SHOW NOTES: NUMBER SYSTEMS

ALED WALKER

Abstract. In these brief informal notes, we discuss the development of pure mathematics from the perspective of expanding number systems – starting with the integers and the rational numbers, and moving on to discuss the real numbers, complex numbers, algebraic numbers, algebraic integers, constructible numbers, transcendental numbers, finite fields, p-adic numbers, the idèles and the adèles.

The Cole prizes are awarded by the American Mathematical Society, once every three years. There are two such prizes – one is for number theory, the other for algebra – and, to cut a long story short, they are quite difficult to win. But who was the eponymous Frank Nelson Cole? A mathematician, obviously, and one who is now most widely remembered for a single lecture he gave at a 1903 meeting of the Society. He walked to the blackboard, saying nothing, and carefully calculated $2^{67} - 1$, i.e. he multiplied $2 \times 2 \times 2...$ and so on, 67 times, and then subtracted 1. On a second blackboard, he then performed a long-multiplication of $193707721 \times 761838257287$. The two final answers matched. He sat down, to rapturous applause.

$$2^{67} - 1 = 147,573,952,589,676,412,927 = 193707721 \times 761838257287.$$

This anecdote was recounted, and probably embellished, by Bell in the 1950s. It survives as a nugget of mathematical culture, in part because of the natural theatricality of the event itself, but also because the topic of Cole's lecture was so unusual. Contrary to certain popular belief, performing ever more complicated multiplications and additions – of the kind that plague primary school classrooms to this day – is not what professional mathematicians spend their time doing.

This is not to say that there isn't a rich and interesting story to be told about why a man of Cole's distinction was engaged in finding the factors of $2^{67} - 1$, why it was difficult for him to do so, and why the audience applauded him for having done so. It is not our story for this podcast, however: the development of pure mathematical research over the last two millennia has not primarily been the story of increasing computational power. It would be too bold of me to claim that any single theme has, to the exclusion of all others, typified mathematical evolution. Yet, a large part of the development, elaboration, and maturing of mathematics can be construed as comprising a single broad endeavour: an expansion of the notion of 'number'.

These notes are about the number systems $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\overline{\mathbb{Q}}$, $\mathbb{F}_p$, $\mathbb{Q}_p$, and $\mathbb{A}_{\mathbb{Q}}$, and their respective mathematical and historical contexts. Some of our discussion will be quite mathematically detailed; however, as the mathematics becomes harder, we have found it necessary to employ more analogies and vague allusions, in lieu of a precise mathematical handling of the number systems involved. Upon reading this, some readers may be disappointed – others relieved!

Arithmetic and the ancient world

Traditionally, Ancient Greek civilisation gets the credit for inventing the discipline of mathematics. Certainly, as far as the archaeological record attests, no other contemporary culture in Western Eurasia seems to have been interested in the concept of mathematical proof (using a chain of logical reasoning to justify a mathematical claim). Athens of the 4^{th} century BC had Plato's Academy, and the Hellenistic cities of 3^{rd} century Alexandria and Syracuse harboured Euclid and Archimedes respectively. It was Euclid who summarised, codified, and extended the Greek geometric project with such an extraordinary diligence that his twelve volume textbook *The Elements* became the standard mathematical treatise for the next two thousand years; whereas Archimedes developed a deeper understanding of solid geometry and infinitesimal processes than would be found anywhere on Earth until the 1600s.

The early Greek mathematicians thought about numbers in a way that feels, to modern sensibilities at least, deeply peculiar. For them, geometry was the central mathematical discipline, not arithmetic, and so, when expressing arithmetical arguments, they did so in a geometric language. They didn't write about the number '1', but rather referred to the 'unit line segment' with respect to which all other lengths would be measured. Rather than talking about the number 5 as an independent entity, they would refer to a line segment which was five times as long as the original unit line segment. Proportion and commensurability were the key concepts; they would say that two lines AB and CD were *commensurable* if there was another line segment EF which could act as a unit line segment between them, so that the length of AB was a whole multiple of the length of EF and the length of CD was also a whole multiple of the length of EF. If no such segment EF existed, the two line segments AB and CD were said to be *incommensurable*.

What does this mean in terms of a modern-day schoolchild's concept of numbers? Well, it means that the Greeks could work with the ordinary whole numbers 1, 2, 3, 4, etc., by taking the unit line segment and then putting multiple copies end to end. For no good reason at all, modern mathematicians have a fancy name for these whole numbers (we call them the *natural numbers* and denote the set of all natural numbers by the letter $\mathbb{N}$). The Greeks could also work with fractions given by whole numbers. Say AB and CD are commensurable line segments with lengths given by 5 units and 7 units respectively. Then the ratio of the length of AB to the length of CD (which is a kind of concept which comes up again and again in Euclid's Elements) is $\frac{5}{7}$, or in others words the fraction $\frac{|AB|}{|CD|} = \frac{5}{7}$. Modern mathematicians have a fancy name for this kind of fraction too: we call them *rational numbers* (note how the root word 'ratio' survives in this nomenclature). Rational numbers are just fractions, but where the top and the bottom of the fraction – the numerator and denominator – are required to be whole numbers.

The Greeks did not accept negative numbers into their mathematics.[1] After all, for them numbers corresponded to geometrical configurations, and what line segment has a length *minus* 7?! This limitation persisted in Western mathematics right up until the Renaissance, when Italian algebraists made their work substantially more difficult by refusing to treat negative quantities in the same way as positive ones. The number zero, too, doesn't really feature in Greek mathematics, and it seems they possessed no special symbol for it. Such a symbol was invented many times independently throughout the world; the root of the current Arabic numerical 0 starts in Ancient Sumerian tax accounting, travels via 1^{st} millenium India and then back to the Islamic empire of the Middle Ages.

Nowadays, we consider negative numbers on an equal footing to positive numbers. The collection of all whole numbers (positive and negative), together with zero, is called the *integers*, and goes by the symbol $\mathbb{Z}$ (after the German *Zahlen*). So

$$\mathbb{Z} = \{..., -2, -1, 0, 1, 2, \dots\}.$$

Regarding the rational numbers, you can divide any integer by any other integer except 0. In other words, the rational numbers are all fractions, positive and negative, which are

[1] The Han Chinese did express negative numbers around 200BC – roughly contemporary with Euclid – in their treatise 'The nine chapters of the mathematical arts'.

expressed by a whole number divided by another whole number. We give them the symbol $\mathbb{Q}$, for *quotient*.

$$\mathbb{Q} = \{\frac{a}{b} : a, b \text{ integers}, b \neq 0\}.$$

The discussion above has skipped over a great many important developments, of course. The period of Hellenistic civilisation lasted over 500 years, and naturally there were changes in perspective during so great a span of time. Diophantus, writing in about 250AD, wrote a 13 volume work *Arithmetica* which considered (positive) rational numbers in largely the same way as we do now. This book would not be bettered as a number theory text until the publication of Gauss's *Disquisitiones Arithmeticae* in 1801. The type of equations that Diophantus considered were largely algebraic equations (such as the quadratic equation $ax^2 + bx + c = 0$) in which the coefficients (a, b, c) were rational numbers. We now call these kind of equations *diophantine equations* in his honour.[2]

I have also given short shrift to the mathematics of other ancient civilisations – which is most unfair. The Sumerians of Mesopotamia, as far back as 3500BC, had a well-developed government bureaucracy and taxation system, documented by many surviving tablets, indicating facility with mathematical calculation. Intriguing material evidence survives from the later Babylonian civilisation, such as *Plimpton 322* (which is so called because of the collection at Columbia University in which it is stored). This is a small clay tabel, written (mostly) in Akkadian, that gives a list of triples of whole numbers a, b, c for which $a^2 + b^2 = c^2$, i.e. for which a mutliplied by itself, added to b multiplied by itself, equals c multiplied by itself. For example $(3, 4, 5)$ or $(7, 15, 17)$. We would now call such things 'Pythagorean triples', as they can form the sides of a right-angled triangle. What's more, it seems as if the scribe who made the tablet (despite making a few errors) was using a method for generating Pythagorean triples, rather than just approaching via trial and error. This tablet dates from about 1800BC, or about 1300 years before Pythagoras lived!

