Mathematical Musings

Everything you ever wanted to know about
Mathematics but were too embarrassed to ask

~ ~ ~ ~ ~ ~ ~ ~ ~ ~

Oliver Linton

Most of the illustrations in this book were generated using programs written by the author. Many of these can be found on the author's website www.jolinton.co.uk.

Contents

Numbers

When I was taught mathematics at school I was told lots of things which I just had to take on trust. But every time I asked my teacher to explain why this was true or why I had to do that she invariably replied 'well, that's just the way it is.' Perhaps after all these years, you can help me. For example, here is a simple question: Some numbers like 12 and 91 can be expressed as the product of two (or more) smaller numbers but others like 13 and 93 cannot. Why is this?

Well I could just say that 'that's the way it is.' but I suspect I won't get many thanks. In fact, your question deserves a better answer than that. What you are really asking is this: given a certain number, is there any way of *predicting* whether it will be prime or composite? I may as well confess at the outset that in spite of 2000 years of trying, we still don't know the answer – but we do know a lot about the properties and distribution of primes so lets go back to the beginning.

Do you remember being given a tray of 12 bricks when you were very young, neatly packed in a box of 4×3? Perhaps your parents encouraged you to find some other ways of ordering them. You will, no doubt, have discovered several: you could place them in a long line of 12 bricks; or in two rows of 6. You may even have discovered that the 12 bricks could be stacked in a box shape with width 2, length 3 and height 2. All good mathematical training.

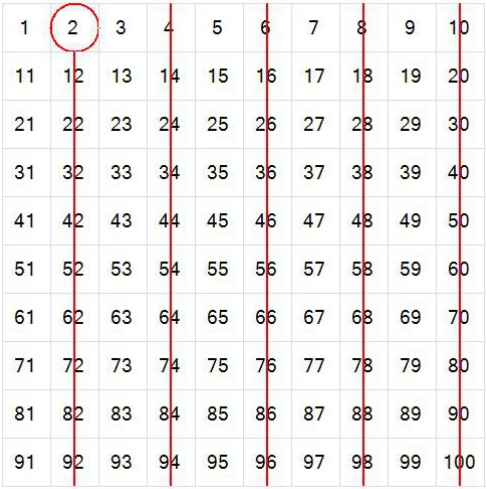
Perhaps, one day, one of the bricks went missing. Try as you might, the remaining 11 brick stubbornly refused to be placed in any kind of regular arrangement except, of course, one long line. In time you came to accept that certain numbers were fundamentally different from the majority. These numbers are the *prime* numbers and they have no factors (other than themselves and 1).

Is there an easy way of finding prime numbers?

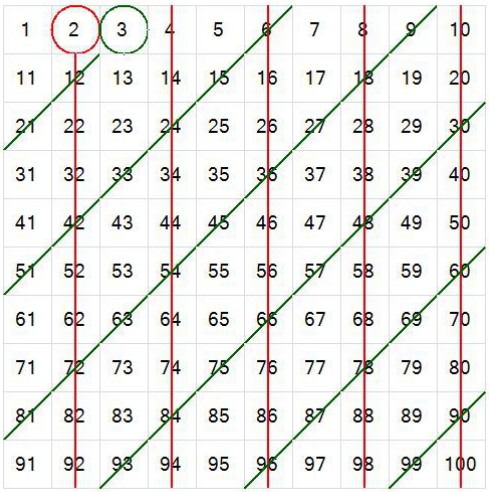
Not really. One way of discovering new prime numbers is by using

Eratosthenes' Sieve

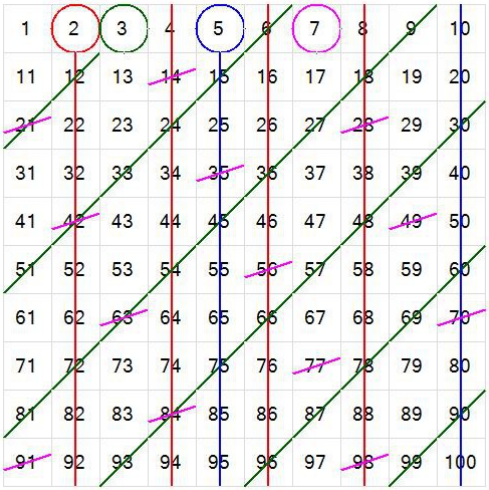
(Eratosthenes, by the way, was a Greek mathematician born in Cyrene in about 276 BC). First we write out all the numbers from 1 to 100, say, in a rectangular grid; then we strike out all those numbers which are divisible by 2 – except 2 itself, of course, because this is the first prime number. (Of course, if we started by striking out all the numbers which are divisible by 1 we would strike out all of them! This is why 1 is not considered to be a prime number.)



The next number which has not been struck out is 3 so we adopt this as the next prime number and strike out all the numbers divisible by 3.

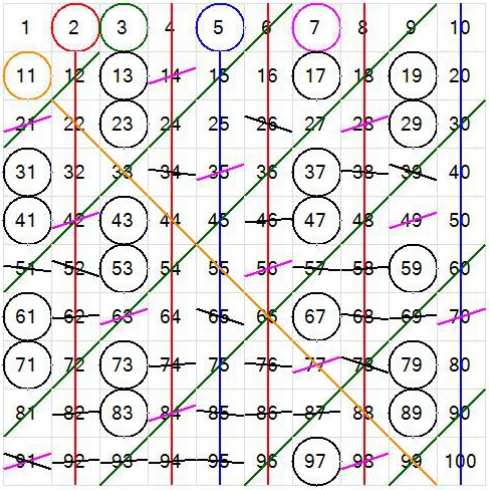


Since 4 has already been removed, the next prime to strike out is 5. And since 7 is not divisible by 5 we might as well strike out all those numbers as well. This gives us:



Now you might think that we would have to continue this process striking out all those numbers which are divisible by 11, 13, 17 etc. but

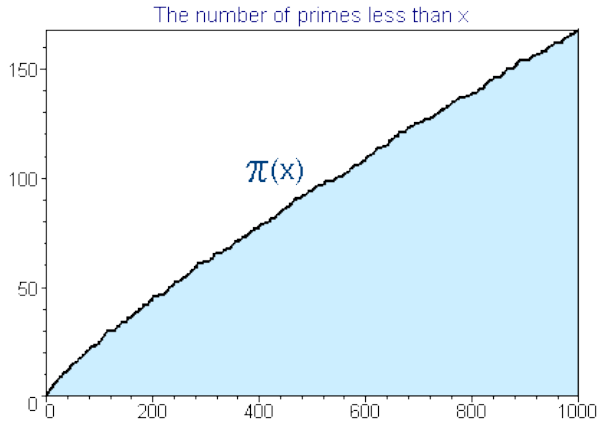
in fact we do not really have to do this because all these numbers are greater than 10 – the square root of 100; and any number which is divisible by one of these numbers will also be divisible by one of the smaller primes which we have already struck out. It is useful, however, to highlight all the remaining numbers which have not yet been struck out. These are the prime numbers less than 100. There are 25 of them.



Armed with these numbers, we can test any number up to 10,000. If we do this we find that there are 1,229 of them.

Hang on a minute – I have thought of another really interesting question: Is there a largest prime? After all, the bigger the numbers get the more potential primes there are which could divide it so, perhaps, there comes a point when every number is divisible by some prime or other.

The standard way (indeed the *only* way we know of) of testing whether or not a given number *N* is prime is by trying to divide it by all the prime numbers less than the square root of *N*. Computers are good at this sort of thing and tables of prime numbers up to a million million million million (10^{24}) and more have been published and there seems to be no end to them. A graph of the number of primes below any given number (up to 1000) looks like this:



https://primes.utm.edu/howmany.html

Well it certainly looks as if the number of primes goes on increasing for ever but how can you prove something like that – after all, you can't test every number!

Amazingly it was that mathematical genius Euclid who proved that the number of primes is infinite some 300 years before the birth of Christ by the following simple but elegant argument:

Euclid's proof of the infinity of primes

Suppose that the list of primes is finite $p_1, p_2, \ldots p_n$. Consider the number $N = p_1 \times p_2 \times \ \ldots \ \times p_n$. This huge number is divisible by all the primes. Now add 1 to this number. Obviously this number cannot be divisible by any of the primes we have listed because it will leave a remainder of 1 behind.

We are now faced with two possibilities. Either this number is itself prime (in which case our original list was not complete and the supposition is false) or it is composite, in which case it must be the product of two or more numbers, all of which are larger than p_n. (Any number smaller than p_n is either one of the listed primes or divisible by one of them and we know that $N + 1$ is not divisible by any of them. All these factors of $N + 1$ must therefore be new primes, larger then p_n.

This is one of the most remarkable proofs in all of mathematics. It is

so simple that any one can understand it and yet its implication is almost impossible to comprehend. What the proof is saying is that there exists at least one prime (in fact an infinite number of them) which is bigger than any number I care to write down! It seems incredible that I can say with 100% confidence a fact about numbers so huge they are way way way beyond our imaginings, and yet we can. Of course, while the proof tells us that such huge primes exist – it doesn't give us a single clue as to how to find them!

That's really amazing. I never realised that you could prove facts about numbers you can't even write down!

There is another thing about prime numbers which has always bugged me: from a very early age 1 was told that all composite numbers can be factorized into their prime components and we were frequently given numbers like 60 to factorize as homework. But I often wondered why there was only one possible answer. Are there any numbers which can be factorised into prime components in two different ways?

God, your teachers must have hated you! But you are absolutely right to ask the question. The prime factors of 60 are 2, 2, 3 and 5. There is apparently no other way of multiplying some other set of (prime) numbers to make 60. 60 is *made up of* 2's, 3's and 5's; these are the *prime components* of 60. To make a cake you need flour, butter and sugar. You can't make a cake out of chalk, cheese and salt. Numbers, we are told, are just the same. To make 60 you need two 2's, a 3 and a 5. Nothing else will do.

This idea is sometimes called the Fundamental Theorem of Arithmetic and it is so ingrained into our psyche that few people ever question it or ask for a proof. And yet it is a truly remarkable statement. How do we *know* for sure that there isn't some huge number which can actually be factorized into prime components in two different ways? Well, it turns out that the proof is not trivial. Not that it is difficult either, but as soon as the question is asked, you *have* to find a proof. The alternative is mathematical suicide.

That sounds painful!

S0, assuming that you do not want to commit mathematical suicide, here is a fairly simple proof.

The Fundamental Theorem of Arithmetic

Everything hinges on the fact that if the product of two numbers a and b is divisible by a prime number p, then either a or b (or both) must be divisible by p. For example: 60 is divisible by 5. Now $60 = 4 \times 15$ so what we are saying is that either 4 or 15 must be divisible by 5. (Note that the number we are trying to divide by must be prime. 60 is also divisible by 6 but neither 4 nor 15 are divisible by 6.)

OK – so we know what we are trying to prove. But how to go about it? Let us start by assuming that a is not divisible by p. We shall attempt to prove that if this is the case, the b must be divisible by p. (Of course if a is divisible by p then the theorem is proved already.) Let us suppose by way of a concrete example that $a = 24$, $b = 15$, $a \times b = 360$ and $p = 5$.

Now since p is prime, it is only divisible by itself or 1. It follows that the only factors which any number can have in common with p are p or 1. But since a is supposedly not divisible by p, the highest common factor of a and p is 1. (In our example, $a = 24$ – a number not divisible by 5 – and $p = 5$, the highest number which divides both numbers must be 1.)

Now what about the highest common factor of $a \times c$ and $p \times c$ (where c is any integer)? Can you see that the highest common factor of these two numbers must be c? (For example, the highest common factor of 144 ($= 24 \times 6$) and 30 ($= 5 \times 6$) must be 6 because we have already established that 24 and 5 have no common factors other than 1.)

This step deserves closer scrutiny as it is crucial to the argument. In particular, it is only true if p is prime. Suppose that $p = 15$. a (which = 24) is still not divisible by p but the highest common factor of 144 ($= 24 \times 6$) and 90 ($= 15 \times 6$) is not 6 but 18. This is because, while 15 does not *divide into* 24, it does *share a factor* with 24 namely 3.

What is true of any number c is also true of the number b so what we have proved is that the highest common factor of $a \times b$ and $p \times b$ is b. (In our example, the highest common factor of 360 and 75 is 5)

Here comes the crunch line. We know that the product of a and b is divisible by p – i.e. $a \times b = p \times n$ where n is some integer. Substituting this into the previous statement we deduce that the highest common factor of $p \times n$ and $p \times b$ is b. (i.e. the HCF of 360 = 5 × 72 and 75 = 5 × 15 is 15) Now since both $p \times n$ and $p \times b$ are divisible by p, it follows that the highest common factor of n and b is b/p. (i.e. since 360 and 75 are both divisible by 5 the HCF of 72 and 15 must be 15/5 = 3) But since a highest common factor of two integers must be an integer, b must be divisible by p. Q.E.D.

That's the difficult bit over. To complete the proof of the Fundamental Theorem of Arithmetic, suppose that there exists a number N_1 which has two different factorizations i.e.:

$$N_1 = p_1 \times p_2 \times p_3 \times ... = q \times [\text{some other primes}]...$$

where $p_1, p_2, p_3 \ldots$ and q are all prime with q not equal to any of the p primes.

Now by the theorem we have just proved, since
$N_1 = p_1 \times \left[p_2 \times p_3 \times ...\right]$ and p_1 is not divisible by q, then
$N_2 = p_2 \times p_3 \times ...$ must be divisible by q.

But we can repeat this process until we have divided out all the p primes leaving nothing left! This is a contradiction so the number N_1 cannot exist.

Well, I am not sure I followed every step there but it is nice to know that the proof exists and that my teachers were not telling porkies!

Yes. I have tried to explain the proof of this important theorem in as non-mathematical way as possible but even so, it is difficult, even for a mathematician, to grasp the whole proof at once – which is why the theorem was never proved to you when you were at school. It is probable, even, that your teacher would have been unable to supply you with a proof had you the wit to ask for it.

If that is the case, I wonder just how how many other mathematical statements there are which we have imbibed from birth, as it were, without ever being given the opportunity of

asking for a proof? What about proving that 2 + 2 = 4?

Well I don't think there is much difficulty about that one. Take two stars **. Cut and paste them alongside the original and what do you get? Four stars ****.

OK but what about this one: Why does minus times a minus equal a plus?

This is a bit more difficult because this is not so much a fact as part of the definition of what we mean by a negative number. One of the things we want negative number to do is to obey what is called the distributive law – the law that says $a(b + c) = ab + ac$. Now consider the expression $-a(b + -b)$ where a and b are positive numbers. This must be equal to zero because any number multiplied by zero is zero. Applying the distributive law we get $–ab + [– –ab] = 0$. This only makes sense if $[– –ab] = +ab$.)

I get that – but what is this thing called the 'Distributive Law'. Surely you have to prove that as well.

Right. Lets go back to the box of 12 bricks. You could arrange them like this:

■■■ ■■■
■■■ ■■■

There are two ways of interpreting this arrangement: either as 2 rows of 6 bricks – i.e. $2(3 + 3)$; or 2 rows of 3 bricks plus another 2 rows of 3 bricks – i.e. $2\times3 + 2\times3$.

Of course, this is not really a proof, just a demonstration. Mathematics is founded upon logic and the Distributive law in algebra follows directly from logical axioms such as 'if A and either B or C then either A and B or A and C'.[1]

Well, I can't really argue with that!

1 If you want to know a bit more about the properties of numbers, go to http://www.jolinton.co.uk/Mathematics/Some_Properties_of_Numbers/Text.pdf

Squares and roots

Here's something else I have always puzzled about. What is the exact value of √2?

The square root of 2 is approximately 1.414 but 1.414 × 1.414 = 1.999396 and 1.415 × 1.415 = 2.002225 so obviously √2 lies somewhere between these two numbers. Now it turns out that however many decimal digits you calculate, you never get the exact value of √2. This is because √2 is what is called an *irrational* number. Its decimal expansion is infinitely long. Here are the first 100 digits:

√2 = 1.4142135623 7309504880 1688724209 6980785696 7187537694 8073176679 7379907324 7846210703 8850387534 3276415727 ….

Another way of saying this is that no fraction of the form a/b where a and b are integers will be exactly equal to √2. Could you prove this? The Pythagorean Greeks could – and the proof that they devised is a model of elegance and ingenuity.

The irrationality of 2

Suppose that $\sqrt{2} = a/b$ where a and b are two integers which have no common factors. Applying a little simple algebra we get:

$$(\sqrt{2})^2 = 2 = \frac{a^2}{b^2}$$

$$2b^2 = a^2$$

What this means is that a must be an even number. If this is the case we can write $a = 2c$. This means that

$$(2c)^2 = 4c^2 = 2b^2$$

$$2c^2 = b^2$$

But this means that b must be even too! This contradicts the supposition that a and b had no common factors so √2 cannot be a simple fraction.

Now why didn't I think of that!

In fact the n^{th} root of any integer is either an integer or it is irrational.

That reminds me: why is √2 sometimes written as $2^{½}$?

I am sure you know that $2^2 = 2 \times 2$ and that $2^3 = 2 \times 2 \times 2$ etc. It follows that $2^2 \times 2^3 = 2^5$. In general $2^a \times 2^b = 2^{a+b}$ and we will assume that this holds for all values of a and b both negative and fractional.

On this basis, $2^{½} \times 2^{½}$ must be equal to 2. It follows that $2^{½}$ must be equal to the square root of 2.

Lets take this a bit further:

The meaning of x^n

First we must establish that $x^0 = 1$. Here is the sequence of powers of 2

$$2^1 = 2 \quad 2^2 = 4 \quad 2^3 = 8 \quad 2^4 = 16 \quad 2^5 = 32 \quad \ldots$$

Adding 1 to the power means multiplying the previous term by 2. Similarly. Moving from right to left, subtracting 1 from the power means dividing each term by 2. This observation allows us to extend the meaning of 2^n to zero and the negative integers:

$$\ldots \quad 2^3 = 8 \quad 2^2 = 4 \quad 2^1 = 2 \quad 2^0 = 1 \quad 2^{-1} = \frac{1}{2} \quad 2^{-2} = \frac{1}{4} \quad 2^{-3} = \frac{1}{8} \quad \ldots$$

This pretty well establishes that $x^0 = 1$ and that $x^{-n} = \dfrac{1}{x^n}$

But what if n is fractional (or even irrational)? Again, we are not so much looking for a proof here as a definition of x^n which will enable us to use all the usual rules which govern the behaviour of integral powers. One of these rules is the fact that $x^{a\,b} = (x^a)$ to the power of b.

