Numbers and proof

Tom Elliott

July 23, 2020

Contents

Part I

Numbers and proof

Chapter 1

Integers

Integers

The *natural* or counting numbers which everyone learns very early in life are $1, 2, 3$ and so on.

One can get hung up on the question of whether the natural numbers would exist without the problem of counting a dozen sheep or all twenty of our fingers and toes. Leopold Kronecker famously said "God made the integers; all else is man's handiwork".

We will not worry about where they come from.

Mathematicians refer to the *set* of natural numbers and give that set a special symbol, $\mathbb{N}$. We write

$$\mathbb{N} = \{1, 2, 3 \dots\}$$

The brackets contain between them the elements or members of the set. The dots mean that this sequence continues forever.

How can we decide whether a particular n is in the set if we can't enumerate all of its members? We can tell by its form whether some n is a natural number or not.

If this seems problematic, you might call $\mathbb{N}$ a class instead (Hamming); we carry out *classification* to decide whether n is a natural number.

The notion of an unending sequence can be unnerving upon first encounter.

construction of N

To construct the set $\mathbb{N}$, start with the smallest element, 1. Then

$$1 + 1 = 2$$

$$2 + 1 = 3$$

$$3 + 1 = 4$$

$$\ldots$$

Add successive elements by forming $a_n + 1 = a_{n+1}$.

$\mathbb{N}$ is an infinite set.

We say there is no largest number in $\mathbb{N}$, no largest $n \in \mathbb{N}$. The symbol $\in$ means "in the set" or "is a member of the set".

Proof:

Suppose $\mathbb{N}$ did have a largest member, M.

Well, what about $M + 1$? By the definition we can construct it and it is clearly a member of the set, but $M + 1 > M$ so M is not the largest number in the set.

This is a proof by contradiction that $\mathbb{N}$ is infinite.

□

set membership

Sometimes people say that

$$0 \in \mathbb{N}$$

(0 is a part of the set) but most do not, and we will follow the definition given above. If you wanted to be explicit about this you could write

$$0 \notin \mathbb{N}$$

What do we mean by infinity? We mean an upper bound on the natural numbers, and later, all rational and indeed all real numbers.

All numbers $n \in \mathbb{N}$ have the property that n is contained in the interval $[1..\infty)$. However, ∞ is *not* considered part of the interval, and that is the meaning of the the right parenthesis.

∞ is not a number so it probably doesn't even make sense to write $\infty \notin \mathbb{N}$.

least element

$\mathbb{N}$ does not have a greatest number, but it does have a smallest or least one. If pairwise comparisons are carried out, a single element, the number 1, has the property that $1 \leq n$ for all numbers $n \in \mathbb{N}$. As we go on, we will find that other types of numbers (rationals and real numbers), do not have a least positive number.

well-ordered property

Since we can also find the least member of the set excluding 1, written $\mathbb{N} \setminus 1$, we can order every number in $\mathbb{N}$.

This property is called the **well-ordered** property.

the Integers

The set $\mathbb{Z}$ contains all the members of $\mathbb{N}$ plus their negatives, as well as the special number 0, often called the additive identity since $0 + n = n$ for all $n \in \mathbb{N}$.

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

$\mathbb{Z}$ stands for the German word *Zahlen*, number. The set $\mathbb{Z}$ is usually referred to as the integers.

$\mathbb{Z}$ is also an infinite set and also has the well-ordered property. To show this simply order all numbers $n > 0$ with respect to zero using $<$, and all the numbers $n < 0$ using $>$.

Chapter 2

Arithmetic

divisibility

In this chapter we're working with integers or just natural numbers.

Doing division, computing n/d, means we want to find q such that:

$$n = q \cdot d + r$$

and we want $r \geq 0$ but $r < d$.

An important special case is when $r = 0$. We are concerned to know whether the division is *even* or not:

$$n = q \cdot d, \qquad \frac{n}{d} = q$$

If $r = 0$, then the result of dividing n/d is an integer.

There is a special symbol for that: if d divides n evenly, we write $d|n$. But don't mix up the new symbol | with the division symbol seen in the line before (i.e. n/d).

Now suppose we have another number m that is also divisible by d

$$m = p \cdot d$$

Then

$$m + n = (p + q) \cdot d$$

$$m - n = (p - q) \cdot d$$

- If $d|m$ and $d|n$, d also divides their sum and difference.

In general if

$$a + b = c$$

and d divides any two of them, it divides the third.

In the proof above, we used the principle that additional factors don't matter.

- If $d|n$ then $d|an$, where a is any whole number.

In particular, powers of 10 don't matter. If $d|n$, then d divides any multiple of n, such as $100 \cdot n$.

divisibility by 3 or 9

It is very helpful to decide quickly whether a number n is divisible by a given small prime number like $2, 3, 5, 7, 11$.

Clearly, if n is even, its last digit is one of 02468.

And n is divisible by 5 if its last digit is one of 05.

The first trick is that n is divisible by 3 if its digits add to a multiple of 3.

Example:

91 is not divisible by 3, but 912 is.

When processing digits, we can just subtract and forget any multiples of 3 along the way. With 91, recognize that 9 is divisible by 3, so just forget it and move on.

That leaves 1, which is not divisible by 3, or $1 + 2$, which is.

Proof.

Suppose the number is:

$$\begin{aligned} abcd &= a \cdot 10^3 + b \cdot 10^2 + c \cdot 10^1 + d \\ &= a(999 + 1) + b(99 + 1) + c(9 + 1) + d \end{aligned}$$

As we proved above, if $x|(y + z)$ and $x|y$ then it must be that $x|z$. Since 9 times anything is divisible by 3, it follows that 3 must divide $a + b + c + d$ for $abcd$ to be divisible by 3.

□

A similar thing is true of 9 except that the digits must add only to 9. The proof is the same.

How about 123456789. Add the digits:

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = 9 + 9 + 9 + 9 + 9$$

Do you see the trick there?

So 123456789 is evenly divisible by both 3 and by 9.

casting out nines

As an aside, this trick with digit sums and 9 is the basis for an error-checking method. If two integers are multiplied together $m \cdot n = p$, then the product of the digit sums on the left-hand side must be equal to the digit sum on the right. Two consecutive primes are 11677 and 11681. The digit sums are 4 and 8. The product is then 32, whose sum is 5.

We do the multiplication and obtain 136399037. Since that digit sum is also 5, it is likely that we did not make an error, since if we had, the odds are that the digit sum would be different.

divisibility by 7

There is also a test for 7. Take the last digit of n away from the number. Then double it and subtract. Repeat if necessary. If you reach a multiple of 7, then 7 also divides the larger number.

Example.

Let $n = 3101$. The last digit is 1, double it and subtract:

$$310 - 2 \cdot 1 = 308$$

Repeat

$$30 - 16 = 14$$

So yes, 3101 is divisible by 7. We can check pretty easily now that we know it's worth it.

$$7 \cdot 400 = 2800$$

Subtracting, I get 301.

$$7 \cdot 40 = 280$$

Subtracting, I get $21 = 7 \cdot 3$.

So $7|3101$ and $3101/7 = 400 + 44 + 3 = 443$.

Proof.

Write

$$m = n - 21b$$

This is the subtraction step. We subtract one of b from the last place and twice b from the next to last place.

We use our three rules from the beginning. First, the 0 in the last place goes away, because of what we said about factors of 10 above. Ignore the 0.

Then, we know that $7|(21 \cdot b)$ because $7|21$.

So now we know that $7|n$ if and only if $7|m$.

□

divisibility by 11

Working from *right to left*, add up the digits in odd positions and separately, the digits in even positions. Subtract the larger from the smaller. If the difference is a multiple of 11 (including 0), the number is divisible by 11.

For any number where the multiplication did not require you to carry a 1, this is obvious.

Examples: 77, 121, 198, 1441.

If you do have a carry:

```
11 x 256 = 2816
6 + 8 = 14
1 + 2 = 3
14 - 3 = 11
```

See the end for this special one:

```
11 x 123456789 = 1358024679
9 + 6 + 2 + 8 + 3 = 28
```

```
7 + 4 + 0 + 5 + 1 = 17
28 - 17 = 11
```

We won't prove this rule.

distributive law

The the *distributive property* of multiplication over addition says that, for any numbers a, b, c:

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

We used this principle above in trying compute 3101/7 in our head. Write:

$$3101 = 7 \cdot (400 + \dots)$$

That's the first move. Then what's left is $3101 - 2800$. So we need to find 301/7. And so on.

I can use this principle to check numbers less than 100 that end in 3 for whether they are prime. To begin with, we have:

```
13 23 33 43 53 63 73 83 93
```

We use the 3's trick to remove multiples of 3:

```
13 23 43 53 73 83
```

Since $7 \cdot 10 = 70$, we know that none of $53, 73, 83$ is divisible by 7. I use the times table for the rest: $3 \cdot 7 = 21$ and $6 \cdot 7 = 42$. So none of the numbers in our list is divisible by 7.

Therefore, they are all prime, as we will see soon. The smallest number that is divisible by p but not by primes smaller than p is p^2. So we do not need to check for divisibility by 11.

times table

I know it's boring, but it is extremely helpful to be proficient with multiplication of small numbers:

```
    2  3  4  5  6  7  8   9  10  11  12
 2  4
 3  6  9
 4  8 12 16
 5 10 15 20 25
 6 12 18 24 30 36
 7 14 21 28 35 42 49
 8 16 24 32 40 48 56 64
 9 18 27 36 45 54 63 72  81
10 20 30 40 50 60 70 80  90 100
11 22 33 44 55 66 77 88  99 110 121
12 24 36 48 60 72 84 96 108 120 132 144
```

And it wouldn't hurt to go up as far as 20. The 9's trick can be helpful for that row.

There are all kinds of tricks. Here is one for multiplying n by 11.

- write the first digit of n

- add the first two digits of n and write them

- continue with each pair of digits to the end

- write the last digit of n

The thing that makes it hard is you must "carry" the ones. Consider $n = 123456789$

```
 123456789
1234567890
----------
1357913579
```

That's *almost* right.

But we need another 1 in the columns with $9+8, 8+7, 7+6, 6+5$, and that extra 1 will have to get carried again to the column with $4+3$. Compare

```
1357913579
    1111
1358024679
```

A bit too easy to make a mistake for my taste.

Chapter 3

Primes

prime numbers

As you know, the positive integers larger than 1 are of two types:

- a prime number p has only two factors, p itself and 1
- a composite number has at least one additional factor. Either the number is a perfect square of a prime, or it has additional factors: $n = p_1, p_2 \ldots p_k$.

The first ten primes are:

```
2 3 5 7 11 13 17 19 23 29 ...
```

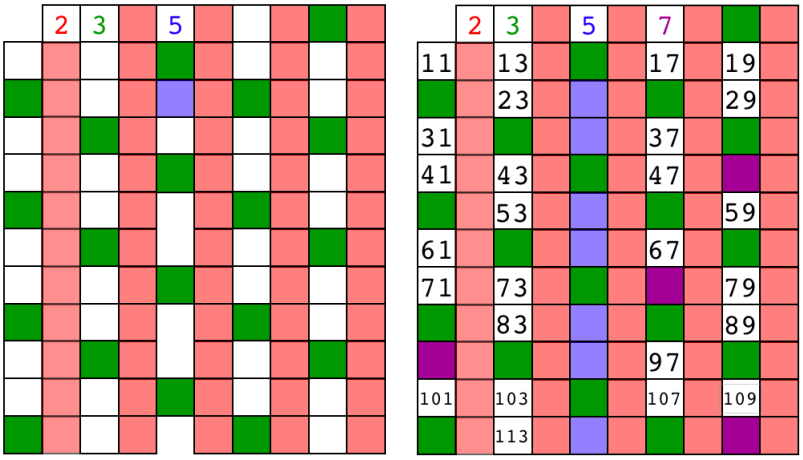
The sieve of Eratosthenes

Eratosthenes is famous in mathematics for his "sieve" which allows one to determine which numbers are prime in an economical fashion.

We will take note of him again in talking about the circumference of the earth. He was a contemporary of Archimedes and became the chief librarian at the Library of Alexandria when he was only about 35 years old.

The sieve operates by first writing down all the integers to some upper limit (here 120). To carry out the process manually it is convenient to use rows with 10 values, so there are 12 rows in all here. Most of the boxes have not yet been numbered (below, left).

Starting with the first prime number, 2, eliminate all the numbers divisible by 2 (all the red numbers, or even numbers). Here this has been done by coloring red all squares with numbers ending in $2, 4, 6, 8, 0$.



Next, do the same thing with 3 (green). 6 was already eliminated previously, but odd multiples of 3 like 9, 15 and 21 go away at this step.

The next larger number that still has a white square is 5. All the squares eliminated at this step are white ones in the fifth row, starting with 25. Continue with 7, eliminating $49, 77, 91$ and 119.

Notice that the smallest number to be eliminated with 7 is $7^2 = 49$, similarly with 5 the first was 25. The first number to be eliminated with q is q^2. This is always true.

The sieve now ends (for this upper bound of 120).

The rule is that at the beginning of a round, test the next candidate prime q by squaring and comparing with the upper limit L. If $q^2 > L$, we terminate. So after that round using 7 , the smallest remaining integer is 11, but we terminate since $11^2 = 121 > 120$.

The graphic shows all the numbers which have yet to be eliminated after the round of 7. All of these numbers, 11, 13, 17, and so on, as well as those used as divisors for each round of the sieve $(2, 3, 5, 7)$, are prime numbers.

By testing for division by $2, 3, 5$ and 7, we have found the first 30 prime numbers.

```
 2   3   5   7  11  13  17  19  23  29
31  37  41  43  47  53  59  61  67  71
```

```
73  79  83  89  97 101 103 107 109 113
```

From a performance standpoint, it is important that we do not need to carry out division. All that is really needed is repeated addition. Coding this algorithm in, say, Python is a good challenge.

A bigger challenge is to come up with a method to *grow* the list of primes on demand. This can be done by keeping track of the first value to be tested above the limit, for each prime in the current list.

recognizing primes quickly

We did a bit of this in the previous chapter, we'll leave this here as review.

There are school problems that require you to factor numbers at least up to 100, maybe more, quickly. It can be helpful to learn to recognize the primes in this range.

- First, primes end in one of the digits: $1, 3, 7, 9$.
- Test quickly for 3 as a factor by the digit sum trick.
- Test 7 as a factor by trial multiplication or the 7 trick we talked about before.

For trial multiplication, let's do the first row, for an example:

```
11 21 31 41 51 61 71 81 91 101 111
```

You should recognize 11 as prime, immediately. Then remove the numbers whose digits add up to 3 or a multiple, leaving

```
31 41 61 71 91 101
```

Then, trial multiplication by 7 to get a number ending in 1.

The way I do this is by computing the difference with a number that is divisible by 7. If the difference is not divisible by 7, then neither is the candidate number.

For example, since we're in the ones column, I'm thinking $7{\cdot}3 = 21$ and so $7{\cdot}(10{+}3) = 70 + 21 = 91$, so we can eliminate 91 as a possible prime.

Furthermore, by comparison with 91:

```
 91 - 71 = 20
 91 - 61 = 30
 91 - 41 = 50
101 - 91 = 30
```

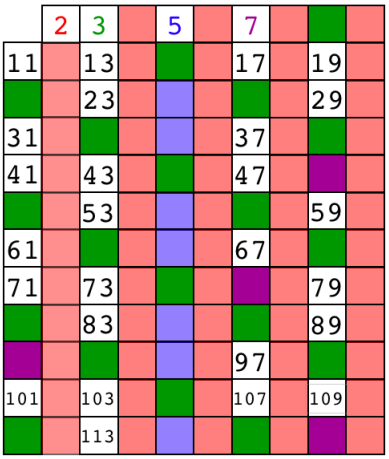
None of these differences is divisible by 7, so the numbers themselves are not, either.

I trust you recognize that 31 is 3 more than $7 \cdot 4$.

We do not need to test any primes larger than 11. All multiples of 11 are repeated double digits $(22, 33 \ldots)$, until 110.

That leaves:

```
31 41 61 71 101
```



infinite primes

Euclid has a theorem and a proof that the number of primes is infinite.

Proof:

By contradiction.

Suppose the set of primes is finite, and that p_1, $p_2 \ldots p_k$ are all of the primes. Construct the following numbers:

$$P = (p_1 \cdot p_2 \cdot \ldots p_k)$$

$$Q = P + 1$$

For a prime number p to evenly divide Q, it must divide the difference between Q and P. But that difference is 1 and so can't be divided evenly by any prime.

Therefore, none of the known primes divides Q and at least one of these is true:

- Q is a prime not in the set of known primes
- the set was originally incomplete

The assumption that the set of primes is finite leads to a contradiction.

□

Even for a relatively small number of primes, we may encounter the second situation. Start with the first prime: 2:

$$2 + 1 = 3 \text{ (prime)}$$

$$2 \cdot 3 + 1 = 7 \text{ (prime)}$$

$$2 \cdot 3 \cdot 7 + 1 = 43 \text{ (prime)}$$

$$2 \cdot 3 \cdot 7 \cdot 43 + 1 = 1807$$

1807 is *not* prime. $(1807 = 13 \cdot 139)$.

testing primality

This is a pretty deep subject. However, a simple filter to apply first is to ask:

- Is the last digit one of $\{0, 2, 4, 6, 8\}$, i.e. is the number even?
- Does the number end in 5?
- Or is the number divisible by 3 or 9?
- Is the number divisible by 7.

We looked at these tricks in a previous chapter ().

A more general observation is that all primes greater than 3 are of the form $4k+1$ or $4k+3$, for integer k. That's because $4k$ and $4k+2$ are even, and $4k+4 = 4(k+1)$.

Any composite number n has a unique prime factorization. Its smallest prime factor p has the property (easily proved):

$$p^2 \leq n$$

Therefore, it suffices to check whether the prime numbers less than or equal to the square root of n divide n. If the square root is not an integer, we need check only

the next smallest integer, what is called the *floor* of the value. If no prime less than that divides n, then n is a prime.