Some intriguing remains survive from Egypt too. Most famously, the Rhind papyrus, dating from around 1550 BC, contains a wealth of mathematical calculations; these range from the geometric (such as computing the volumes of granaries) to the arithmetic (expressing rational numbers of the form $\frac{2}{n}$ as sums of rational numbers of the form $\frac{1}{k}$ for some other integers k). For example,

$$\frac{2}{15} = \frac{1}{10} + \frac{1}{30}.$$

Manipulating with rational numbers of the form $\frac{1}{k}$ is still known as doing 'Egyptian fractions'.

Real numbers

I left a loose end in my discussion of the Greeks: what did they do about incommensurable lengths? They knew such incommensurable ratios existed; for example, take the ratio of the length of the diagonal of a square to the length of one of the sides of the same square. Yet their early theory of proportion didn't account for this possibility, either working only with commensurable ratios or somewhat fudging the point and avoiding an explicit discussion of whether the ratios involved were commensurable or not.[3]

The challenge was taken up by Eudoxus (who is the greatest mathematician you've never heard of). We know rather little about his life, except that he lived around the 4^{th} century

[2] Fermat, the infamous amateur French mathematician of the 1600s, had a copy of Diophantus, in which he penned the marginalia that posed what is now known as 'Fermat's Last Theorem'.

[3] There is a (no-doubt apocryphal) story of how the follower of Pythagoras who discovered that incommensurable ratios existed, and who then shared his great discovery with Pythagoras himself, was summarily thrown overboard from the boat in which the posse was travelling. A funny tale, but, given that the major sources about the life of Pythagoras come from over 500 years after his death, probably myth. Yet it is certainly true that Pythagoras would not have known how to rigorously manipulate incommensurable ratios.

BC, and none of his writing survives directly. Yet his theory of proportion (as well as his other great contribution, the method of exhaustion) survive by having been included by Euclid in his treatise. Book V of The Elements contains the theory of proportion.

The main issue is how to decide if two ratios are equal or not. In the case of commensurable ratios, the task is easy – in modern parlance, you find a common denominator for all the fractions involved, and then just directly compare numerators. For example, is $\frac{8}{10}$ equal to $\frac{5}{7}$? Putting these fractions over a common denominator, we find we are comparing $\frac{56}{70}$ with $\frac{50}{70}$, which are clearly not equal because 56 and 50 are not equal. This makes the rational numbers into an ordered set, i.e. we can place them on a number line, as you might have seen on your classroom wall. Easy – but seemingly impossible to do with incommensurable ratios.

Eudoxus' solution is startlingly modern. Although Euclid phrases everything in confusing prose, the idea seems straightforward enough: Eudoxus says that two ratios x and y should be treated as equal if and only if, for every positive rational number $\frac{a}{b}$, either

- $a \leqslant bx$ and $a \leqslant by$; or
- $a \geqslant bx$ and $a \geqslant by$.

In fact, this definition has an extraordinarily high level of sophistication. Note immediately that Eudoxus has conditioned over an infinite set, namely the set of all rational numbers. The consequence of his work is that the collection of all ratios (both commensurable and incommensurable) can be completely ordered, i.e. all put on a number line together. For any two ratios x and y there is a rigorous theoretical way of determining whether $x < y$, $x = y$, or $x > y$ (namely $x < y$ if we can find a positive rational number $\frac{a}{b}$ with $a > bx$ and $a < by$, in which case we have the full inequalities

$$x < \frac{a}{b} < y).$$

Modern mathematics has a different name for this collection of all ratios: we call it the *real numbers*. The reason for this nomenclature will be explained in the next section (when real numbers will be contrasted with so-called *imaginary numbers*). In concrete terms, real numbers are just decimals – any decimals. So 1.414, that's a real number (in fact it is a rational number, being equal to $\frac{1414}{1000}$). But 1.41421356237..., where I keep picking different whole numbers between 0 and 9, and carry on forever – that is also a real number. I'll return to decimal representation at the end of this section.

The real numbers overcome an annoying property of the rational numbers: the set of rational numbers are 'incomplete'. What does this mean? Consider the sequence of rational numbers $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, and so on. These fractions are getting smaller and smaller (imagine getting half a cake one day, but then the next day only getting a third of a cake, the next only a quarter, and so on – each day you are getting less and less cake!). Intuitively, it seems this sequence is getting closer and closer to zero. We say that the *limit* is 0. Since 0 is a rational number, we've managed to stay within the set of rational numbers while taking the limit, and this certainly seems to be a convenient fact. After all, it would be a pain if one took a sequence of commensurable ratios which were all close to each other, took the limit, and then ended up with an incommensurable ratio.

But that's exactly what happens in certain other examples. Here is another (more complicated) sequence:

$$\frac{1}{1}, \frac{7}{5}, \frac{41}{29}, \frac{239}{169}, \dots ,$$

where the rule from getting from one term in the sequence to the next term is

$$\frac{a}{b} \mapsto \frac{3a+4b}{2a+3b}.$$

If I square each element in this sequence, I end up with

$$\frac{1}{1}, \frac{49}{25}, \frac{1681}{841}, \frac{57121}{28561}, \ldots,$$

Notice how for each of these fractions, the numerator is very close to twice the denominator!

The fractions in the second list get closer and closer to 2 (this can be rigorously proved). So we would like to say that the fractions in the first list have a limit of the square-root of 2 (denoted $\sqrt{2}$), which is nothing more or less than some number x such that $x^2 = 2$.

But $\sqrt{2}$ is not a commensurable ratio! Indeed, it is the same as the ratio between the length of a diagonal of a square and the length of the side length of the same square, and the Pythagoreans already knew that this ratio was incommensurable. This is the way in which we mean that the rational numbers are incomplete; you can take limits, but the limit you end up with might not be a rational number – there might not be a crock of gold at the end of the proverbial rainbow.

However, the advantage of considering commensurable and incommensurable ratios together is that this combined collection *is* complete. The incommensurable ratios 'fill in the gaps' that the commensurable ratios leave behind. This property is the basis for most of calculus, differential equations, mathematical physics, and of the whole edifice of mathematical analysis which surrounds these fields.