Now suppose that $b = 1/a$. Now we have $x^{a/a} = x^1 = (x^a)$ to the power of $1/a$. This only makes sense if 'raising to the power of $1/a$' is the same as 'taking the a^{th} root'. i.e. $x^{1/a} = \sqrt[a]{x}$. In general therefore:

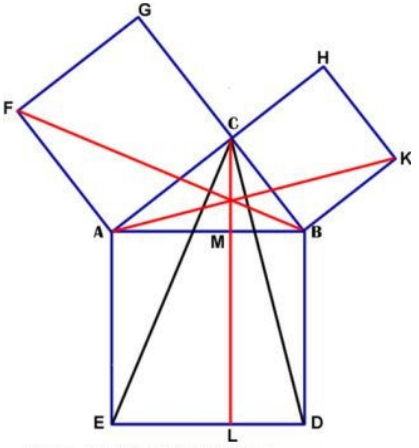
$$\boxed{x^{a/b} = \sqrt[b]{x^a}} \qquad (1)$$

Well, that makes a bit more sense now. Thanks.

Pythagoras' Theorem

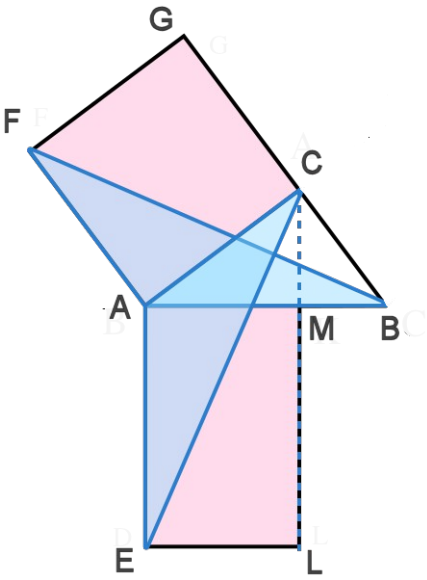
There is one proof which I do remember being shown when I was at school and that was a proof of Pythagoras' theorem. I seem to remember a diagram which looked rather like a windmill but I can't for the life of me remember the proof. What was it?

This was the diagram you were shown and it was first drawn by Euclid and appears in his famous textbook on geometry – Elements.



It looks confusing because the proof is in two separate halves. It consists in proving that the area of the square AFGC is equal to the area of the rectangle EAML and likewise that the area of the square CHKB is equal to the area of the rectangle LMBD. This implies that the sum of the areas of the two small squares is equal to the area of the large square thus proving that AC^2 n+ $CB^2 = AB^2$.

It is much easier to see how the proof works if we just concentrate on the first half of the proof and rub out all the lines we don't need like this:



The main part of the proof requires us to show that the two triangle shaded in blue are congruent – i.e. they are exactly the same size and shape. The reason for this is easy to see once we realize that the obtuse angles of the two triangles at A are the same and that the adjacent sides are equal in length.

Now we have to use another theorem (which I won't prove here unless you insist) which is that all triangles with the same base length and altitude have the same area. The upshot of this is that the area of the triangle ABF is exactly half the area of the square AFCG and that the area of the triangle AEC is exactly half the area of the rectangle EAML. All of which implies that the area of the square is equal to the area of the rectangle.

What is true of the left hand side of the diagram is true of the right hand side and the proof is complete.

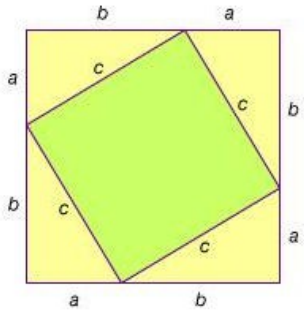
It all sounds a bit complicated to me. And you seemed to have proved the theorem without using the fact that the triangle ABC is a right angled triangle. How is that?

Good point. But if the angle ABC was not a right angle, then the line

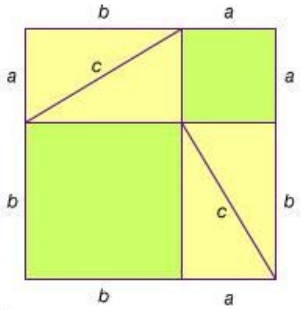
BCG would not be straight and the area of the triangle theorem would not work. The same goes for the line ACH in the original diagram.

I see. But it still sounds over complicated to me. Isn't there a simpler proof?

Well, there are probably more proofs of Pythagoras' Theorem than any other theorem in mathematics. See if you prefer this one. Take 4 copies of a right angled triangle with sides *a*, *b* and *c* and place them in a square like this:



Obviously the green square has area c^2. Now rearrange the triangles like this:



The green areas now add up to $a^2 + b^2$.[2]

Cool!

2 If you want a few more proofs of Pythagoras' Theorem go to http://www.cut-the-knot.org/pythagoras/

Pythagorean Triplets

Everyone knows that $3^2 + 4^2 = 5^2$ – and that a triangle with sides equal to 3, 4 and 5 units is a right angle triangle too. But are there any other right angle triangles with integral sides or is this the only one? Or, to put it another way, are there any other triplets of numbers such that $a^2 + b^2 = c^2$?

Well, actually there are lots of them (not counting examples like $6^2 + 8^2 = 10^2$) and Euclid devised a means of generating them all 300 years before the birth of Christ.

Euclid's formula for generating all the primitive pythagorean triplets (PPT's)

We are looking for three integers a. b and c such that

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

Since a, b and c have no common factors, either a or b or both must be odd.

Suppose that a and b are both odd (and c is even). We can therefore write:

$$(2p + 1)^2 + (2q + 1)^2 = (2r)^2$$

Where p, q and r are integers.

In which case

$$4p^2 + 4p + 1 + 4q^2 + 4q + 1 = 4r^2$$
$$2(p^2 + p + q^2 + q) = 2r^2 - 1$$

Now the left hand side of this equation is even and the right hand side is odd. This is a contradiction so our assumption that a and b are both odd is wrong.

Let us suppose that a is odd and b is even. This means that c must be odd. It also implies that $c + a$ and $c - a$ must both be even. We can therefore write

$$p = (c + a)/2$$
$$q = (c - a)/2$$

Where p and q are integers.

In which case

$$a = p - q$$
$$c = p + q$$

and

$$b = \sqrt{4pq}$$

Now if p and q have a common factor, a, b and c will share the same factor and the triplet will not be primitive. p and q must therefore be co-prime.

If that is the case, then in order for $4pq$ to have an integer square root, both p and q must be perfect squares. Suppose $p = m^2$ and $q = n^2$. Then

$$\boxed{\begin{aligned} a &= m^2 - n^2 \\ b &= 2mn \\ c &= m^2 + n^2 \end{aligned}} \qquad (2)$$

where m and n are co-prime, $m > n$ and one only is odd.

Here is a table of possible values of m and n with the corresponding Pythagorean triplet.

m	n	a	b	c
2	1	3	4	5
3	2	5	12	13
4	1	15	8	17
4	3	7	24	25
5	2	21	20	29
5	4	9	40	41
6	1	35	12	37
6	3	27	36	45
6	5	11	60	61
7	2	45	28	53
7	4	33	56	65
7	6	13	84	85
8	1	63	16	65
8	3	55	48	73
8	5	39	80	89
8	7	15	112	113

(The entry for (6, 3) has been greyed out because 6 and 3 are not co-

prime.)

Isn't here some story about a French mathematician who found a proof to do with this that wouldn't fit into a margin?

That's right. It is called **Fermat's Last Theorem** and it states that there are *no integers* such that $a^3 + b^3 = c^3$ – or indeed there are no integers such that $a^n + b^n = c^n$ where n is any integer greater than 2. In spite of intense searching, no one has ever found such a triplet. This led Fermat (among others) to claim that he had found a simple proof that such triplets do not exist. This is unlikely. The proof, when is was eventually published by Andrew Wiles in 1993 proved to be 129 pages long – far far too big to be written in the margin of any book. In all probability Fermat discovered one of the many many shorter 'proofs' which actually contain an error or an unproved assumption. Still, the search for a shorter proof goes on.

Curiously, there do exist quartets of numbers such that $a^3 + b^3 + c^3 = d^3$. Amazingly, one of these is $3^3 + 4^3 + 5^3$ which is equal to 6^3. Wouldn't it be nice if $3^4 + 4^4 + 5^4 + 6^4$ was equal to 7^4. Unfortunately, it doesn't. Many other quartets and quintets of numbers with similar properties are known but as far as I know, no general formulae exist for generating them.[3]

3 Euclid's is by no means the only formula which can generate Pythagorean triplets. Further information can be found in the following article: http://www.jolinton.co.uk/MatheEmatics/Pythagorean_Triplets/Text.pdf

Cyclic Numbers

The other day I came across the number 142857 in a book. The author said that if you keep adding this number to itself again and again you keep on getting the same digits. Why is this?

The number 142857 is quite remarkable. When added to itself repeatedly, it generates the following sequence:

1	142857
2	285714
3	428571
4	571428
5	714285
6	857142
7	**999999**

It seems astonishing that the same digits should reappear over and over again and then suddenly crystallise into a series of nines – but that behaviour gives us the clue that we need. 999999 is divisible by 7. 142857 is what is known as a cyclic number (in the decimal system at any rate).

So how exactly do cyclic numbers work and how are there any other cyclic numbers?

Consider the process of dividing unity by a prime number (which is not a factor of the number base we are working in i.e. 10).

For example, let us divide 1 by 7. The sum goes like this:

10 over 7 = 1 remainder 3 → 0.1
30 over 7 = 4 remainder 2 → 0.14
20 over 7 = 2 remainder 6 → 0.142
60 over 7 = 8 remainder 4 → 0.1428
40 over 7 = 5 remainder 5 → 0.14285
50 over 7 = 7 remainder 1 → 0.142857
10 over 7 = 1 remainder 3 → 0.1428571
30 over 7 = 4 remainder 2 → 0.1428571 4
etc. etc.

It is obvious that the remainders can only be selected from the

numbers 1 to 6[4] and that when we reach the same remainder as one we had before (3 in this case) the cycle will repeat over and over again. This means that 1/7 = 0.142857 recurring.

Now what about 2/7?

20 over 7 = 2 remainder 6
60 over 7 = 8 remainder 4
40 over 7 = 5 remainder 5
50 over 7 = 7 remainder 1
10 over 7 = 1 remainder 3
30 over 7 = 4 remainder 2
20 over 7 = 2 remainder 6
etc. etc.

Now you can see why you keep on getting the same digits. It is because the sequence of remainders must necessarily be the same each time. The only difference is in the starting point. The recurring sequence has order 6 (i.e. it is 6 digits long).

Now if you take any of these recurring decimals, multiply by a million and lop off the remainder, you get the sequence listed earlier.

All this strongly suggests that we can search for cyclic numbers by dividing 1 by various prime numbers *p* and looking at the pattern of repeating decimals that we obtain. Here is a short list.

2	(factor of 10)
3	0.3 recurring
5	(factor of 10)
7	0.142857 recurring
11	0.09 recurring
13	0.076923 recurring
17	0.0588235294117647 recurring
19	0.052631578947368421 recurring

This suggests that the number 076923 (derived by dividing unity by 13) should be cyclic too.

Lets try it:

4 If we got a remainder of 0 it would mean that 10 was exactly divisible by 7 – which contradicts our assumption that the number we are dividing by is not a factor of the base number.)

1	076923	7	538461
2	153846	8	615384
3	230769	9	692307
4	307692	10	769230
5	384615	11	846153
6	461538	12	923076
		13	**999999**

A close look at this list will reveal that it is indeed cyclic but that it consists of two interleaved cycles, one with the digits 076923 and the other with the digits 153846. We say that it has *order* 6 (because each cycle has 6 numbers in it) and *multiplicity* of 2 (because there are 2 different cycles) and it is clear that the *order* × the *multiplicity* must be 1 less than the *prime divisor*.

The reason why 076923 has a order 6 rather than 12 becomes clear when we do the division:

10 over 13 = 0 remainder 10
100 over 13 = 7 remainder 9
90 over 13 = 6 remainder 12
120 over 13 = 9 remainder 3
30 over 13 = 2 remainder 4
40 over 13 = 3 remainder 1
10 over 13 = 0 remainder 10
etc. etc.

After 6 divisions, we come across the same remainder.

If we insist on all cyclic numbers having a multiplicity of 1, then the next one is

0588235294117647 = 9999999999999999 ÷ 17

That's fascinating – but I still don't understand why 13 produces two sequences while 7 and 17 only one. And what about 11? Does that generate a cyclic sequence?

Well the first thing to notice is that when you add the cyclic number to itself *p* times you always get a string of 9's.

142857 × 7 = 999999
09 × 11 = 99
076923 × 13 = 999999

Actually it would be better to write this list as follows:

$$142857 \times 7 = 999{,}999 = 10^7 - 1$$
$$0909090909 \times 11 = 9{,}999{,}999{,}999 = 10^{11} - 1$$
$$076923076923 \times 13 = 999{,}999{,}999{,}999 = 10^{13} - 1$$

Note that the number of 9's in the string is always one less that the prime p. The pattern is clear. For any prime number p, $10^p - 1$ is *always* divisible by p[5]. For example, the number 9999...9999 (with 30 9's in it) must be divisible by 31 because if we divide 1 by 31, the remainders must repeat themselves after at most 30 digits. This will generate a cyclic number with 30 digits which, like all cyclic numbers resolves into a string of 9's when multiplied by 31.

Now it sometimes happens that p divides into a shorter string of 9's. In this case, the cyclic sequence will have a multiplicity which is greater than 1. For example, while 999,999 is divisible by 7, 13 divides both 999,999 and 999,999,999,999.

The number 11 divides all the even strings of 9's up to 9,999,999,999 and therefore generates a cyclic sequence with multiplicity 5. It looks like this:

1	0909090909
2	1818181818
3	272727272
4	3636363636
5	4545454545
6	5454545454
7	6363636363
8	7272727272
9	8181818181
10	9090909090
11	**9999999999**

If you want to know a bit more about cyclic numbers, have a look at the article of the same name on my website:
http://www.jolinton.co.uk/Mathematics/Cyclic_Numbers/Text.pdf

5 This amazing result is a special example of what is known as Fermat's Little Theorem. We shall come across it again several times.

Pi

At school I was told that the ratio of the circumference to the diameter of a circle was a number called PI and that it was approximately equal to 22/7. Why did my teacher use the word approximately*? What is its* exact *value?*

I can't answer that; nor could your teacher. The truth is – nobody knows!

What do you mean – nobody knows? Surely there must be a way of calculating it?

Yes, of course there is – but the calculation is infinitely long. You can calculate π to any desired degree of accuracy but you cannot calculate its exact value.

One of the most famous formulas for π is as follows:

$$\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \ ...\right)$$

but this is of little use in actually calculating π because it converges so slowly. It takes more than 300 terms to calculate π to 2 decimal places!

I see. π is like the square root of 2 – it is irrational.

Absolutely. But in a way, the situation is even worse. With a number like √2 we can at least square it to obtain a rational number. Numbers like this are called algebraic numbers because there is always an algebraic formula which will turn them into an integer.

Take another example. The golden ratio ς is equal to (1 + √5)/2 and it has the value 1.618033989.... To turn it into a rational number, all you have to do is double it (3.23606...), subtract 1 (2.23606...) and square the result to get 5.

But with π it is different. It turns out that whatever you do to π in the way of multiplying, dividing, adding or subtracting with itself or other rational numbers, the result is always irrational. π is not only irrational, it is also what is called *transcendental*.

Hang on a minute. I thought all irrational numbers would be like

√2. In fact I can almost see a way of proving that. After all, if you start with a number like √2 (i.e. any algebraic number) and change any digit, you must generate another algebraic number. Right?

Right.

So starting from √2 you could change it into π just by changing the digits one by one. Therefore π is an algebraic number too.

Wrong. The crucial thing is that you would have to change an *infinite number* of digits. Algebraic numbers are defined by a *finite* set of arithmetical operations but transcendental numbers are defined by an *infinite* number of operations. This observation led a mathematician called Georg Cantor in 1883 to an astounding conclusion. The number of algebraic numbers (between 0 and 1, say) is obviously infinite – but the number of transcendental numbers is even greater than that.

Now you are just pulling my leg. How can the number of transcendental number between 0 and 1 be greater than infinity? That just doesn't make sense.

Well it is true that his conclusion didn't make a lot of sense to his contemporaries at the time either but his proof is perfectly simple and quite watertight.

Cantor's proof that the number of transcendental numbers is greater than infinity

His crucial insight was to see that the process of counting was simply a case of placing the items to be counted into a one-to-one correspondence with the integers. For example, to answer the question how many primes are there less than 10, we place the primes (2,3,5 & 7) with the integers:

1	2
2	3
3	5
4	7

and we see that there are four of them.

To answer a question like which is greater, the number of planets in the solar system or the number of traditional colours in the rainbow, we again pair them off like this (in any order we like):

blue	Mars
violet	Venus
yellow	Neptune
orange	Mercury
indigo	Earth
red	Jupiter
	Saturn
	etc.

Since we run out of colours before we run out of planets, we conclude that there are more planets than colours.

Now let us list all the algebraic numbers between 0 and 1. We shall use the method that you yourself have suggested, changing one digit at a time so that we can be sure we have left none of them out. The list starts with zero:

0.00000000000000...

Now we list the 9 numbers which differ from zero in the first decimal digit:

0.10000000000000...
0.20000000000000...
0.30000000000000...
...
0.90000000000000...