This can be improved still more.

https://en.wikipedia.org/wiki/Primality_test

Chapter 4

Prime factorization

We will prove that every integer has a unique *prime factorization*. This is also called *the fundamental theorem of arithmetic*.

$$n = p_1 \cdot p_2 \ldots p_k$$

In the list of the prime factors of n, a factor may be repeated.

Example:

$$12 = 2 \cdot 2 \cdot 3$$

To compare two factorizations for uniqueness, we suppose they are sorted (say, from smallest to greatest).

More examples for relatively small numbers:

```
 39  = 3.13
144  = 2.2.2.2.3.3
210  = 2.3.5.7
2310 = 2.3.5.7.11
```

Sometimes the factors can be hard to find.

Example:

Let's try 123456789. By using the digit addition trick, we can tell that this number is divisible by 9 (I get $9 + 9 + 9 + 9 + 9$).

At first it seems easy. I found two factors of 3, leaving 13717421.

Then my luck ran out. The smallest prime factor was too large for me to find by hand. So I used Python.

https://gist.github.com/telliott99/3043a0d9ddc44f8503c83c848b2f8382

3607 and 3803 are the two prime factors of that number, which once found, are easily confirmed. Factoring is a hard problem.

background

We will prove the unique prime factorization theorem.

But before starting on that, when we say that one integer *evenly divides* another one, written as $a|n$ or a is a factor of n, we mean there exists another integer k such that

$$a \cdot k = n$$

a times k is exactly equal to n with no remainder.

Take care to distinguish $a|b$ (a divides b) from a/b (a divided by b).

If there is also a number m where a evenly divides m, we write $a|m$ and mean that

$$a \cdot j = m$$

Addition or subtraction of $m + n$ gives

$$m + n = a \cdot j + a \cdot k = a(k + j)$$
$$m - n = a \cdot j - a \cdot k = a(k - j)$$

- If $a|m$ and $a|n$, a also divides their sum or difference.

We rely on this fact below.

Also, we can manually check all the numbers up to some reasonable lower limit, like 100. They all have unique prime factorizations. Therefore, if there is a number with two such factorizations, it is larger than 100, and there must be a smallest such number.

abnormal numbers

Hardy and Wright (*Theory of Numbers*, sect. 2:11) have a proof of prime factorization which I find quite elegant.

Proof.

By contradiction.

Hardy:

> Let us call numbers which can be factored into primes in more than one way, *abnormal*, and let n be the smallest abnormal number.

Start by supposing that there are two different factorizations of n:

$$n = p_1 \cdot p_2 \ldots p_k$$

and

$$n = q_1 \cdot q_2 \ldots q_j$$

where the p's and q's are all primes.

Different factorizations

As a preliminary result, consider the possibility that some factor appears in both factorizations, that some p is equal to a q.

Let us rearrange if necessary so the shared factor is listed first: let $p_1 = q_1$ and

$$n = p_1 \cdot p_2 \ldots p_k$$

$$n = p_1 \cdot q_2 \ldots q_j$$

But now, n/p_1 (n divided by p_1) is abnormal, because it has two different prime factorizations.

That is impossible, because n is the smallest abnormal number.

Therefore, no p is a q and no q is a p. If there exist abnormal numbers with two factorizations, those factorizations must be completely different.

inequality

We may take p_1 to be the least p and q_1 to be the least q. In this part, we establish that $p_1 \cdot q_1 < n$.

Since n is composite, either

- $p_1 \cdot p_1 = n$, or
- p_1 times some number larger than p_1 is equal to n.

In the second case, $p_1 \cdot p_1 < n$.

A similar result holds for q_1.

But, since $p_1 \neq q_1$, only one of p_1 or q_1 at most, can be squared to give n. Either $p_1 \cdot p_1 < n$ or $q_1 \cdot q_1 < n$ or perhaps both are true.

From this it follows that

$$p_1 \cdot q_1 < n$$

the contradiction

Let

$$N = n - p_1 q_1$$

We know that $0 < p_1 q_1 < n$ because of the last section, and because neither is equal to zero.

Therefore $0 < N < n$ also.

We know that N is not abnormal (because n is the smallest abnormal number).

We're given that $p_1 | n$ and so, from the above equality

$$N = n - p_1 q_1$$

and our preliminary result about what factorization means, it must be that $p_1 | N$. The same is true for q_1, namely $q_1 | N$.

Hence both p_1 and q_1 appear in the unique factorization of N, so $p_1 q_1 | N$.

Certainly, $p_1 q_1$ divides itself, so it follows that $p_1 q_1 | n$, using our preliminary result about factorization.

Hence, $q_1 | (n/p_1)$.

But n/p_1 is less than n and has the unique prime factorization $p_2 \cdot p_3 \ldots p_k$.

Since q_1 is not a p, this is impossible.

Hence there cannot be any abnormal numbers.

□

As we said, this is *fundamental theorem of arithmetic*, so it's worth a bit of effort for the proof.

Chapter 5

Induction

the problem

Suppose we have some theorem that we think *might be true* for all numbers n, because we've tried it on a few different values of n and the theorem is true for all of them.

A classic example (Courant and Robbins) is this prime number generator:

$$p(n) = n^2 - n + 41$$

The remarkable function $p(n)$ produces a prime number for integer $0 < n < 41$.

```
  41   43   47   53   61   71   83   97
 113  131  151  173  197  223  251  281
 313  347  383  421  461  503  547  593
 641  691  743  797  853  911  971 1033
1097 1163 1231 1301 1373 1447 1523 1601
1681
```

But, for $n = 41$, the last two terms cancel in

$$p(n) = n^2 - n + 41$$

and then n^2 is divisible by n, thus the result cannot be prime.

By testing them all, I found that 41 is the largest prime smaller than 2000 with this property (I don't know of a proof that no more exist). The primes with this property are:

```
2  3  5 11 17 41
```

Hamming has some other examples of theorems with many true candidates, but which are false. Here is one:

$$f(n) = n(n-1)(n-2)\dots(n-k)$$

$f(n) = 0$ for all $0 \leq n \leq k$, but will never be zero for any other $n > k$.

That is because there are only k zeroes of a kth degree polynomial. (As an aside, this is a consequence of the *fundamental theorem of algebra*).

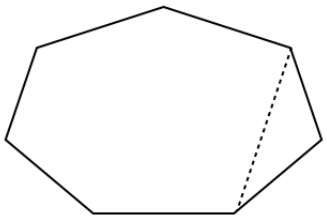
By choosing k large and expanding the definition, we can generate as many true cases as you have patience for.

Furthermore, for any function $g(n)$, $f(n) + g(n)$ will have the same property.

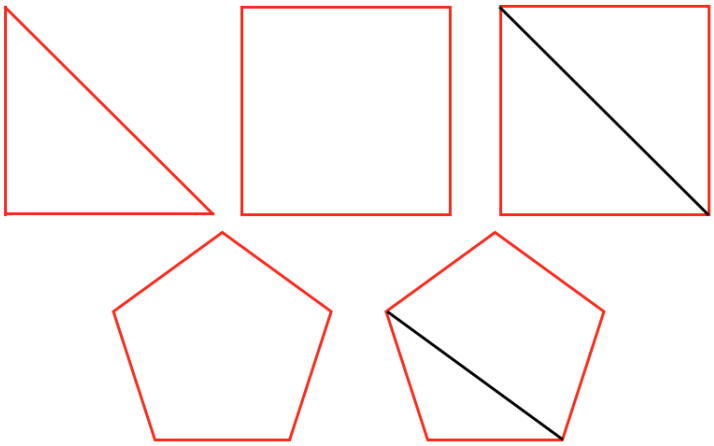
induction in geometry

In the figure below is a polygon—an irregular heptagon. Actually, there are three polygons altogether, there is the heptagon with $n + 1$ sides, the hexagon with only n sides that would result from cutting along the dotted line, and the triangle that is cut off.

We want to find a formula for the sum of the internal angles that depends only on the number of sides or vertices.



The first part of the answer is to guess.



We know that for a triangle ($n = 3$), the sum of the angles is 180°, and the sum does not depend on whether the triangle is acute, right or obtuse.

Continuing with the square ($n = 4$), we can draw the diagonal and observe that the sum of all the angles is twice 180° or 360°. The partition into two triangles can be carried out with any quadrilateral, it does not require any sides being equal.

From this we guess that the formula may be:

$$S_n = (n - 2) \cdot 180$$

And indeed, in going from $n = 4$ to $n = 5$ sides we can think of the pentagon as being a quadrilateral with an extra triangle.

And in the first figure, you can see that by adding the extra vertex to go to the $n+1$-gon, we added a triangle, or perhaps you'd rather say than in going from $n+1$ to n we lost a triangle.

In all cases, the difference between n and $n + 1$ is 180°.

The formula *seems* to work.

We can use induction to *prove* that it is correct.

The proof has two parts. We must verify the formula for a base case like the triangle, which we've done. You may wish to check that it works for the square as well, but that's not strictly necessary.

The second part of the proof is to verify that in going from n to $n+1$, we add another $180°$. The formula for n sides is $(n-2)180°$, adding another triangle gives:

$$(n-2)180° + 180°$$

That must be equal to what the formula gives for $n+1$ sides:

$$((n+1)-2)180°$$

Substituting x for $180°$ and equating the two, we have

$$(n-2)x + x = ((n+1)-2)x$$

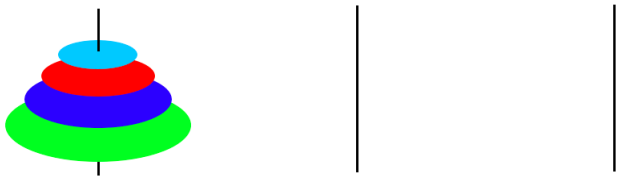
$$n-2+1 = n+1-2$$

$$n = n$$

which is certainly correct.

□

Towers of Hanoi



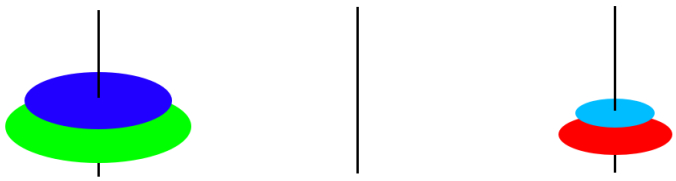
In this famous game the goal is to move a set of disks from one peg to another. Let us choose the one on the right as the target.

https://en.wikipedia.org/wiki/Tower_of_Hanoi

The rules are:

- Only one disk may be moved at a time.
- Each move consists of taking the upper disk from one of the pegs and sliding it onto another peg, on top of the other disks that may already be present on that peg.
- No disk may be placed on top of a smaller disk.

Here is an intermediate stage of the game:



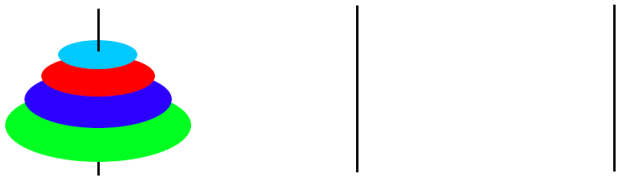
The next move is to place the blue disk on the middle peg. I think you can take it from there.

We can solve the puzzle for any number of disks n.

Proof:

By induction.

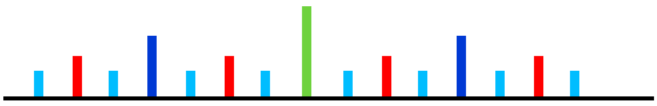
Start from the first position:



Suppose we know how to move $n-1$ disks from one peg to another. Move them to the middle peg, then move the nth disk to the right peg, then place all the $n-1$ disks on top. We have moved n disks.

The base case is to move the single light blue disk. That's trivial. The only thing to watch is if the number of disks is even or odd. If even, choose peg 2, otherwise peg 3.

Which peg is to be moved at each stage is shown in this graphic:



> The puzzle was invented by the French mathematician Édouard Lucas in 1883. There is a legend about a Vietnamese temple which contains a large room with three time-worn posts in it surrounded by 64 golden disks. The monks of Hanoi, acting out the command of an ancient prophecy, have been moving these disks, in accordance with the rules of the puzzle, since

that time. The puzzle is therefore also known as the Tower of Brahma puzzle. According to the legend, when the last move of the puzzle is completed, the world will end.

summary

We can visualize an inductive proof as a kind of chain. We show that the base case is true, for some value of n. Then we show that if the formula works for n , it must work for $n + 1$.

> Mathematical induction proves that we can climb as high as we like on a ladder, by proving that we can climb onto the bottom rung (the basis) and that from each rung we can climb up to the next one (the step).

- Graham, Knuth and Patashnik

[There is a variant called *strong* induction where we know some statement is true for *all* $0 < k \leq n$ and use it to prove something about $n + 1$.]

A few more examples:

sum of digits and divisibility

It is very easy to check whether any number n is divisible by 9. Simply add up all the digits of say, 234783738:

$$2 + 3 + 4 + 7 + 8 + 3 + 7 + 3 + 8$$

$$= 5 + 1 + 1 + 1 + 1 + 1 + 0 + 8$$

$$= 1 + 8 = 9$$

Yes, 234783738 is a multiple of 9.

We propose that

$9|(10^n - 1)$ for all integers $n \geq 0$.

The statement $9|n$ means "9 divides n".

Suppose we know that $9|10^k - 1$ for some n. We mean that

$$10^k - 1 = 9x$$

for some x. Multiply by 10:

$$10 \cdot (10^k - 1) = 10 \cdot 9x$$

$$10^{k+1} - 10 = 9 \cdot 10x$$

$$10^{k+1} - 1 = 9 \cdot 10x + 9 = 9(10x + 1)$$

The right-hand side is clearly divisible by 9, and then so is the left-hand side.

The base case is $9|0$ which is true by definition but may be confusing. Try $n = 1$, then $9|(10 - 1)$ is certainly correct.

□

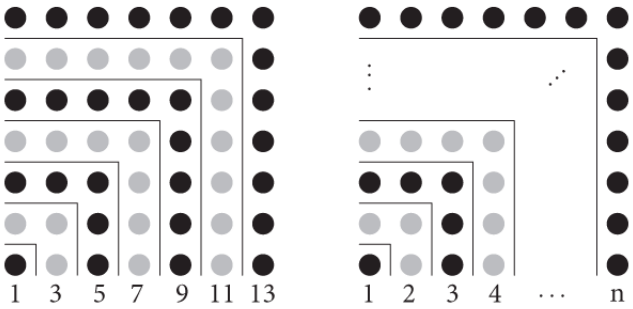
Given this, it is easy to show that the sum of digits method always works.

We demonstrated that this works for 3 previously.

Odd number theorem

Here is a simple but very useful inductive proof.

The *odd number theorem* says that the sum of the first n odd numbers is equal to n^2. Here is a "proof without words".



We prove this by induction.

$$(0 \cdot 2 + 1) + (1 \cdot 2 + 1) + \ldots \cdots + ((n - 1) \cdot 2 + 1) = n^2$$

Notice that the nth odd number is $2 \cdot (n - 1) + 1$.

Our formula says that

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2$$

If you like the summation style:

$$\sum_{k=0}^{n} 2k - 1 = n^2$$

As an example, the first five odd numbers are

$$1 + 3 + 5 + 7 + 9 = 25 = 5^2$$

So, if we consider the next odd number, n changes to $n + 1$. The left-hand side gets another term: we add $2 \cdot (n + 1) - 1$ to it. That is equal to $2n + 1$.

To maintain the equality, add the same quantity to the right-hand side:

$$n^2 + 2n + 1 = (n + 1)^2$$

Rearrange the result, and that's our formula back again. We have proved the inductive step.

To finish, note that the base case is simply

$$1 = 1^2$$

□

The binomial theorem gives the cofactors for a binomial expansion like:

$$(a + b)^5 = a^5 + 5ab^4 + 10a^2b3 + 10a^3b^2 + 5a^4b + b^5$$

We will prove this theorem using induction **here**.

proof of induction

According to Hamming, if you are not convinced by the ladder analogy, here is another proof that induction works:

Suppose the statement is not true for every positive (non-negative) integer. Then there are some false cases. Consider the set for which the statement is false. *If* this is a non-empty set, then it would have a least integer, which is m. Now consider the preceeding case, which is $m-1$. This $(m-1)$th case must be true by definition, and we know that there is such a case because as a basis for the induction we showed that there was at least one true case. We now apply the step forward, starting from this true case $m-1$, and conclude that the next case, case m, must be true. But we assumed that it was *false*! A contradiction.

Therefore, there are no false cases.

□

Chapter 6

Sum of integers

In calculus, we will compute Riemann sums, and to do that we need to find formulas for the sum of squared integers, cubed integers, and so on. To keep it simple, let's start with the integers from 1 to n.

In a previous chapter we introduced the method called induction. Probably the most famous example of an inductive proof is that for the sum of integers.

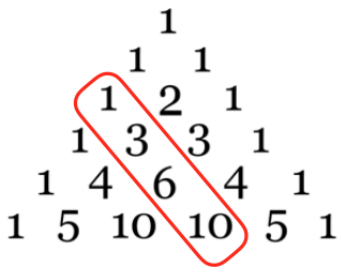
$$S_n = 1 + 2 + \cdots + n$$

Proof.

The numbers we seek are called the triangular numbers. These are

$$1, 3, 6, 10 \cdots$$

These are generated as the third diagonal of Pascal's triangle.:



Suppose someone has sent us, anonymously, a formula which they claim gives the sum of the first n integers, namely

$$S_n = \frac{n(n+1)}{2}$$

Assume the formula is correct for S_n. Add $n+1$ to both sides. The left-hand side becomes S_{n+1}, so we have:

$$S_{n+1} = \frac{(n)(n+1)}{2} + (n+1)$$

Rearranging:

$$= \frac{n(n+1) + 2(n+1)}{2}$$

$$= \frac{(n+1)(n+2)}{2}$$

which is exactly what we'd get by substituting $n+1$ for n in the original formula.

Alternatively, sometimes it's clearer to assume the $n-1$ case and prove the formula is correct for n:

$$S_{n-1} = \frac{n(n-1)}{2}$$

$$S_n = \frac{n(n-1)}{2} + n$$

$$= \frac{n(n-1) + 2n}{2}$$

$$= \frac{n(n-1+2)}{2} = \frac{n(n+1)}{2}$$

So we have proven that if the S_n formula is correct, then so is the one for S_{n+1}.