We shouldn't leave this section before cleaning up a few points. Firstly, these days we don't talk about about commensurable and incommensurable ratios: we refer instead to rational numbers and irrational numbers respectively (so $\sqrt{2}$ is an irrational number). Secondly, Eudoxus's method for putting an order theory on irrational numbers is doubly remarkable in that it foreshadows the 19^{th} century work of Dedekind, who used almost exactly the same device to logically define a real number x as synonymous with set of rational numbers q with $q \leqslant x$. This is the theory known as 'Dedekind cuts', and is lectured to undergraduates today.

Finally, I promised to return to the issue of decimals. The decimal expression of a real number x, say $x = 1.414421356237...$ as earlier, is a way of recording a sequence of rational numbers whose limit is x. In other words, we're saying that x is quite close to $\frac{14}{10}$, even closer to $\frac{141}{100}$, closer still to $\frac{1414}{1000}$, even closer to $\frac{14142}{10000}$, and so on.

Addendum: there is a postscript to be added about the real numbers, one which I am scared to mention as it runs so counter to our everyday intuition about the world. But mention it I will, as one cannot fully appreciate the real numbers without it.

Let's ask a strange question: are there more rational numbers or integers? This seems silly for a variety of reasons, not least because both these sets are infinite. However, there is a certain way in which we can pair-off integers and rational numbers in a one-to-one fashion. Such a pairing is called a *bijection*, and although the details a little technical, the overall idea is that you can list all the rational numbers with numerator and denominator between -10 and 10 (say) – this is a finite list – followed by all the rational numbers with numerator and denominator between -100 and 100, and so on. By making such a list, you are giving a first element, a second element, and third element of the list, and so on, which is assigning an integer to each rational number. Fiddling with the details, you can pair each integer with exactly one rational number and vice versa. We say that the rational numbers are *countable*, because we can list them in this way.

But is there such a pairing between the integers and the real numbers? It turns out that there is *not*. This was a wild, epoch defining result of the French mathematician Cantor, and led to a fracturing of the mathematical community in the late 19^{th} century from which it took decades to recover. In 1886, an exasperated Kronecker was moved to exclaim, "God created the integers: all else is the work of man!" (Of course he said it in German, "*Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk!*") What Cantor was saying was that, although the set of all integers $\mathbb{Z}$ and the set of all real numbers (denoted $\mathbb{R}$) are both infinite sets, $\mathbb{R}$ is 'more infinite' then $\mathbb{Z}$. Infinities can have different sizes!

Cantor's main contribution to mathematics was to help usher in the era in which the logical foundations of mathematics were plumbed for inconsistencies and contradictions. Another worthy story, for another time.

Algebraic numbers

The square-root of 2, denoted $\sqrt{2}$, has already played a role in our discussion. One way of considering $\sqrt{2}$ is as a solution to a certain kind of equation, called a *polynomial equation*, namely $x^2 - 2 = 0$. To spell it out, if you multiply $\sqrt{2}$ by itself and then subtract 2, you get 0.

Quadratic equations of the kind one might remember from school, namely $ax^2+bx+c = 0$, are types of polynomial equation. They have 'degree 2', which means that the highest power of x which appears is x^2. Yet one can consider cubic equations, which are of the form $ax^3 + bx^2 + cx + d = 0$ (and so are of degree 3), or quartic equations (of degree 4), or in general a polynomial of degree n for any natural number n. We call the numbers a, b, c etc. the *coefficients* of the polynomial.

This brings us to the definition of algebraic numbers. If x is the solution to a polynomial equation where the coefficients are integers, we call x an *algebraic number*. Informally, this means x can be constructed by algebra – although there some important subtleties underlying what we might mean by 'constructed' here.

Let's have some examples. Every rational number is also an algebraic number, since if $x = \frac{a}{b}$ is rational (with integers a and b) then it is the solution to the degree 1 polynomial equation

$$bx + a = 0.$$

The square-root of 2 is also algebraic, as it is a solution to $x^2 - 2 = 0$. More exotically, a number such as $4^{1/3} + 1$ (that is, the cube-root of 4, add 1) is algebraic, as it is the solution to the degree 3 equation

$$x^3 - 3x^2 + 3x - 5 = 0.$$

Algebraic numbers are the solutions to the manner of problem that Diophantus liked to pose. We give them the symbol $\overline{\mathbb{Q}}$.

I promised to mention some of the subtleties, and here they are. We might remember from school the quadratic formula, namely that the two solutions to $ax^2 + bx + c = 0$ are

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

In other words, this formula gives an expression for all degree 2 algebraic numbers in terms of the coefficients of the relevant polynomial and square-roots.

In Renaissance Italy, formulae for degree 3 and degree 4 polynomial were discovered. These are both significantly more complicated than the quadratic case: the degree 3 formula gives a solution to the equation $ax^3 + bx + cx + d = 0$ as

$$\begin{aligned} x = &\Big(\Big(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\Big) + \Big((\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\Big)^2 + \Big(\frac{c}{3a} - \frac{-b^2}{9a^2}\Big)^3\Big)^{1/2}\Big)^{1/3} \\ &+ \Big(\Big(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\Big) - \Big((\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\Big)^2 + \Big(\frac{c}{3a} - \frac{-b^2}{9a^2}\Big)^3\Big)^{1/2}\Big)^{1/3} - \frac{b}{3a} \end{aligned}$$

(appearing in Cardano's book from 1545). However, the structure of this formula is broadly similar to the quadratic formula: we can write the algebraic number x as a combination of the coefficients, involving square-roots and cube-roots.

The natural question is then: is there a quintic formula? Given a polynomial of degree 5 (of more generally of degree n with $n \geqslant 5$), can we write the solutions to the polynomial equation as some combination of the coefficients involving just square-roots, cube-roots, fifth-roots etc.? If so, then this would give an alternative characterisation of algebraic numbers.

The answer turns out to be no! This was first proved, more-or-less, by Italian mathematician Ruffini in the late 18^{th} century. But the European mathematical community took astonishingly little notice of his treatise (reception was hindered by the Napoleonic wars, and by the French master mathematician Lagrange seeming to lose his copy of Ruffini's manuscript, or maybe never even receiving it). The great length of Ruffini's proof also hindered its comprehensibility. It was the Norwegian prodigy Abel who provided the first full proof in the 1820s which was universally accepted. The French revolutionary Galois – another prodigy, who died in a duel at the age of 20 – also considered this problem. His work was largely lost until its rediscovery by Liouville in the 1840s. Taken together, these works heralded the arrival of modern abstract algebra: group theory. This is the mathematics of symmetry, which underpins everything from computer graphics to sub-atomic physics.

Transcendental numbers

Are all real numbers algebraic? What about the classical constant of antiquity, π (the ratio between the circumferences of a circle and its diameter) – is that algebraic? What about that central constant underpinning exponential growth, e?

These questions were asked and settled in the 19^{th} century. Liouville was the first to show the existence of real numbers which are not algebraic. These are called *transcendental numbers.* Liouville's number was

$$0.110001000000000000000001...,$$

in which 1s were included at places 1, 2, 6, 24, 120, 720, etc. (spot the pattern). Hermite proved in the 1870s that e was transcendental, then Lindemann adapted the proof to show that π was transcendental. (It might seem bizarre that the same proof can work for two such different numbers, but of course these numbers are connected by Euler's famous relation $e^{i\pi} = -1$.)

