternions
11. Octonions
12. Incidence geometry
Glossary
Bibliography
Biographical Sketch

Summary

A survey of the number concept, from its prehistoric origins to its many applications in mathematics today. The emphasis is on how the number concept was *expanded* to meet the demands of arithmetic, geometry and algebra.

1. Introduction

The first numbers we all meet are the *positive integers* 1, 2, 3, 4, … We use them for counting – that is, for measuring the size of collections – and no doubt integers were first invented for that purpose. Counting seems to be a simple process, but it leads to more complex processes, such as *addition* (if I have 19 sheep and you have 26 sheep, how many do we have altogether?) and *multiplication* (if seven people each have 13 sheep, how many do they have altogether?). Addition leads in turn to the idea of *subtraction*, which raises questions that have no answers in the positive integers. For example, what is the result of subtracting 7 from 5? To answer such questions we introduce *negative* integers and thus make an extension of the system of positive integers.

The introduction of negative integers is just the simplest of many possible extensions of the number concept. Some of these extensions were made in response to the demands of arithmetic (What is 5 minus 7? What is 5 divided by 7?). Others are responses to the demands of geometry to measure lengths, areas, and volumes. Measuring is quite a different problem from counting, because the range of answers is *continuous* rather than *discrete*. A person can own 6 or 7 sheep, but no number in between, whereas the length of a stick can be 1 metre, 2 metres, or any number in between. So what *are* the numbers between 1 and 2?

Filling the number gap between 1 and 2 amounts to bridging the conceptual gap between discrete and continuous, or between counting and measuring. It is the most profound problem in developing a comprehensive concept of number. Other problems arise along the way, but the problem of continuity stretches through most of the history of mathematics, and continues to affect the subject from top to bottom today. Our story will therefore be dominated by the search for continuity, though that is not our only theme.

The second theme is the search for *arithmetic*, that is, for systems of objects we can add, subtract, multiply and divide, subject to the usual rules of calculation. This of course forces us to decide what *are* the usual rules, and leads to the discovery that systems satisfying with "usual" rules are surprisingly rare.

2. Arithmetic

Since positive integers arise from counting, we can *define* them as the objects obtained by starting with 1 and repeatedly adding 1. The process of adding 1, called the *successor* operation, is not only the genesis of numbers, but of all facts about numbers. To prove a fact about all positive integers it suffices to prove that:

a. The fact is true for 1
b. If the fact is true for any positive integer n then it is true for $n+1$.

This method of reasoning is *proof by induction*, and it is the basis for all proofs in number theory, as was first pointed out in 1888 by the German mathematician Richard Dedekind in his book *Was sind und was sollen die Zahlen?* (roughly, *What are numbers and what are they for?*) .

Alongside proof by induction we have *definition by induction*, an idea that goes back to Dedekind's compatriot Hermann Grassmann (in a textbook of arithmetic intended for high school students!) in 1861. Grassmann observed that a function f is defined for all positive integers as soon as:

a. The value $f(n)$ is defined for $n=1$ (or $n=0$, if that is a better starting point).
b. The value $f(n+1)$ is defined in terms of $f(n)$.

For example, the *sum function* $m+n$ is defined for all positive integers (and zero) by defining $m+0$, then each of the values $m+0$, $m+1$, $m+2$, … in terms of the one before:

a. $m+0 = m$, for all numbers m.
b. $m+(n+1) =$ successor of $m+n = (m+n)+1$.

Then the *product function mn* is defined, with the help of sum, by successively defining $m\cdot 0$, $m\cdot 1$, $m\cdot 2$, … as follows.

a. $m\cdot 0 = 0$.
b. $m(n+1) = mn+m$.

Grassmann then went on to pave the way for Dedekind by *proving* (by induction) the basic rules of arithmetic:

$a+b = b+a,\quad ab = ba,$
$a+(b+c) = (a+b)+c,\quad a(bc) = (ab)c,$
$a+0 = a,\quad a\cdot 1 = a,$
$a(b+c) = ab+ac.$

Probably no one before Grassmann had thought it helpful to prove the rules of arithmetic, and the mind boggles at the thought of teaching such proofs to children. Nevertheless, Grassmann was onto something important: *the rules of arithmetic for positive integers have a common origin in the nature of counting.* This simple foundation also serves for subsequent extensions of the number concept, to which we will soon turn.