How do we know that S_n is correct?

Just check the *base case*:

$$S_1 = \frac{1(1+1)}{2} = 1$$

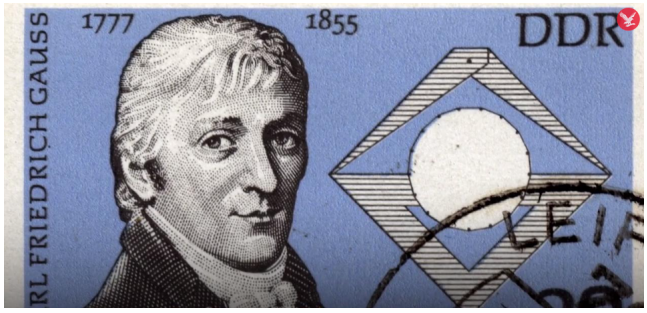
Since S_1 is clearly correct, S_2 must be also, and this continues all the way to S_n.

$$S_1 \Rightarrow S_2 \Rightarrow \ldots S_{n-1} \Rightarrow S_n \Rightarrow S_{n+1}$$

Therefore, it must be true for *every* integer n.

□

There is a famous story about Gauss. As a schoolboy, he "saw" how to add the integers from 1 to 100 as two parallel sums.



Added together horizontally, these two series must equal twice the sum of 1 to 100.

But vertically, we notice that each sum is equal to $n + 1$, and we have n of them.

1	2	...	99	100	S_n
100	99	...	2	1	S_n
101	101		101	101	

So, again

$$2S_n = n(n+1)$$

$$S_n = \frac{1}{2}\, n(n+1)$$

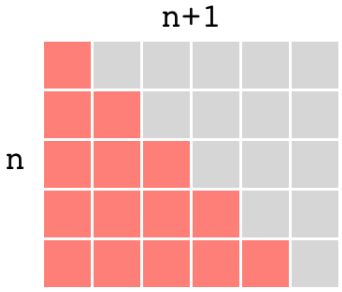
For $n = 100$ the value of the sum is 5050, which is what Gauss wrote on his slate and presented to the teacher immediately on being given the problem as a make-work exercise.

One way of looking at this result is that between 1 and 100 there are 100 representatives of the "average" value in the sequence, which (because of the monotonic steps) is $(100 + 1)/2 = 50.5$.

Or alternatively, view the sum as ranging from 0 to 100 (with the same answer). Now there are 101 examples of the average value $(100 + 0)/2 = 50$).

proof without words

Here is a striking *visual proof* of the formula to obtain T_n, the n^{th} such number. The total number of circles in the figure below is $n \times (n+1)$ and this is exactly two times the sum of the integers from 1 to n.



$$2S = n(n+1)$$

Derivation using sums

It seems a shame to spoil such a beautiful proof "without words" as the one above by saying anything more, but I can't resist. I'd like to derive the equation we have been using using algebra. The general method will help us later.

For any number, and in particular, any integer k it is true that

$$(k+1)^2 = k^2 + 2k + 1$$

So consider what happens if we sum the values from $k = 1 \rightarrow n$ for each of these terms

$$\sum_{k=1}^{n}(k+1)^2 = \sum_{k=1}^{n}k^2 + \sum_{k=1}^{n}2k + \sum_{k=1}^{n}1$$

If the equation is valid for any individual k, then the sum is also valid, plugging in all k up to n.

Rearranging

$$\sum_{k=1}^{n}(k+1)^2 - \sum_{k=1}^{n}k^2 = \sum_{k=1}^{n}2k + \sum_{k=1}^{n}1$$

Now think about the left-hand side in our equation.

$$\sum_{k=1}^{n}(k+1)^2 - \sum_{k=1}^{n}k^2$$

We have a bunch of terms starting with 2^2:

$$2^2 + 3^2 + \cdots + n^2 + (n+1)^2$$

we also have a bunch of terms to subtract starting with 1^2:

$$1^2 + 2^2 + 3^2 + \cdots + n^2$$

Almost everything cancels. This is called a "collapsing" or "telescoping" sum. We have

$$(n+1)^2 - 1 = n^2 + 2n$$

Bringing back the right-hand side we obtain:

$$n^2 + 2n = \sum_{k=1}^{n} 2k + \sum_{k=1}^{n} 1$$

We can bring the constant factor 2 out of the sum, and also, we recognize that the sum of the value 1 a total of n times is just n.

$$n^2 + 2n = 2\sum_{k=1}^{n} k + n$$

Subtract n from both sides and divide by 2:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

That's it!

Part II

The real numbers

Chapter 7

Rational numbers

The integers are great, they give us an infinite supply of numbers.

However, there is a problem with division. For

$$p \in \mathbb{N}, \quad q \in \mathbb{Z}$$

very often the result of $p \div q$ is not contained in $\mathbb{N}$ or even in $\mathbb{Z}$ — the result is not an integer. We say these sets are not *closed* under division.

For example $3 \div 2 = ?$

So, we just leave the result as

$$\frac{p}{q} = \frac{3}{2}$$

where p/q is in "lowest terms", i.e. they have no common factor other than 1. Of course if p and q have a common divisor, then we can divide both top and bottom by the largest common divisor.

q must not be zero because division by zero is not defined.

theorem

• The arithmetic combinations of two rational numbers are rational numbers.

Proof.

This is just basic algebra:

$$\frac{p}{q} + \frac{r}{s} = \frac{ps + rq}{qs}$$
$$\frac{p}{q} - \frac{r}{s} = \frac{ps - rq}{qs}$$
$$\frac{p}{q} \cdot \frac{r}{s} = \frac{pr}{qs}$$
$$\frac{p}{q} \div \frac{r}{s} = \frac{p}{q} \cdot \frac{s}{r} = \frac{ps}{qr}$$

theorem

• Between *any* two rational numbers it is always possible to find another rational number.

Consider two rational numbers, not equal. Let

$$s = \frac{p}{q} \quad t = \frac{p'}{q'}$$

Suppose $s < t$.

The *average* of these two numbers is:

$$r = \frac{1}{2} \left[s + t \right]$$

Then

$$2r = s + t$$
$$2r - 2s = t - s$$

We have that $s < t$, so $t - s > 0$ and then

$$r - s > 0$$
$$r > s$$

A similar argument will show that

$$r < t$$

so

$$s < r < t$$

□

Thus, one can always find a new rational number that lies between two known rational numbers. In particular, there is no *smallest* positive rational number.

decimal representation

Every rational number can be represented as a decimal, using the method called long division.

Consider 1/2

$$2\overline{)1.000}$$

We say that 2 does not *go into* 1, since $2 > 1$, so we have the first part of our result as 0, followed by a decimal point. But 2 does go into 10 exactly 5 times, giving 0.5. The remainder is zero and so the division process terminates.

Consider 1/8.

$$8\overline{)1.000}$$

- 8 goes into 10 once, leaving 2 as remainder
- 8 goes into 20 twice, leaving 4.
- 8 goes into 40 exactly 5 times with no remainder.

The result is 0.125.

The other possibility is that in going through the process a remainder comes up that has been seen previously. If this happens then the sequence will repeat forever.

If we don't terminate with zero, then this must eventually happen, because there are only as many as q possible remainders.

Thus, for example

$$1/7 = 0.142857142857\ldots$$

which contains 142857, repeating.

decimals to fractions

Conversely, every repeating decimal can be represented as a rational number. For example

```
      1 x r =      0.142857142857...
1000000 x r = 142857.142857...
 999999 x r = 142857
```

$$r = \frac{142857}{999999} = \frac{1}{7}$$

since 7×142857 equals 999999 exactly.

You can do this trick with

```
     r = 0.333
10 x r = 3.33
 9 x r = 3
```

$$r = \frac{3}{9} = \frac{1}{3}$$

or even

```
     r = 0.4999
10 x r = 4.999
 9 x r = 4.5
```

$$r = \frac{4.5}{9} = \frac{1}{2}$$

and

```
     r = 0.9999
10 x r = 9.999
 9 x r = 9
```

$$r = \frac{9}{9} = 1$$

This is one of the subtleties of numbers. In what sense can we say that

$$0.5 = 0.4999\ldots$$

$$1 = 0.9999\ldots$$

Most everyone is OK with the example $1/3 = 0.3333\ldots$ but some may be uneasy with the other two.

Ultimately, we justify the result as defined by evaluation of a limit.

Consider 0.9999. If n is the number of places in the result, then as $n \to \infty$ the number being shown approaches 1 as its limit. We'll come back to this after considering the real numbers.

ordering

For two rational numbers a and b there are only three cases:

$$r_1 = r_2, \qquad r_1 < r_2, \qquad r_1 > r_2$$

It is a property of the integers, that if

$$a < b$$

then for $c > 0$

$$ca < cb$$

using that property

$$\frac{p}{q} < \frac{s}{t} \iff pt < qs$$

p/q is less than s/t if and only if $pt < qs$. (If only one of p and q or s and t is negative, associate the minus sign with the numerator).

Ordering of the integers guarantees ordering of the rational numbers. For any rational numbers, if

$$\frac{p}{q} < \frac{s}{t}$$

then for $c > 0$

$$c \cdot \frac{p}{q} < c \cdot \frac{s}{t}$$

For that matter, it is generally true for real numbers (which include integers and rationals) that if

$$a < b$$

for $c > 0$

$$ca < cb$$

Chapter 8

Infinity

infinity

The symbol for infinity is ∞.

In the old days, they used to write things like

$$\frac{1}{0} = \infty, \qquad \frac{1}{\infty} = 0$$

John Wallis wrote $24/0 = \infty$, in 1656, which is when the ∞ symbol was introduced with its current definition. Even Euler argued that $n/0 = \infty$ when it suited him,

It is claimed that the symbol derives from the Roman symbol for 100 million. That's interesting. I never knew any symbols larger than M, for one thousand. And I'm not sure I believe it, but that's what some people say.

According to

https://notevenpast.org/dividing-nothing/

> On 21 September 1997, the USS Yorktown battleship was testing "Smart Ship" technologies on the coast of Cape Charles, Virginia. At one point, a crew member entered a set of data that mistakenly included a zero in one field, causing a Windows NT computer program to divide by zero. This generated an error that crashed the computer network, causing failure of the ship's propulsion system, paralyzing the cruiser for more than a day.

no division by zero

There is a fundamental problem when we set up a division problem and 0 is in the denominator. What goes wrong when we attempt to divide by zero?

$$\frac{a}{0} = ?$$

Well, what do we mean by an expression such as

$$\frac{a}{b} = c$$

By *definition*, we mean that we will try to find c such that

$$c \cdot b = a$$

For the integers, of course, there is the problem of a possible remainder. Let us leave that aside for a minute.

Suppose we have $c \cdot b = a$ but then take b to be very small though not 0. In that case, the number c may get very large. That's OK.

We can make b as small as we wish by making c large enough or vice versa. And we can say that as $b \to 0$, then $c \to \infty$.

But we can't say $a/0 =$ some number.

If there were such a number (say $a/0 = \infty$, infinity), then what about

$$\frac{b}{0} = ??, \qquad \frac{c}{0} = ??$$

It would mean that whatever the expression $b/0$ is equal to, when multiplied by zero, we would obtain any number whatsoever. This makes no sense.

Here is another, perhaps silly, example.

$$0 \cdot 1 = 0$$

$$0 \cdot 2 = 0$$

so

$$0 \cdot 1 = 0 \cdot 2$$

but then

$$1 = 2$$

By definition, we do not allow division by zero.

infinity is not a number

And we can't answer the question what is $2 \cdot \infty$? If we allowed multiplication by ∞ then the only reasonable answer would be

$$2 \cdot \infty = \infty$$

so then also

$$n \cdot \infty = \infty$$

where n is any number. But then say

$$2 \cdot \infty = 3 \cdot \infty$$

so, cancelling

$$2 = 3$$

This would be a mess.

By definition, *infinity is not a number* and division by 0 is *undefined*.

limits

Often people say that calculus is all about limits, and they are certainly where you start in proving the theoretical basis of the field.

We will keep the discussion of limits and ϵ-δ formalism to a minimum for the reasons explained in the Introduction. But let us try to establish an intuitive idea about what we mean when we say "in the limit as $N \to \infty$".

Above we had that there is no greatest integer.

A corollary of that is the limit

$$\lim_{n \to \infty} \frac{(n+1) - n}{n} = 0$$

Why? As n increases without bound, the difference between successive numbers, as a fraction of n, tends to zero.

To get an idea about this, first simplify by multiplying by $1/n$ on top and bottom. Then we have

$$\lim_{n \to \infty} \frac{(1 + 1/n - 1)}{1} = \frac{1}{n}$$

We say that $1/n$ *tends* to zero as $n \to \infty$, and so does $[(n+1) - n]/n$.

Chapter 9

Euclid's algorithm

Consider two natural numbers a and b. Usually a is allowed to be an integer (i.e., it can be negative), but to keep things simple here we will say that $a, b \in \mathbb{N}$, a and b are positive integers.

We can find their *greatest common divisor*, written (a, b). First we write the unique prime factorization of a and b:

```
180 =          2 x 2 x 3 x 3 x 5
140 =          2 x 2 x         5 x 7
gcd(140,180) = 2 x 2 x         5 = 20
```

Pick out the common factors and the $\gcd(a, b)$ will be their product. It is important that we do not need to actually factor a and b.

(We will develop a theorem on unique prime factorization in another chapter).

The algorithm works like this. Find integers $r \geq 0$ and $q > 0$ such that

$$a = b \cdot q + r$$

• If $r = 0$ we are done: b divides a equally. Otherwise

◦ switch $a = b$ and $b = r$ and repeat.

Then b is the gcd of the original a and b.

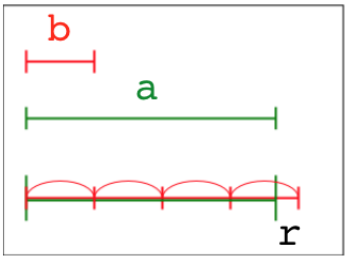
In our example

```
180 = 140 x 1 + 40
140 =  40 x 3 + 20
 40 =  20 x 2 + 0
gcd =  20
```

Here is the reason this works. First, we can always find q and r such that

$$a = b \cdot q + r$$

This is a version of the Archimedean property for positive integers.



It may be paraphrased by saying

> given a bathtub full of water and a teaspoon, it is possible to empty the bathtub.

Either $a = b \cdot q$ and we are done or:

$$b \cdot q < a < b \cdot q + b$$

So then

$$a - bq > 0$$

$$a - bq < b$$

With $r = a - bq$, we obtain $0 < r < b$.

Let u be the largest integer that divides both a and b (the greatest common divisor)

$$a = su$$

$$b = tu$$

Then

$$su = q \cdot tu + r$$

$$r = su - q \cdot tu$$

$$r = u(s - q \cdot t)$$

So u divides r.

Hence every common divisor of a and b is also a divisor of b and r.

recursive program

Here are two examples of programs in different styles that implement the algorithm (with no error checking):

```
def gcd(a,b):
    r = a % b
    if r == 0:
        return b
    return gcd(b,r)
```

```
def gcd(a,b):
    r = a % b
    while r != 0:
        a,b = b,r
        r = a % b
    return b
```

The first version is *recursive*, it may call itself. The second uses a **while** loop to accomplish the same thing.

Chapter 10

Real numbers

There is a big problem with rational numbers which you probably know: some numbers cannot be expressed as the ratio of two integers, as a first example, the number which when multiplied by itself is equal to 2, written $\sqrt{2}$.

The discovery that one cannot find integer p and q such that

$$(\frac{p}{q})^2 = 2$$

is due to the Pythagorean school and was most unwelcome since it screwed up their cherished theory of the universe.

Some say that they drowned the guy who discovered it by throwing him overboard, and that his name was Hippasus. Like most stories about Greek mathematicians, the truth is unknown.

We will see that there is a similar problem (called irrationality) with $\sqrt{3}$, $\sqrt{5}$, $\sqrt{7}$, etc., as well as with $3^{1/3}$ and so on.

Proof.

For $\sqrt{2}$:

We assume that there does exist a rational number p/q such that

$$\frac{p}{q} = \sqrt{2}$$

We will show that this assumption leads to a contradiction.

A crucial part of the proof is that we suppose p/q to be in lowest terms and in particular, that p and q are not both even. It would be easy to recognize the case if they were both even, for then each would have their terminal digit in the set $\{0, 2, 4, 6, 8\}$.

Another fact we will need is that every odd number, when squared, gives an odd result. Proof: every odd number can be written as $2k + 1$ (for non-negative integer k) and then

$$(2k + 1)^2 = 4k^2 + 4k + 1$$

which is an odd number. Therefore, if n^2 is even, n is also even.

So go back to

$$\frac{p}{q} = \sqrt{2}$$

Move the q term to the right-hand side and square both sides:

$$p^2 = 2q^2$$

This implies that p^2 and p are even, using the result from above. So we can write that $p = 2m$. But now

$$(2m)^2 = 2q^2$$
$$2m^2 = q^2$$

which implies that q is *also* even.

We started with the assumption that p and q are not both even, but now we've reached a contradiction. We conclude that there do not exist two integers p and q such that $p/q = \sqrt{2}$.

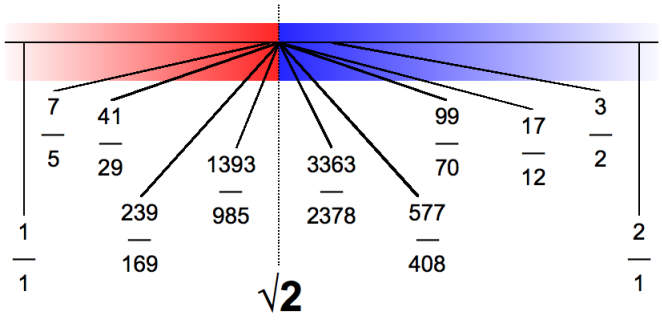
discussion

To quote Hardy (*A Mathematician's Apology*):

> The proof is by reductio ad absurdum, and reductio ad absurdum, which Euclid loved so much, is one of a mathematician's finest weapons. It is a far finer gambit than any chess gambit: a chess player may offer the sacrifice of a pawn or even a piece, but a mathematician offers *the game*.