Cantor worked more indirectly. As the rational numbers are countable, he also showed that the algebraic numbers are also countable. Since the real numbers are uncountable, this means that there must exist transcendental numbers – also Cantor's argument doesn't construct a single explicit transcendental number!

These issues have long shelf-lives. In the 1960s, the British matematician Alan Baker showed that a broad class of numbers were transcendental, namely numbers x of the form

$$a_1^{b_1} a_2^{b_2} \cdots a_n^{b_n},$$

where $b_1, b_2, \ldots, b_n$ are all algebraic and irrational[4] and the a_i are all algebraic and not 0 or 1. His theorem, going via 'linear forms in logarithms',1 attacked a variety of other problems, including the class number problem of Gauss. For this he was awarded the Fields medal in 1970

[4] and satisfy a technical side-condition about how rational multiples of the b_i interact, called *linear independence*

Algebraic integers

There are special kinds of algebraic numbers called *algebraic integers*. These are defined as follows. Given an algebraic number x, it has a so-called 'minimal polynomial'. This is the polynomial equation with integer coefficients that has the smallest degree amongst all those which have x as a solution. For example, $x^2 - 2$ is the minimal polynomial of $\sqrt{2}$, even though $\sqrt{2}$ is also a solution of $x^3 - 2x = 0$ (this second polynomial has higher degree, so can't be the minimal polynomial). We say that x is an algebraic integer if the leading coefficient (i.e. the first coefficient) of its minimal polynomial is 1. So $\sqrt{2}$ is an algebraic integer, but $(3/5)^{1/3} - 1$ is not (since its minimal polynomial is $5x^3 + 15x^2 + 15x + 2$).

Amongst rational numbers, only the integers are algebraic integers (this explains the name). Indeed, we've already seen how the minimal polynomial for $\frac{a}{b}$ is $bx - a$ (at least if a and b share no common factors), which means that we only get an algebraic integer if $b = 1$, i.e. if $x = a$ is an integer.

It turns out that many properties of what we might call 'usual arithmetic', i.e. arithmetic with integers, carry over to arithmetic with algebraic integers. For example, if x and y are both algebraic integers then $x + y$ and xy are also both algebraic integers. Yet, there are differences too, most notably surrounding prime factorisation. The arithmetic of algebraic integers lies at the heart of a great deal of higher number theory; for example, to understand the integer solutions (x, y) to the equation $y^2 = x^5 - 10$, it is necessary to understand the structure of the algebraic integers involving $\sqrt{-10}$. (One might ask how one can take the square-root of a negative number – we'll come back to that in the next section.)

Here's a lighter example of the influence of algebraic integers. If you can get a computer to calculate $e^{\pi\sqrt{163}}$, you'll see that this number it is exceptionally close to an integer (the decimal expansion has twelve 9's after the decimal point).

$$e^{\pi\sqrt{163}} = 262537412640768743.99999999999925...$$

The reason for this seeming coincidence turns out to have a deep connection to the algebraic integers related to $\sqrt{-163}$.

Imaginary numbers, complex numbers

If any fundamental concept from higher mathematics has ever truly caught the popular imagination, it is the concept of *imaginary numbers*. "Adam and Eve are like imaginary numbers, like the square root of minus one," says Philip Pullmann, through Lord Asriel's mouth, somewhat cryptically. In Lacan's analysis, "the erectile organ can be equated with the square root of minus one." Hmm. (Several books of argument and counterargument have been devoted to trying to work out what on earth Lacan might have actually meant here.) As a metaphor, the square root of minus one is frequently used as the archetype of an object that is both there and yet not-there. No familiar physical quantity – neither length, weight, nor volume – can be multiplied by itself to give a negative number; the square root of minus one seems then to be the diabolical discovery-cum-invention-cum-creation of mathematicians run amok, and its devilish usefulness in the mathematical arts no compensation for its offence to our philosophical sensibilities.

But one should not think like this. The square root of minus one first came into mathematics in a serious way through the work of those Italian mathematicians of the 16th century – Tartaglia, Cardano, Scipione del Ferro and Bombelli – who were considering the solutions to algebraic equations. As we've already mentioned, they discovered a method for finding a solution to cubic equations (such as $x^3 - 15x - 4 = 0$). Yet, even when the final solution was a perfectly familiar number (the number 4 in the above example), their method sometimes necessitated, during intermediate steps of the calculation, the use and manipulation of the square roots of negative quantities (the square root of minus 11 in this instance). These square roots did not exist in the number system as the thinkers of the day knew it, nor as any present-day secondary school pupil would know it. A quandary, no doubt, and one that took mathematicians about another 200 years to properly get themselves out of, after these first Italian forays into the hitherto unknown. But the solution was ultimately very simple: in the face of a mathematical pathology, one simply extends the number system to include the pathology.

And so imaginary numbers take their place in humanity's well-trodden investigation of the following simple truism: life is the sweeter for equations having solutions. Consider the state of the human being who knows, as a given, only the existence of the numbers $0, 1, 2, 3, 4$, etcetera, i.e. *non-negative integers.* They want the equation $x + 2 = 0$ to have a solution? Well then, they're going to need a new number (let's call it minus two, and give it a symbol, say -2). Thus they construct the integers $\mathbb{Z}$. They want the equation $5x - 3 = 0$ to have a solution? Well then, they're going to need a new number (let's call it three-fifths, and give it a symbol, say $\frac{3}{5}$). Thus they construct the rational numbers $\mathbb{Q}$. They start getting interested in the hypotenuses of right-angled triangles, and want the equation $x^2 - 2 = 0$ to have a solution? Well then, they're going to need a new number (let's call it the square root of 2, and give it a symbol, say $\sqrt{2}$). Thus they construct the real algebraic numbers.

But none of the numbers they have constructed so far will solve the equation $x^2 + 1 = 0$. So, if our friend wants this equation to have a solution, they will need a new number (let's call it the square root of minus one, and give it a symbol, i say). Since i doesn't correspond to any physical quantity, the whizz-kid in marketing (René Descartes, in this instance) decides to call it an 'imaginary number', and all the previously constructed numbers 'real numbers'; the name sticks, the mystery cult grows, and the whizz-kid gets a pay-rise.

Imaginary numbers are not they terribly significant in and of themselves, despite the publicity afforded them by the likes of Pullman and Lacan. The numbers that mathematicians really care about are not the purely imaginary numbers (i, $2i$, $3i$, and so on), but the so-called *complex numbers* ($3+4i$, $1-7i$, $\sqrt{2}+\frac{3}{5}i$, and so on), numbers where one has a purely real part plus a purely imaginary part. We give these numbers the symbol $\mathbb{C}$.