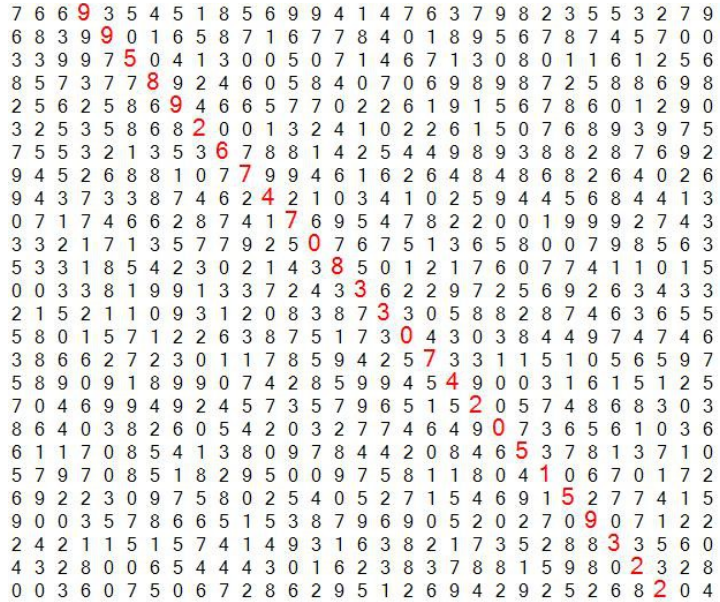
Then we go through all the possibilities changing the second digit. There will be 90 of them. Then we add the 900 numbers which differ from all of these in the third digit. And so on and so on.

Because of the way we have organised the list it is not too difficult to write down a formula for the n^{th} number in the list and because our method is algebraic, all the numbers listed are algebraic.

Now common sense would indicate to you that since we have changed every digit in a systematic and exhaustive way, we must have listed all the numbers which actually exist. Not so! Here is a simple way of writing down a number which is definitely not in the list. It starts

with a zero and a decimal point. The first digit after the decimal point is chosen to differ from the first digit of the first number in the list. Since this number is 0.000000000000 we could choose, say, 0.7.The second digit after the decimal point is chosen to differ from the second digit of the second number in the list. Since this number is 0.100000000000 we could choose, say, 3 and our number is now 0.73. We do the same for the third, fourth, fifth, indeed, every digit.

Here is a picture of a small fragment of the list. The digits picked out in red are NOT the digits of the number we are looking for. The digits of the number we are after differ from every one of these.



so the number we have generated is not in the list! The method we used to generate the list did indeed produce every *algebraic* number but it failed to generate any of the *transcendental* numbers.

Hang on a minute; let me get my head round this. Are you saying that although there are an infinite number of algebraic numbers between 0 and 1, there are infinitely more transcendental numbers as well?

Yes I am saying exactly that.

I don't see how there is room for them all.

Fair point. But it remains the case that while you can approximate π to any degree of accuracy you like (355/113 is accurate to 6 decimal places!) no algebraic number exactly equals π.

How do you know that?

Regrettably there is no simple proof for the transcendental nature of π but a proof was discovered in 1882. Many other numbers have been proved to be transcendental but there are many more which are thought to be transcendental but are not proven – for example π^{π}. (Wikipedia lists this number as one of those which is not even known to be irrational, let alone transcendental!)[6]

I find that truly astonishing

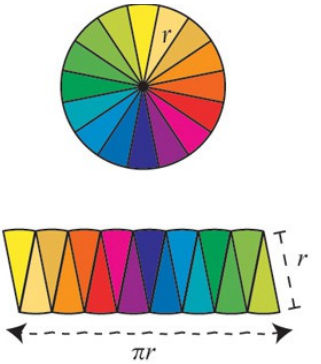
6 For a slightly tongue-in-cheek article on Pi try:
http://www.jolinton.co.uk/Mathematics/2pi_or_not_2pi/Text.pdf

Areas and Volumes

While we are on the subject of Pi – can you prove to me why the area of a a circle is πr^2?

In a way this is surprisingly difficult. The calculation of areas and volumes is the province of a branch of mathematics called the integral calculus – a subject which has caused much grief amongst students of mathematics, myself included. But basically the idea is simple enough.

You cut the circle up into pieces and rearrange them in a way which makes the area easy to calculate. For example, suppose you cut a circle into a large (even) number of pie slices like this:

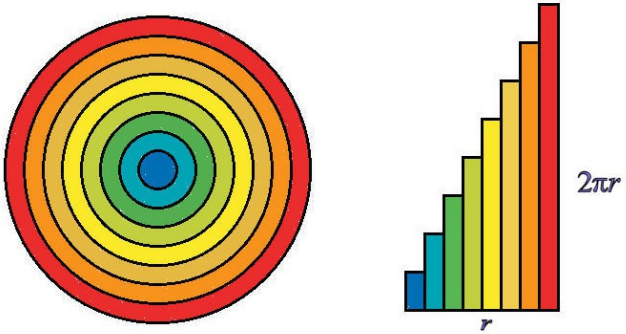


Rearrange them to make an approximate rectangle. The slices all have the same length r and the curvy edges will each be approximately πr in length so the area of the rectangle will be approximately πr^2.

Obviously the more slices you use the more accurate the calculation and if you use an infinite (!) number of slices the calculation will be exact.

Another way of slicing up the circle is to cut it into concentric rings. If the radius of one of the rings is x then its length will be $2\pi x$.

Now lets stack the side by side like this:

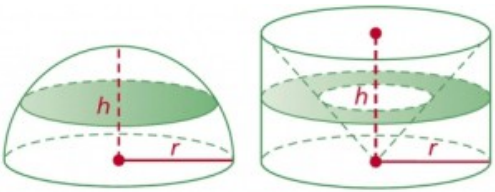


They make a triangle whose longest strip is $2\pi r$. The area of this triangle is ½ base × height = πr^2.

That seems clear enough. Now what about the volume of a sphere?

I am afraid this is surprisingly difficult. You start by slicing the sphere into discs and working out a formula for the area of each disc; then you add up all the areas to get the volume. Unfortunately, unlike the case of the area of a circle, there is no easy way to draw a simple picture showing how the result is obtained.

Incredibly, Archimedes figured out the answer 1800 years before Newton and Leibniz invented the integral calculus. Using his quite extraordinary powers of perception and intuition, he realized that the area of each disc was the same as the area of the ring formed at the same height when a cone is cut out of a cylinder. Quite how he came to this conclusion is beyond me but he is perfectly correct.



Since the volume of a cylinder of height r is πr^3 and the volume of a cone of the same height is $\frac{1}{3}\pi r^3$ [7], the volume of a hemisphere is $\frac{2}{3}\pi r^3$.

7 Even this is not obvious!

e

There is another number which I never understood. Something to do with logarithms, I think

Yes, the number you are thinking of is called *e* and has the value 2.7182818285... It is also transcendental.

So what's so special about it? And how come it equals 2.7 something?

Well, suppose you have £100 to invest and you find a bank who is offering 10% per annum (some hope!). After 1 year you will have £110 in the account. After 2 years you will have £110 × 1.1 = £121. After 3 years, £133.10 etc. Every year your balance is multiplied by 1.1 so after 10 years you will have £100 × 1.1^{10} = £259.37.

But one day you realise that you have been cheated. The interest was calculated by the bank only at the end of each year, so for 11 months of the year, your money was effectively gaining no interest at all. So you complain to the bank and insist that your interest is calculated every month, not every year. In other words, your balance is multiplied by (1 + 1/120) every month rather than (1 + 1/10) every year. Your final balance will be £100 × 1.008333^{120} = £270.70. an extra £11 and 33p!

For a while this satisfies you but the next time you have £100 to invest, you insist that the interest be calculated every day – no, every minute – no, every second – no, why not *continuously*! Surely if you calculate the interest more and more often you will get more and more interest.

Well, that may be true but the interest you get will not be a lot more. If you divide the 10 year period of the investment into smaller and smaller intervals your final balance will be £100 × $(1 + 1/n)^n$ where n is the number if intervals.

Now there is an important theorem called the Binomial Theorem which says that

$$(1 + x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \ldots$$

where 2! stands for factorial $2 - 1 \times 2 = 6$, and $3! = 1 \times 2 \times 3 = 6$ etc.

Now if we substitute $x = 1/n$ we get

$$\left(1 + \frac{1}{n}\right)^n = 1 + \frac{n}{n} + \frac{n(n-1)}{2!}\frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!}\frac{1}{n^3} + \ldots$$

Now as n gets larger and larger, the little bits which we are supposed to subtract from n on the top of the fraction get less and less significant and in the end it all boils down to

$$e = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots$$
$$= 1 + 1 + \frac{1}{2} + \frac{1}{6} + \ldots = 2.7182818285\ldots \qquad (3)$$

so even if your money is compounded continuously, you will never get more than £271.83 at the end of 10 years.

Well that's all very interesting but it doesn't answer my question. Why is this number so special and what has it got to do with logarithms?

Fair point. Lets go back to the bank and calculate how fast you balance is growing. When you were offered 10% per year, your balance started by growing at a rate of £10 a year during the first year; £11 a year in the second and so on.

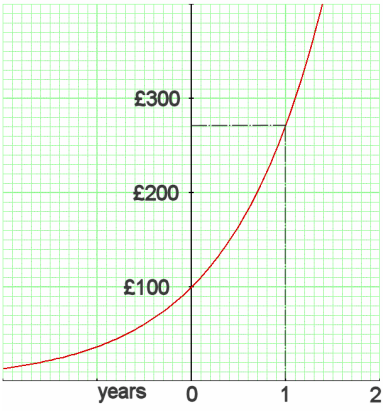
When you insisted on calculating the interest every month, the *rate of increase* of your balance was still £10 per year at the start but the rate was increased bit by bit every month rather than just at the end of the year. When the balance was compounded continuously, the *rate of increase* was equal to *current balance*/10 at every instant. A graph of your balance against time would show it growing smoothly and at every instant the *gradient* of the graph would be equal to the *current balance* divided by 10.

Now let us suppose you found a bank which offered you a rate of interest of 100% per year compounded instantaneously[8]. This time your

8 This doesn't mean that every instant you get 100% interest – it means simply that at every instant your balance is growing at a rate equal to your current balance per year. Think of a car acceleration in such a way that at all times its acceleration (in mph per second) is equal to its speed (in mph). If it is travelling at 1 mph then it

balance will grow to £271.83 in one year and every year thereafter it will multiply by this magic number 2.7182818285...

A graph of the balance in your account against time will be one in which the gradient of the graph is everywhere equal to the current balance. This is what it looks like:



At the start (year 0) the slope of the line is 45° showing that it is growing at a rate of 100% per year). As the balance grows, the gradient grows too and after exactly 1 year the balance is £271.83.

We can write down a simple formula for this graph. Since every year the balance increases by a factor of e (= 2.7182818285...) the formula is

$$balance = e^n \quad \text{(where } n \text{ is the time elapsed in years)}$$

This is the *only* graph which has the unique property that the gradient equals the y coordinate.

OK – but what has this got to do with logarithms?

I think logarithms deserve a chapter on their own.

---

will be accelerating at 1 mph/second. Less than a second later it will be travelling at 2 mph and its acceleration will have increased to 2mph/second etc. etc.

Logarithms

Of all the things you were taught at school I suspect that the subject of logarithms was the most detested – and you can see the problem when you try to describe what a logarithm is. The conventional description is this: the logarithm of a number is the number to which a certain number (called the *base)* is raised to equal the number which you first thought of.

No wonder you were confused!

Lets try again.

Any number x can be regarded as the sum of *two* other numbers b and y (eg $10 = 6 + 4$) and given x and b it is easy to calculate y by subtracting b from x (= 10 - 6). Likewise x can be regarded as the *product* of two numbers b and y (eg $24 = 6 \times 4$). This time to calculate y you need to *divide* x by b (= 24 / 6).

But there is another operator which we are getting used to. What if you regard x as the *power*[9] of b and y – ie $x = b^y$. Given x and b how do we calculate y? And what words shall we use to describe the action of calculating y from x and b?

For historical reasons we say that y is the *logarithm of x to the base b* and we write this as $y = \log_b(x)$. (If I had my way, I would have adopted the following notation: if $x = b^y$ (pronounced x equals b to the power of y) then $y = x_b$ (pronounced y equals x from the power of b) and much grief would have been avoided!)

Now everyone knows how to subtract and most people can do a bit of division, but calculating the logarithm of a number x using the base b (in my terminology, calculating x from the power b) is not easy. Fortunately we now have pocket calculators to do the job for us so that doesn't matter much.

Yes. But why are logarithms important?

9 Part of the problem is that we don't have a simple word to use for the action of 'raising to the power of'.

Hold on to your hat – I am coming to that. There are really two reasons. The first was of great historical importance but is less so nowadays while the second is purely mathematical.

First let us prove a few simple theorems in logarithms.

Suppose $x = b^{p+q}$. This means that $x = b^p \times b^q = y \times z$

Now we have defined the logarithm function so that $\log_b(x) = p + q$, $\log_b(y) = p$ and $\log_b(z) = q$. ie:

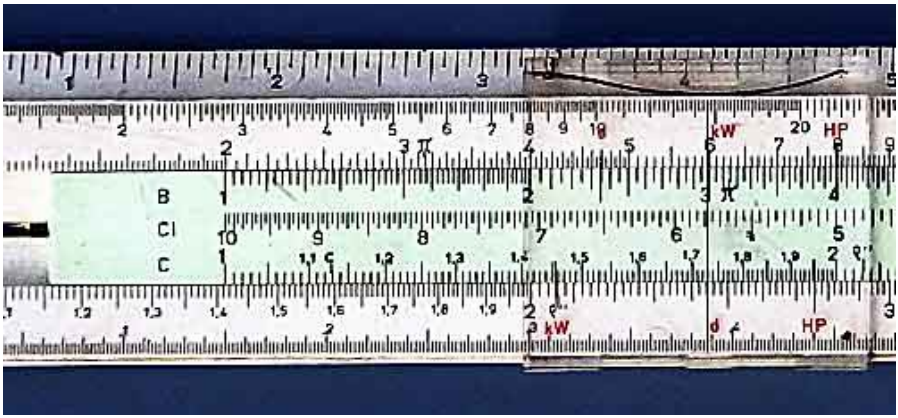
$$\boxed{\log_b(y \times z) = \log_b(y) + \log_b(z)} \qquad (4)$$

In other words to calculate the logarithm of a product, you add the logarithms of the two original numbers (using any base you like).

It follows from this that the logarithm of x^n is simply n times the logarithm of x.

$$\boxed{\log_b(x^n) = n \log_b(x)} \qquad (5)$$

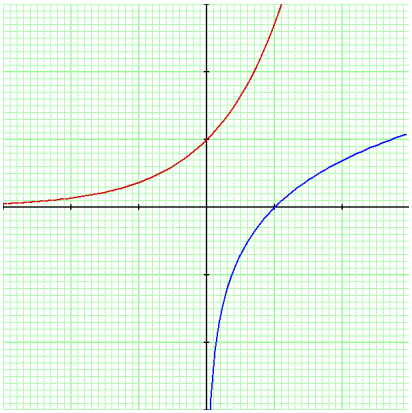
Before we had calculators, these relations were widely used to turn complicated multiplication and division sums into relatively easy addition and subtraction. This is how a slide rule worked. The scales are marked off in logarithms and to multiply two numbers (eg 1.41 and 1.73) you placed the end of the slide against the first number on the stock, place the cursor over the second number on the slide and read off the answer on the stock (2.44)



I see – but what is the real reason why logarithms are so important?

It is all to do with the shape of the logarithmic graph.

If you recall, the logarithm of x to the base b is the power to which b must be raised to equal x. More simply if $y = \log_b(x)$ then $x = b^y$. What this means is that, just as dividing is the inverse of multiplying and subtracting is the inverse of adding, 'taking the logarithm of' is the inverse of 'raising to the power of'. Have a look at the graphs of the two functions (using e as a base of course)



The red graph of the graph of e^x while the blue graph is the graph of $\log_e(x)$ (usually written just as $\log(x)$ or $\ln(x)$ on a calculator). You can see that they are exactly the same shape only with the axes swapped over.

This has the following profound consequence. Since the *gradient* of the graph $y = e^x$ is equal to y, the *gradient* of the graph $z = \log_e(x)$ is equal to $1/x$. Just look. When x is small the gradient of the blue line is steep; at $x = 1$ the gradient is 45° and as x increase so the gradient gets less and less. We shall see just how important this observation is when we talk about calculus.

Calculus! Oh no! We're not going to talk about calculus are we?

Of course. But not just yet.

Probability

Let's talk about something a bit simpler. I remember one day at school – it was my birthday, and I was surprised to find that it was also the birthday of another member of the class too. I remarked to my maths teacher that this was a remarkable coincidence because the probability of two people having the same birthday is only 1 in 365. He said I was wrong and that in fact the probability was better than evens.

Thinking about it later I realised that, because it could have been any of the other 24 boys in the class, the probability that one of them had the same birthday as me was not 1/365 but 24/365. But that still isn't any where near 50% is it?

I am not sure if probability is simpler than logarithms but it is certainly an interesting topic and the 'paradox of the birthdays' is a famous one. Let's do the calculations.

As you say, the probability that two specific people have the same birthday is 1/365. In other words, in a class of 2, the probability that both of them have the same birthday is 1/365.

Now what about a class of 3? Now there are four possibilities: A could have the same birthday as B; B could have the same birthday as C; C could have the same birthday as A; or all three could share the same birthday. Working out the probability that at least two of them share a birthday is going to be a nightmare.

So let's turn the problem on its head. What is the probability that *none* of the boys and girls share a birthday? This is much easier.

The probability that 2 people do *not* share a birthday is 364/365. The probability that a third member of the class does *not* share a birthday with *either* of the other two is 363/365.

Why is that?

Because there are only 363 days left. So the probability that none of the three classmates share a birthday is 364/365 × 363.365.

Why did you multiply the probabilities together?

This is the way probabilities work. If the probability of event A happening is $p(\mathrm{A})$ and the probability of event B happening is $p(\mathrm{B})$, then, *providing the two events are completely independent of one another*, the probability that *both* events will happen is $p(\mathrm{A}) \times p(\mathrm{B})$.

Why?

OK. Suppose you have a tombola with N tickets in it of which n win prizes. What is the probability of winning a prize?

n/N obviously.

Let's call this probability $p(\mathrm{A})$.

Now suppose you have a second tombola with M tickets and m prizes. The probability of winning this one is $p(\mathrm{B}) = m/M$.