The numbers like $\sqrt{2}$ are said to be *irrational* numbers and the set of these, plus all the other numbers is called the set of real numbers $\mathbb{R}$.

This led Dedekind to formulate the famous Dedekind cut. Visualize the standard number line as an infinite line on (an infinite) piece of paper.



Each real number corresponds to a cut, a knife-edge coming down somewhere on this number line. Every other number that is not equal to this one, is either $>$ or $<$ the number specified by the cut.

One position is $\sqrt{2}$, another is $3/2$ and so on.

proof using prime factors

The fundamental theorem of arithmetic says that any positive integer greater than 1 can be expressed as a product of its prime factors

$$n = p_1 \cdot p_2 \ldots p_k$$

where this factorization is unique (if the factors are sorted first), and multiple copies allowed. For example

$$60 = 2 \cdot 2 \cdot 3 \cdot 5$$

A corollary says that the square of any integer (a perfect square) has an even number of prime factors since

$$n^2 = p_1^2 \cdot p_2^2 \ldots p_k^2$$

In the expression from above

$$p^2 = 2q^2$$

the number of prime factors on the left is therefore even, but the number on the right is odd. This is a contradiction. Therefore p and q cannot both be integers.

continued fractions

Square roots can be represented as continued fractions. Some smart person figured out that we can write this:

$$(\sqrt{2}-1)(\sqrt{2}+1) = 2-1 = 1$$

Now, rearrange to get a substitution we will use repeatedly

$$\sqrt{2}-1 = \frac{1}{\sqrt{2}+1}$$

Add one and subtract one on the bottom right:

$$\sqrt{2}-1 = \frac{1}{2+\sqrt{2}-1}$$

And substitute for $\sqrt{2}-1$:

$$= \frac{1}{2+\frac{1}{\sqrt{2}+1}}$$

Lather, rinse, and repeat:

$$= \frac{1}{2+\frac{1}{2+\sqrt{2}-1}} = \frac{1}{2+\frac{1}{2+\frac{1}{\sqrt{2}+1}}}$$

Clearly, this goes on forever.

$$\sqrt{2}-1 = \cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\ldots}}}$$

Add 1 to the value of the *continued fraction* to get an expression for the square root of 2.

The numerators are all 1, so this is called a simple continued fraction. The continued fraction representation of $\sqrt{2}$ is usually written as $[1:2]$, meaning that there is an initial 1 followed by repeated 2's.

This fraction goes on forever (since $\sqrt{2}$ is irrational). One can view the existence of the infinite continued fraction as a proof of irrationality.

We can turn the above into an approximate decimal representation of $\sqrt{2}$, by truncating the infinite expansion at the .... Then the last fraction is 5/2. Invert and add, repeatedly:

```
2 + 1/2 = 5/2
2 + 2/5 = 12/5
2 + 5/12 = 29/12
2 + 12/29 = 70/29
2 + 29/70 = 169/70
2 + 70/169 = 408/169
```

To terminate we need to use that initial 1:

```
1 + 169/408 = 577/408 = 1.414216
```

To six places, $\sqrt{2} = 1.414213$. We have five places, and can easily get more.

geometric proof

There are many other proofs of the irrationality of the square root of 2.

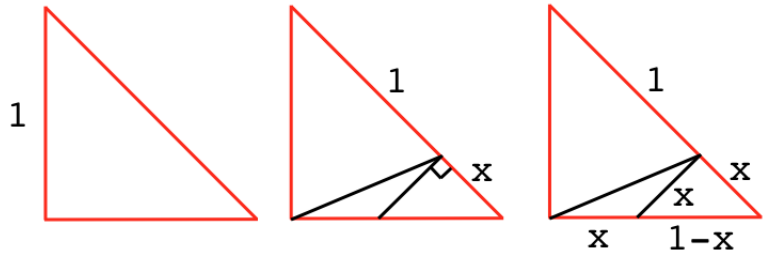
https://www.cut-the-knot.org/proofs/sq_root.shtml

Here we will look at one more, before considering a more general proof for all non-perfect squares. This one is from Tom Apostol (see the link). A more elaborate exposition is:

https://jeremykun.com/2011/08/14/the-square-root-of-2-is-irrational-geometric-proof/

Draw an isosceles triangle with side length 1, then Pythagoras tells us that the hypotenuse is equal in length to $\sqrt{2}$ (left panel).

Our hypothesis is that this length is a rational number, and its ratio to the side is in "lowest terms".



Mark off the length of the side (length 1) on the hypotenuse, and erect a perpendicular (middle panel). Also draw the line segment to the opposite vertex of the original triangle.

The new small triangle that is formed containing the right angle and with side length x in the middle panel is isosceles, because it is a right triangle, and it contains one of the complementary angles of the original right triangle.

By hypothesis, its side length x is the difference of two rational numbers, so x is a rational number.

Furthermore, the *other* small triangle is also isosceles. Its base angles, when added to the equal angles of an isosceles triangle, form right angles. This allows us to mark the side along the base as having length x as well.

Therefore, the hypotenuse of the new, small right triangle is a rational number, since it is equal to $1 - x$.

We are back where we started, with an isosceles triangle that has all rational sides.

It is clear that this process can continue forever. The sides will never be in "lowest terms" because we can always form a new similar but smaller right triangle, which amounts to evenly dividing both the sides and the hypotenuse by a rational number.

general proof

I found a long algebraic proof of the general irrationality of roots and I wrote it up the big calculus book. But then I came upon simple elegant proof based on the fundamental theorem of arithmetic.

We suppose that there exist two integers a and b such that

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

Both a and b have a unique prime factorization. Suppose that gives $a = a_1 \cdot a_2 \ldots a_i$ and likewise for b so:

$$(\frac{a_1 \cdot a_2 \ldots a_i}{b_1 \cdot b_2 \ldots b_j})^2 = n$$

There must be at least some b_j which are not a_i, otherwise we could cancel all of them and so a/b would be an integer.

Call that factor or factors q, so in lowest terms we have

$$(\frac{a_1 \cdot a_2 \ldots}{q_1 \ldots q_k})^2 = n$$

But then, after squaring, we will have q_1^2 (for example) in the denominator and no corresponding factor of either q_1 or q_1^2 in the numerator.

The q_k cannot be canceled and so the result cannot be an integer.

This proves that the only n with rational square roots are perfect squares with integer roots.

The proof also applies generally to other powers like cube and the fourth and fifth power and so on.

other irrational numbers

There are many other irrational numbers besides these square roots.

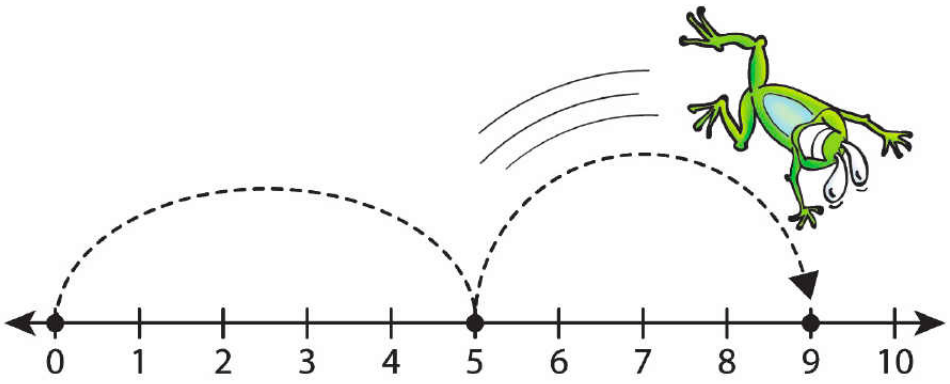
The proof that e is irrational is easy, but since we haven't introduced the exponential yet we need to wait. The proof that π is irrational is bit harder, we will skip that.

density

number line

A simple tool to visualize all of the real numbers is the familiar number line. Here is the number line with numbers marked from $\mathbb{N}$, but obviously we could also draw one for $\mathbb{Z}$ or $\mathbb{Q}$.

We explore the application of the number line to $\mathbb{R}$ as we proceed.



We might simply assume that to every point on the number line there corresponds a rational or irrational number, and that this total collection obeys the same laws of arithmetic as the rational numbers do.

As mentioned above, the need for the real numbers is indicated by empty "holes" in the number line corresponding to the irrational numbers like $\sqrt{2}$.

A problem that arises is how to specify an irrational number non-geometrically and other than as the solution to an equation such as $r^2 = 2$. We saw above a method involving continued fractions.

approximations

In all cases we write particular real numbers as *approximations*. For example, the square root of 2 lies between 1 and 2 because

$$1^2 = 1 < 2$$

$$2^2 = 4 > 2$$

Implying that $\sqrt{2} < 2$. At the second place:

$$1.4^2 = 1.96 < 2$$

$$1.5^2 = 2.25 > 2$$

Implying that $\sqrt{2} < 1.5$. At the third:

$$1.41^2 = 1.9881 < 2$$

$$1.42^2 = 2.0164 > 2$$

Implying that $\sqrt{2} < 1.42$.

This process may be continued for as long as desired.

We can never write down the decimal value of $\sqrt{2}$ exactly, but only approximate it to greater and greater precision. It goes on forever.

In carrying out this recursive process, suppose we know 1.41 and we seek the next digit. One way is to just try all the digits in order starting with 1.

https://gist.github.com/telliott99/79178d6752b6f7b9325476a64bc82953

In this Python code, I have dispensed with the decimal point to make the math easier. We generate a sequence of integers starting with

```
14
141
1414
14142
..
```

This code generates 1000 digits of the value instantly.

However, rather than try all the digits in order starting with 1, there is a better way, which is to try to estimate the error

For example $141^2 = 19881$ so we are short of 20000 by 119.

$142^2 = 20164$ so the difference is 283 and the fraction of the difference that we're under is $119/283 = 0.4205$. In fact, the next two digits of the approximation to $\sqrt{2}$ are 42.

Rather than implement this idea, we will see a better method for obtaining the value in calculus, called the Newton or Newton-Raphson method.

At the seventh place

$$1.414213^2 = 1.9999984093689998.. < 2$$

$$1.414214^2 = 2.0000012377960004 > 2$$

Because any repeating decimal can be written as a fraction, we know that the sequence cannot repeat (any apparent repeat will be illusory).

It is a curious fact that all the digits of π, *to whatever accuracy you desire*, can be found in the correct order, somewhere within the digital expansion of e or ϕ or indeed, any irrational number. The converse is also true.

Another way to say the same thing is that *any* finite sequence can be found within *any* infinite sequence, and in as many copies as you have the patience to discover. The sequence 271828 is found starting around digit 33,790 of π, but 2718281 (adding the next digit of e) is not found within the first million digits of π. You just need more.

limit of a sequence

The real number $\sqrt{2}$ is defined to be the limit of the sequence

$1.4, 1.41, 1.414, \ldots 1.414214 \ldots$

as the number of terms $n \to \infty$.

In a similar way, the number e can be viewed as

$$\lim_{n\to\infty}(1+\frac{1}{n})^n$$

And the number π can be viewed as the limit of the method of exhaustion applied to the area of a unit circle.

density of numbers

We showed previously that between any two rational numbers, including 0 and the *smallest* positive number, one can find another rational number which lies between them.

Three related statements are also true.

- for any two rational numbers one can find a real number which lies between them
- for any two real numbers one can find a rational number which lies between them
- for any two real numbers one can find a real number which lies between them

Proofs of these are readily accessible but we'll do only one, for the sake of brevity.

theorem

• Between any two rational numbers it is always possible to find a real number.

Proof.

We will find a *particular* irrational in the interval (a, b), where a and b are rational. For $a < b$, we simply add to the number a the following

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

c is smaller than $b - a$ (because $\sqrt{2}/2 < 1$) so the result $a + c$ lies between a and b.

We also know that c is irrational, because $\sqrt{2}$ times any rational number is irrational. Finally, $a+c$ is irrational because adding $\sqrt{2}$ times a rational number to any rational number produces an irrational number.

Proof of the first preliminary requirement: $\sqrt{2}$ times a rational is irrational. Suppose for integer p, q, r, s we have

$$\sqrt{2}\ \frac{p}{q} = \frac{r}{s}$$

then

$$\sqrt{2} = \frac{rq}{ps}$$

But the right-hand side is rational, so this is a contradiction.

For the second requirement, again by contradiction suppose

$$\sqrt{2}\ \frac{p}{q} + \frac{s}{t} = \frac{u}{v}$$

for integer p, q, r, s, u, v. But the right-hand side of

$$\sqrt{2} = \frac{q}{p}(\frac{u}{v} - \frac{s}{t})$$

is rational, so this is a contradiction.

□

Note in passing that powers are different. Consider:

$$r = \sqrt{2}^{\sqrt{2}}$$

It is hard to imagine that this is a rational number (and it is actually known to be irrational). But how about

$$r^{\sqrt{2}}$$

then

$$r^{\sqrt{2}} = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^2 = 2$$

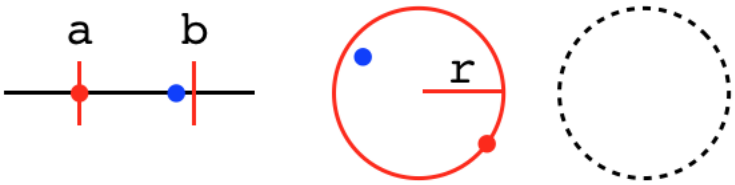
intervals

Not only the real but also the rational numbers, have the property that there is *no closest number* to any given number.

This accounts for virtually all of the theoretical difficulties in calculus which are solved by the use of limits and the apparatus of δ and ϵ or alternatively, neighborhoods.

We denote the numbers greater than a and less than b as lying in the interval (a, b). With parentheses, the interval described is *open*, which means that it does not include the boundary values.

To describe a *closed* interval, write $[u, v]$. This interval includes all the previous values, plus it also includes a and b.



On the number line, the red point is the value a which is the left boundary of the intervals (a, b) and $[a, b]$. a is included in the second but not the first.

For a region in the plane, say a circle of radius r, the red point is on the boundary and the blue point is inside. Such regions are called disks. An open disk is shown with a dashed line as the boundary.

Because of the density property described below, any interval such as

$$I = [0, 1]$$

contains an *infinite* quantity of rational numbers.

Furthermore, it is impossible to specify a number that is as close as possible to 0. There is no smallest possible number.

Proof.

Suppose there does exist a smallest positive number, call it n.

$$n = \frac{p}{q}, \qquad p, q \in \mathbb{N}$$

Consider

$$m = \frac{n}{2} = \frac{p}{2q}$$

$m > 0$ (since all of $p, q, 2$ are > 0) but $m < n$ since $2m = n$.

This is a contradiction.

Therefore, there is no smallest positive number.

□

If someone says to specify a positive number very close to 0, the question you must ask them is "how close do you want to be"? For any ϵ, no matter how small, it is possible to find a δ such that

$$0 < \delta < \epsilon$$

That's the whole game.

density

Consider the set of all points

$$x = \frac{p}{10^n}$$

for all natural numbers n and integers p.

It is clear that simply by increasing the value of n, we can construct a set of equally spaced rational numbers as tightly clustered as we wish.

The rational numbers are said to be *dense* on the number line.

Part III

Algebra

Chapter 11

Basic algebra

There is very little algebra that needs to be memorized for where we're going. We just finished a chapter about the sum of integer squares. If you can follow that, you're in good shape. If not, go back and work through it again, carefully.

Here is a bit more:

inequality

You have surely seen and used the symbols $>$ (greater than), and $<$ (less than) before we used them a second ago.

Among the axioms of the number systems is the collection of *order axioms*. A few definitions:

- $x < y$ means that $y - x$ is positive
- $y > x$ means that $x < y$

For arbitrary numbers a and b only one of three statements is true:

- $a < b$
- $a = b$
- $a > b$

There is no attempt or need to be systematic here. Let us just mention that these properties (and their kin) are true not just for natural numbers, but also for the

rational numbers and the real numbers, as we will see in due course.

Here are a few important theorems about order which we will use often:

- If $a < b$, and c is any number, then $a + c < b + c$
- If $a < b$, then $-b < -a$
- If $a < b$ and $c > 0$, then $ac < bc$

The first one above implies the second and the third.

algebraic operations

- addition: $a + b$
- subtraction: $a - b = a + (-b)$

The negative integers and 0 solve the problem of how to evaluate $a - b$ when $b \geq a$.

- multiplication: $a \cdot b$, also often written ab (but not $a \times b$, at this level).

And then:

- division a/b, equivalent to finding a number c such that $c \cdot b = a$.

algebra

As you know, the basic axioms of algebra include the following:

- Commutativity for addition and multiplication:

$$a + b = b + a, \qquad a \cdot b = b \cdot a$$

- Associativity for addition and multiplication:

$$(a + b) + c = a + (b + c), \qquad (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

- Distributivity of addition over multiplication:

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

- Additive identity: $0 + a = a$.
- Multiplicative identity: $1 \cdot a = a$.

binomial theorem

One basic idea from algebra is the binomial theorem:

$$(a+b)^2 = (a+b)(a+b) = a^2 + 2ab + b^2$$

so then

$$\begin{aligned}(a+b)^3 &= (a+b)(a^2+2ab+b^2) \\ &= a^3 + 3a^2b + 3ab^2 + b^3\end{aligned}$$

And in general, to get the next $n+1$ power, go through the expansion for the n power and multiply each term separately by a and b.

You will find that the cofactors are given by Pascal's triangle.

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & 1 & & 1 & & \\ & 1 & & 2 & & 1 & \\ 1 & & 3 & & 3 & & 1 \\ 1 & 4 & & 6 & & 4 & 1 \\ 1 & 5 & 10 & & 10 & 5 & 1 \end{array}$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

$$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$$

$$\ldots$$

If you substitute $-b$ for b you will find that everything is exactly the same, except those terms with b raised to an odd power have acquired a minus sign.

$$(a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$$

Hence

$$(a-b)(a^2-2ab+b^2) = a^3 - 3a^2b + 3ab^2 - b^3$$

The binomial theorem is usually stated and worked with in terms of positive integers.