One could fill several libraries with books on the theory of the complex numbers (and people have). So what should the non-specialist know? Most immediately, the elementary algebra of complex numbers, based directly on the fact that $i^2 + 1 = 0$, turns out to be an extremely convenient way of encoding the algebraic relationships between the trigonometric functions sine and cosine. De Moivre knew this in the 1730s, and today, from fluid mechanics to electrical engineering, complex numbers are an invaluable tool for calculations, calculations that could be done using the usual trigonometric functions, with no mention of i or complex numbers, but which are greatly simplified by their introduction. It is in this part of the discipline that one finds Euler's famous relation $e^{i\pi} = -1$.

But the complex numbers also have a life of their own. It turns out that they enjoy a rich geometry – discovered initially by that peerless mathematician Gauss, and independently by a Parisian bookkeeper named Argand – in which the purely imaginary numbers may be placed at right angles to the purely real numbers. This forms the so-called *complex plane*, on which one may consider the point with coordinates (a, b) as if it were the complex number $a + ib$. Amongst other deep consequences, this representation can be used to show that every single algebraic equation with coefficients in the complex numbers can be solved using complex numbers. That's every equation, not just the equation $x^2 + 1 = 0$ (which, you'll recall, was the only equation we used to introduce i). In terms of the story above – adding symbols when algebraic equations don't have solutions – this theorem says that, once we include the symbol i in our mathematics, we don't need to add any other symbols. This result is called the Fundamental Theorem of Algebra, and it is, well, pretty fundamental. Any theory based on linear algebra – quantum mechanics, for a start – needs this remarkable theorem somewhere in its construction. Far from being philosophical aberrations, it is with complex numbers that mathematics achieves its full splendour.

Finite fields

Consider an analogue clock, and answer me the following question: what time will it be nine hours after four o'clock? Well, one o'clock, obviously. Now, by saying this, you aren't saying that $4 + 9 = 1$: that would be silly. What you are saying is,"Discounting multiples of 12, $4 + 9$ is the same as 1." You perform a mental conjuring act, without even realising it: you pretend that any multiple of 12 is actually equal to 0, and then you carry on the addition calculation regardless. Mathematicians gave this a name: 'modular arithmetic'. The term was introduced by Gauss in 1801, with the etymology seeming to be from the Latin 'modulus', meaning a small measure or interval. We say that $4 + 9$ is *congruent* to 1 *modulo* 12, or in other words that 12 divides $(4 + 9 - 1)$. Similarly 27 is congruent to 3 modulo 12 and $10 + 23$ is congruent to 9 modulo 12. But there's no reason to stick with 12. Using the same language, 14 is congruent to 23 modulo 9, and 6 is congruent to 20 modulo 7.

You might be forgiven for thinking that this is all unnecessarily complicated, given that calculating the time on a analogue clock is not a difficult task – why invent all this highfalutin jargon? The point is, by codifying the key aspects of the calculation, mathematicians have given this operation a name, and, having given it name, one can refer to this operation much more freely in further discussion. As an immediate (though rather tame) example, with modular arithmetic in hand one can easily show those high-school tricks, that an integer is a multiple of 9 if and only if the sum of its digits is a multiple of 9, and that an integer (written $a_1a_2...a_k$) is a multiple of 11 if and only if $-a_1 + a_2 - \cdots + (-1)^k a_k$ is a multiple of 11. For example, 1089 is a multiple of 11, since $-1 + 0 - 8 + 9 = 0$.

Rather more seriously, begin by taking your favourite prime number p (Andrew Wiles's is 37, so let's use that). Take any number x that isn't a multiple of p. Let's pick 2. Then compute x^{p-1} modulo p, so in our example we're computing 2^{36} modulo 37. Unless you are very good at your 37 times tables you might struggle to do this in your head, but point it at a computer and you will get the answer 1. In fact no matter what x you pick, and no matter what prime p you pick, you will always get the answer 1. This is called Fermat's Little Theorem, and the proof is easily comprehensible to a bright A-Level student, but that same student might have really struggled to see this fact just by staring at a clock for a while. And far from being a triviality, Fermat's Little Theorem is what underlies RSA encryption, the algorithm that is one of the ways in which your credit card data remains secure when it is being used for an internet payment. Giving things names is useful, if you know the names.

So what is a 'finite field'? Well, consider the collection of numbers 1, 2, 3, ..., 12 again. This is a finite set. We can define a notion of addition on this set, namely addition modulo 12. We can multiply elements together too, just by performing multiplication modulo 12. For example $5 \times 7 = 35 \equiv 11$ modulo 12. (Here $\equiv$ stands for 'is congruent to'). So we've created an arithmetic structure which looks a lot like the integers, except that it is finite. We call such a thing a *finite ring.*

But there are certain awkward properties about this particular ring. In particular, you can multiply two non-zero numbers together and get zero; for example $3 \times 4 = 12 \equiv 0$ modulo 12. It turns out that if, instead of using 12, one works modulo p for a prime number p, e.g. $p = 37$, the problem goes away. In this world, you can even define a notion of division. For example 1 divided by 2 (modulo 37) is equal to 17, since $17 \times 2 = 38 \equiv 1$ modulo 37.

Such a structure looks quite like the rational numbers, or the real numbers, where we can add, subtract, multiply and divide. The name for such a structure is a *field.* I guess the idea of this name is that, in the middle of an open field, you can go anywhere! Whereas if you are stuck on a ring, your movements are constrained.

Here we have it then, a *finite field.* We give this the name $\mathbb{F}_p$.

p-adic numbers

In these two final sections, I'll try give a briefest of glimpses into modern algebraic number theory. Actually, we'll only get as far as around 1950. But still, alongside everything, one can see how mathematics truly changed from the domain of gentleman amateurs to a professionals-only endeavour.

We saw earlier how useful it was to 'complete' the rational numbers, and thereby generate the real numbers. There were no holes, and limits were always well-defined. I mentioned in passing that this is what enables the machinery of calculus to flourish.

Let's return to this topic with a little extra notation in hand. When talking about limits, I spoke informally about some rational numbers being 'close together'. I meant this in the sense of the usual number line: the rational numbers $\frac{1681}{841}$ and 2 are very close together because their difference, i.e. $2 - \frac{1681}{841} = \frac{1}{841}$, is very small. Lying behind this is a notion of the distance between two rational numbers x and y, namely

$$\text{distance between } x \text{ and } y = \max(x-y, y-x) = |x-y|.$$

(This means that distance is never negative.) We call $|x|$ the *absolute value* of x; we either have $|x| = x$ or $|x| = -x$, whichever is positive.

There are two particular properties of the absolute value which I would like to draw out. Firstly, it satisfies something we call the 'triangle inequality', which is the inequality

$$|x+y| \leqslant |x| + |y|$$

for all x and y. Geometrically, this says that the sum of lengths of two of the sides of a triangle is at least the length of the third side. That makes intuitive geometric sense – if you lived in a featureless desert, you would rather go home in a straight line than go via another location first, as the direct route would be shorter. Secondly, there is the algebra property

$$|xy| = |x||y|.$$

This means that the absolute value respects the multiplicative structure of the rational numbers.