Now what is the probability of winning prizes in both tombolas? As you are clearly aware, the probability is equal to

$$\frac{\textit{the total number of ways you can win}}{\textit{the total number of way you can pull two tickets out of the tombolas}}$$

Now since there are N ways of pulling out a ticket from the first tombola and M ways of pulling a ticket out of the second, there are $N \times M$ ways of pulling two tickets out of the two tombolas. Similarly, if there are n ways of winning the fists tombola and m ways of winning the second, there are $n \times m$ ways of winning both. The probability of winning both tombolas is therefore:

$$\frac{nm}{NM} = \frac{n}{N} \times \frac{m}{M} = p(\mathrm{A}) \times p(\mathrm{B})$$

Now lets go back to the birthday problem. We have seen that the probability the three classmates do *not* share a birthday is

$$\frac{364}{365} \times \frac{363}{365}$$

What about a class of four?

Well, there are now only 362 days left, I guess the answer will be:

$$\frac{364}{365} \times \frac{363}{365} \times \frac{362}{365}$$

Excellent. All we have to do now is to carry on multiplying by smaller and smaller fractions until we get to a figure less than 0.5. Using a computer we can easily show that this happens when there are 23 people in the class. How many boys and girls were there in your class?

25

Well, there you are. Your teacher was right.

I still don't quite get it. What was wrong with my argument?

Nothing at all. But what you calculated was the probability that another member of the class had the same birthday *as you*. What the teacher was referring to was the probability that *any two members of the class* had the same birthday. Not the same thing at all.

I see.

By the way; I seem to remember that there are some circumstances in which you add *probabilities rather than multiply them. Why is that?*

Yes. That is when you have two *mutually exclusive* events. For example, what is the probability that you will throw *either* a 2 *or* a 3 on a single throw of a die?

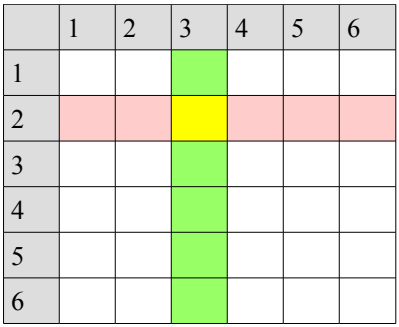
I suppose it will be 1/6 + 1/6 = 2/6 = 1/3

Correct. But this is because you c*annot throw both at the same time*. The two events are *mutually exclusive*. If I gave you two dice and asked you what was the probability of *either* throwing a 2 with the first die *or* throwing a 3 with the second (*but not both*) the answer would be different.

Why? Surely that is basically the same thing isn't it?

Well there are 36 ways of throwing two dice. There are 5 ways of throwing a 2 with the first die (but something other than a 3 with the second). Similarly there are 5 ways of throwing a 3 with the second die (but something other than a 2 with the first). The answer is therefore 10/36, not 1/3.

You can best see why by using a grid of possibilities with the first throw down the side and the second throw along the top.



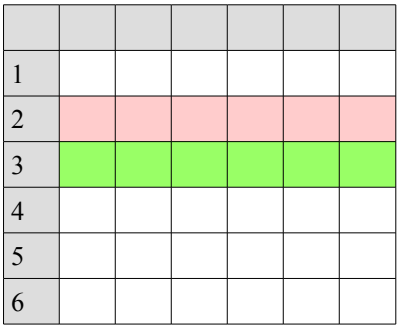
The red cells show the cases where you throw a 2 but not a 3, the green ones are where you throw a 3 but not a 2. You are not permitted to count the yellow square.

I get it. Even if you do count the yellow square (i.e. you get a prize when you throw both *a 3* and *a 2), the probability of winning is still not 1/3, it is only 11/36!*

Well done! You are really getting the hang of this probability business now.

Perhaps. But I still don't quite understand why ruling out the case where you throw both winners is not the same as throwing one die

Well if you are only give one die and one throw the possibilities look like this:



The red and the green bars do not cross. There is no yellow square. The events are *mutually exclusive*.[10]

10 For a discussion of some of the more difficult paradoxes in probability see: http://www.jolinton.co.uk/Books_and_Articles/Other_Essays/Paradoxes_in_Probability.pdf

Permutations and Combinations

Talking of pulling tickets out of a tombola reminds me of a lesson which was about the different ways you could pull lottery numbers out of a drum. What was that about?

In the National Lottery there are 49 balls in the drum numbered 1 to 49 and in each draw, 6 balls are drawn out. Obviously they come out in a certain order, but this does not matter in the end. In order to win you only have to choose the right numbers. (For ease of checking, they put the chosen numbers into numerical order.)

Obviously we want to know what are the chances of winning.

Now if you have ever watched the draw on television you will know that in order to keep up the suspense, after each ball is drawn out, the drum spins again. But this is not really necessary. Suppose that when the drum has stopped spinning, all 49 balls drop into 49 hoppers, only the first 6 of which are transparent – lie this:

23	5	19	18	31	15	

The first thing to ask ourselves is – how many different ways are there for 49 balls top fall into 49 hoppers?

I think I know how to do that.

Go on.

Well, there are 49 ways of filling the first hopper, 48 ways of filling the second, 47 ways of filling the third etc. etc. So the total number of ways is $49 \times 48 \times 47 \times 46 \times \ldots \times 2 \times 1$.

Well done. This number is called factorial(49) and is written using an exclamation mark 49! It is a stupendously big number!

How big?

Well, it has approximately 60 digits and is considerably greater than the number of atoms in the Sun.

In general, the number of different *permutations* of N objects is $N!$

Now the Lottery rules divide the 49 numbers into two sets of 6 and

43 and they state that it doesn't matter what order either of these sets are in. In other words we can *permutate* (i.e. rearrange the order of) both sets without changing the result.

Now there are 6! Ways of permutating 6 objects and 42! ways of permutating 42 objects so the total number of ways of arriving at a specific result like the one illustrated is:

$$\frac{49!}{6! \times 42!} = \frac{49 \times 48 \times 47 \times \ldots \times 1}{(6 \times 5 \times \ldots \times 1) \times (42 \times 41 \times \ldots \times 1)}$$

Now you can see that 49! contains 42! so we can cancel all these factors out leaving

$$\frac{49!}{6! \times 42!} = \frac{49 \times 48 \times 47 \times 46 \times 45 \times 44}{6 \times 5 \times 4 \times 3 \times 2 \times 1}$$
$$= 49 \times 8 \times 47 \times 23 \times 3 \times 11$$
$$= 13983816$$

What this means is that you have a 1 in 13,983,816 chance of winning the National Lottery with a single ticket. Since at least 15 million tickets are sold each week, the jackpot will usually be shared each week by several people.

In general, the number of ways of separating a heap of x items from a total of $N = x + y$ is:

$$\boxed{{}_N C_x = \frac{N!}{x! \times (N - x)!} = \frac{(x + y)!}{x! \times y!}} \qquad (6)$$

where ${}_N C_x$ is pronounced "*N choose x*" and describes the number of *combinations* of x items out of a total of N.[11] This formula crops up in Pascal's Triangle. Have you heard of that?

Yes, but I have forgotten how to construct it.

11 For a much more detailed discussion of permutations and combinations have a look at the following article on my website:
http://www.jolinton.co.uk/Mathematics/Permutations_and_Combinations/Text.pdf

Pascal's triangle

That's easy. Here it is:

					1						1
				1		1					2
			1		2		1				4
		1		3		3		1			8
	1		4		6		4		1		16
1		5		10		10		5		1	32
	1		6		15		20		15		6		1		64
1		7		21		35		35		21		7		1	128
1		8		28		56		70		56		28		8		1	256

Every number is the sum of the two numbers above it. But make a note of two remarkable patterns. First, the sum of all the numbers in the n^{th} row is equal to 2^n. And second, all the numbers in the 7^{th} row (except the two 1's at the ends) are divisible by 7. Why is this?

Consider the action of multiplying a number like 1331 by 11. The sum goes like this:

$$\begin{array}{r} 1\,3\,3\,1\,0 \\ \underline{\quad 1\,3\,3\,1} \\ 1\,4\,6\,4\,1 \end{array}$$

You can see how the number is shifted one place to the left and then added to itself giving rise to the same pattern of numbers.

Now what has this got to do with Pascal's triangle?

Consider the task of multiplying $(a + b)$ by $(a + b)$. The sum looks like this:

$a \times (a + b) \; + b \; \times (a + b) \; = (a^2 + ab) + (ab + b^2)$

If we write this sum down in the same way that we did before you

can see the parallel:

$$\begin{array}{r} a^2 + \ ab \\ \underline{\qquad\quad ab + b^2} \\ a^2 + 2ab + b^2 \end{array}$$

The next sum looks like this:

$$\begin{array}{r} a^3 + 2a^2b \ + \ ab^2 \\ \underline{\qquad\quad a^2b \ + 2ab^2 + b^3} \\ a^3 + 3a^2b \ + 3ab^2 + \mathrm{b}^3 \end{array}$$

Again you see how the same coefficients keep recurring. In general the numbers in the n^{th} row of the triangle are the coefficients of the expression $(a + b)^n$. Now you can see why the numbers in the rows add up to a power of 2. Just put a and b equal to 1.

That's very interesting – but why are the numbers in row 7 all divisible by 7?

To answer that question we need to find an formula which will tell us the number in any particular box. This is best done by writing the triangle in a rectangular XY array like this:

8	1								
7	1	8							
6	1	7	28						
5	1	6	21	56					
4	1	5	15	35	70				
3	1	4	10	20	35	56			
2	1	3	6	10	15	21	28		
1	1	2	3	4	5	6	7	8	
0	1	1	1	1	1	1	1	1	1
	0	1	2	3	4	5	6	7	8

What we need is a formula relating each number in the boxes to its XY coordinate. Here it is:

$$C = \frac{(x + y)!}{x! \times y!}$$

Isn't that the formula for the number of combinations of x items out of N?

Exactly so. To give an example, the number in the box (3, 4) will be

$$C = \frac{(3 + 4)!}{3! \times 4!} = \frac{7!}{3! \times 4!} = \frac{1.2.3.4.5.6.7}{1.2.3 \times 1.2.3.4} = \frac{5.6.7}{1.2.3} = 35$$

It is worth noting two things. First, swapping x and y makes no difference to the formula. This explains the symmetry in the triangle. Secondly – and this is of profound significance – if $x + y$ is prime (for example 7) there will be a 7 in the numerator but there can never be a 7 in the denominator (because x and y must necessarily be less than 7) so it can never cancel out. (A number like 6 can cancel because the denominator can contain a 2 and a 3.) The only time the 7 cancels is if either x or y is itself equal to 7. The upshot of this is that whenever n is prime, *all the coefficients in the n^{th} row (except the 1's) will be divisible by n*.

Now since the sum of all the coefficients (including the two end 1's) is equal to 2^p, it follows from this that

for every prime p, $2^p - 2$ must be divisible by p

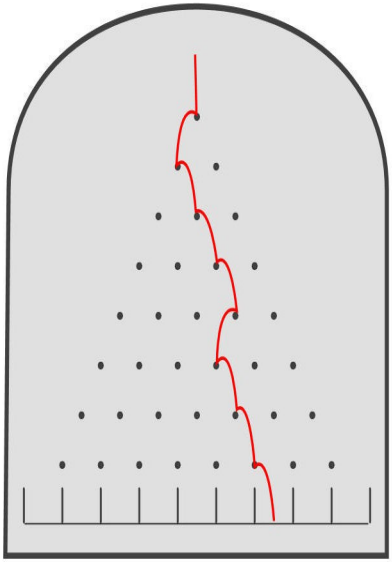
This is a truly remarkable fact and it is known as Fermat's LITTLE theorem (NOT Fermat's LAST theorem which is of much less significance.) Try some numbers. $2^5 = 32$ and 30 is divisible by 5; $2^{11} = 2048$ and 2046 is divisible by 11 (yes it is!). I haven't a clue what 2^{97} is but since 97 is a prime I know that the 30 digit number $2^{97} - 2$ *must* be divisible by 97.

Don't you get that wonderful feeling of POWER!

That is amazing. But I am a bit puzzled as to where all those factorials came from – after all, to construct the triangle all you do is add up a few numbers.

Yes it is rather surprising. You can think of Pascal's triangle a s

mathematical pin ball machine:



The idea is that whenever the ball strikes a pin it has a 50:50 chance of going one way or the other. Here a ball is shown bouncing eight times, going LRRRLRRR, ending up in the 7th of the nine possible bins. We want to know what proportion of balls will end in each of the bins.

It is obvious that the only way for a ball to end in bin number 1 is if it bounces to the left 8 times in a row. On the other hand, there are lots of ways in which the ball could end in bin number 7. All that is needed is for the ball to bounce to the left twice and to the right 6 times.

In other words, the total number of ways the ball can end in bin number 7 is equal to the total number of ways in which 8 objects can be divided into two piles of 2 and 6. This is, of course ${}_8C_2 = 28$

I see. But I thought that the number in the triangle was the sum of the two numbers above it. What has that got to do with it also being the number of ways in which the pin ball could get there?

Good question. Have a look at the 7th number in the last row of the triangle (highlighted in red). The two numbers above the number we are interested in are 21 and 7. Now the only way in which the pin ball can

get to bin number 7 is by hitting one or other of these pins. So the total number of ways of reaching the final bin must be the sum of the two numbers above it!

Hang on a minute though – how can the sum of those two horrendous factorials be another horrendous factorial?

Good point. We need to calculate the following sum:

$$N = \frac{((x-1)+y)!}{(x-1)! \times y!} + \frac{x+(y-1)!}{x! \times (y-1)!}$$

and show that it equals the equation on page 39.

First we make the denominators equal by multiplying the top and bottom of both fractions by x and y respectively, then add them together:

$$N = \frac{x \times (x+y-1)!}{x! \times y!} + \frac{y \times (x+y-1)!}{x! \times y!}$$
$$= (x+y) \times \frac{(x+y-1)!}{x! \times y!} = \frac{(x+y)!}{x! \times y!}$$

I am impressed!

So you should be. Not only have we explained many of the remarkable properties of Pascal's Triangle, we have also essentially proved the Binomial Theorem too which we used to calculate e on page 29!

How come?

Earlier we showed that $(a+b)^3$ was equal to $a^3 + 3a^2b + 3ab^2 + b^3$. In other words, the coefficients of the various terms of a binomial expansion are just the figures which appear in the corresponding rows in the triangle. Spelling it out in more detail:

$$(a+b)^3 = \frac{(0+3)!}{0! \times 3!}a^3 + \frac{(1+2)!}{1! \times 2!}a^2b + \frac{(2+1)!}{2! \times 1!}ab^2 + \frac{(3+0)!}{3! \times 0!}b^3$$

If we put $a = 1$, $b = x$ and generalise to the n^{th} degree we get:

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \ldots$$

which is what we needed to prove.[12]

12 For an extension of these ideas have a look at:
http://www.jolinton.co.uk/Mathematics/Pascals_Pyramid/Text.pdf

Prime Numbers

One of the things I learnt at primary school was that, if you wanted to know if a certain number was divisible by 3, all you had to do was add up the digits and see if that was divisible by 3 but I was never told how the method worked. Why does it work?

That's easy. Consider the number $N = DCBA$ where D, C, B and A are the digits. I am sure you realise that this means that

$$N = 1\times A + 10\times B + 100\times C + 1000\times D$$

Now lets split this sum as follows:

$$N = 1\times A + (9+1)\times B + (99+1)\times C + (999+1)\times D$$
$$N = (A+B+C+D) + 9\times B + 99\times C + 999\times D$$

Now it is obvious that the numbers 9, 99, 999 etc. are all divisible by 3 so if A+B+C+D is divisible by 3, so is N.

Why didn't I think of that? Its obvious really. In fact I can see that the test will just as well work for divisibility by 9 as it does for 3.

Yes, absolutely. But it is not an entirely trivial task to prove that every power of 10 minus 1 is divisible by 3 (or 9).

I can!. The Binomial Theorem says that $(1 + 9)^n = 1 + n.9 +$ *other powers of 9 – so, subtract 1 and all you have left are powers of 9!*

Now it really is my turn to be impressed!

What about tests for division by other numbers? I can see that numbers which end in 0 or 5 are divisible by 10 and 5 and I know that all even numbers end in an even digit – but what about division by 4 for example?

Even numbers end in a even digit because all multiples of 10 are even. But 10 is not divisible by 4. On the other hand, 100 is so if the last two digits are divisible by 4 then the whole number is. Similarly to test for division by 8 you need to take the last 3 digits.

What about dividing by 7?

Here is one approach: the number $DCBA$ can be written as:

$$N = 1\times A + (7+3)\times B + (98+2)\times C + (994+6)\times D$$

where 7, 98, 994 etc are the nearest numbers to powers of 10 which are divisible by 7.

To test for divisibility by 7 you have to do the following sum:

$$R = 1\times A + 3\times B + 2\times C + 6\times D$$

and for numbers with more digits you have to use the following sequence of multipliers 1 3 2 6 4 5 repeated. (It is no accident that this sequence is the same as the sequence of remainders which we met on page 19.

In practice, in the case of division by 7 it is probably easier just to do the division – or find a calculator but division by 11 produces a more usable algorithm. *DCBA* is equal to

$$N = 1\times A + (11-1)\times B + (99+1)\times C + (1001\pm 1)\times D$$

Now it turns out that any number of the form 100...001 (which contains an even number of zeros in the middle) is divisible by 11 and, of course any number of the form 999...9 (with and even number of 9's) is also divisible by 11 so the remainder is

$$R = A - B + C - D + \text{ etc.}$$

For example: to test whether the number 123456 is divisible by 11, add up all the odd digits (6 + 4 + 2 = 12) and subtract the even ones (5 + 3 + 1 = 9) to get the remainder which in this case is 3.

That's interesting – but I can see that methods like these are going to be pretty useless when it comes to testing for divisibility by larger numbers. How do they go about testing large numbers to see if they are prime or not?