But it actually works for negative integers and fractional powers as well. The big difference is that the series of terms is *infinite*.

Newton discovered that, though he didn't prove it. He just used it as a tool. He found that

$$\frac{1}{1+x} = (1+x)^{-1}$$
$$= 1 - x + x^2 - x^3 + x^4 + \cdots$$

Newton checked this by multiplying:

$$(1+x)(1 - x + x^2 - x^3 + x^4 + \cdots)$$
$$= 1 - x + x^2 - x^3 + x^4 + \cdots$$
$$+x - x^2 + x^3 - x^4 + \cdots = 1$$

But be careful! What happens if $x = -1$?

factoring

Here's something we will use often: the difference of two squares.

$$a^2 - b^2 = (a+b)(a-b)$$

The classic quadratic equation is often written

$$y = ax^2 + bx + c$$

This is a parabola that opens up ($a > 0$).

Depending on the values of the *coefficients* a, b, c, this equation may or may not have solutions when $y = 0$

$$ax^2 + bx + c = 0$$
$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0$$

Suppose that r and s are such solutions, then

$$(x-r)(x-s) = 0$$

and so

$$x^2 - (r+s)x + rs = 0$$

A fair amount of effort in algebra goes into guessing values of r and s that work, that have the appropriate sum and difference to match the equation we are given.

However, the answers frequently are not integers, and in that case, this approach is doomed.

A formula that always works is the quadratic formula:

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

I said "always works". To be more precise, it works if there are any solutions. If the discriminant $D = b^2 - 4ac$ is negative, then we're trying to take a square root of a negative number, and we don't know how to do that in the real numbers.

We will develop this more in the chapters on analytic geometry. That's really all you will need.

Chapter 12

Fibonacci sequence

Continuing with the topic of induction, let's introduce the Fibonacci numbers and Binet's formula.

These are numbers in a series formed by adding together the two previous numbers in the series:

$$F_n = F_{n-2} + F_{n-1}$$

or

$$F_{n+1} = F_{n-1} + F_n$$

It remains to choose the first two numbers, which are 1 and 1. Thus the first ten Fibonacci numbers are

$$1 \quad 1 \quad 2 \quad 3 \quad 5 \quad 8 \quad 13 \quad 21 \quad 34 \quad 55$$

Fibonacci example

We will use induction to prove that the sum of the first n Fibonacci numbers is

$$1 + 1 + 2 + \cdots + F_n = F_{n+2} - 1$$

We assume that the formula is correct for F_{n-1}:

$$1 + 1 + 2 + \cdots + F_{n-1} = F_{n+1} - 1$$

Add F_n to both sides

$$1 + 1 + 2 + \cdots + F_n = F_n + F_{n+1} - 1$$

$$= F_{n+2} - 1$$

This completes the induction.

The base case is

$$1 = 2 - 1$$

□

Another way to check this is to write the sum as

```
1 + 1 + 2 + 3 + 5 + ..       .. + F_{n}   + F_{n+1}
    1 + 1 + 2 + 3 + 5 + ..   .. + F_{n-1} + F_{n}
-------------------------------------------------
1 + 0 + 1 + 1 + 2 + ..       .. + F_{n-2} + F_{n-1}
```

Subtracting the second sum from the first, we obtain the third:

$$\sum F_{n+1} - \sum F_n = \sum F_{n-1} + 1$$

$$F_{n+1} = \sum F_{n-1} + 1$$

which rearranges to give the formula.

Binet

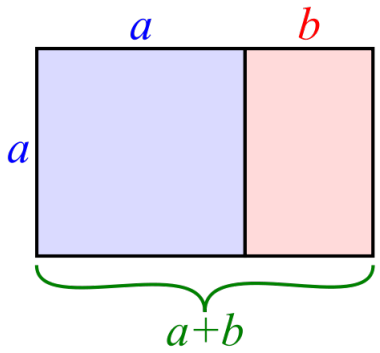
Binet's formula is an explicit formula for F_n which saves us from calculating all the intermediate numbers:

$$F_n = \frac{\phi^n - \psi^n}{\phi - \psi}$$

where ϕ is the Golden Ratio $(1 + \sqrt{5})/2$ and ψ is its conjugate $(1 - \sqrt{5})/2$.

Golden ratio

The basic definition involves the following construction:

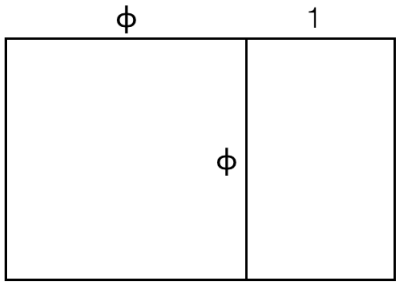


We start with a square of side length a and then extend one side by length b, forming two rectangles. When the ratios of side lengths for these two rectangles are the same, then that ratio is the golden ratio:

$$\phi = \frac{a}{b} = \frac{a+b}{a}$$

(ϕ is often written Φ).

Rescaling of the figure in both dimensions doesn't change the ratios, so let $b = 1$. Then (changing to the symbol ϕ):



$$\frac{\phi}{1} = \frac{\phi+1}{\phi}$$

$$\phi^2 = 1 + \phi$$

We will need only this last result below.

The equation can be solved numerically using the quadratic formula. Put everything on the left-hand side

$$\phi^2 - \phi - 1 = 0$$

Recall that the solutions are

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

We obtain

$$\phi = \frac{1+\sqrt{5}}{2} \qquad \psi = \frac{1-\sqrt{5}}{2}$$

This gives values of $\phi \approx 1.61803$ and $\psi \approx -0.61803$.

It seems like $\phi = 1 - \psi$. Proof:

$$2\phi = 1 + \sqrt{5}$$
$$2\psi = 1 - \sqrt{5}$$

Adding them together

$$2(\phi + \psi) = 2$$

What a nice symmetry:

$$\phi + \psi = 1$$
$$\phi = 1 - \psi$$
$$\psi = 1 - \phi$$

Note that since ψ is a solution it is also true that

$$\psi^2 = 1 + \psi$$

An alternative proof of that is:

$$\psi^2 = (1-\phi)^2$$
$$= 1 - 2\phi + \phi^2$$
$$= 1 - 2\phi + 1 + \phi$$
$$= 2 - \phi = 2 - (1 - \psi) = 1 + \psi$$

We can also do the arithmetic

$$\phi^2 = \frac{1+\sqrt{5}}{2} \cdot \frac{1+\sqrt{5}}{2} = \frac{6+2\sqrt{5}}{4} = 1 + \phi$$

$$\psi^2 = \frac{1-\sqrt{5}}{2} \cdot \frac{1-\sqrt{5}}{2} = \frac{6-2\sqrt{5}}{4} = 1 + \psi$$

back to the proof

Wikipedia says that you can prove Binet's formula using induction.

$$F_n = \frac{\phi^n - \psi^n}{\phi - \psi}$$

https://en.wikipedia.org/wiki/Mathematical_induction#Example:_Fibonacci_numbers

That is an entertaining challenge. For induction we assume that the formula is correct for F_{n-1} and F_n and must prove that:

$$F_{n+1} = F_n + F_{n-1}$$

$$= \frac{\phi^n - \psi^n}{\phi - \psi} + \frac{\phi^{n-1} - \psi^{n-1}}{\phi - \psi}$$

$$= \frac{(\phi^n + \phi^{n-1}) - (\psi^n + \psi^{n-1})}{\phi - \psi}$$

Although one could imagine it is more complicated, the simple idea is to try to show that

$$\phi^{n+1} = \phi^n + \phi^{n-1}$$

and the same for ψ, and that will complete the proof of the inductive step.

Write

$$\phi^{n+1} = \phi^2 \ \phi^{n-1} = (1 + \phi) \ \phi^{n-1} = \phi^{n-1} + \phi^n$$

similarly

$$\psi^{n+1} = \psi^2 \ \psi^{n-1} = (1 + \psi)\psi^{n-1} = \psi^{n-1} + \psi^n$$

This shows that the inductive step is valid.

Now we just need to verify the base cases. We should check at least the first two of them, because there are two values in the recursion formula $F_n + F_{n-1}$.

It's a matter of convenience whether we consider the series to start with $n = 1$ or $n = 0$. If the latter, then the zeroth Fibonacci number is 0, and the first is 1 and we obtain the same series.

For $n = 0$ we have $\phi^0 - \psi^0$ which is just zero.

For $n = 1$, we have

$$\frac{\phi^1 - \psi^1}{\phi - \psi} = 1$$

We decide to continue with $n = 2$

$$\phi^2 = \phi + 1$$

$$\psi^2 = \psi + 1$$

so

$$\frac{\phi^2 - \psi^2}{\phi - \psi} = \frac{\phi + 1 - \psi - 1}{\phi - \psi} = 1$$

This completes the proof.

□

At this point we note the curious pattern

$$\phi^2 = \phi + 1$$

$$\phi^3 = \phi^2 + \phi = 2\phi + 1$$

$$\phi^4 = \phi^2 + 2\phi + 1 = 3\phi + 2$$

$$\phi^5 = 2\phi^2 + 3\phi + 1 = 5\phi + 3$$

$$\phi^6 = 4\phi^2 + 4\phi + 1 = 8\phi + 5$$

so it looks like

$$\phi^n = F_n\ \phi + F_{n-1}$$

The coefficients *are* the Fibonacci numbers.

The same is true for ψ since $\psi^2 = \psi + 1$.

We could prove this by induction, and it would be a proof of Binet's formula as well because the coefficient of ϕ^n or ψ^n is equal to F_n so

$$\phi^n = F_n\ \phi + F_{n-1}$$

$$\psi^n = F_n\ \psi + F_{n-1}$$

and then

$$\frac{\phi^n - \psi^n}{\phi - \psi} = \frac{(F_n\ \phi + F_{n-1}) - (F_n\ \psi + F_{n-1})}{\phi - \psi}$$

$$= F_n\ \frac{\phi - \psi}{\phi - \psi} = F_n$$

Chapter 13

Exponential

Principal and interest

Suppose I put 100 dollars in the bank, and the people at the bank say that after one year, they will give me an additional \$10 at that time. They will pay 10% interest for the year on the principal P of \$100.

However, suppose I bargain with them. I get them to promise to pay me half the interest (5%) at the six-month mark, and the rest after one year. My account will hold \$105 after six months, and the interest due for the second half will be 5% of \$105, which is \$5.25 for a total of \$10.25.

The equation to describe this situation is that if the rate of interest for the year is r and the year is broken up into n periods when interest will be paid, the total amount at the end will be:

$$A = P(1 + \frac{r}{n})^n$$

In the example, we have $r = 0.10$ and $n = 2$ so

$$A = 100(1 + 0.05)^2 = 110.25$$

This is compound interest. If there are additional years t, the exponent will be nt rather than n.

And now we start wondering what happens if the bank pays every month so that $n = 12$ or every day so $n = 365$ or even every second. What happens if the interest

is compounded *continuously*?

$$A = \lim_{n\to\infty} P\ [\ (1+\frac{r}{n})^n\ ]$$

Now it turns out that in the limit as n approaches ∞ these two expressions are equal

$$(1+\frac{r}{n})^n = \ [\ (1+\frac{1}{n})^n\ ]^r$$

The same factor r can be either in the numerator of the second term inside or up in the exponent outside.

A quick proof is:

$$\lim_{n\to\infty}(1+\frac{r}{n})^n$$
$$= \lim_{n\to\infty}(1+\frac{r}{n})^{(n/r)r}$$

Define $m = n/r$ and so as $n \to \infty$, so does $m \to \infty$ and then we have

$$\lim_{m\to\infty}(1+\frac{1}{m})^{(m)r}$$

and the r is outside. m is just a dummy variable so we write:

$$\lim_{n\to\infty}\ [\ (1+\frac{1}{n})^n\ ]^r$$

□

Therefore, going back to what we were working on, let us bring out the factor r and obtain

$$A = P(1+\frac{1}{n})^{nr}$$
$$A = P\ [\ (1+\frac{1}{n})^n\ ]^r$$

Thus, the important question is, what is the value of this expression?

$$A = \lim_{n\to\infty}(1+\frac{1}{n})^n$$

It does not depend on r. It will turn out that this limit is equal to the number e.

$$e = 2.71828\ 18284\ 59045\ldots$$

That's really all we need to worry about with respect to e, for now.

As far as general exponents go, I'm sure you know that:

$$x^a \; x^b = x^{a+b}$$

$$(x^a)^e = x^{ae}$$

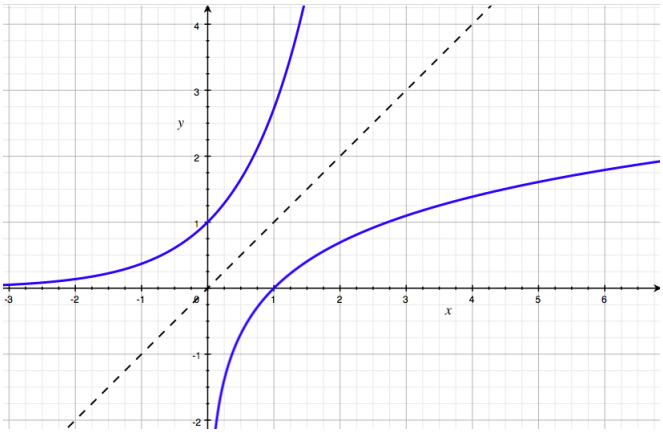
$$x^{-a} = \frac{1}{x^a}$$

$$x^{1/2} = \sqrt{x}$$

Chapter 14

Logarithms

The logarithm and the exponential are inverse functions, we can see that if we plot them together:



The upper curve is $y = e^x$ and the lower one is $y = \ln x$. As inverse functions, they are symmetric about the line $y = x$.

If we have that

$$y = b^x$$

for some $b > 0, b \neq 1$, then we say that

$$x = \log_b y$$

Putting them together

$$y = b^{\log_b y}$$

The usual bases are

- 10 (common logarithm, $\log_{10}$, or just log)
- e (natural logarithm or ln)
- 2 (binary logarithm, $\log_2$).

The rules for exponents are simple, if p and q are two numbers and we know the logarithms of p and q to base b

$$p = b^u$$

$$q = b^v$$

then their product can be computed as:

$$pq = b^u \cdot b^v = b^{u+v}$$

To multiply two numbers, *add* their logarithms.

It helps if we can actually compute b^{u+v}. In the old days there were tables of logarithms, so you just looked up the answer in the table.

LOGARITHMS, BASE 10 $\log_{10} x$ **or** $\lg x$

x	0	1	2	3	4	5	6	7	8	9
10	.0000	0043	0086	0128	0170	0212	0253	0294	0334	0374
11	.0414	0453	0492	0531	0569	0607	0645	0682	0719	0755
12	.0792	0828	0864	0899	0934	0969	1004	1038	1072	1106
13	.1139	1173	1206	1239	1271	1303	1335	1367	1399	1430
14	.1461	1492	1523	1553	1584	1614	1644	1673	1703	1732
15	.1761	1790	1818	1847	1875	1903	1931	1959	1987	2014
16	.2041	2068	2095	2122	2148	2175	2201	2227	2253	2279
17	.2304	2330	2355	2380	2405	2430	2455	2480	2504	2529
18	.2553	2577	2601	2625	2648	2672	2695	2718	2742	2765
19	.2788	2810	2833	2856	2878	2900	2923	2945	2967	2989
20	.3010	3032	3054	3075	3096	3118	3139	3160	3181	3201
21	.3222	3243	3263	3284	3304	3324	3345	3365	3385	3404
22	.3424	3444	3464	3483	3502	3522	3541	3560	3579	3598
23	.3617	3636	3655	3674	3692	3711	3729	3747	3766	3784
24	.3802	3820	3838	3856	3874	3892	3909	3927	3945	3962

$\log 2 \approx 0.3010$.

https://en.wikipedia.org/wiki/Mathematical_table#Tables_of_logarithms

The second rule is for exponentiation:

$$(b^u)^v = b^{uv}$$

And in terms of logarithms we can write

$$\begin{aligned} uv &= \log_b (b^u)^v \\ &= \log b^{uv} = v \log_b (b^u) \end{aligned}$$

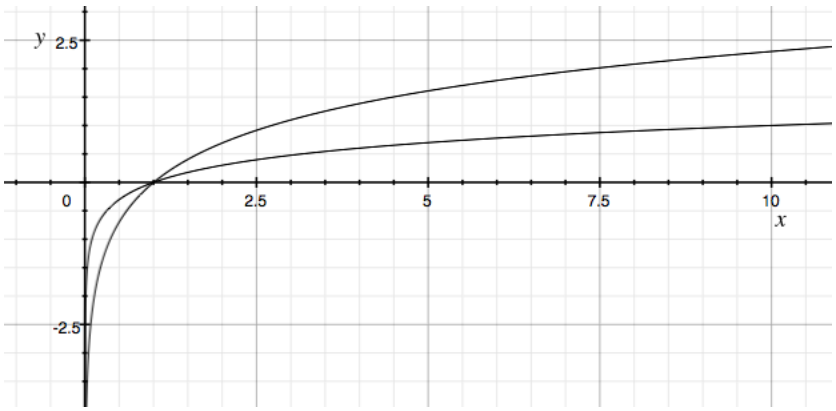
For example

$$\begin{aligned} 2^2 &= 2 \cdot 2 = 4 \\ 2^3 &= 2 \cdot 2 \cdot 2 = 8 \\ 4 \cdot 8 &= 2^2 \cdot 2^3 = 2^{2+3} = 2^5 \\ &= 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 32 \end{aligned}$$

and

$$(2^2)^3 = 4^3 = 64 = 2^6 = 2^{2 \cdot 3}$$

Here is a plot of $\log_{10}(x)$ and $\ln x$:



The first function reaches the value 1 when $x = 10$ and the second reaches the value 1 when $x = e$. Both have the value 0 at $x = 1$ because $b^0 = 1$ for any base, so the logarithm to any base of 1 is equal to 0.