However, the absolute value is not the only way of defining the distance between two rational numbers. Here's another way. Start by picking your favourite prime number p. For ease of exposition we'll pick a smaller prime number than 37 this time around – let's choose $p = 7$ instead. Let $x = \frac{a}{b}$ be a rational number, written in 'lowest terms' in the sense that a and b share no common factors. Informally speaking, we say that x is small if a large power of 7 divides a; we say that x is large is a large power of 7 divides b. So, a rational number like $\frac{686}{5}$ is very *small*, since $686 = 2 \times 343 = 2 \times 7^3$ meaning that a large power of 7 divides the numerator. Conversely, $\frac{5}{686}$ is very large.

We can be more precise. We define the *p-adic valuation v_p* by saying that $v_p(a)$ is the largest integer such that $p^{v_p(a)}$ divides a. So $v_7(686) = 3$, and $v_7(5) = 0$. But $v_5(5) = 1$, say – the valuation depends on the prime you pick! We then define $v_p(\frac{a}{b})$ to be $v_p(a) - v_p(b)$. Finally, the p-adic distance between two rational numbers $\frac{a}{b}$ and $\frac{c}{b}$ is given as

$$\left|\frac{a}{b} - \frac{c}{d}\right|_p := p^{-v_p(\frac{a}{b} - \frac{c}{d})}.$$

So $|\frac{a}{b}|_p$ is very small if the p-adic valuation of $\frac{a}{b}$ is large, i.e. if many powers of p divide a.

We call $|\cdot|_p$ the p-adic distance. This terminology is supposed to be in reference to the classical word dyadic, meaning in relation to powers of 2.

All this might seem like, at best, a baroque and irrelevant obfuscation. This is not so. Firstly, the p-adic metric has some convenient properties: it also satisfies the 'triangle inequality'

$$|x+y|_p \leqslant |x|_p + |y|_p,$$

as well as the algebra property

$$|xy|_p = |x|_p|y|_p.$$

Broadly speaking, these properties mean we can consider number-theoretic arguments as if the were geometry! In fact, there are classical arguments from calculus which, when run with the absolute value $|\cdot|$, give a geometric consequence, but when run with the p-adic distance $|\cdot|_p$ have important number-theoretic consequences.

The standard example is the so-called Newton–Raphson method for finding where a graph of a function f crosses the x-axis. It gives an iterative procedure for approximating the intersection, where the terms of the iteration are calculated using the derivative f' of the function, namely

$$x_n := x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}.$$

This is taught to A-Level students. Yet, what isn't taught (even to many university mathematics students) is that if you run exactly the same argument with the p-adic metric, you end up showing a fundamental result known as Hensel's Lemma (after Hensel, who introduced p-adics at the end of the 19^{th} century). You show that, if a polynomial equation has a solution modulo p then (under mild technical conditions) it also has a solution modulo p^2, modulo p^3, in fact, modulo any power p^n. So you get more solutions 'for free'.

To close, we should talk about 'completeness' again. We saw how the rationals $\mathbb{Q}$ were incomplete with respect to the absolute value, and then we completed them to form the real numbers $\mathbb{R}$. Well, you can complete $\mathbb{Q}$ with respect to the p-adic metrics too, creating a complete field $\mathbb{Q}_p$ known as the p-adic numbers. The p-adics are what you get by working modulo p, modulo p^2, modulo p^3 etc. all at the same time. In the same way that you have decimal expansion of real numbers, which writes a real number x as

$$x = a_0.a_1a_2a_3a_3\cdots = a_0 + 10^{-1}a_1 + 10^{-2}a_2 + 10^{-3}a_3 + 10^{-4}a_4 + \cdots$$

the p-adic numbers have a p-adic expansion

$$y = b_0 + pb_1 + p^2b_2 + p^3b_3 + p^4b_4 + \cdots .$$

If this looks strange, remember, powers of p are *small* in the p-adic metric.

As an example, the 5-adic expansion of $\frac{2}{3}$ is

$$\frac{2}{3} = 4 + 1\cdot 5 + 3\cdot 5^2 + 1\cdot 5^3 + 3\cdot 5^4 + 1\cdot 5^5 + \cdots$$

where the coefficients alternated between 1 and 3 for ever.

In passing, I want to mention another connection to geometry. In geometry, a central notion is the 'vanishing order' of a function at a given point. Informally speaking, this is a measure of how flat a curved surface is: the flatter the surface, the greater the vanishing order (which can be any integer). Now, the theory used to classify functions according to vanishing order is the same as the theory used to understand the p-adic numbers! [The buzzword is 'local rings'.]

Idèles and Adèles

To finish, I want to tell you about one of the most brilliant PhD theses ever written. The American John Tate was the author, graduating from Princeton in 1950, and he went on to have an illustrious career at Harvard, culminating the award of the Abel prize in 2010.

The idea is as follows. Rather than considering the absolute value and the p-adic valuations separately, it might be profitable to consider them all simultaneously. This idea in itself was not new when Tate came to the problem. Hasse (active from the 1920s onwards) had thought this way; in fact, a certain 'local-to-global' principle for diophantine equations is still called the Hasse principle, in his name. It was also known (Ostrowski's theorem) that the only metrics which satisfied the triangle inequality and the algebra property were the absolute value and the p-adic metrics. However, what Tate did was to demonstrate the full geometric implications of this perspective.

French mathematician Chevalley (1936) introduced objects that he called ‘ideal elements’ – the name eventually was shortened to *idèles.* These are sequences of numbers

$$(a_\infty; a_2, a_3, a_5, a_7, a_{11}, \dots, a_p, \dots)$$

where a_∞ is a non-zero real number and, for each p, a_p is a non-zero p-adic number. There is one extra condition, namely that all but finitely many of the a_p must satisfy $|a_p|_p = 1$. Two such sequences can be multiplied together: if I have another such sequence

$$(b_\infty; b_2, b_3, b_5, b_7, b_{11}, \dots, b_p, \dots)$$

then the product of the two sequences is

$$(a_\infty b_\infty; a_2b_2, a_3b_3, a_5b_5, a_7b_7, a_{11}b_{11}, \dots, a_pb_p, \dots),$$

i.e. we do the product pointwise. However, you can’t necessarily add two idèles together, because you might end up with elements of your sequence that are 0 (and this is forbidden).

Precursors of Tate (in particular André Weil) expanded Chevalley’s definition to define the *adèles*, the name coming from a contraction of ‘additive idèles. These were sequences

$$(a_\infty; a_2, a_3, a_5, a_7, a_{11}, \dots, a_p, \dots)$$

where a_∞ is a real number (possibly zero), and for each p a_p is a p-adic number (possibly zero). The extra condition this time around is that all but finitely many of the a_p must satisfy $|a_p|_p \leqslant 1$.

Tate’s contribution was to give the adèles (which we will denote $\mathbb{A}_\mathbb{Q}$ to emphasise the fact that we’ve started with the rational numbers underneath everything) something called a *topology.* Now, I certainly don’t have time to fully explain what a topology is. In its simplest incarnation, a topology on a set X is an abstract way of ‘doing geometry’ on the set X. It tells you which points are close together, and which ones are far apart (although topologies don’t always come with an explicit notion of distance). Tate used something called the ‘restricted product topology’, building from the geometries of the p-adic numbers discussed above.