Testing for primes

You probably know that a common method of encrypting sensitive information such as credit card numbers and the like is to use a code which uses a couple of large prime numbers as 'keys'. These numbers are typically 100 or more digits long. Now if you want to know whether or not a certain large number *N* is prime, you could obviously just try

dividing it by every number smaller than N. You could save a lot of time by restricting yourself to numbers less than $\sqrt{N}$ – and you could save yourself even more time if you restricted yourself to all the *prime* numbers less than $\sqrt{N}$ (provided you had a list of all these numbers handy). The trouble with this method is that the time taken to do all these calculations increases steadily with increasing N. In fact, a number 10 times bigger will take approximately $\sqrt{10}$ times longer to test.

On my PC it takes typically 10 seconds to test a 15 digit number. On this basis it would take 100 seconds to test a 17 digit number, 1000 seconds (a quarter bof an hour) to test a 19 digit number etc. etc. To test a 100 digit number would take longer than the age of the universe. (In fact it would take far far longer than the age of the universe but you get my point!)

Fortunately there are faster ways of testing for primality and there are algorithms which can churn out 100 digit primes to order.

You mean there is a formula which churns out all the primes in sequence?

How I wish! No, its not as simple as that. There is a surprisingly simple formula with 26 variables a, b, c, … z. When you substitute 26 integers for the 26 variables, the expression is either negative or prime. Now you might think that this formula will be an easy way to generate huge primes: just plug in some huge numbers and see what you get. The trouble is, the formula almost always produces negative numbers and it turns out that the easiest way to generate a large prime number is simply to test some randomly generated numbers until you hit lucky. If only we had a formula $P(n)$ which spat out the n^{th} prime number – but there are good reasons for supposing that such a formula does not exist.

But I thought you said it would take longer than the age of the universe to test a 100 digit number?

Yes – that's the brute force method. But there are subtler tests. Do you remember Fermat's little theorem?

You mean – for every prime p $2^p - 2$ is divisible by p?

Yes – that's the one. Now it is important to remember that although

we know this works for every prime, it is not true that every number for which this works is prime. Numbers which satisfy Fermat's test but which are not prime are called pseudo-primes. Of the 169 numbers less than 1000 which pass the test, only 3 are not prime and the larger the numbers get, the smaller the proportion of pseudo-primes to real primes.

Here is a 48 digit number which almost certainly has never been computed by any one before – ever. It satisfies Fermat's test and my computer generated it in less than a second. I would (almost) be willing to bet my life that it is prime!

260174 469621 671560 555932 780692 628594 970045 782113

That's saying something! But how on earth can a computer handle such large numbers?

Well it is not difficult to program a computer to handle 48 digit numbers. On the other hand, if p has 48 digits (ie is something like 10^{48}), then 2^p will have something like 10^{47} digits and no conceivable computer could handle a number like that! But you have to remember that we do not actually have to calculate 2^p – we only need to calculate the remainder when divided by p.

How can you calculate $2^p/p$ without calculating 2^p first?

Like this. Suppose we want to calculate the remainder when 2^7 is divided by 7.

We use the simple fact that if you double a number, the remainder will also double.

Can you explain that?

Sure. The number $(np + r)$ where n is an integer, has remainder r when divided by p . So the number $2(np + r) = 2np + 2r$ will have a remainder of $2r$ when divided by p.

So going back to our example:

$2 \div 7$ has remainder 2
$(2\times2) \div 7$ has remainder $2\times2 = 4$
$(2\times2\times2) \div 7$ has remainder $4\times2 = 8 = 1$ (modulo 7)
$(2\times2\times2\times2) \div 7$ has remainder $1\times2 = 2$
$(2\times2\times2\times2\times2) \div 7$ has remainder $2\times2 = 4$

$(2\times2\times2\times2\times2\times2) \div 7$ has remainder $4\times2 = 8 = 1$ (modulo 7)
$(2\times2\times2\times2\times2\times2\times2) \div 7$ has remainder $2\times1 = 2$

(and of course $2^7 = 128$ and $128 \div 7$ is 18 remainder 2)

By continually multiplying the remainder by 2 and throwing out the 7's, we never actually have to calculate 2^7 . Similarly we only ever have to multiply our 48 digit number by 2 and do a few subtractions to calculate the eventual remainder[13]. Cunning eh?!

That is quite staggering. You mean that you can prove that $2^{260174\ 469621\ 671560\ 555932\ 780692\ 628594\ 970045\ 782113}$ *has a remainder 2 when divided by* 260174 469621 671560 555932 780692 628594 970045 782113 *with just a few dozen 48-digit operations? That is quite amazing.*

By the way – how big is $2^{260174\ 469621\ 671560\ 555932\ 780692\ 628594\ 970045\ 782113}$*? You said it would have about* 10^{47} *digits. Could you write a number with* 10^{47} *digits down in a book?*

Good heavens no! 10^{47} is much bigger then the number of stars in the universe. In fact it is approximately equal to the number of atoms in the Earth.

My mind is reeling! I am feeling quite dizzy!

13 I have simplified things a bit here. In fact the algorithm involves squaring the number rather than doubling it but the basic principle is the same

Fermat's Little Theorem

We first came across Fermat's little Theorem on page 21 in the context of cyclic numbers. There I said that any number of the form $10^{(p-1)} - 1$ was divisible by p (where p is prime, of course). For example, we used the fact that 999,999 was divisible by 7 to generate the cyclic number 142857.

I remember.

We have also proved using Pascal's Triangle that $2^p - 2$ is divisible by p. (page 43). I would like to show you how these two theorems are both special cases of a more general theorem.

Go on then.

We start with the (unproved) proposition that $a^p - a$ is divisible by p.

Now consider the number $b^p - b$ where $b = a + 1$:

$$b^p - b = (a + 1)^p - (a + 1)$$

Now let's look carefully at the expression $(a + 1)^p$. We can use the Binomial Theorem to expand it into the following:

$$(a + 1)^p = a^p + 1 + \text{multiples of } p$$

Notice that all the middle terms are multiples of p. In other words:

$$\begin{aligned} b^p - b &= a^p + 1 + \text{multiples of } p - (a + 1) \\ &= a^p - a + \text{multiples of } p \end{aligned}$$

But we have assumed that $a^p - a$ is divisible by p. So $b^p - b$ must be divisible by p as well!

Hang on a minute. Put some figures in for me. All this algebra is confusing.

OK. we know that $2^7 - 2$ is divisible by 7. What we have proved is that therefore $3^7 - 3$ must also be divisible by 7.

Is it?

Well 3^7 is 2187 and $3^7 - 3$ is 2184 which equals 312×7, so there's your answer!

And I suppose that if $3^7 - 3$ is divisible by 7, then so is $4^7 - 4$ and $5^7 - 5$ etc. etc.

Correct.

In other words we have proved that for all integers a, and all primes p:

$$a^p - a \text{ must be divisible by } p \qquad (7)$$

So why is 999,999 divisible by 7? 999,999 doesn't sound like a number of the form $a^p - a$.

No – but $10^7 - 10$ is. This means that $10^7 - 10$ (= 9999990) is divisible by 7. So, provided that 10 and 7 are co-prime (i.e. provided that 10 is not divisible by 7) we can divide by 10 to prove that $10^6 - 1$ is divisible by 7. In general (provided b and p are co-prime)

$$b^{(p-1)} - 1 \text{ is divisible by } p$$

which means that all numbers of the form 9999....9 with one fewer nines than a prime number p will be divisible by p and will generate a cyclic number[14]!

That's quite cool. But I have one more question – why was 999,999 also divisible by 13, giving rise to a cyclic number with a multiplicity of 2?

You really do ask them, don't you? I think, fundamentally, the answer is that it is just chance. Lets try to predict the multiplicity of the cycle generated by the prime 19. We know from Fermat's Little theorem that 999,999,999,999,999,999 is divisible by 19. Now 18 9's can be grouped into 2 groups of 9, 3 groups of 6, 6 groups of 3 or 9 groups of 2. If 999,999,999 is divisible by 19, then the cycle will have multiplicity of at least 2. As it happens, 999,999,999 is not divisible by 19. But perhaps 999,999 is divisible by 19. If so the cycle will have multiplicity 3. But 999,999 is not divisible by 19 either, nor is 999 nor is 99 so we can conclude that the 19 cycle has multiplicity of 1.

14 Except, of course for 9 which is not divisible by 2 and 9999 which is not divisible by 5; 2 and 5 being factors of 10.

Graphs and gradients

One of the few things I enjoyed about maths at school was graphs. I used to love plotting out the various shapes – you know, parabolas and the like.

I wonder if you were taught about the gradients of those graphs.

The gradient? That's the slope of the graph, isn't it?

Yes that's right. The gradient of a graph is like the slope of a hill – a 1 in 7 hill goes up 1 m for every 7 m on the map.

And a 1 in 1 hill is vertical.

No. A 1 in 1 hill would be at 45°. A vertical hill would have infinite gradient.

That's not what I heard.

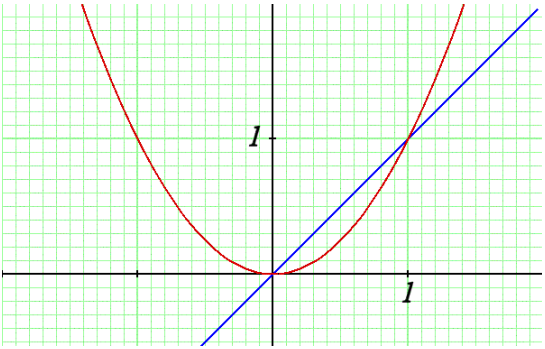
Well I won't argue with you. Surveyors use the convention you have suggested because it is easier to measure distances along the ground rather than horizontally but mathematicians define the gradient as the vertical height gained / horizontal distance travelled. i.e. gradient is the tangent of the angle of the slope.

I guess that makes sense.

The simplest possible function of a real variable x is $y = x$. This is just a straight line through the origin at 45° and its gradient is constant and equal to 1. The gradient of the graph $y = 2x$ will be 2 and the gradient of the graph $y = -x$ will be – 1. OK?

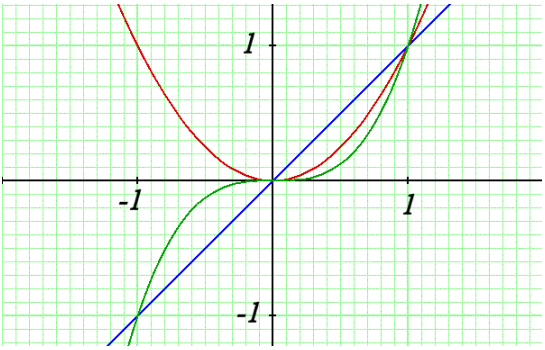
OK

Now, as you know, the graph of $y = x^2$ is a parabola. Its gradient is negative when x is negative (because it is sloping down); becomes zero at $x = 0$ and is positive when x is positive. We might even hazard a guess that the gradient of the graph is actually equal to x but if you take some measurements on a piece of graph paper, you will find that this is not quite correct. The gradient is actually equal to $2x$.



This is obvious at the point (1, 1) where the red parabola is clearly a lot steeper that the blue line.

Now lets consider the graph of $y = x^3$.



When x is negative, the green graph ($y = x^3$) has a positive gradient; dips down to zero at $x = 0$ and then increases again.

Are you saying that the gradient of the cubic curve is a parabola?

Yes – pretty much – except that, as before, at the point (1, 1) it is even steeper. It turns out that the gradient of the graph of $y = x^3$ is equal to $3x^2$. Can you guess what the gradient of the graph of $y = x^4$ will be?

Might it be $4x^3$?

Fantastic! Now we need just one more idea and then you know everything you need to know about differential calculus.

Calculus? Is that what we are doing? I thought calculus was difficult!

Well there is another calculus called integral calculus which, I have to admit, is often difficult but we won't be needing that here.

Thank goodness!

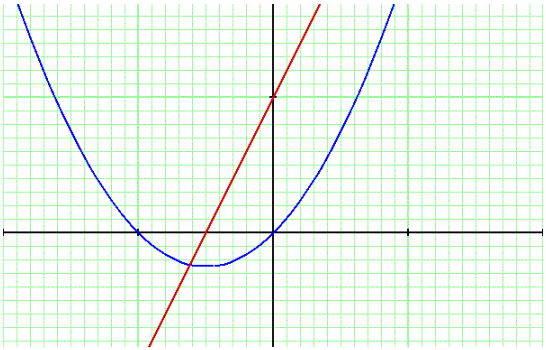
Suppose we add two functions together. For example $y = x^2 + x$. What do you think the gradient of this graph will be?

Well, the gradient of x^2 is 2x and the gradient of x is 1 so could the answer be 2x + 1?

Really – you are on a roll today! You are absolutely right. When you have terms added together, you simply differentiate each term separately.

Differentiate? What's that?

Sorry. That's what we have been doing all along. To 'differentiate a function' you simply calculate its gradient. Here is the graph (in blue) of $y = x^2 + x$ and its gradient in red.



As you can see, to the left of the point (-½, 0) the blue line is sloping down and the red line is negative; then at (-½, 0) the gradient of the blue line is zero and the red line crosses the axis; and after that the gradient increases steadily in the positive direction.

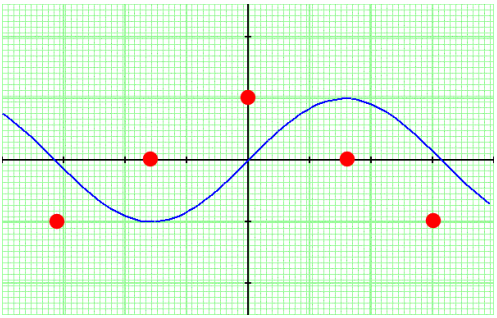
I think I have got that

Now do you remember what we said about the gradient of the graph of $y = e^x$? (You can find it on page 31)

Oh yes, it says that the graph of e^x is the only one whose gradient is equal to the y coordinate.

That's right. What this means is that differentiating e^x gets you to e^x. It is this property of e^x which makes this function so special.

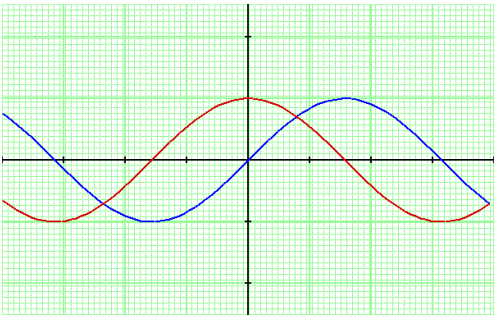
Now take a look at the sine function. Lets see if we can plot out its gradient.



The red spots show where the gradient is either –1, zero or +1. Now, bearing in mind that the graph of the gradient is going to be smooth, not angular, what do you think the function looks like?

Could it be the cosine curve?

Spot on. Differentiate $\sin(x)$ and you get $\cos(x)$. Here it is:



What about differentiating $\cos(x)$?

I don't know. Differentiating sin shifts the curve 90° to the left so I suppose it will shift 90° further.

And what does that give you?

It looks as if it will be –sin(x).

Well done. Differentiating a third time will give you $-\cos(x)$ and differentiating once more gets you back where you started $\sin(x)$!

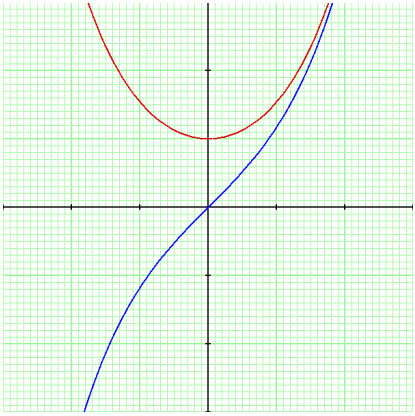
That's quite cool. I never knew that.

Wait a minute. I have just thought of something. Differentiating e^x gives you e^x; differentiating sin(x) four times gets you sin(x); is there a function which returns to itself after two or three differentiations?

The answer is yes to the first but no to the second.

That's odd. What does the first function look like?

Here it is in blue with its gradient in red:



It is a bit like the straight line and the parabola, but the blue line is kinked in such a way as to make the red line reflect the gradient of the blue line as well as the blue line being the gradient of the red line.

Well I would never have thought of that. Do these graphs have names?

Yes. They are called *hyperbolic sine and cosine*, shortened to sinh(x) (the blue line) and cosh(x) (the red line) respectively.

Well I never! I have heard those names but I had no idea what they were.

They are interesting functions but you will be glad to hear that we do not need to pursue their properties just now.

That's a relief!

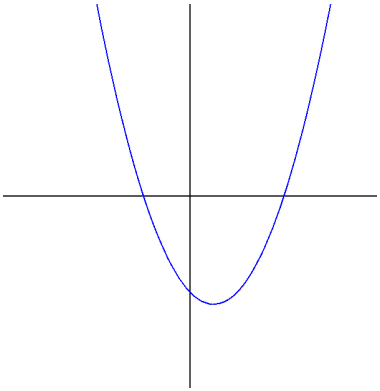
Complex Numbers

I was delighted when I discovered that the graph of a quadratic equation was a parabola and that was why quadratic equations had two solutions.

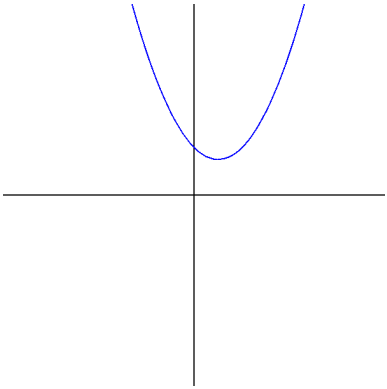
Yes. In general a parabola will intersect the X axis in two places; a cubic in three places and so on.

That's right. But I got really confused when my teacher said that, if the parabola didn't cut the axis, there were still two 'solutions' but that they were 'imaginary'. That's bonkers. If the graph doesn't cut the axis there simply aren't any solutions. You can't just imagine some.

Quite so. Here are two equations, one with 2 real solutions and one without.



$y = x^2 - x - 2$



$y = x^2 - x + 1$

Exactly – so what was my teacher talking about?

Well – do you remember being told about a formula for the solutions to quadratic equations like the ones illustrated above?

Like hell I do! I hated that formula! I could never get the right answers. I always seemed to end up having to take the square root of a negative number!

Do you remember the formula?

Not really. But I know it had a square root in it somewhere.

That's right. Here it is:

if $$a x^2 + b x + c = 0$$

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

Ugh! How ghastly!