It turns out that if we take the logarithm of x (where x is any number > 1) to two *different* bases, the ratio of the logarithms is a constant, independent of the value of x. And it is not hard to imagine that the ratios of the two values for any x is a constant, in the plot above.

change of bases

This relationship is nicely shown by the change of bases formula.

$$\log_b x = \frac{\log_a x}{\log_a b}$$

Derivation.

Start with an expression with b as the base:

$$y = b^x$$

and by the definition of the logarithm

$$x = \log_b y$$

To derive the formula, take the logarithm to the base a on both sides of the first expression:

$$\log_a y = \log_a(b^x)$$

Now, just invoke the second rule on the right-hand side

$$= x \log_a b$$

and then substitute for x from the second expression above

$$\log_a y = \log_b y \log_a b$$

We're basically done.

□

y can be any value, so replace it by x

$$\log_a x = \log_b x \log_a b$$

Rearranging:

$$\log_b x = \frac{\log_a x}{\log_a b}$$

Three ideas for remembering the formula.

(1) Learn the derivation.

(2) Logarithms of x to different bases b and a are connected by some constant k

$$\frac{\log_b x}{\log_a x} = k$$

$$\log_b x = k \log_a x$$

and we substitute for k the inverse of the log to the *same* base as we have in the numerator:

$$\log_b x = \frac{1}{\log_a b} \cdot \log_a x$$

that is, I remember that we want $\log_a$ something *over* $\log_a$ something on the right.

(3) You might look at the other formula

$$\log_a x = \log_a b \log_b x$$

and imagine the b's canceling in some way.

One other thing we can do is to set $x = a$ in the above formula. We start from

$$\log_b x = \frac{\log_a x}{\log_a b}$$

then with $x = a$

$$\log_b a = \frac{\log_a a}{\log_a b}$$

but $\log_a a = 1$ so

$$\log_b a = \frac{1}{\log_a b}$$

And that makes perfect sense. If we multiply by some factor k to convert from the logarithm in base a to base b, we must multiply by the inverse of the same factor to convert back again.

For the figure above of the common log (base 10) and the natural logarithm, $\ln 10 = 2.303$, and that looks about right, when $x = 10$ the first function is 1.0 and the second one is about 2.3.

fractional exponents

The introduction above dealt mainly with integer exponents, but of course you know that the practical use of logarithms depends on fractional values. The simplest way to see how this works is to consider the square root.

$$\sqrt{2} \times \sqrt{2} = 2$$

If we think about what the exponent u to the base 2 would be such that

$$2^u = \sqrt{2}$$

We observe that by the rules for exponents

$$\sqrt{2} \times \sqrt{2} = 2^u \times 2^u = 2^{u+u} = 2^1$$

That is

$$u + u = 1$$

so $u = 1/2$. By the same logic the n^{th} root of b is $b^{1/n}$. And of course

$$(b^2)^{1/2} = b^{2\times 1/2} = b^1$$

fractional exponents

Feynman has a nice description of how logarithms were calculated (see Lectures, volume 1, Chapter 22, Algebra)

http://www.feynmanlectures.caltech.edu/I_22.html

The basic idea is to take repeated square roots of the base (10), and then combine those to form the required value.

So for example

$$10^{1/2} = 3.1622776602$$

this has been rounded up, the next term in the expansion is $3.162277\ldots$. To obtain $10^{1/4}$, compute the square root of $10^{1/2}$.

$$10^{1/4} = 1.7782794100$$

$$10^{1/8} = 1.3335214321$$

$$10^{1/16} = 1.1547819847$$
$$10^{1/32} = 1.07460782832$$
$$10^{1/64} = 1.03663292844$$
$$10^{1/128} = 1.0181517217$$
$$10^{1/256} = 1.009035044841$$
$$10^{1/512} = 1.004507364254$$
$$10^{1/1024} = 1.002251148293$$

and so on. Eventually (around 10^{1024}) we stop, because there is a neat trick that saves us from calculating. It's a result from basic calculus so we won't say more.

Having these powers of 10, now we want to compute a logarithm, say $\log_{10} 2$. The first thing is that 2 is smaller than $10^{1/2}$.

So we start with

$$2 = 10^{1/4} \cdot \; ??$$

By trying the various powers of 10, we settle on

$$2 = 10^{1/4} \cdot 10^{1/32} \cdot 10^{1/64} \cdot 10^{1/256} \cdot r$$

We won't worry about the r, it is very close to 1.

To get the logarithm of 2 just add the logs of those powers:

$$\log_{10} 2 = 0.25 + 0.03125 + 0.015625 + 0.00390625 = 0.30078125$$

The actual value is 0.30103, to five places.

You can see that to get good accuracy, we will need to figure out the correction term.

Less than 1

Fractional exponents leads to consideration of $0 < x < 1$. Write

$$x \cdot \frac{1}{x} = 1$$

Take the logarithm of both sides

$$\log(x \cdot \frac{1}{x}) = \log 1 = 0$$

$$= \log x + \log \frac{1}{x}$$

Thus

$$\log \frac{1}{x} = -\log x$$

Chapter 15

Sum of squares

We want to find a formula for the sum of the squares of the first n integers, which is written variously as

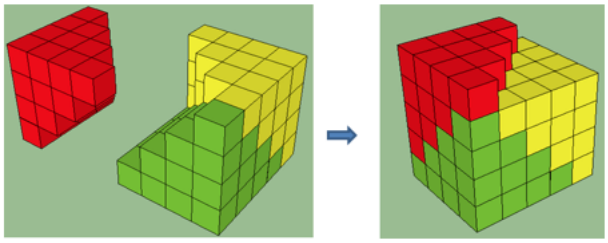
$$\frac{n\ (n+1)(2n+1)}{6}, \qquad \frac{n(n+1)}{2} \cdot \frac{(2n+1)}{3}$$

$$\frac{1}{6}\ (2n^3+3n^2+2n), \qquad \frac{n^3}{3}+\frac{n^2}{2}+\frac{n}{3}$$

and my favorite:

$$\frac{1}{3}\ n \cdot (n+1)(n+\frac{1}{2})$$

which explains the proof without words:



sum of squares

We us use exactly the same method as for the sum of integers to determine the formula for

$$S_n = 1^2 + 2^2 + \ldots n^2$$

Since the formula for the sum of integers has a square, we expect that this will be a cubic. Write

$$(k+1)^3 = k^3 + 3k^2 + 3k + 1$$

$$\sum_{k=1}^{n}(k+1)^3 = \sum_{k=1}^{n} k^3 + \sum_{k=1}^{n} 3k^2 + \sum_{k=1}^{n} 3k + \sum_{k=1}^{n} 1$$

Moving the first term from the right-hand side to the left, we obtain a telescoping sum as the difference, just like before:

$$(n+1)^3 - 1 = \sum_{k=1}^{n} 3k^2 + \sum_{k=1}^{n} 3k + \sum_{k=1}^{n} 1$$

Now it's just some messy algebra. The second term on the right-hand side includes our previous result:

$$n^3 + 3n^2 + 3n = 3\sum_{k=1}^{n} k^2 + 3\ \frac{n(n+1)}{2} + n$$

$$6\sum_{k=1}^{n} k^2 = 2(n^3 + 3n^2 + 3n) - 3n(n+1) - 2n$$

$$6\sum_{k=1}^{n} k^2 = n\ [\ 2(n^2 + 3n + 3) - 3(n+1) - 2\ ]$$

$$6\sum_{k=1}^{n} k^2 = n\ [\ 2n^2 + 3n + 1\ ]$$

$$6\sum_{k=1}^{n} k^2 = n\ (n+1)(2n+1)$$

$$\sum_{k=1}^{n} k^2 = \frac{n\ (n+1)(2n+1)}{6}$$

which can be re-written in various ways including:

This formula is also written variously as

$$\frac{n(n+1)}{2} \cdot \frac{(2n+1)}{3}$$

$$\frac{1}{6}\ (2n^3 + 3n^2 + 2n)$$

$$\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{3}$$

But I find it easiest to remember the first version.

We can check it by induction. The base case is easy

$$\frac{1(2)(3)}{6} = 1$$

Now for the induction step:

$$\frac{n(n+1)(2n+1)}{6} + (n+1)^2$$

$$= \frac{n+1}{6}\ [\ (n)(2n+1) + 6(n+1)\ ]$$

Look at what's in the brackets

$$(n)(2n+1) + 6(n+1)$$

$$= 2n^2 + 7n + 6$$

$$= (n+2)(2n+3)$$

$$= (n+1+1)(2(n+1)+1)$$

So altogether we have

$$= \frac{(n+1)(n+1+1)(2(n+1)+1)}{6}$$

which indeed, is the formula we had above, substituting $n+1$ for n.

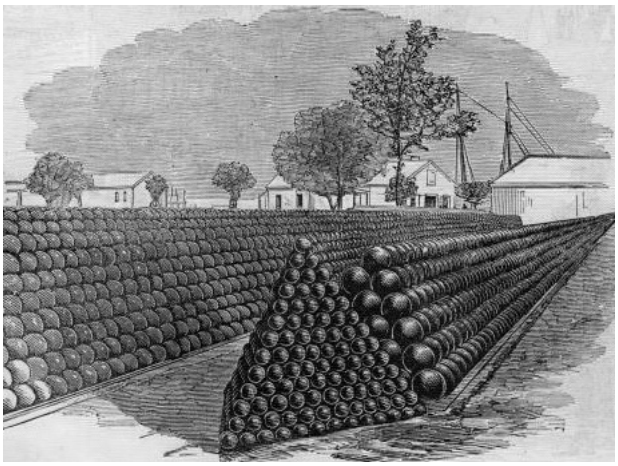
□

Strang's proof

Here is another approach, from Strang's *Calculus*. He says "the best place to start is a good guess". So again, our goal is to find a formula for:

$$S = \sum_{k=1}^{n} k^2$$

Perhaps we visualize a pile of cannonballs



Each layer contains a square number of cannonballs (1, then 4, then 9, etc.). The shape is a pyramid with dimensions $n \times n \times n$. We know the formula for the volume of a pyramid, and guess

$$S_n = \frac{1}{3}n^3$$

To test it, check whether this difference is n^2 (as it should be):

$$S_n - S_{n-1} = \frac{1}{3}n^3 - \frac{1}{3}(n-1)^3$$

Now

$$(n-1)^2 = n^2 - 2n + 1$$
$$(n-1)^3 = (n-1)(n^2 - 2n + 1)$$

$$= n^3 - 3n^2 + 3n - 1$$

So

$$S_n - S_{n-1} = \frac{1}{3}(n^3 - n^3 + 3n^2 - 3n + 1)$$

We see that our guess is off by the residual terms

$$\frac{1}{3}(3n^2 - 3n + 1)$$
$$= n^2 - n + \frac{1}{3}$$

Strang says: the guess needs *correction terms*. To cancel $1/3$ in the difference, subtract $n/3$ from the sum. And to add back n in the difference, add back $1 + 2 + \cdots + n(n+1)/2$ to the sum. Our new guess is

$$S_n = \frac{1}{3}n^3 + \frac{n(n+1)}{2} - \frac{n}{3}$$
$$= \frac{n}{6}(2n^2 + 3(n+1) - 2)$$
$$= \frac{n}{6}(2n+1)(n+1)$$
$$= \frac{n(n+1)(2n+1)}{6}$$

which may be easier to remember as

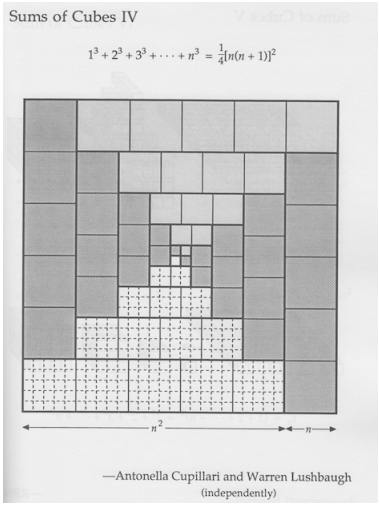
$$S_n = \frac{n(n+1)}{2} \cdot \frac{2n+1}{3}$$

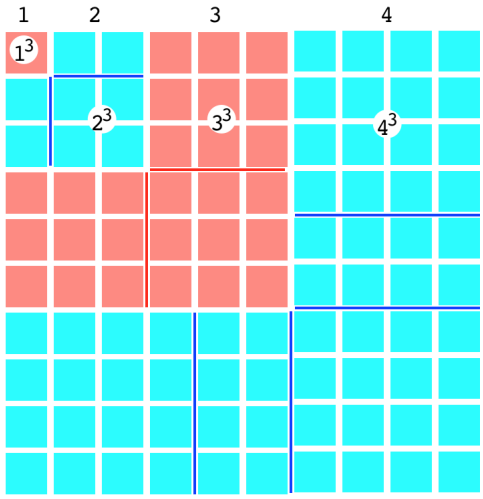
Chapter 16

Sum of cubes

The formula is

$$\sum_{k=1}^{n} k^3 = [\ \sum_{k=1}^{n} k\ ]^2$$





Proof.

By induction.

The "base case" is pretty simple. For $n = 2$

$$1^3 + 2^3 = 1 + 8 = 9$$

and

$$\frac{n^2(n+1)^2}{2^2} = \frac{2^2(3^2)}{2^2} = 3^2 = 9$$

Now for the induction step what we need to show is that what we get assuming the formula for n is correct and then adding the term $(n+1)^3$

$$\frac{n^2(n+1)^2}{2^2} + (n+1)^3$$

is equal to what we get by plugging $n+1$ into the formula.

$$\frac{(n+1)^2(n+2)^2}{2^2}$$

We need to show that eqn 2 is equal to eqn 3.

$$\frac{n^2(n+1)^2}{2^2} + (n+1)^3 = \frac{(n+1)^2(n+2)^2}{2^2}$$

First, we can factor out and cancel $(n+1)^2$ from both sides. So then we have

$$\frac{n^2}{2^2} + (n+1) \stackrel{?}{=} \frac{(n+2)^2}{2^2}$$

$$n^2 + 4(n+1) \stackrel{?}{=} (n+2)^2$$

That looks correct!

□

derivation by collapsing sum

We proceed exactly as before

$$(k+1)^4 = k^4 + 4k^3 + 6k^2 + 4k + 1$$

Sum each term from $k = 1 \to k = n$

$$\sum_{k=1}^{n}(k+1)^4 = \sum_{k=1}^{n} k^4 + \sum_{k=1}^{n} 4k^3 + \sum_{k=1}^{n} 6k^2 + \sum_{k=1}^{n} 4k + \sum_{k=1}^{n} 1$$

Rearrange and compute the collapsing sum.

$$\sum_{k=1}^{n}(k+1)^4 - \sum_{k=1}^{n} k^4 = \sum_{k=1}^{n} 4k^3 + \sum_{k=1}^{n} 6k^2 + \sum_{k=1}^{n} 4k + \sum_{k=1}^{n} 1$$

$$(n+1)^4 - 1 = \sum_{k=1}^{n} 4k^3 + \sum_{k=1}^{n} 6k^2 + \sum_{k=1}^{n} 4k + \sum_{k=1}^{n} 1$$

Substitute for the right-hand sum

$$(n+1)^4 - 1 = \sum_{k=1}^{n} 4k^3 + \sum_{k=1}^{n} 6k^2 + \sum_{k=1}^{n} 4k + n$$

Rearrange some more

$$\sum_{k=1}^{n} 4k^3 = (n+1)^4 - 1 - \sum_{k=1}^{n} 6k^2 - \sum_{k=1}^{n} 4k - n$$

Expand the term $(n+1)^4$ and pick up the $-1-n$:

$$\begin{aligned} &(n+1)^4 - 1 - n \\ &= n^4 + 4n^3 + 6n^2 + 4n + 1 - 1 - n \\ &= n^4 + 4n^3 + 6n^2 + 3n \end{aligned}$$

Factor out an n

$$= (n)(n^3 + 4n^2 + 6n + 3)$$

And another $n+1$

$$= (n)(n+1)(n^2 + 3n + 3)$$

Recall our previous results:

$$\sum_{k=1}^{n} 6k^2 = 6 \sum_{k=1}^{n} k^2$$
$$= 6 \ \frac{n(n+1)(2n+1)}{6}$$
$$= n(n+1)(2n+1)$$

Similarly

$$\sum_{k=1}^{n} 4k = 4 \sum_{k=1}^{n} k$$
$$= 4 \ \frac{n(n+1)}{2}$$
$$= 2n(n+1)$$

Substitute all three of these results (and pull out the factor of 4 from the sum):

$$4 \sum_{k=1}^{n} k^3 = (n)(n+1)(n^2+3n+3) - n(n+1)(2n+1) - 2n(n+1)$$

Just a bit more algebra. See that we have $n(n+1)$ in each term. We have

$$= n(n+1) \ [\ (n^2+3n+3) - (2n+1) - 2 \ ]$$
$$= n(n+1) \ [\ n^2+3n+3-2n-1-2 \ ]$$
$$= n(n+1) \ [\ n^2+n \ ]$$
$$= n(n+1) \cdot n(n+1)$$

So all together we have

$$4 \sum_{k=1}^{n} k^3 = n(n+1) \cdot n(n+1)$$
$$\sum_{k=1}^{n} k^3 = \frac{n(n+1)}{2} \cdot \frac{n(n+1)}{2}$$

$$\sum_{k=1}^{n} k^3 = \left[\frac{n(n+1)}{2} \right]^2$$

A remarkable simplification!

Looking deeper

$$\sum_{k=1}^{n} k^3 = \left[\sum_{k=1}^{n} k \right]^2$$

We want to try to understand something more about why this is true.

A web search revealed the answer. Here's an interesting pattern for the cubes of integers

$$1^3 = 1$$

$$2^3 = 8 = 3 + 5$$

$$3^3 = 27 = 7 + 9 + 11$$

$$4^3 = 64 = 13 + 15 + 17 + 19$$

$$5^3 = 125 = 21 + 23 + 25 + 27 + 29$$

If you want a formula for n^3, notice that the first term is $n^2 - n + 1$ and the last term is $n^2 - n + 2n - 1$, and the number of terms for each sum equals n. (There are n odd numbers between 1 and $2n - 1$).

In other words, the sum of all the cubes of integers from 1^3 to n^3 is equal to the sum of all the odd numbers up to $n^2 - n + 2n - 1 = n^2 + n - 1$.