Almost all of the central theorems of 19^{th} century algebraic number theory can be converted into topological statements about the adèles and idèles. Something called the ‘finiteness of the class number’ (which is basically a measure of how badly unique prime factorisation can fail in more exotic settings); this translates to the geometric statement that a certain quotient of the idèles is ‘compact’ (i.e. informally speaking it ‘looks like a beach ball, with a hard edge’ rather than ‘a cloud with a fuzzy edge’). I still find this perspective extraordinarily beautiful – number theory and geometry becoming essentially the same endeavour.

You can add adèles, and there is a notion of geometry. When phrased precisely, this makes the adèles into something called a Locally Compact Abelian Group (LCAG). And on any LCAG, you can perform harmonic analysis (which is essentially an abstract form of breaking down sound into its constituent pure waves). Tate did this, and found that properties of harmonic analysis on this space elucidated other number-theoretic results, both old and new. For example, the proof of the functional equation for Hecke L-functions arises from the Poisson summation formula (a standard formula for any harmonic analysis set-up). The beginning of the Langland’s programme lies here.

To explain these points would need a whole other set of notes. It’s probably best to stop here.

Real and Complex Numbers

Starting with Real numbers: The “red thread” through this would be the question “What is a number?”

- We’ve all come across them, from learning to count in school (1,2,3, .. the natural numbers) to the change at the supermarket (12.67$ … rational numbers, finite decimals) to pi and e (irrational numbers … infinitely long decimals)
- Something interesting to cover here (although philosophy is not generally my forte) is the question “What actually is a number?”
- Let’s touch on 3 interesting thoughts on this: formalism, intuitionism, and logicism, which came about in the early 20th century

Part 1: formalism. The idea that all these numbers and equations and symbols have no actual meaning. They’re more of a game - we have certain strings of syntax and we get to rearrange and mess with them according to established rules.

Part 2: intuitionism. The idea that numbers are purely a mental exercise, so are constructed by people in their minds, and these numbers don’t reveal properties of the physical world around us but are just used by the human mind to analyse more complex mental constructs. So in essence it’s the idea that numbers only exist when there are humans around to think them up. Billions of years ago, numbers didn’t exist - there was no concept of “There are 3 trees in a valley”, “ 30 days since since the extinction”.

Part 3: logicism. It’s roughly the idea that maths can be reduced to logic. After being initialised by Frege it found some of its biggest proponents in Russell and Dedekind after Dedekind concluded that natural numbers were reducible to sets and mappings so could be “reduced” to logic.

QUESTION: So none of these ideas posit that numbers exist in their own right in the physical world - is that right?/ are there theories that do say numbers are a concept independent of the human mind?

Part 4: Yes - to mention a last thought, platonism does say that numbers always exist in their own right.
Numbers are real non-physical things. What do we mean by this? We say that numbers are “real” meaning they exist outside of our minds, so independent of a human being around to count a certain quantity or talk about it or think about it. But of course they are non physical, i.e. we can’t bump into the number 24 in the street, I can’t pick up the number 5. In a sense you can think about this as using the logic ‘Something has the potential to exist, and therefore it does exist.’ So millions of years ago there were 10 dinosaurs on a hill, and even though no one was around to count them, if someone were there, they could have been counted, there would have been exactly 10, so the number 10 has always existed, because the potential for someone to describe that quantity with the number 10 existed.

(I like this idea because it helped me at the beginning accept that we can define things into existence in maths. At the start of an undergraduate degree, or even later in this podcast when we talk about complex numbers, we say ‘Let there exist a number such that.. Xyz’ or ‘let there be a group with these properties'. Accepting that the potential for something to exist means that it does helped me say ok, let's go with this, I can imagine such a number so it's real, we're going to use it.

QUESTION: what are the different kinds of real numbers?

(Aled has some very nice history on these, I don't think I can compete with that detail here)
Naturals (counting numbers)
Integers (...,-3,-2,-1,0,1,2,...)
Rational numbers (numbers that can be written as a fraction)
Irrationals (eg pi or e) Interestingly enough there is a theory that a Greek philosopher named Hippasus was killed by Pythagoras' followers after he discovered irrational numbers - a belief that defied Pythagoreans' belief that every number could be expressed as the ratio between two numbers.
Transcendental numbers - which can never be a solution to a polynomial with integer coefficients.

QUESTION: How do you take the next step to complex numbers?

Here I would talk about the question "Are complex numbers real?" covering the following points:

- Thinking about it initially the number makes no sense - the square root of -1 cannot exist according to the "real" maths we meet at school. This is why they are also called "imaginary numbers" - the mathematician Descartes called them imaginary because he believed they were "wrong" in some way and didn't exist.
- BUT now we have an issue)to just name one here). There's a theorem called the fundamental theorem of algebra proven by Gauss in 1799, that states that any non-constant polynomial with complex coefficients has a root in the complex numbers. Sounds complicated but basically take it to mean, if we have a polynomial, so an equation like x^2+1=0 or x^3+3x=7, anything like that, then there will exist at least one solution for x. But hang on, x^2= -1 doesn`t have a solution in the reals… the solution would have to be the square root of -1. Every root of a polynomial can be expressed as some real plus some complex component. So what is this square root of minus one and is it less "valid" or "real" than integers, rationals, irrationals, reals in general?
- Well, we use polynomials and other equations *all the time* in maths, in physics, in engineering. They're extremely useful, and there are even situations where by using complex numbers you can then find some real-number solutions.
- One of the most helpful formulas for this is Euler's formula - which says e^(ix)=cos(x)+i sin(x). That sounds a little bizarre… exponentiating i times a number gives cos of that number plus i times sin of that number - Maybe the special case e^(i pi)= -1 is more familiar. We use this equation to find the sums of sines and cosines which are out of phase for example. Complex numbers are used routinely all across applied mathematics, in fluid dynamics, in electrical engineering, in quantum mechanics (and also in pure maths for that matter).
- Imagine you have a pendulum. You can model its swinging motion using equations that describe the forces acting on the pendulum. We want to find a solution to these equations, so an equation that says x (distance from equil. point) = some function of time. That's a solution - then we have a way of seeing how this pendulum is moving. Now we have a few

options for what happens. The first possibility is that the system is over damped, meaning that the system tends to the equilibrium position into infinity, not crossing it again. Imagine a very very unoiled hinge at this pendulum, so you lift the pendulum to the side, drop it, and it slowly slowly returns to the equilibrium position of hanging straight down. In this case the solution to this equation we're using to describe the motion is real. But (as is most often the case) we can have under-damped motion - the pendulum will swing back and forth until it eventually comes to rest at the equilibrium position. In this case if we solve these equations we will get *complex solutions to our equations.* So complex numbers arise in everyday situations to describe motion - they appear in the solutions to simple equations of motion and can be reformulated to yield real results that tell us something about the physical world.