But first I want to point out that sum and product of positive integers are already very sophisticated operations. In fact, they are the source of some of the hardest problems in mathematics. To see how this comes about, let us recall the concepts of *divisibility* and *prime numbers*, both of which stem from the concept of product for positive integers. If a, b, c are positive integers and

$a = bc$

then we say that a is *divisible* by b, or that b *divides* a, or that a is a *multiple* of b. For example, the number 30 is divisible by 1, 2, 3, 5, 6, 10, 15, and 30. On the other hand, the number 31 is divisible only by 1 and 31. A positive integer greater than 1, and divisible only by itself and 1, is called a *prime* number. Thus the first 25 prime numbers are

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97.

It is hard to see any pattern in this sequence. Indeed the prime numbers are defined by their *avoidance* of simple patterns; they avoid the multiples of 2 (after 2 itself), they avoid the multiples of 3 (after 3 itself), and so on. Because of this lack of pattern, it is not even clear whether the sequence of prime numbers continues forever, though in fact it does, as Euclid proved in his *Elements* around 300 BCE. An even harder question is whether there are infinitely many pairs of primes (called *twin primes*) that differ by 2. One can see several such pairs in the list above, for example 5, 7 and 29, 31 and 71, 73, but it is not known whether such pairs keep appearing forever.

Now the concept of prime number is defined in terms of divisibility, and hence in terms of the product operation. For the concept of twin prime we need the idea of adding 2, and hence the sum operation. Thus the question whether there are infinitely many twin primes is a question about sums and products of positive integers. In fact, there is an inexhaustible supply of hard questions about sums and products, but I think the twin primes question illustrates the point well enough: arithmetic is not trivial!

It is, however, possible to make arithmetic a little easier to manage by extending the concept of integer to include *negative* integers and *zero*. We can then solve any equation of the form

$x+a = b$

for x when a and b are any integer, namely $x = b - a$. If, say, $b = 1$ and $a = 5$ then

$x = 1 - 5 = -4.$

The negative integer -4 is the reverse or *inverse* of 4 in the sense that

$4 + (-4) = 0,$

and in general every positive integer m has an inverse $-m$ such that $m + (-m) = 0$.
The negative integers and zero "complete" the positive integers with respect to the *subtraction* operation, because $m - n$ is an integer for any positive integers m and n.

Of course, these new integers do not arise from counting the number of objects in a collection, as the positive integers did, but they have other everyday uses. In cold countries one needs negative numbers to measure temperature relative to freezing point, and in all countries there are people with a negative bank balance.

There is a further extension of the integers that completes them with respect to division.
For each nonzero integer a we introduce the *reciprocal* or *multiplicative inverse* of a, $1/a$. The reciprocal enables us to solve any equation of the form

$ax = b,$

where a and b are integers and a is nonzero. The solution $x = b(1/a) = b/a$ is called a *rational* number because it is a *ratio* of integers. The system of rational numbers is ideal for arithmetic because it is closed under the operations of addition, subtraction, multiplication, and division (by nonzero numbers). Along with the calculation rules for positive and zero integers listed above, it satisfies the following rules for inverses:

$a + (-a) = 0, \quad a(1/a) = 1.$

In some sense, mathematicians have been aware of the rational numbers and their rules of calculation since ancient times, but the concept of a *system* of numbers and its rules did not emerge until the 19th century. Until this time, there was no dispute about the rules (though there was some dispute about the existence of negative numbers), and hence no need to discuss rules in general, because different concepts of "addition" and (especially) "multiplication" had not yet come to light. Also, until the 19th century, questions about the nature of numbers were overshadowed by questions from geometry: what are length, area, and volume, and can these quantities be regarded as numbers?

This explains the surprisingly late realization (by Grassmann and Dedekind) that the rules of arithmetic have their logical basis in the inductive nature of the positive integers, that is, in counting. Before the 19th century, all branches of mathematics were dominated by geometry, and one did not expect counting to be a basis for serious mathematics. In the next section, we discuss how geometry became dominant.