How many of these numbers are there? A little thought should convince you that the answer is $(n^2+n)/2$. For example, with $n = 5$, our last odd number is $5^2+5-1 = 29$, and we have $(25 + 5)/2 = 15$ terms.

We want the sum of the first $(n^2 + n)/2$ odd numbers.

Let's look at another pattern

$$1 = 1$$

$$2^2 = 4 = 1 + 3$$

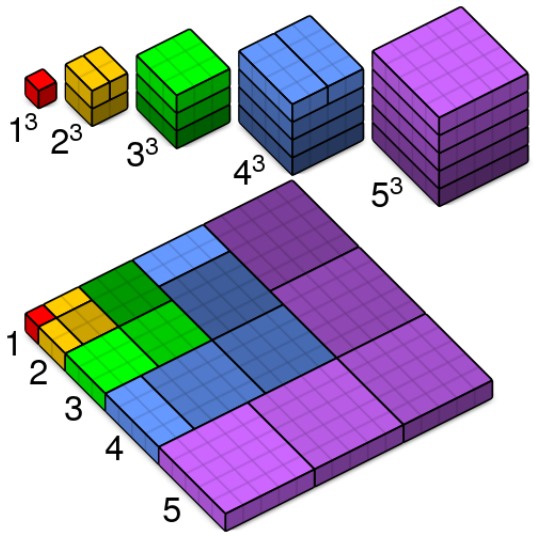
$$3^2 = 9 = 1 + 3 + 5$$

$$4^2 = 16 = 1 + 3 + 5 + 7$$

$$5^2 = 25 = 1 + 3 + 5 + 7 + 9$$

The *odd number theorem* says that the sum of the first n odd numbers is equal to n^2. We want the sum of the first $(n^2 + n)/2$ odd numbers, so that's $((n^2 + n)/2)^2$. And that's how we get our formula.

Here is another beautiful proof without words:



The length of the bottom pattern is a triangular number, which is itself a sum of squares. When squared it equals the sum of cubes.

another method

Here is another approach that I found very early in Hamming (Chapter 2) and have not seen in other books. It is called the method of *undetermined coefficients*.

We observe that the sum of integers formula has order n^2, while the sum of squares has order n^3, so we expect the sum of cubes would have n^4.

$$\sum_{k=0}^{k=n} k^3 = an^4 + bn^3 + cn^2 + dn + e$$

and if $n = 0$ the sum is zero so $e = 0$.

The right-hand side is

$$an^4 + bn^3 + cn^2 + dn$$

The inductive step is to write the formula for $m-1$, and then add m^3 to it.

The right-hand side is just the formula, writing m for n

$$am^4 + bm^3 + cm^2 + dm$$

The left-hand side is the formula for $(m-1)$, plus m^3 from the induction step:

$$a(m-1)^4 + b(m-1)^3 + c(m-1)^2 + d(m-1) + m^3$$

We work with the left-hand side. Expand each term using the binomial theorem:

$$a\ [\ m^4 - 4m^3 + 6m^2 - 4m + 1\ ]$$

$$b\ [\ m^3 - 3m^2 + 3m - 1\ ]$$

$$c\ [\ m^2 - 2m + 1\ ]$$

$$d\ [\ m - 1\ ]$$

Next, group the cofactors by the corresponding powers:

$$[\ a\ ]\ m^4$$

$$[\ -4a + b + 1\ ]\ m^3$$

$$[\ 6a - 3b + c\ ]\ m^2$$

$$[\ -4a + 3b - 2c + d\ ]\ m$$

$$a - b + c - d$$

Now to the point. The cofactors for *each power* of m must cancel exactly.

am^4 cancels on left and right, likewise bm^3, cm^2 and dm. That leaves four equations.

$$-4a + 1 = 0$$

$$6a - 3b = 0$$

$$-4a + 3b - 2c = 0$$

$$a - b + c - d = 0$$

We find that $a = 1/4$, $b = 1/2$, $c = 1/4$, $d = 0$. So then finally the formula is

$$an^4 + bn^3 + cn^2 + dn$$

$$= \frac{n^4 + 2n^3 + n^2}{4}$$

$$\frac{(n^2 + n)^2}{2^2} = \left[\frac{n(n+1)}{2} \right]^2$$

which is exactly what we will have from other approaches.

Hamming uses this method to get a general formula, but we will not need that, because we will show how to use the binomial theorem to get what is necessary.

Part IV

Extra

Chapter 17

Pythagorean triples

In the previous chapter we derived what are called the double-angle formulas:

$$\sin 2s = 2 \sin s \cos s$$

$$\cos 2s = \cos^2 s - \sin^2 s$$

We will manipulate these to find expressions in terms of the same variable, using the following identity:

$$\sin^2 \theta + \cos^2 \theta = 1$$

$$\tan^2 \theta + 1 = \frac{1}{\cos^2 \theta}$$

sine

$$\begin{aligned} \sin 2s &= 2 \sin s \cos s \\ &= 2 \frac{\sin s}{\cos s} \cos^2 s \\ &= 2 \tan s \ \frac{1}{1 + \tan^2 s} \end{aligned}$$

Let $a = \tan s$, then

$$\sin 2s = \frac{2a}{1 + a^2}$$

cosine

$$\cos 2s = \cos^2 s - \sin^2 s$$
$$= \left[\frac{\cos^2 s}{\cos^2 s} - \frac{\sin^2 s}{\cos^2 s} \right] \cos^2 s$$
$$= \left[\frac{1 - \tan^2 s}{1 + \tan^2 s} \right]$$

so

$$\cos 2s = \frac{1 - a^2}{1 + a^2}$$

triples

In general, a can be anything. But if a is a rational number, then we can obtain the corresponding sides of a right triangle with rational lengths as well.

The sides are: $2a, 1 - a^2$ with the hypotenuse:

$$\sqrt{4a^2 + (1 - 2a^2 + a^4)}$$
$$\sqrt{1 + 2a^2 + a^4)}$$
$$= 1 + a^2$$

Suppose $a = \frac{2}{3}$. Then, we have side lengths: $\frac{4}{3} = \frac{12}{9}, \frac{5}{9}$, and $\frac{13}{9}$, which can be converted to integers: $12, 5, 13$.

In general, if $\tan s = p/q$ then the sides are

$$\frac{2p}{q}, \qquad 1 - \frac{p^2}{q^2}, \qquad 1 + \frac{p^2}{q^2}$$

which as integers will be

$$2pq, \qquad q^2 - p^2, \qquad q^2 + p^2$$

This formula was found by Euclid.

https://en.wikipedia.org/wiki/Pythagorean_triple

If p and q are two odd integers the sum and difference of squares is even so we can write

$$pq, \qquad \frac{q^2 - p^2}{2}, \qquad \frac{q^2 + p^2}{2}$$

As another example, let $q = 5$, $p = 3$, and we have $15, 8, 17$, another triple.

Chapter 18

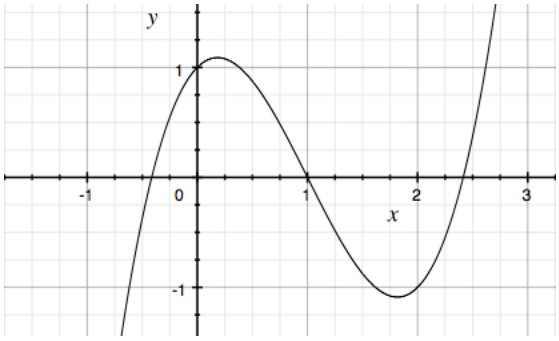
Cubics

Every cubic polynomial equation has at least one term containing x^3 but lacks any higher powers of x such as x^4.

The general equation is

$$y = ax^3 + bx^2 + cx + d$$

and a typical graph ($x^3 - 3x^2 + x + 1$) looks something like this:



There is an axis of symmetry, here at $x = 1$, and the left half is the negative reflection of the right half about the value of $y = f(1) = 0$.

roots

By the *roots* of the equation, we mean those values of x giving $y = 0$, that is, we are solving

$$ax^3 + bx^2 + cx + d = 0$$

In this case we can always multiply through by $1/a$ so the term x^3 has a coefficient of 1, and if we do that then the coefficients are often renamed as:

$$x^3 + ax^2 + bx + c = 0$$

The cubic is an odd function, so the sign of x carries through in x^3. Since the x^3 term dominates the value of the function for extreme values of x, when $x << 0$, y is large and negative, while for $x >> 0$, y is large and positive. This is clearly seen for the above plot.

As a result, the graph of the function must cross the x-axis at least once, and thus every cubic has at least one real root, where $f(x) = 0$.

From this we conclude that every cubic can be factored into

$$(x - r)(x^2 + sx + t)$$

where $x = r$ is the guaranteed real root, although it isn't always the case that r is an integer, of course.

This expression is equal to zero either when $x = r$ or when the quadratic term is zero.

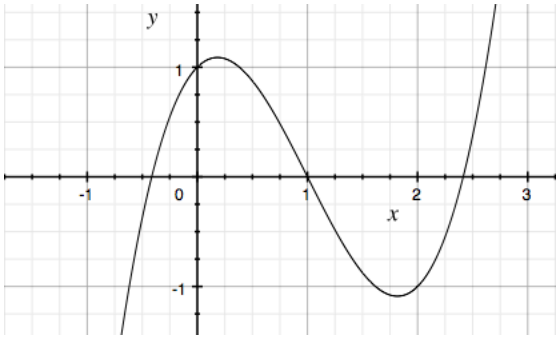
The roots of a quadratic $x^2 + sx + t$ are given by the familiar

$$\frac{-s \pm \sqrt{s^2 - 4t}}{2}$$

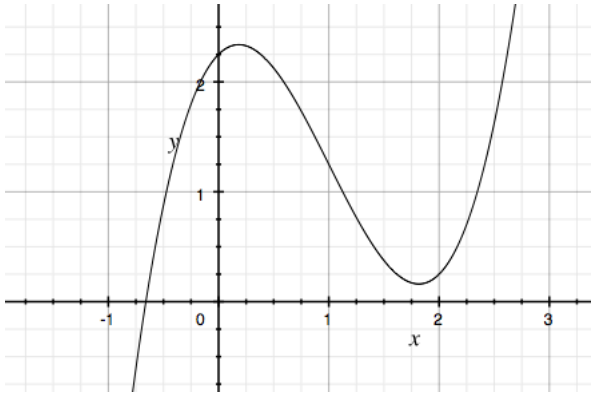
We know that quadratics have either two real roots or none depending on the value of the discriminant under the square root. We consider the case of repeated roots (when the discriminant is zero) as *two* roots.

Therefore, every cubic has either one real root or three of them.

Graphically, we can easily see the general truth of this statement.



In this example from before, near $x = 1.8$, no matter how far the graph goes below the x-axis before it turns the second time, we can add that much to the constant c. The result will be that the graph just touches the x-axis at this point. A tiny bit more and it will not cross at all.



The above graph is the same equation but with 1.25 added to c: $(x^3 - 3x^2 + x + 2.25)$.

factoring

Every cubic has at least one real root and as a consequence, the solution where $y = 0$ can be written as

$$(x - r)(x^2 + sx + t) = 0$$

Multiplying out

$$= x^3 + (s - r)x^2 + (t - rs)x - rt$$

Thus $a = s - r$ and $c = -rt$.

Suppose we know, or have guessed r. We can find s and t by comparison with the original equation.

As an example, consider:

$$(x - 1)(x + 1)(x + 2) = 0$$

$$= (x - 1)(x^2 + 3x + 2)$$

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

We plot this, guess that $x = 1$ is a root, check and find that $x = 1$ solves the equation. Thus $r = 1$ and since

$$c = -2 = -rt$$

then $t = 2$. Also

$$a = 2 = s - r = s - 1$$

and $s = 3$.

Alternatively, there is a formalism called synthetic division for deriving s and t. I have a simple version of this I like better than the complete formal approach. Consider

$$x^3 - 5x^2 - 2x + 24 = 0$$

Suppose we are given that $x = -2$ is a solution, which is easily checked. Now write:

$$x^3 - 5x^2 - 2x + 24$$
$$= (x + 2)(x^2 + \_\_\_\ x + \_\_\_) = 0$$

The cofactor of x^2, the first term on the right, is clearly just 1, so that we get the desired x^3 in the product.

Then we see that, multiplying by 2 we get $2 \times x^2 = 2x^2$, where the desired result is $-5x^2$. We need another $-7x^2$. Therefore the cofactor of x on the right must be -7 so that $-7x^2 + 2x^2 = -5x^2$:

$$(x + 2)(x^2 - 7x + \_\_\_) = 0$$

Then we see that, multiplying by 2 we have $2 \times -7x = -14x$, where the desired result is $-2x$. We need another $12x$. Therefore, the constant term on the right must be 12 so that $-14x + 12x = -2x$.

$$(x + 2)(x^2 - 7x + 12) = 0$$

And finally, we check the whole thing, multiplying 2×12 to give the desired constant, 24. This must work out if we've done the rest correctly and $x = -2$ is really a solution.

Inspired guessing can help. Consider

$$x^3 - 4x^2 - 9x + 36 = 0$$

You may notice that $a \times b = c$. This tells us that $1 - 4$ is a factor of $-9 + 36$. We guess that $x = 4$ is a solution:

$$(x - 4)(x^2 + \_\_\_\ x + \_\_\_) = 0$$

We don't need anything more than $-4x^2$ so the cofactor of x on the right is 0 and write

$$(x-4)(x^2+\_\_\_) = 0$$

For the constant, we guess -9 so as to get $-9x$ and then check $-4 \times -9 = 36$.

$$(x-4)(x^2-9) = 0$$

Finally, we can factor the second term

$$(x-4)(x+3)(x-3) = 0$$

relating roots to cofactors

Suppose a cubic has three distinct real roots, meaning there are real numbers p, q, r such that

$$(x-p)(x-q)(x-r) = 0$$

Multiplying out we would obtain for the constant term: $-pqr$.

$$(x^2 - qx - px + pq)(x-r) = 0$$
$$x^3 - qx^2 - px^2 + pqx - rx^2 + qrx + prx - pqr = 0$$
$$x^3 - (p+q+r)x^2 + (pq+qr+pr)x - pqr = 0$$

So

$$d = -pqr$$

Furthermore

$$a = -(p+q+r)$$
$$b = pq + qr + pr$$

If there are three real roots, they multiply to give the constant term.

$$(x-1)(x+2)(x+1)$$
$$= (x^2 + x - 2)(x+1)$$
$$= x^3 + 2x^2 - x - 2$$

$1 \times -2 \times -1 = -2$.

Actually, a related statement is true even if two of the roots are not real. In principle, we can factor out the single real root, leaving a quadratic.

We said that the roots of a quadratic $x^2 + sx + t$ are given by

$$\frac{-s \pm \sqrt{s^2 - 4t}}{2}$$

If the second and third roots are imaginary they are so because what is under the square root is negative, so the above can be written as

$$z = u \pm iv$$

These consist of two complex numbers, which are complex conjugates. Their product is a real number:

$$(u + vi)(u - vi) = u^2 + v^2$$

When we plug the three roots into $(x - p)(x - q)(x - r)$ we get

$$(x - u + vi)(x - u - vi)(x - r)$$

$$= x^2 - ux - vix - ux + u^2 + uvi + vix - uvi + v^2)(x - r)$$

The imaginary terms with i cancel.

$$= (x^2 - 2ux + u^2 + v^2)(x - r)$$

$$= x^3 - 2ux^2 + u^2x + v^2x - rx^2 + 2urx - u^2r - v^2r$$

$$= x^3 - (2u + r)x^2 + (u^2 + 2ur + v^2)x - r(u^2 + v^2)$$

The result is all real coefficients.

$$d = -r(u^2 + v^2)$$

and d is the product of the three roots.

extreme points

The points where the graph turns around can be found by taking the derivative and setting it equal to zero.

$$y = ax^3 + bx^2 + cx + d$$

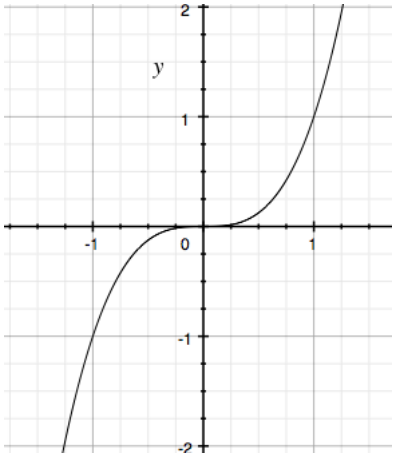
$$\frac{dy}{dx} = 3ax^2 + 2bx + c = 0$$

$$x = \frac{-2b \pm \sqrt{4b^2 - 12ac}}{6a}$$

Not all cubics have a downward sloping segment. This happens when the quadratic for the slope has no real roots, i.e. when the square root term is less than or equal to zero. A simple example of this is when $b = 0$, such as

$$y = x^3$$

This obviously has 3 equal real roots, all zero.



The slope is $2x^2$, which is never negative and equal to zero only at $x = 0$.

repeated roots

A cubic can have repeated roots. Suppose

$$(x - p)^2(x - q) = 0$$

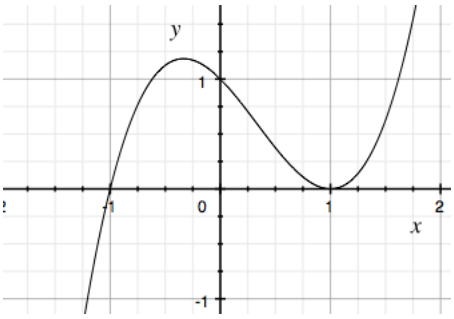
$$= (x^2 - 2px + p^2)(x - q)$$

$$= x^3 - qx^2 - 2px^2 + 2pqx + p^2x - p^2q$$
$$= x^3 - (q + 2p)x^2 + (2pq + p^2)x - p^2q$$

And, as before, the product of the roots is $-d$.