Yes it does look a bit ugly – but you will have to lodge your complaints with the Great Mathematician in the Sky; the formula wasn't invented by a malicious exam board!

Okay – so why is it so ugly?

Well, let's look at a slightly simpler equation:

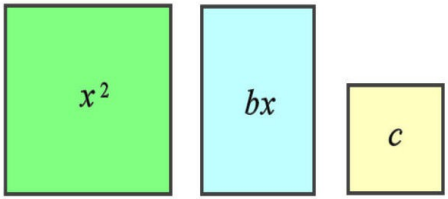
$$x^2 + b x + c = 0$$

(All we have done is remove the coefficient *a* by dividing it out.)

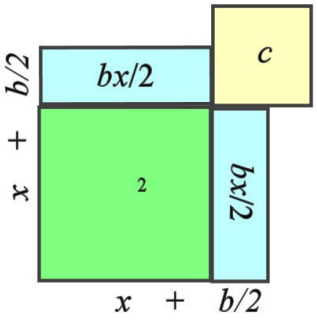
Now to solve this equation we have got to get rid of the square term somehow. One way to do this would be to take the square root of the equation but we can only do this if *c* has exactly the right value. To see why this is, lets translate the equation into pictures.

That sounds like a good idea.

x^2 is a square; bx is a rectangle and c is a little bit left over:

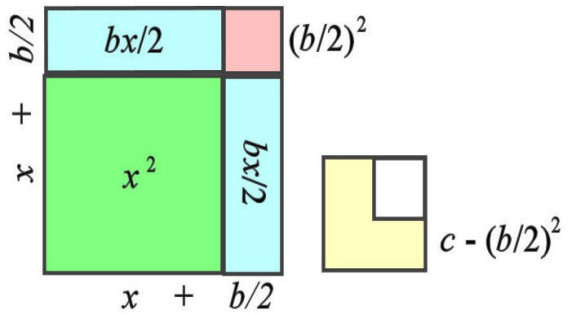


Now lets try to make a square out of these bits. First we split the bx rectangle down the middle and place it beside the square; then we try to fill in the missing bit with the bit left over c.



(As you can see. It will only fit if c happens to equal $(b/2)^2$)

Lets take the bit that we need to complete the square:



Now we can write our original equation in the following form:

$$(x + b/2)^2 + c - (b/2)^2 = 0$$

and a bit of simple algebra gives us:

$$(x + b/2)^2 = (b/2)^2 - c$$

$$x + b/2 = \pm\sqrt{((b/2)^2 - c)}$$

$$x = \frac{-b \pm \sqrt{(b^2 - 4c)}}{2}$$

which (except for the coefficient a) is the formula we are looking for.

Yes, well, I can see where the complexities come from – but what has this got to do with the missing solutions?

It is all to do with that square root sign. If b^2 is greater than $4c$ then all is well and the equation will have 2 perfectly respectable roots. But if

$4c$ is greater than b^2 then you end up having to take the square root of a minus number and, as you know, this is impossible.

So that just proves my point. When 4c is greater than b^2 there are NO SOLUTIONS!

You are perfectly entitled to take that line but just supposing that you *could* take the square root of a negative number? What then?

That's just ridiculous.

Bear with me. Suppose there was a number – lets call it **i** – which was the square root of –1. The square root of, say, –9 could be calculated as follows:

$$\sqrt{(-9)} = \sqrt{(9 \times -1)} = \sqrt{(9)} \times \sqrt{(-1)} = 3\mathbf{i}$$

Lets use this idea to calculate the 'imaginary' solutions to the equation $x^2 - x + 1 = 0$. Here $b = -1$ and $c = 1$ so the 'solutions 'are:

$$x = \frac{-b \pm \sqrt{b^2 - 4c}}{2} = \frac{1 \pm \sqrt{1 - 4}}{2} = 0.5 \pm 0.87\mathbf{i}$$

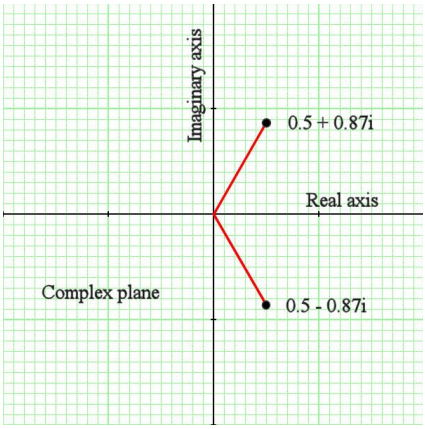
Notice that the solution has two parts. The 0.5 is called the 'real' part of the solution while the 0.87**i** is called the 'imaginary' part. ($0.87 = \sqrt{3}/2$) Together they form what is known as a 'complex' number.

I don't know why you should call it 'imaginary'. I simply can't imagine a number which is the square root of a negative number. They simply don't exist!

Well there was a time when many people thought the same about negative numbers. But we have got used to the idea because negative numbers turn out to be really useful. The same is true of these so-called 'imaginary' numbers. They are really useful.

Okay. Show me. Where are these 'imaginary' numbers?

You are, of course familiar with the idea that all real numbers can be put on a line stretching from $-\infty$ to $+\infty$. Lets plot the imaginary numbers on a line *at right angles* to the real number line. In effect we are plotting the real part of any complex number on the X axis and the imaginary part on the Y axis. The whole graph is called the 'complex plane' and here are the two solutions to the equation we have solved:



I still don't get it. In what sense are these points places where the parabola that we drew originally 'crosses the X axis'

Well, in truth, as you have been constantly pointing out, the graph doesn't cross the X axis – but that is because we have only really been plotting the 'real' bit of the graph.

Have a look at both the parabolas we drew originally. The only difference between them is in the constant *c*. As we increase c from -2 to +1 then parabola rises higher and higher. What happens to the two solutions as we, as it were, pull the parabola up?

I suppose, to start with, there are two spots on the X axis.

That's right. They start at +2 and -1.

Then as you pull the parabola, the spots will get closer together until, they merge into one?

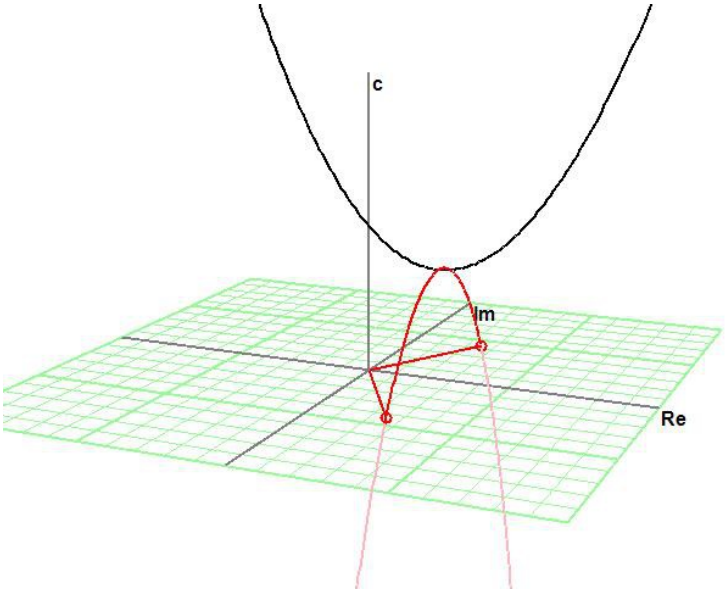
Right again. This happens when $c = 0.25$ and the two solution are $x = 0.5$ and $x = 0.5$. (Yes there are *two* solutions – they just happen to be equal!)

Now pull the parabola up a bit more.

I suppose they split apart again and diverge up and down the 0.5 line.

Fantastic!. And the higher you pull the parabola the more they diverge. It is just as if, underneath the real parabola there is another

'imaginary' one looking like this.



The imaginary half is like a mirror image but rotated by 90°. You can easily see where the two imaginary roots lie now.

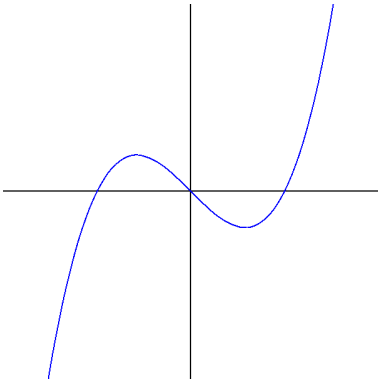
That's really cool. You mean to say that all that time I was enjoying myself drawing graphs of quadratic equations and the like, I was really only drawing half of them?

I suppose you could say that.

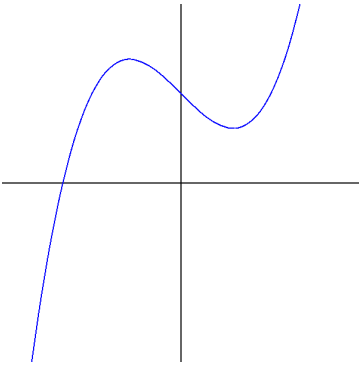
What do the roots of a cubic equation look like?

This is where things get really pretty.

Take the equation $x^3 - x = 0$. It is easy to verify that this has three real roots, -1, 0 and 1. This is what it looks like:



$y = x^3 - x$

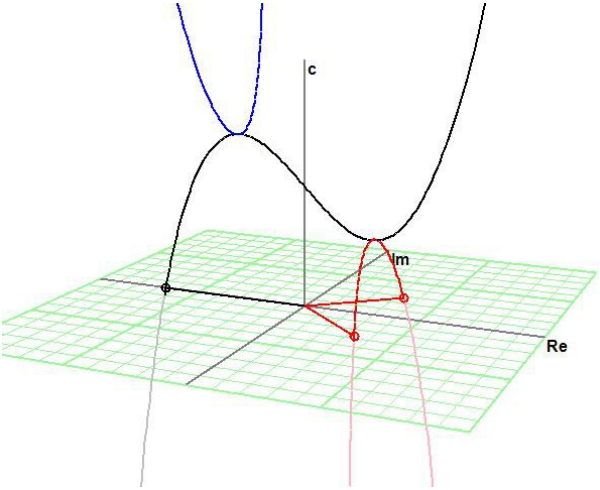


$y = x^3 - x + 1$

Also illustrated is an equation which appears to have only one root – but being a cubic equation, it ought to have three. Where are the other two?

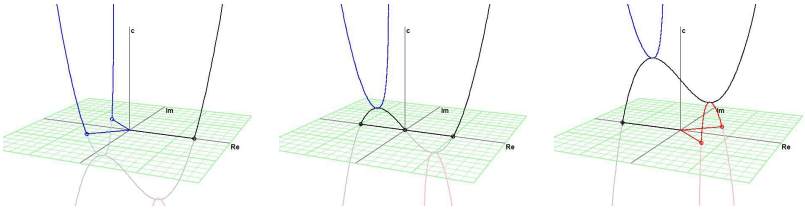
Well, I guess they are at right angles on a kind of parabola hanging down underneath the dip in the curve.

Absolutely right! Here they are in 3D.



Solving cubic equations is much more difficult than solving quadratics but reading off the graph we can see that the three roots of this equation are approximately -1.35, 0.65 + 0.55**i** and 0.65 – 0.55**i**.

Note the other (blue) curve too which represents more imaginary solutions. Here is a sequence showing what happens as the constant increases from -1 through to +1.

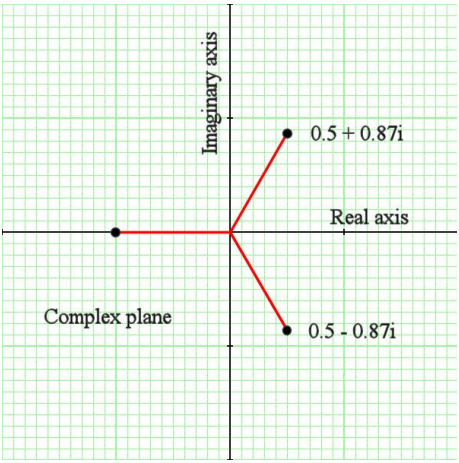


Hm.. Very pretty; but I am a bit surprised that the roots of a cubic equation could be expressed in terms of **i** *which is the square root of -1, not the cube root.*

Well this is one of the amazing things about complex numbers. We don't have to invent a new number **j** which is the cube root of –1; ordinary complex numbers are sufficient to solve equations of any order.

So what is the cube root of –1?

Well, as you would expect, there are three of them. One of them is –1 because $-1 \times -1 \times -1 = -1$. The other two are $0.5 + 0.87\mathbf{i}$ and $0.5 - 0.87\mathbf{i}$. This is what they look like in the complex plane.



That's odd. Don't I recognise those numbers?

Yes, you do. It so happens that the roots of the equation

$x^2 - x + 1 = 0$ also satisfy the equation $x^3 = -1$. Here is the simple proof:

$$x^2 - x + 1 = 0$$

multiply by x: $x^3 - x^2 + x = 0$

add the two equations: $x^3 + 1 = 0$

hence: $x^3 = -1$

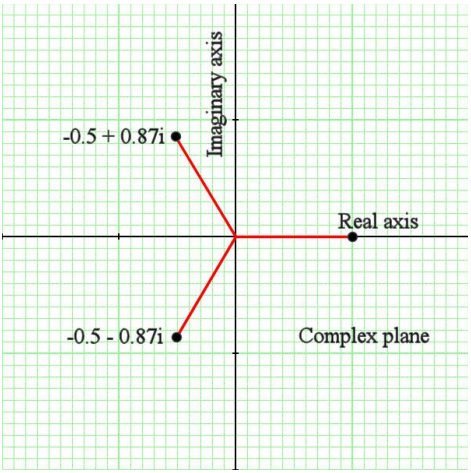
But that is not the most interesting thing about the three cube roots of –1. Can you see what it is?

Are the three angles all equal?

Yes – well done!. The angles are all 120°. Now, I have another question for you. If the three cube roots of –1 are –1, 0.5 + 0.87**i** and 0.5 – 0.87**i**. What do you think the cube roots of 1 are?

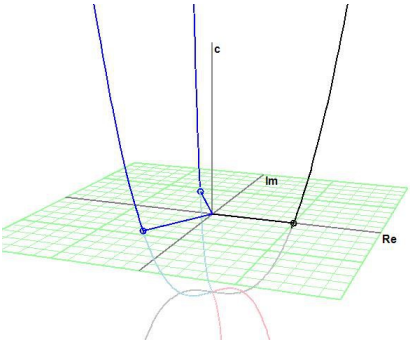
Well one of them is obviously 1 itself. Could the other two be at 120° on the other side?

Right again! Here they are:

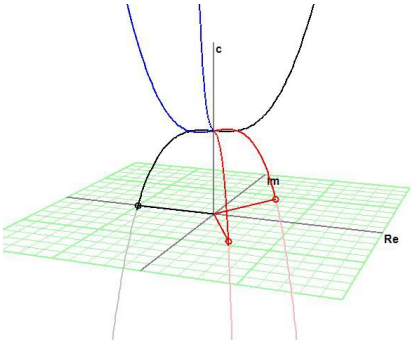


What do the cube roots of 1 and –1 look like in 3D?

Here they are:



Cube roots of 1



Cube roots of -1

As you 'pull the curve upwards' the three roots converge on zero and then diverge again, always maintaining an angle of 120° between them.

I think I am getting the hang of this. Would I be right in thinking that the roots of the equation $x^4 = -1$ *would lie at 90°?*

Absolutely.

And the roots of $x^5 = -1$ *at 360/5 = 72°?*

Right again!. The roots of $x^4 = -1$ are, of course, 1, **i**, –1 and –**i**.

Now consider this. 1 × **i** = **i;** **i** × **i** = -1; -1 × **i** = **-i** and **-i** × **i** = 1. In other words, every time you multiply by **i** you move round the sequence by 90° and when you have multiplied by **i** four times you are back where you started!

So what you are saying is that when you multiply complex numbers together it is like rotating them – is that it?

You know when Howard Carter poked a hole in the sealed entrance to King Tutankhamen's tomb and glimpsed 'wonderful things'? You have just been granted a similar glimpse into the wonderful treasure house of complex numbers. Just wait until we open it up!

I think I will have a stiff drink first!

Complex Arithmetic

OK. Sock it to me. What can you tell me about complex numbers that I will be able to understand?

Well let's consider how complex numbers are added and multiplied.

What do think will be the result if we add $(1 + 2\mathbf{i})$ and $(3 - \mathbf{i})$?

I don't know. Do you just add the real and imaginary bits together?

Correct. $(1 + 2\mathbf{i}) + (3 - \mathbf{i}) = (4 + \mathbf{i})$

Well, I suppose that makes sense – but where is the justification for this?

Well I could point out that you would have no difficulty in accepting that $(1 + 2\boldsymbol{x}) + (3 - \boldsymbol{x}) = (4 + \boldsymbol{x})$. All we have done is to extend the idea of addition so that the same method works even when x is a complex number.

OK. But what about $(1 + 2\boldsymbol{i}) \times (3 - \boldsymbol{i})$?

Let's remind ourselves how we do ordinary multiplication first.

$$(1 + 2\boldsymbol{x}) \times (3 - \boldsymbol{x}) = 1 \times 3 + -1 \times x + 2x \times 3 + 2x \times -1$$
$$= 3 + 5x - 2x^2$$

Now all we do is put $x = \mathbf{i}$. This gives us

$$(1 + 2\mathbf{i}) \times (3 - \mathbf{i}) = 3 + 5\mathbf{i} - \mathbf{2i^2} = 5 + 5\mathbf{i}$$

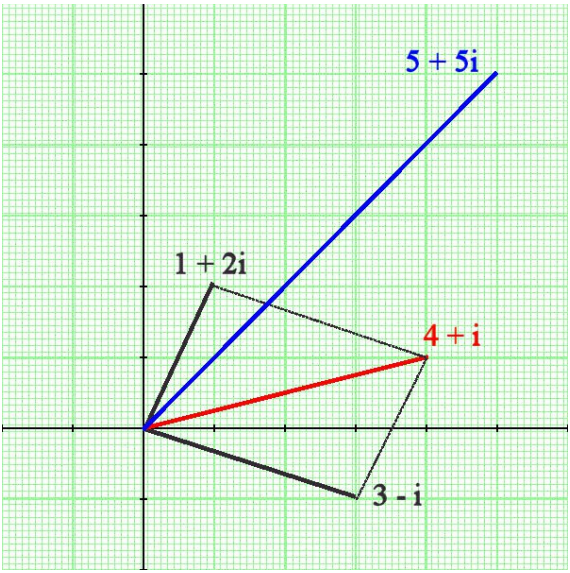
Hang on. How did you do that last bit?