3. Length and area

The first bridge between numbers and geometry is the Pythagorean theorem – but it is a slippery bridge that the Pythagoreans were reluctant to cross. As all readers will recall from school, the Pythagorean theorem states that *the sum of the squares on the two sides of a right-angled triangle equals the square on the hypotenuse*. The theorem is illustrated by Figure 1.

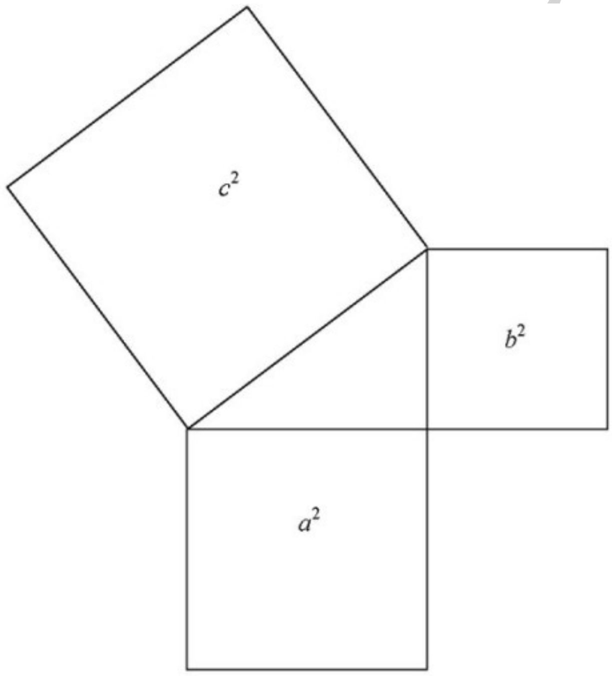

Figure 1: The Pythagorean theorem.

It is also expressed by the equation

$a^2 + b^2 = c^2$,

where a^2, or "*a* squared,'' denotes the square on side *a*, b^2 denotes the square on side *b*, and c^2 denotes the square on the hypotenuse, *c*. However, we also interpret "*a* squared'' as "*a* times *a*," because we compute the area of a square (or indeed, any rectangle) by multiplying the lengths of its perpendicular sides. For reasons that will soon become apparent, the Pythagoreans did not think this way. For them, a "square'' was just a square.

The Pythagorean theorem was discovered in several ancient cultures – in particular, in Mesopotamia, China, and India – and in some cases a long time before the Pythagoreans, who lived around 500 BCE. However, it was the Pythagoreans or their Greek successors who discovered the most remarkable consequence of the theorem: the existence of *irrational lengths*, such as $\sqrt{2}$.

The simplest right-angled triangle, with perpendicular sides of length 1, has squares of unit area on these sides, and hence a square of area $1 + 1 = 2$ on its hypotenuse. The length of the hypotenuse is therefore $\sqrt{2}$, because $\sqrt{2}$ is the side of a square of area 2, by definition. Thus very simple geometry demands the existence of the length $\sqrt{2}$. However, $\sqrt{2}$ *cannot be expressed as a ratio of positive integers, m/n*, since this assumption leads to a contradiction.

We do not know exactly when this contradiction was discovered, but it was common knowledge by the time of Aristotle, around 350 BCE. In his *Prior Analytics*, Aristotle feels it necessary only to mention that irrationality of $\sqrt{2}$ follows from consideration of even and odd numbers. Presumably, the assumption $\sqrt{2} = m/n$ was known to contradict some simple facts, such as these:

- The square of an odd number is odd.
- Therefore, an even square is the square of an even number.
- If $\sqrt{2} = m/n$, then $n^2 = 2m^2$ is even, hence n is even, say $n = 2k$.
- Then $2m^2 = n^2 = 4k^2$, so $m^2 = 2k^2$, hence m is also even.
- But we can take the fraction m/n in lowest terms, with m and n *not* both even.

But if $\sqrt{2}$ is not a rational number, what is it? The Greek view was that $\sqrt{2}$ is a length but *not* a number, and that lengths do not even "behave" like numbers. In particular, the Greeks viewed the product of two lengths a and b not as a length, but as a rectangle with perpendicular sides a and b. Thus "$\sqrt{2}$ squared" was literally a square with sides of length $\sqrt{2}$, not a line of length 2. And "$\sqrt{2}$ cubed" was, you guessed it, literally a cube with sides of length $\sqrt{2}$.