- Now why does this mean that they are "real"? Why does this give them any more weight? Let's make a comparison here to negative numbers. Imagine you're living in ancient Greece, where numbers are considered in terms of "units", so the number 5 is five times the unit length, 7 is seven times the unit length, 1 isn't a number, it's just the unit used to define other quantities. Everything is considered geometrically. In this context negative numbers make no sense. In fact the Greeks really didn't accept negative numbers, much like they didn't accept irrational numbers. This held for centuries, and Descartes, who lived in the 17th century, even rejected negative roots of equations as "false", since they represented numbers less than nothing.
- And if you think about it from the perspective that maths is useful for counting your cattle, how many bags of grain you harvested, how many fingers you have - then negative numbers intuitively don't make sense. What is less than nothing? Two less than nothing?
- But we learn about negative numbers so early in school, and they're useful all over the place in equations to describe motion, when looking at vectors and negative distances mean "in the opposite direction to initial motion", when looking at a situation where someone loses something - that we have accepted negative numbers as real because they appear in equations we use to describe nature all the time. In a similar way, complex numbers appear in natural equations and their solutions, and are immensely helpful to us. You can then try to take the same "jump" as with negative numbers and say "since they appear in maths everywhere, they must be in some sense real".

Number systems - Álvaro

What is a number?

A **number** is a mathematical object used to count, measure and label. Though if we are honest, this is the kind of definition that does not tell us much by itself. Numbers are better understood by examples, we could all think of many examples of numbers, some simple such as 5 or 13; some trickier such as *pi* or the number *e*.

Can you give us a short summary of how numbers came to be what they are right now?

The first instances of use of numbers have been found as tally marks in bones and stones about 44 thousand years ago. It was not until the year 3400 BC that a place value number system was developed in Mesopotamia. This was similar to the decimal system we currently use, but instead of being of base 10, it had base 60. The numbers we use to count (1,2,3 and so on) are what we mathematicians call **natural numbers**.

But… That's not all, isn't it?

No! The first non-natural numbers that humans considered were the **negative numbers** (-1, -2…) which were first used in China around the year 100 BC. The combination of natural numbers, negative numbers and zero is the set of numbers known as the **integers**.

Then, there is the **rational numbers**, which are what we commonly know as fractions. For a long time, people thought that this was all, that every number could be expressed as a fraction. But a disciple of Pythagoras, the mathematician famously known by his theorem for triangles, proved that the square root of 2 could not possibly be rational, and so, the irrational numbers - numbers that could not be expressed as fractions- where discovered.

The union of both rational and irrational are what mathematicians call **real numbers** simply because… well… they are real. These numbers are the ones that you will find when you do pretty much any computation describing things like your height, the probability that it will rain tomorrow or the age of the Queen of England.

If real numbers are all that we need, why do mathematicians need to think about other number systems?

Well, let’s go back to high school for a second. If we had any positive number x, we could press the square root button in our calculator to find a number that when we multiplied it by itself would give us x again. However, when we start with a negative number and try to find it its square root, we find out that it is not possible, as the product of either two positive numbers or two negative numbers is always positive.

Mathematicians have realised that, while real numbers are the numbers that “made sense”, we could imagine other kinds of number system that helped us solve problems, even if these new systems were as fictitious as fairy tales. Going back to the problem of the square root, someone thought what if we call the square root of -1 i (like the letter after h) and try to develop this logically? This is the story of how **imaginary numbers** were invented, which, together with the real numbers, conform the **complex numbers**.

In what sort of problems are complex numbers useful?

Complex numbers give us a very convenient way to express and work with the trigonometric functions: the sine, the cosine, the tangent… These are useful to describe signals and rotations and therefore complex numbers are used in signal processing, electrical engineering, and fluid dynamics.

Furthermore, from a more abstract point of view, complex numbers are fundamental in areas of pure mathematics such as geometry, number theory or dynamical systems.

You mentioned that complex numbers had applications in geometry, but this does not seem to be a field where numbers would play a role. How does it work?

When we generally think about geometry, we think about it in a visual way, we think about circles, triangles, quadrilaterals… and we sometimes forget that behind the formulas that give their areas and perimeters, there are mathematical quantities that can be studied.

For instance, the Greeks were particularly interested in how to construct polygons by only using a compass and a straight edge. This, at first, had nothing to do with numbers, but when mathematicians were later interested in studying it in a more rigorous way, it contributed to the development of what is known as **constructible numbers**.

Let’s suppose we are given a line segment of unit length (for example, of 1 inch). We say that a number n is **constructible** if we can draw a segment of n inches by only using a compass and a ruler without markings.

Everyone can take a piece of paper and have a go. It is easy to check that all natural numbers are constructible as we can easily use the compass to duplicate the length of the segment. Similarly, by constructing parallel lines, we can always divide the original segment in as many parts as we want to, so all rational numbers are constructible.

But also, there are irrational numbers that can be constructed with a ruler and a compass. With them we can draw a right-angled triangle whose legs (the shorter sides) have unit length, for example, and then the hypothenuse would as length the square root of 2, which is an irrational number.

The possibilities seem endless, with ruler and compass we can draw things such as equilateral triangles, regular pentagons and hexagons…

So, one may think that all real numbers are constructible…

It would be beautiful if it were true, but unfortunately it is not. There are real numbers that are not constructible. And even though the Greeks did not know that, they were the first ones that had the suspicion that some things simply could not be constructed with ruler and compass.

Historically there were three problems that they did not know how to solve using only a ruler and compass:

- The quadrature of the circle, that is, given a circle, to find a square with the same area.
- Given a cube, to draw a cube with the double of volume.
- Given any angle, to divide it into three equal angles.

The Greeks thought that they had not discovered the answer to these problems because they were too difficult, but in the eighteenth century, it was proven that these problems were, in fact, impossible to solve.

The way this was proven was by translating these problems to the language of constructible numbers: if these problems were solvable, it would imply that the quantities square root of pi, cubic root of 2 and the cosine of pi over 9 are all constructible. The development of an area of mathematics currently known as **Galois theory** helped to prove that none of those numbers were constructible and thus, these ancient problems could not be proven.

So do we now know everything that can be constructed with a ruler and a compass?

We are far from knowing everything, but at least with Galois theory, we now have an abstract framework that allow us to tackle this kind of problems.

For example, the Greeks knew how to construct regular polygons with 3, 4, 5, 6 and 8 sides with ruler and compass, but they did not know how to construct a regular heptagon. Now we know that this is not possible, and we have discovered a connection between the number of sides of a polygon and whether they are constructible or not. This has to do with a very cool family of numbers called **Fermat primes**, which are the primes of the form 2 to the 2 to the n plus one where n is a natural number.

It is funny to think that we have been able to show how to draw a polygon with 65537 sides, with a ruler and a compass, yet it is impossible to draw one with only 9 sides.

It shows how mathematics defies intuition on which things are possible and which ones are not.

How does someone invent a system of numbers?

Well, in mathematics inventions are motivated by its usefulness; by how they are helpful to solve problems. Anyone can think of a curious family of numbers, for example, all numbers whose decimal expansion does not contain a 3.

But in general, we ask system of numbers to verify some properties: that we can define operations with them such as the sum or the product; that the sum of two numbers in the system is another number in the system, that they have a zero… things of that sort that makes them similar to the systems of numbers that we already know.

How could someone learn more about number systems?

One option is, obviously, pursuing studies in mathematics! But for those who are not willing to make such sacrifice, and want to learn about number systems for the fun of it, there are many books for non-specialists. Even though is a little old, I would recommend *The book of numbers*, by John Conway and Richard Guy; it has a lot of drawings in it that makes some of these number systems much easier to understand.