I used the fact that $\mathbf{i}^2 = -1$ so $-2\mathbf{i}^2 = +2$.

Let's summarise what we have found out.

$$(a + \mathbf{i}b) + (c + \mathbf{i}d) = (a + c) + \mathbf{i}(b + d)$$
$$(a + \mathbf{i}b) \times (c + \mathbf{i}d) = (ac - bd) + \mathbf{i}(bc + ad)$$

Since it is always nice to put things in pictures, here is a representation of the two sums we worked out on the complex plane: the result of the addition sum is shown in red and the result of the multiplication in blue.



I see. The addition sum reminds me of the addition of vectors which we did in Physics.

Absolutely correct. It is exactly the same. I must warn you though; the multiplication is totally different.

That's all right. I always thought that multiplying vectors was a stupid idea anyway.

Can you see any geometric relation between the black lines and the blue line?

Frankly, No.

I agree. It is not at all obvious. But there must be one.

Go on then – I can see you are itching to tell me.

First I must remind (tell?) you about a couple of remarkable trig formulae which you have probably forgotten (or perhaps never even knew). Here they are:

$$\sin(A + B) = \sin(A)\cos(B) + \cos(A)\sin(B)$$
$$\cos(A + B) = \cos(A)\cos(B) - \sin(A)\sin(B)$$

They tell you how to work out the sine and cosine of the sum of two angles A and B.

Now I am going to replace cos(A) with a; sin(A) with b; $cos(B)$ with c and sin(B) with d.

$$\sin(A + B) = bc + ad$$
$$\cos(A + B) = ac - bd$$

Now look back at the yellow box on the previous page. Amazingly, exactly the same arrangement of letters appears in the multiplication formula! I can't really emphasise enough how absolutely stunning this discovery is. It appears quite out of the blue. But what does it mean?

Obviously, the first thing we must do is rewrite our multiplication formula in terms of sines and cosines:

$$(\cos(A) + \mathbf{i}\ \sin(A)) \times (\cos(B) + \mathbf{i}\ \sin(B)) =$$
$$(\cos(A)\cos(B) - \sin(A)\sin(B)) + \mathbf{i}\ (\sin(A)\cos(B) + \cos(A)\sin(B)) =$$
$$\cos(A + B) + \mathbf{i}\ \sin(A + B)$$

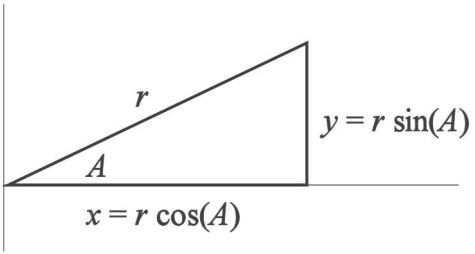
but this cannot be quite right because while a, b, c and d can be any number you like, cos(A) and sin(A) etc. must lie between -1 and +1.

We can easily remedy this defect by multiplying each of the two original numbers expressions on the left hand side by factors r and s. We now have:

$$r(\cos(A) + \mathbf{i}\ \sin(A)) \times s(\cos(B) + \mathbf{i}\ \sin(B)) =$$
$$rs(\cos(A + B) + \mathbf{i}\ \sin(A + B))$$

Now – pause for drum roll – we can pull the rabbit out of the hat and reveal the secret!

Any complex number ($x + \mathbf{i}y$) can be written in the form r(cos(A) + **i** sin(A)) where r is the 'length' of the vector (called the **modulus**) and A is the angle which the vector makes with the X axis (called the **argument**). I am sure you are familiar with this picture:



Yes – I remember that we used it in Physics and the x and y were the 'components' of the vector r

This is just the same. All we are saying here is that instead of using Cartesian coordinates to represent a complex number, we can use polar coordinates. To understand complex addition we need to use the former but to understand complex multiplication we need to use the latter.

So here, at last, is the secret of complex numbers and and the reason why they are so incredibly useful:

> To multiply two complex numbers together you **multiply** their **moduli** and **add** their **arguments**!

I don't understand.

Let's check this out using our numerical example. The modulus of $(1 + 2\mathbf{i})$ is $\sqrt{(1^2 + 2^2)} = \sqrt{5}$ and the modulus of $(3 - \mathbf{i})$ is $\sqrt{(3^2 + 1^2)} = \sqrt{10}$.

The modulus of the result should therefore be $\sqrt{5} \times \sqrt{10} = \sqrt{50}$ and a glance at the graph on page 69 will show that this is indeed the case.

Now what about the angles? The argument of $(1 + 2\mathbf{i})$ is the angle whose tangent is 2. This is 63.4°. The argument of $(3 - \mathbf{i})$ is the angle whose tangent is minus 1/3. this is –18.4°. Add these two together (not forgetting the minus sign!) and what do you get – 45°!

To put it another way, if we represent a complex number in the notation $\{r, A\}$ where r is the modulus and A is the angle then

$$\{r_1 , A_1\} \times \{r_2 , A_2\} = \{r_1 r_2 , A_1 + A_2\}$$

Phew! That was more like getting blood out of a stone than pulling rabbits out of a hat!

Yes – but the result was worth it wasn't it?

I suppose so. But I still don't see where it gets us really.

Just you wait and see!

Complex Functions

So where are we?

A mathematician called Leopold Kronecker once said “God made the integers, all else is the work of Man.” I disagree. If we accept that 'God made the integers' (i.e. if we accept the Platonic stance that integers exists independently of the material universe) and if we also concede that integers may be added and subtracted, we have to conclude that negative integers also exist.

Likewise, as soon as we have realised that positive and negative integers can be multiplied and divided, we have to accept that fractions exist. Every time we invent a new operation (addition, multiplication etc.) we find that its inverse leads to a new extension of the idea of number.

Yes – when you try to take the square root of a number like 2, you have to admit the existence of irrational numbers.

And when you try to take the square root of a negative number you discover complex numbers.

And I suppose you are going to tell me that another inverse operation is going to need super-complex numbers and that another will lead to super-super-complex numbers and so on ad infinitum.

Actually, no. I am not going to say that. One of the most remarkable things about the system of complex numbers which we have described is that it is complete. The process stops here. You can perform any mathematical operation you like on a complex number and the answer will always be just another complex number (or numbers).

But surely, if adding imaginary terms turns the number line into a complex plane, couldn't we just add a super-imaginary term and turn the plane into a solid?

What you are saying is – couldn't we invent a system of three dimensional numbers like (3 + 2**i** + 7**j**)? Well, yes, we could; but there

are two objections to this. First, there is no operation which we can perform on a complex number which *necessitates* a number like this, and second, these numbers (which we might call ternions) just don't *work*. You can define addition and subtraction easily enough but it proves impossible to invent a consistent definition of multiplication and division bearing in mind that the definition must reduce to the complex version whenever the coefficients of either of the imaginary components are zero.

In 1843 Rowan Hamilton,professor of Astronomy at Dublin university, discovered how to make quaternions – e.g. numbers like $(3 + 2\mathbf{i} + 7\mathbf{j} + 4\mathbf{k})$ – work and although these numbers have proved to be very useful in many different fields, they are not strictly speaking, necessary.

If Kronecker had said "God made the complex numbers, all else is the work of Man" I would be more tempted to agree.

Well I have to admit I am quite relieved to hear that we don't have to go any further along the road to more and more complex numbers as I am having enough difficulty as it is to picture the ordinary variety.

I fully understand. And the difficulty becomes extreme when we start to try to picture a complex *function*.

What do you mean?

Well, you know what the function $y = x^2$ looks like, don't you?

Yes, of course.

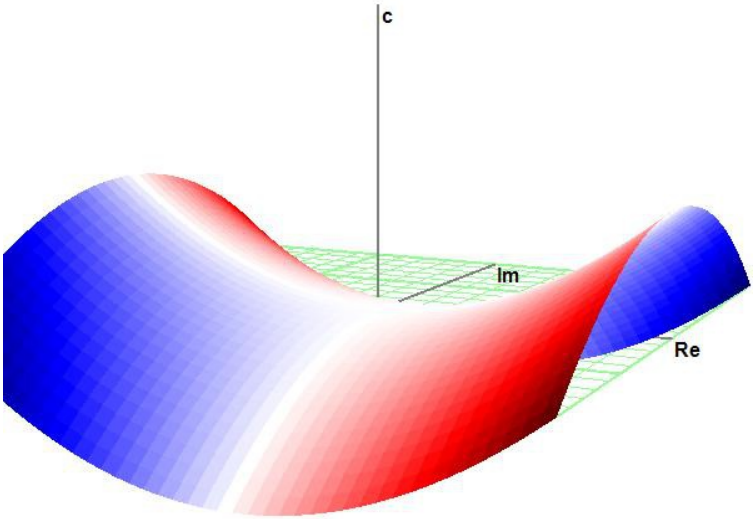
So what do you think the function $\boldsymbol{w} = \boldsymbol{z}^2$ looks like (where $\boldsymbol{w}$ and $\boldsymbol{z}$ are complex numbers)?

I really haven't a clue

The problem is this. Since x and y can be represented by numbers on a line, a graph of $y = x^2$ can be plotted on a two dimensional piece of graph paper; but since both $\boldsymbol{w}$ and $\boldsymbol{z}$ are two dimensional numbers to start with, we need a piece of four dimensional graph paper to plot the function! It was difficult enough plotting the complex roots of x^2 in 3D. This is even worse.

So how can we visualise a complex function?

There are many ways of tackling this difficulty; one of them it to plot a 3D graph of just one aspect of the function using colour to suggest the other. Here is the function $\boldsymbol{w} = \boldsymbol{z}^2$ (where $\mathbf{z} = \mathrm{x} + \mathbf{i}\mathrm{y}$). Below we plot the real part of the function on the vertical axis while the imaginary part is suggested by the colour using the following approximate scale: negative numbers are red shading to white at zero; positive numbers are blue.



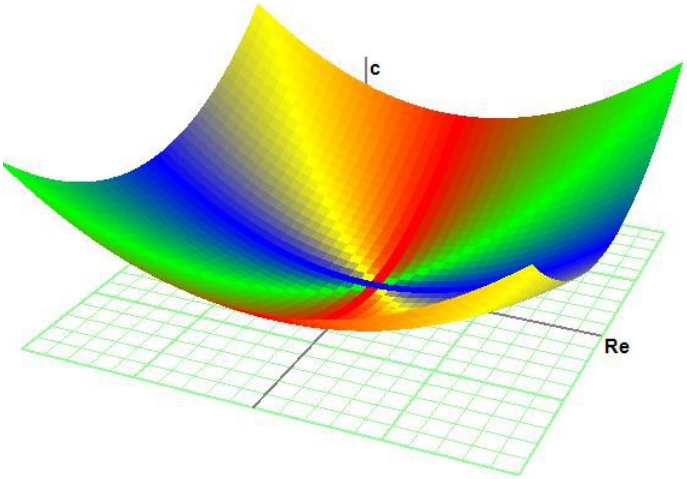
$\boldsymbol{w} = \boldsymbol{z}^2$*: Real part (imaginary part shown in colour)*

On the graph you can see the parabola along the real axis. Here the imaginary part is zero, hence the white stripe. There is also a white stripe along the imaginary axis because the square of any purely imaginary number is real and, of course, the parabola along this axis bends downwards because the square of any purely imaginary number is negative: $(-2\mathbf{i})^2 = (-2)^2 \times (\mathbf{i})^2 = 4 \times -1 = -4$.

The first and third quadrants are blue because the imaginary part of the result is $2xy$. This is positive when x and y have the same sign but negative when they have different signs.

The graph of the Imaginary part is a similar saddle shape but rotated by 45°.

Of even greater interest is the graph of the modulus (i.e. the 'length') of $\boldsymbol{w}$ shown on the vertical axis with the argument (the 'angle') shown in colour. Here the colour scheme is as follows: as before, positive (real) numbers are shown in blue and negative numbers in red; positive imaginary numbers are shown green and negative imaginary numbers in yellow.

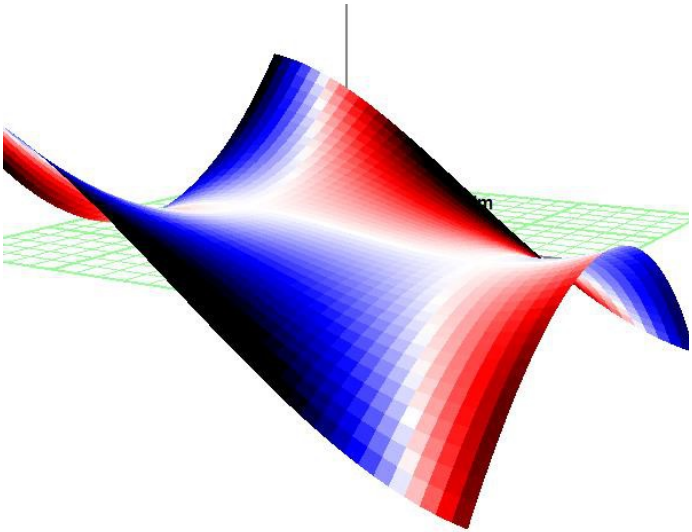


$\boldsymbol{w} = \boldsymbol{z}^2$: Modulus (argument shown in colour)

The real axis is now blue because the result is a wholly real positive number. The imaginary axis is red because the result here is always real and negative. The line at 45° is green because when $x = y$ the real part of the result ($x^2 - y^2$) is zero. The line at 135° is yellow because the real part is also zero here but the imaginary part is negative.

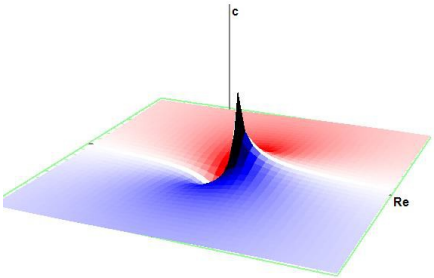
All very colourful! What does z^3 look like?

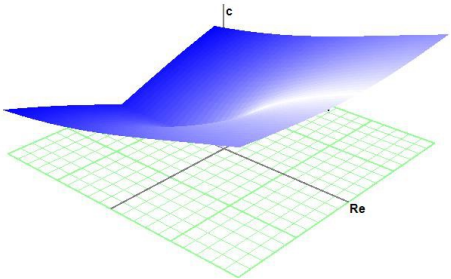
As you might expect, it is a kind of double saddle shape. (I have thrown in an extra term to emphasise the wiggle in the function. You might like to compare this image with the image on page 64. You can see the roots of the equation picked out in white where the imaginary component is zero.)



$\mathbf{w} = \mathbf{z}^3 - 3\mathbf{z}$ *: Real part (imaginary part shown in colour)*

Just for the record, here are graphs of $1/\mathbf{z}$ and $\sqrt{\mathbf{z}}$ too:



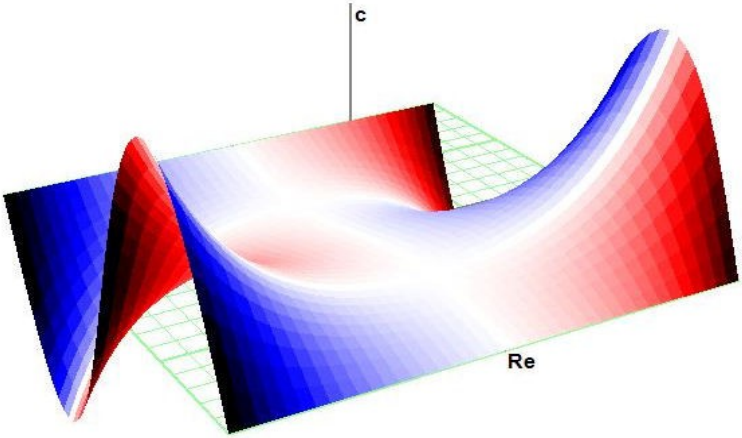


w *= 1/**z** : Real (imag. in colour)*
Note the singularity at z = 0

w *= √**z** : Real (imag. in colour)*
Note this is only the positive half

Well that is pretty amazing. I think you said that you could do pretty much anything to a complex number. Can you show me what sin(z) looks like?

You want sin(***z***) ? Here it is:



w *= sin(**z**) : Real (imag. in colour)*
The wings on each side grow exponentially

Hang on – that was supposed to be a joke! How can you possibly take the sin of a complex number? I thought the sine function was to do with angles?

You are right, of course. But you must admit the graph looks rather pretty so it would be a shame to throw it away and say it is not possible.

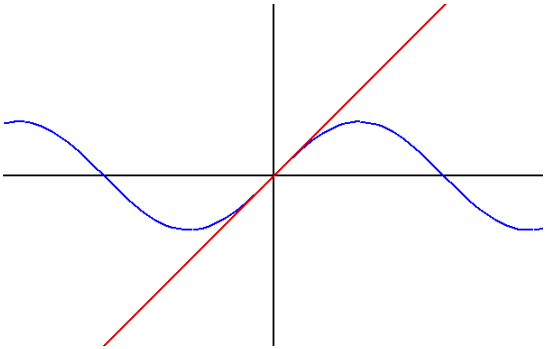
OK – so how did you calculate it?

Normally, if you want to know the sine of an angle you just tap the number into your calculator and up pops the answer. But have you ever thought how the calculator does this? What is the *formula* for sin(x)?

I haven't a clue. Is there one?

Obviously a simple polynomial like $x^2 - 2x + 1$ won't do. That's a parabola and doesn't look anything like a sine curve. On the other hand, a cubic curve has one wiggle a bit like a portion of a sine curve. Perhaps if we add a sufficient number of terms we can *approximate* to a sine curve.

The first thing to notice about sin(x) is that a) it goes through the origin and b) it has a gradient of 1 at the origin. The first approximation to sin(x) is therefore y = x.