This is magnificently visual, but it is not arithmetic. Trying to decide when two products are equal is a headache – in what sense does the rectangle with sides $\sqrt{2}$ and $\sqrt{3}$ "equal" the rectangle with sides 1 and $\sqrt{6}$? – and the product of four or more lengths has no meaning at all, because we cannot visualize more than three dimensions. However, these are disadvantages only with hindsight.

The Greeks were able to develop a vast amount of geometry without a proper arithmetic of lengths, so they never missed it. For example, they succeeded in defining equality of rectangles by "cutting and pasting": rectangles are equal if it is possible to cut one of them into pieces that can be reassembled to form the other.

In this sense the rectangle with sides $\sqrt{2}$ and $\sqrt{3}$ *is* equal to the rectangle with sides 1 and $\sqrt{6}$. It was only after arithmetic itself became more sophisticated, in the form of *algebra*, that it became possible to gain geometric advantages from arithmetic of lengths.

-
-
-

Bibliography

Dedekind, R. (1963). *Essays on the Theory of Numbers*, Dover. [Two essays by Dedekind: one from 1872, on the nature of irrational numbers as cuts in the rational numbers; the other from 1888, on the role of counting and induction in the arithmetic of positive integers].

Descartes, R. (1954). *The Geometry of René Descartes*, Dover. [A facsimile of the 1637 edition of Descartes' *La Geométrie*, together with an English translation and notes by David Eugene Smith].

Diophantus (2003). *Diophantus of Alexandria* (edited by Sir Thomas Heath) Martino Publishing (reprint of 1910 Cambridge University Press edition). [A translation of the first six books Diophantus' *Arithmetica*, together with copious notes by Sir Thomas Heath, documenting the influence of Diophantus on later mathematicians, such as Fermat.]

Ebbinghaus, H.-D, et al (1990). *Numbers*, Springer-Verlag. [A comprehensive modern treatment of real and complex numbers and their generalizations, including quaternions and octonions].

Euclid (1956). *The Thirteen Books of Euclid's Elements* (3 volumes), Dover. [Euclid's geometry and number theory complete, with notes by Sir Thomas Heath. Heath's notes are helpful in interpreting the archaic style of Euclid and relating it to later developments, though they miss the developments that originate with Hilbert].

Hilbert, D. (1971). *The Foundations of Geometry*, Open Court. The 1950 edition is available free from www.gutenberg.org/ebooks/17384. [Hilbert shows the algebraic significance of the Pappus and Desargues theorems, but misses the geometric aspects of quaternions and octonions].

Viète, F. (2006). *The Analytic Art*, Dover. [A translation of Viète's works on algebra, which introduce most of the notation we are familiar with today].

Biographical Sketch

John Stillwell was born in Australia in 1942 and educated at Melbourne High School and the University of Melbourne. He went to the US for graduate studies, obtaining a Ph.D. in mathematics from MIT in 1970. For 30 years he taught mathematics at Monash University in Melbourne, before joining the University of San Francisco in 2002.
During this time he has also held visiting positions at MIT and the University of Cambridge, and has given talks in many countries, including Canada, UK, France, Germany, Denmark, India, Malaysia, and New Zealand.

As a student, his main interest was in mathematical logic, but after his Ph.D. his interests expanded to group theory, topology, geometry, number theory, and the history of mathematics. He is interested in propagating mathematics to a wide audience and has written a number of books and expository papers. His efforts in this direction have included invited addresses at the International Congress of Mathematicians (Zürich 1994) and the Joint Meeting of the American and Australian Mathematical Societies (Melbourne 1999), and have been recognized by the Chauvenet Prize for mathematical exposition (Mathematical Association of America, 2005).

His books include translations of classic works by Dirichlet, Dedekind, Poincaré, and Dehn, and by modern masters such as Serre and Brieskorn. Among his own books, the best-known are *Mathematics and Its History* (2nd edition, Springer 2002), *The Four Pillars of Geometry* (Springer 2005), and *Yearning for the Impossible* (A K Peters 2006).

The paper that won the Chauvenet prize is *The Story of the 120-cell* (Notices of the American Society, January 2001).