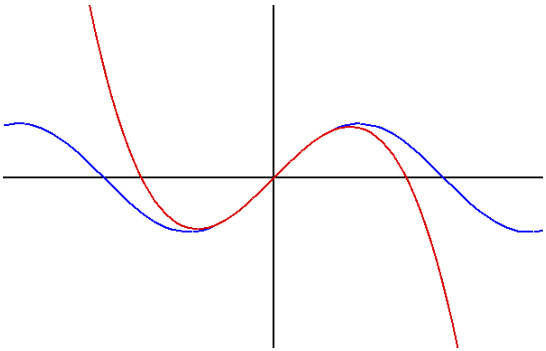


Well I don't think much of that. What about the wiggles?

OK so let's put some wiggles in. A parabola doesn't bend in the right way but a cubic does. The best we can do turns out to be

$$y = x - \frac{x^3}{3!}$$

which looks like this:



That's better. But where did the factorial come from?

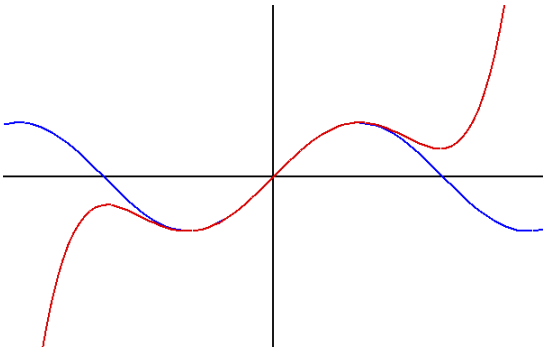
Do you remember the binomial theorem? There were a lot of factorials in that formula and, as we shall see, there is a close link with that formula and what we are discussing today. Bear with me, then, and see if you can guess the next term in the sequence.

Well, I suppose it will be a term in x^5. Could it have a factorial 5 in it?

Not bad. Yes the next term looks like this

$$y = x - \frac{x^3}{3!} + \frac{x^5}{5!}$$

and its graph is a pretty good fit from -90° to +90°



The complete series looks like this. It only contain odd powers of x and the signs alternate.

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \qquad (8)$$

Well that can't be right because the sin of 90 degrees is 1 and if I put x = 90 I can't see it coming out to equal 1

That's because you are using the wrong unit of angle. The formula only works if you use radians instead of degrees.

What are radians?

1 radian is approximately 57° and there are exactly 2π radians in a circle. A right angle is therefore π/2 = 1.571 radians and if you were to substitute x = 1.572 into the above formula it would indeed home in on the answer 1.

OK. I'll take your word for it!

Now what about cos(x). The cosine graph goes through the point (0, 1) and has zero gradient. The first approximation is therefore y =1.

Since the cosine graph looks a bit like an inverted parabola, we can guess that the next term will a function of $-x^2$. This is it:

$$y = 1 - \frac{x^2}{2!}$$

and it should not be too difficult to guess the whole thing:

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \quad (9)$$

I like that. There is a pleasing symmetry in the pair of formulae. They seem to fit like Yin and Yang. But I am not clear where all this is leading. What is the use of this idea?

Well, we are eventually going to use it to calculate the sine of a complex number but first I want to show you something quite remarkable. What I want you to do now is to differentiate sin(x) for me.

*I know the answer to that: cos(*x*)!*

No, that's not what I meant. I meant differentiate the *expansion* of sin(*x*) term by term.

Oh – I see. Well differentiating x gives us 1;

differentiating $\frac{x^3}{3!}$ *gives us* $3\frac{x}{3!}$

which is ?

$$\frac{x}{2!} .$$

Differentiating $\frac{x^5}{5!}$ *gives us* $5\frac{x^4}{5!}$ *which equals* $\frac{x^4}{4!}$

I get it! Every time you differentiate each term, the power reduces by 1 and the factorial reduces by 1 too!

Exactly. So the gradient of sin(*x*) is

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots$$

which is...?

cos(x)!

Isn't that truly amazing! A result like that gives me goose bumps and shivers down my spine! Yes – the gradient of the sine curve is the cosine curve!

Now what about the gradient of the cosine curve?

Hang on – let me guess. What happens to the initial 1 when we differentiate?

The graph of $y = 1$ has zero gradient so this just disappears.

So the result must be:

$$-x + \frac{x^3}{3!} - \frac{x^5}{5!} + \frac{x^7}{7!} + \ldots$$

Wait a minute. That's –sin(x) isn't it?

Correct. And to cut a long story short, differentiating $-\sin(x)$ gives you $-\cos(x)$ and differentiating a fourth time gets you back to $\sin(x)$.

That's amazing. And I can see why you need those factorials too!

Right. I hope all this has convinced you that the series expression we have guessed really are accurate formulae for $\sin(x)$ and $\cos(x)$, at least when x is real.

We are now going to make a big leap and simply assume that what goes for real numbers also goes for complex numbers too. So let's see how to calculate $\sin(\mathbf{i})$.

$$\sin(\mathbf{i}) = \mathbf{i} - \frac{\mathbf{i}^3}{3!} + \frac{\mathbf{i}^5}{5!} - \frac{\mathbf{i}^7}{7!} + \ldots$$

Now we can simplify this because $\mathbf{i}^2 = -1$:

$$\sin(\mathbf{i}) = \mathbf{i} + \frac{\mathbf{i}}{3!} + \frac{\mathbf{i}}{5!} + \frac{\mathbf{i}}{7!} + \ldots \approx 1.175\mathbf{i}$$

*and cos(**i**)?*

$$\cos(\mathbf{i}) = 1 - \frac{\mathbf{i}^2}{2!} + \frac{\mathbf{i}^4}{4!} - \frac{\mathbf{i}^6}{6!} + \ldots$$

$$= 1 + \frac{1}{2!} + \frac{1}{4!} + \frac{1}{6!} + \ldots \approx 1.543$$

That's cool. I love the way the two formulae sort of fit inside each other.

Once again, your intuition is not letting you down because I have just one more series formula which I want to tell you about – and this one is REALLY IMPORTANT!

Do you remember when we were talking about the number *e* we came across the following formula?

$$e \;=\; \text{limit of } \left(1+\frac{1}{n}\right)^{n} \text{ as } n \text{ tends to infinity}$$

and by applying the binomial theorem we found that

$$e \;=\; 1 \;+\; 1 \;+\; \frac{1}{2!} \;+\; \frac{1}{3!} \;+\; \frac{1}{4!} \;+\; \ldots$$ (see page 30)

Now what I want to do is to find a series formula for e^x. This time we have:

$$e^{x} \;=\; \text{limit of } \left(1+\frac{1}{n}\right)^{nx} \text{ as } n \text{ tends to infinity}$$

which becomes

$$e^{x} \;=\; 1 \;+\; x \;+\; \frac{x^2}{2!} \;+\; \frac{x^3}{3!} \;+\; \frac{x^4}{4!} \;+\; \ldots \qquad (10)$$

See those factorials appearing again?

Just to check that this formula does fit the bill, what happens if we differentiate it term by term?

Well, the 1 disappears, x becomes 1, $\frac{x^2}{2!}$ *becomes* $\frac{2x}{2!} = x$,

$\frac{x^3}{3!}$ *becomes* $\frac{3x^2}{3!} = \frac{x^2}{2!}$ *etc. etc.*

Yes it all checks out. Because of those wonderful factorials, every term differentiates into the next term down and the result is that nothing changes. Amazing!

I am glad you find it pleasing.

I shall finish this chapter by calculating $e^{\mathbf{i}}$. I know we have only proved the formula for real x but, as usual, we shall just go ahead and assume that anything you can do the a real number you can do to a complex one too. So here goes...

$$e^{\mathbf{i}} = 1 + \mathbf{i} + \frac{\mathbf{i}^2}{2!} + \frac{\mathbf{i}^3}{3!} + \frac{\mathbf{i}^4}{4!} + \ldots$$
$$= 1 + \mathbf{i} - \frac{1}{2!} - \frac{\mathbf{i}}{3!} + \frac{1}{4!} + \ldots$$

but I shall leave you to gather the terms together and calculate the answer. I think the result will surprise you[15].

15 The answer is $e^{\mathbf{i}} = 0.5403 + 0.8414\,\mathbf{i} = \cos(1) + \mathbf{i}\sin(1)$ where the angle is in radians, of course.

Euler's Equation

We have hacked down the sealed entrance to the house of treasures and been amazed at some of the wonderful things we have discovered so far. But the best is yet to come. We now we approach the holy of holies – the most amazing discovery in the whole history of mathematics; equal in every way to the discovery of Tutankhamen's death mask or Einstein's discovery of the equivalence of matter and energy.

Let us review where we are.

> Complex numbers have the form $x + \mathbf{i}y$ where **i** is the square root of –1.
>
> Complex numbers can also be specified by stating the *modulus r* and the *argument* θ where $r = \sqrt{(x^2 + y^2)}$ and θ = the angle whose tangent is (y/x)
>
> Complex numbers can be added by simply adding the real parts together and adding the imaginary parts.
>
> Complex numbers can be multiplied either using the standard rules of algebra on the real and imaginary parts or by *multiplying their moduli* and *adding their arguments.*
>
> It is possible to calculate transcendental functions like $e^{\mathbf{i}}$ and sin(**i**) by using a suitable series expansion.

Now I want to have a close look at the function $e^{(x + \mathbf{i}y)}$

The first thing to notice is that since $e^{(a + b)} = e^a \times e^b$ it follows that $e^{(x + \mathbf{i}y)} = e^x \times e^{\mathbf{i}y}$.

Now e^x is not a problem. It is simply a real number. And we now know how to expand $e^{\mathbf{i}y}$:

$$e^{\mathbf{i}y} = 1 + \mathbf{i}y + \frac{(\mathbf{i}y)^2}{2!} + \frac{(\mathbf{i}y)^3}{3!} + \frac{(\mathbf{i}y)^4}{4!} + \ldots$$
$$= 1 + \mathbf{i}y - \frac{y^2}{2!} - \mathbf{i}\frac{y^3}{3!} + \frac{y^4}{4!} + \ldots$$

Gathering like terms together...

$$e^{\mathbf{i}y} = 1 - \frac{y^2}{2!} + \frac{y^4}{4!} + \ldots + \mathbf{i}\left(y - \frac{y^3}{3!} + \frac{y^5}{5!}\ldots\right)$$

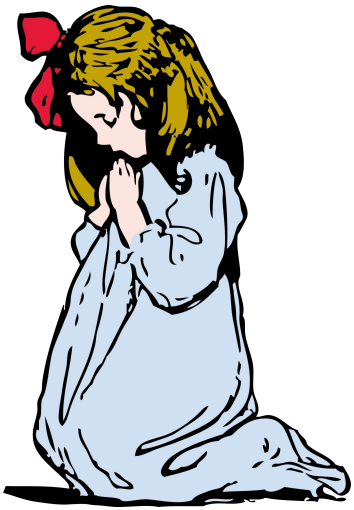
Now look carefully at the real and imaginary parts. Do you recognise them?

Yes I think so. Isn't the first one just the expansion of cosine and the second the expansion of sine?

I am speechless. I can hardly bear to write down the next step. It is like saying the hundredth name of Allah. I am in tears as I write.

$$e^{\mathbf{i}y} = \cos(y) + \mathbf{i}\sin(y)$$

Let us just pause a minute before it in silent prayer.



Well, with the greatest respect and all that, I really don't see what all the fuss is about.

Well perhaps I am overdoing it a bit, but let's look at what we are saying. When we were using the series expansion to calculate $e^{\mathbf{i}}$ I rather glossed over the significance of what we were doing.

What on earth do we mean by '*raising e to the power of a complex number like* **i**'? If you remember we came across exactly the same sort of question when we interpreted expressions like $x^{½}$ and x^{-1}. All we are doing is finding an interpretation which is consistent with everything that we have so far agreed. What is amazing is the *there is only one interpretation* of these expressions which is consistent. The interpretation is *unique*. If there were many ways of interpreting $e^{\mathbf{i}}$ then I wouldn't be nearly so excited about this one. It is the fact that the interpretation which Euler discovered is the only possible one which encourages us to believe that the formula $e^{\mathbf{i}y} = \cos(y) + \mathbf{i}\sin(y)$ is telling us something really important about the actual properties of complex numbers, not merely a definition which we happen to find useful.

Let's see where the equation leads us. Firstly we should put x component back in:

$$e^{(x + \mathbf{i}y)} = e^x \times e^{\mathbf{i}y} = e^x(\cos(y) + \mathbf{i}\sin(y))$$

Now what is the modulus of this number? If you remember, the modulus $r = \sqrt{x^2 + y^2}$ which in this case is:

$$r = \sqrt{(e^x\cos(y))^2 + (e^x\sin(y))^2} = e^x\sqrt{\cos^2(y) + \sin^2(y)} = e^x$$

What about the argument? This is the angle whose tangent is y/x:

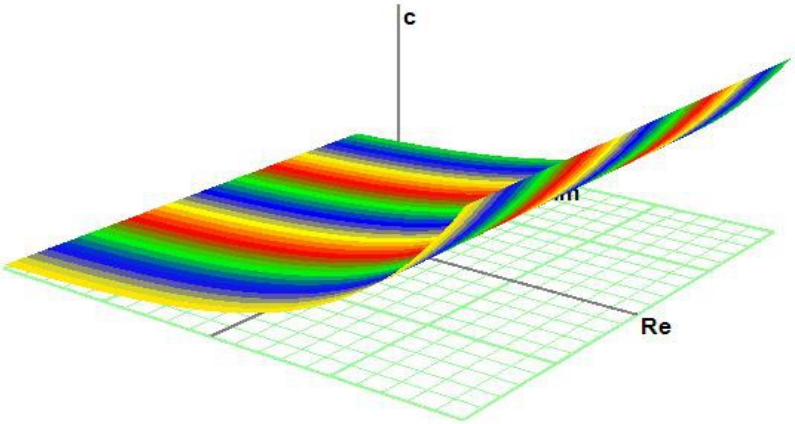
$$\theta = \arctan\left(\frac{e^x\sin(y)}{e^x\cos(y)}\right) = \arctan(\tan(y)) = y$$

The modulus is e^x and the argument it y!!!

So now we have two different ways of expressing a complex number; one in Cartesian coordinates and the other, effectively, in polar coordinates:

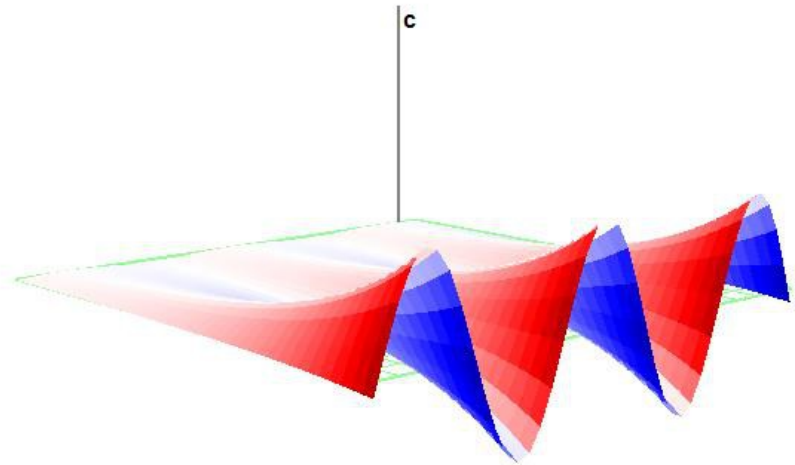
$$z = x + \mathbf{i}y = re^{\mathbf{i}\theta}$$

Here is a 3D graph of the modulus of $e^{(x + iy)}$



It looks a bit like a ski jump. In fact it is simply the familiar exponential curve extended along the imaginary axis and the reason that it is 'flat' is that the modulus *r* only depends on the real part *x* and not on the imaginary part at all.

Now let's look at the real part. This is $e^x \cos(y)$



Can you see the 'ski jump' along the real axis? And the cosine curves

across the imaginary axis?

(The imaginary part looks much the same but the cosine curve is replaced with a sine curve)

Well this is all very interesting but I still don't see why it is so important.

What it signifies is a really deep relation between the exponential function and the trigonometrical functions. Now if you always think of trig function in terms of angles and triangles, Euler's[16] equation (for that is what the equation on page 85 is called) is not a lot of use; but when you consider that the sine and cosine function are fundamental to the analysis of anything that vibrates (and that includes everything from loudspeakers to black holes, you get some idea of the importance of complex analysis. Without Euler's equation it is fair to say you could not design a decent aeroplane or any kind of electronic device. Calculations in quantum mechanics are virtually impossible to do without complex numbers. Essentially it comes down to this. Trig functions are mathematically quite awkward to handle; translate your problem into exponentials using Euler's equation and things often become a whole lot easier.

Take multiplication for example. It took us quite a while to prove that multiplying two complex numbers involved multiplying the moduli and adding the arguments. But if instead of using the Cartesian version $z = x + \mathbf{i}y$ we use the polar version $z = r e^{\mathbf{i}\theta}$ the result is obvious:

$$z_1 \times z_2 = r_1 e^{\mathbf{i}\theta_1} \times r_2 e^{\mathbf{i}\theta_2} = r_1 r_2 e^{\mathbf{i}(\theta_1 + \theta_2)}$$

But, frankly, for the likes of you and me who are never going to be asked to design an aeroplane, or perform quantum calculations, complex numbers are just beautiful. Not knowing about complex numbers would be like seeing in black and white or only hearing the bass drum in an orchestra. Without them, life is not complete. A whole dimension is missing. Complex numbers exists as surely as do negative numbers and fractions and every educated person should be aware of this fact and

16 Leonhard Euler (pronounced 'oiler') was an eighteenth century Swiss mathematician. It is a toss up whether it was Euler or Archimedes who ranks as the greatest mathematician of all time.

know something about them.

As a final thought, let us put $r = 1$ and $\theta = \pi$ into Euler's equation; since $\cos(\pi) = -1$ and $\sin(\pi) = 0$:

$$e^{i\pi} = -1$$

and it is this equation which has often been dubbed 'the most beautiful equation in all of Mathematics' because it contain six of the most important symbols: the exponential constant, the square root of minus one, the trigonometrical constant, equality, negation and unity, all in one elegant formula.

And with that, I will leave you.