Graphically, two repeated roots means that one of the extreme points is also a root.

$$(x + 1)(x - 1)(x - 1)$$
$$= (x + 1)(x^2 - 2x + 1)$$
$$= x^3 - x^2 - x + 1$$



The slope is zero when

$$3x^2 - 2(q + 2p)x + (2pq + p^2) = 0$$

translation

Continuing with $y = x^3$, when we add or subtract a value from y the plot is shifted up or down, similarly, changes to x shift the same curve to the left or right.

For example:

$$(x - 1)^3 = x^3 - 3x^2 + 3x - 1$$

What this means is that the cofactors a and b may be non-zero and the shape still be the same as $y = x^3$. (The fact that a, b, c conform to the cubic expansion is a tipoff, however). The following section describes what is also essentially a horizontal translation.

depressed cubic

Tartaglia discovered that the quadratic term can be removed from a cubic

$$x^3 + ax^2 + bx + c$$

by an inspired substitution, $x = u - a/3$. Actually, I find the arithmetic a bit confusing, so I will further substitute $v = a/3$ and so $x = u - v$.

$$(u - v)^3 + a(u - v)^2 + b(u - v) + c$$

Now, expand each power of $u - v$ in order.

The cubic binomial $(u - v)^3$ has cofactors of 3 for the inner terms

$$\begin{aligned}(u - v)^3 &= (u - v)(u^2 - 2uv + v^2) \\ &= u^3 - 2u^2v + uv^2 - u^2v + 2uv^2 - v^3 \\ &= u^3 - 3u^2v + 3uv^2 - v^3\end{aligned}$$

Switch the order so that the power of u is last in each term

$$= u^3 - 3vu^2 + 3v^2u - v^3$$

The quadratic is

$$\begin{aligned}a(u - v)^2 &= a\ [\ u^2 - 2uv + v^2\ ] \\ &= au^2 - 2avu + av^2\end{aligned}$$

The linear term is just $bu - bv$.

Finally, collecting all the terms and grouping them by powers of u

$$= u^3\ [\ -3v + a\ ]\ u^2 + \ [\ 3v^2 - 2av + b\ ]\ u + \ [\ -v^3 + av^2 - bv + c\ ]$$

The bright idea is that the cofactor of u^2

$$-3v + a$$

is equal to zero by the terms of the substitution ($v = a/3$).

That leaves:

$$= u^3 + \ [\ 3v^2 - 2av + b\ ]\ u + \ [\ -v^3 + av^2 - bv + c\ ]$$

If we write

$$m = 3v^2 - 2av + b$$
$$n = -v^3 + av^2 - bv + c$$

then the cubic is

$$u^3 + mu + n = 0$$

We can reverse the second substitution ($v = a/3$). We have one less term in each formula, which is simplified a bit, but this also makes the formulas more awkward.

$$m = 3\frac{a^2}{3^2} - 2a\frac{a}{3} = \frac{a^2}{3} - 2\frac{a^2}{3} = -\frac{a^2}{3} + b$$

and

$$n = -\frac{a^3}{3^3} + a\frac{a^2}{3^2} - \frac{ab}{3} + c$$
$$= 2\frac{a^3}{3^3} - \frac{ab}{3} + c$$

Here's an example. Consider:

$$x^3 + 3x^2 - x + 1 = 0$$

$$a = 3, \quad b = -1, \quad c = 1$$

So

$$m = -\frac{a^2}{3} + b = -\frac{a^2}{3} - 1 = -4$$

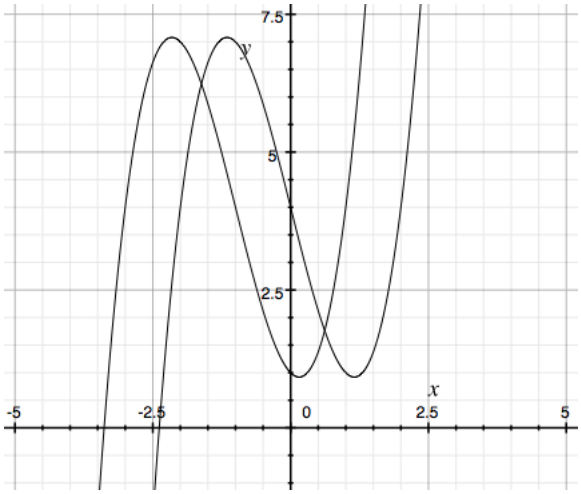
The constant n is

$$n = \frac{2a^3}{3^3} - \frac{ab}{3} + c$$
$$2 + 1 + 1 = 4$$

So the transformed version is

$$u^3 - 4u + 4$$

And this is indeed the same curve, simply displaced to the right by 1 unit, as the substitution $x = u - a/3$ or $x = u - 1$ implies.



The real root is also the same (taking into account the translation).

An example from Nahin is

$$x^3 - 15x^2 + 81x - 175 = 0$$

The coefficient of u is

$$(-\frac{a^2}{3} + b)$$

$$= -\frac{(-15)^2}{3} + 81 = -75 + 81 = 6$$

The constant is

$$\frac{2a^3}{3^3} - \frac{ab}{3} + c$$

$$= \frac{2(-15)^3}{3^3} - \frac{(-15)81}{3} - 175$$

$$= -150 + 405 - 175 = -20$$

Hence we have

$$u^3 + 6u - 20 = 0$$

By trial and error, we find that $u = 2$ is a solution.

Reversing the substitution, $3v = a$. This is the cofactor of x^2, the a from the original equation! So $v = a/3 = -15/3 = -5$.

Thus $x = u - v = 2 - (-5) = 7$. Check by substitution:

$$(7)^3 - 15(7^2) + 81(7) - 175 = 0$$

Factor out 7

$$7^2 - 15(7) + 81 - 25 = 0$$

$$49 - 105 + 81 - 25 = 0$$

which checks.

The resulting equation (lacking a quadratic term), is called a *depressed* cubic.

$$u^3 + mu + n = 0$$

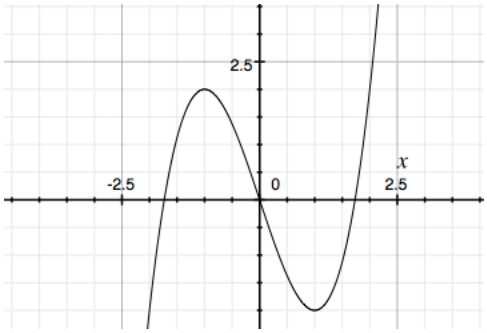
The result is kind of amazing. For any cubic containing ax^2, we can obtain the same curve without any quadratic term.

form of the curve

The constant simply displaces y by some value. The coefficient of x^3 stretches the curve.

From consideration of the depressed quadratic, you can see that the essential form is conferred by the cofactor of x in, say

$$y = x^3 - 3x$$



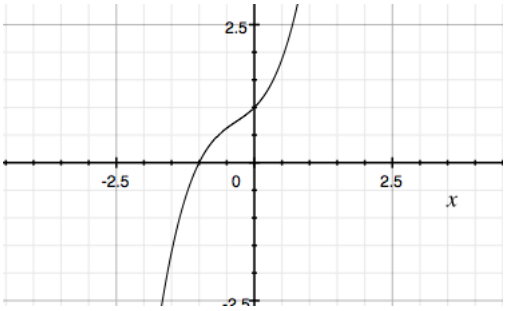
The larger the value of b, the bigger the deviations before the curve turns back. It's curious that the extreme points are exactly $y = 2$ here. That's because

$$\frac{dy}{dx} = 3x^2 - 3 = x^2 - 1$$

Hence they occur at $x = \pm 1$, where $y = \pm 2$.

In an expression like $y = x^3 + ax^2 + bx + c$, increasing a makes the central displacement more pronounced, while increasing b makes it less pronounced. Interestingly, having all three constants equal to 1 makes it go away altogether.

$$y = x^3 + x^2 + x + 1$$



$x = -1$ has solution $y = 0$, and that's the single real root because

$$x^3 + x^2 + x + 1 = (x + 1)(x^2 + 1)$$

and we know $x^2 + 1$ has $i = \pm\sqrt{-1}$ as its solution.

Also, we note for this example

$$y = x^3 + x^2 + x + 1$$

Get the slope as the derivative and set it equal to zero:

$$y' = 3x^2 + 2x + 1 = 0$$

for which the roots are

$$x = \frac{-2 \pm \sqrt{4 - 12}}{6}$$

The discriminant is negative, so there is no x that gives a slope of zero.

The minimum value of y' is $y'' = 0$

$$y'' = 6x + 2 = 0$$

$$x = -\frac{1}{3}$$

$$y' = \frac{3}{9} - \frac{2}{3} + 1 = \frac{2}{3}$$

Solving the depressed cubic

Which brings us finally to Cardano, and the solution of the cubic.

Dunham has a picture of the geometrical division of a cube that Cardano visualized,

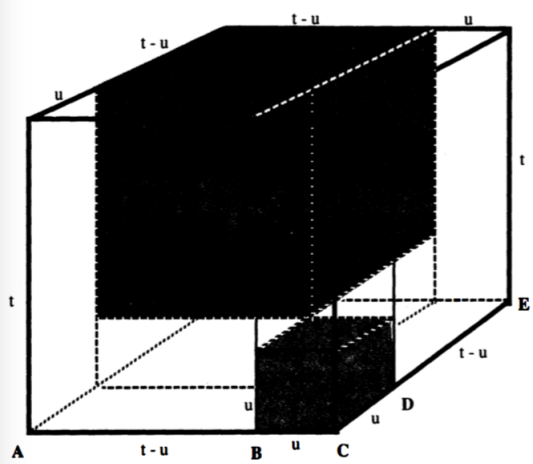


FIGURE 6.1

However, with modern notation, we can get there pretty simply from algebra.

$$(t-u)^3 = t^3 - 3t^2u + 3tu^2 - u^3$$

$$(t-u)^3 + 3t^2u - 3tu^2 = t^3 - u^3$$

$$(t-u)^3 + 3ut(t-u) = t^3 - u^3$$

Now let $x = t - u$

$$x^3 + 3tux = t^3 - u^3$$

Substitute $m = 3tu$ and $n = t^3 - u^3$:

$$x^3 + mx = n$$

This is a depressed cubic.

$$x^3 + mx - n = 0$$

The idea is to start with a depressed cubic we want to solve, and use that to get values for m and n.

If we can then determine values for t and u, $x = t - u$ will be the solution that we seek.

We have the two conditions: $m = 3tu$ and $n = t^3 - u^3$. Solve the first for u

$$u = \frac{m}{3t}$$

and substitute into the second:

$$n = t^3 - \frac{m^3}{(3t)^3}$$

Multiply both sides by t^3

$$nt^3 = t^6 - \frac{m^3}{27}$$

Looks like it's getting more complicated.

But this is a quadratic equation in disguise!

$$t^6 - nt^3 - \frac{m^3}{27} = 0$$

By the quadratic formula:

$$t^3 = \frac{n \pm \sqrt{n^2 + 4m^3/27}}{2}$$

Take the positive square root

$$= \frac{n}{2} + \sqrt{\frac{n^2}{4} + \frac{m^3}{27}}$$

and take its cube root:

$$t = \left[\frac{n}{2} + \sqrt{\frac{n^2}{4} + \frac{m^3}{27}} \right]^{1/3}$$

Since $u^3 = t^3 - n$, just subtract n from the expression for t^3 before taking the cube root.

$$u = \left[-\frac{n}{2} + \sqrt{\frac{n^2}{4} + \frac{m^3}{27}} \right]^{1/3}$$

Then

$$x = t - u$$

$$= \left[\frac{n}{2} + \sqrt{\frac{n^2}{4} + \frac{m^3}{27}} \right]^{1/3} - \left[-\frac{n}{2} + \sqrt{\frac{n^2}{4} + \frac{m^3}{27}} \right]^{1/3}$$

We can write this more simply by pre-computing

$$r = \frac{n}{2}, \qquad s = \frac{m^3}{27}$$

Then

$$x = \ [\ r + \sqrt{r^2 + s}\ ]^{1/3} - \ [\ -r + \sqrt{r^2 + s}\ ]^{1/3}$$

Here is Cardano (recall we are solving $x^3 + mx - n = 0$):

> Cube one-third the coefficient of x; add to it the square of one-half the constant of the equation; and take the square root of the whole. You will duplicate [repeat] this, and to one of the two you add one-half the number you have already squared and from the other you subtract one-half the same . . . Then, subtracting the cube root of the first from the cube root of the second, the remainder which is left is the value of x.

example

How about Cardano's example:

$$x^3 + 6x - 20 = 0$$

Clearly, 2 is a solution to this. It is the only real root.

We have $m = 6$ and $n = 20$, so $n/2 = 10$.

$m = 6$ so

$$\frac{m^3}{27} = \frac{6^3}{27} = \frac{(2 \cdot 3)^3}{3^3} = 2^3 = 8$$

and then

$$x = \ [\ 10 + \sqrt{100 + 8}\ ]^{1/3} - \ [\ -10 + \sqrt{100 + 8}\ ]^{1/3}$$

These two terms are

$$(10 + \sqrt{108})^{1/3} = 2.732$$
$$(-10 + \sqrt{108})^{1/3} = 0.732$$

The difference is indeed very close to 2.

example 2

Another famous example is

$$x^3 - 15x - 4 = 0$$

Guessing, we obtain $x = 4$ as one root.

Now, to factor out $(x - 4)$:

$$x^3 - 15x - 4 = (x - 4)(\_\_\_\ x^2 + \_\_\_\ x + \_\_\_)$$

$$x^3 - 15x - 4 = (x - 4)(x^2 + \_\_\_\ x + \_\_\_)$$

$$x^3 - 15x - 4 = (x - 4)(x^2 + 4x + \_\_\_)$$

$$x^3 - 15x - 4 = (x - 4)(x^2 + 4x + 1)$$

The last multiplication to give the constant works, which provides a check on the whole thing.

We solve the quadratic as

$$x = \frac{-4 \pm \sqrt{16 - 4}}{2} = -2 \pm \sqrt{4 - 1} = -2 \pm \sqrt{3}$$

Check the positive root:

$$(-2 + \sqrt{3})^2 + 4(-2 + \sqrt{3}) + 1$$

$$= 4 - 4\sqrt{3} + 3 - 8 + 4\sqrt{3} + 1$$

$$= 0$$

So we have three real roots. Notice that

$$4 + (-2 + \sqrt{3}) + (-2 - \sqrt{3}) = 0$$

The sum of the roots is zero.

Now use Cardano's solution to solve

$$x^3 - 15x - 4$$

First

$$r = \frac{n}{2} = -2$$

$$s = \frac{m^3}{27} = \frac{-15^3}{27} = -125$$

$$\begin{aligned} x = & \ [\ r + \sqrt{r^2 + s}\ ]^{1/3} - \ [\ -r + \sqrt{r^2 + s}\ ]^{1/3} \\ = & \ [\ -2 + \sqrt{4 + -125}\ ]^{1/3} - \ [\ 2 + \sqrt{4 + -125}\ ]^{1/3} \\ = & \ [\ -2 + \sqrt{-121}\ ]^{1/3} - \ [\ 2 + \sqrt{-121}\ ]^{1/3} \\ = & \ [\ -2 + \sqrt{-121}\ ]^{1/3} + \ [\ -2 - \sqrt{-121}\ ]^{1/3} \end{aligned}$$

That seems strange at first. We have three real roots, but Cardano's solution gives an expression which is the sum of two imaginary numbers.

The resolution is that the two numbers here are complex conjugates. What we have is

$$[\ z\ ]^{1/3} + \ [\ z*\ ]^{1/3}$$

where

$$z = -2 + 11i$$

If we write this in polar form

$$z = re^{i\theta}$$

$$z* = re^{-i\theta}$$

so

$$z^{1/3} = r^{1/3}e^{i\theta/3}$$

$$z*^{1/3} = r^{1/3}e^{i(-\theta/3)}$$

The sum is

$$r^{1/3}\ [\ e^{i\theta/3} + e^{i(-\theta/3)}\ ]$$

The term in the brackets is the sum of a complex number and its complex conjugate, $w + w*$, which is completely real, so the whole thing is completely real.

$$e^{i\theta/3} + e^{i(-\theta/3)} = 2\cos\ (\theta/3)$$

To actually do the calculation

$$z = -2 + 11i$$

$$zz* = (-2 + 11i)(-2 - 11i) = -4 + 121$$

$$r = \sqrt{zz*} = \sqrt{117}$$

$$r^{1/3} = 2.211$$

For the angle

$$\theta = \tan^{-1} -\frac{11}{2} = -1.391$$

$$\theta/3 = -0.46$$

The term in the brackets is

$$2\cos\,(\theta/3) = 2\cos\,(-0.46) = 1.788$$

The whole thing is

$$r^{1/3}\,[\,e^{i\theta/3} + e^{i(-\theta/3)}\,] = 2.211(1.788) \approx 4$$

A much simpler method is to notice that

$$(2+\sqrt{-1})^3 = (2+\sqrt{-1})(4+4\sqrt{-1}-1)$$

$$= (2+\sqrt{-1})(3+4\sqrt{-1})$$

$$= 6+3\sqrt{-1}+8\sqrt{-1}-4$$

$$= 2+11\sqrt{-1} = 2+\sqrt{-121}$$

The same result is obtained with $-2+\sqrt{-1})^3$.

Hence

$$x = \,[\,-2+\sqrt{-121}\,]^{1/3} - \,[\,2+\sqrt{-121}\,]^{1/3}$$

$$= 2+2 = 4$$

example 3

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

It might be simplest to try $m = 3$ and $n = 2$, so $n/2 = 1$.

$m = 3$ so

$$\frac{m^3}{27} = 1$$

Then

$$x = \,[\,1+\sqrt{2}\,]^{1/3} - \,[\,-1+\sqrt{2}\,]^{1/3}$$

These two terms are

$$= 1.3415 - 0.745 = 0.596$$

This doesn't quite match the plot, however (which is closer to 0.7).

The arithmetic is tiresome, so write a script to do it.

script

```
# Cardano's method for solving
# x^3 + mx = n

import sys
from math import sqrt

def simplify(a,b,c):
    f = 1.0*a/3
    g = a*f   # a^2/3
    m = -g + b
    h = f*g   # a^3/3^2
    j = h/3   # a^3/3^3
    return m, (-j + h - (a*b*1.0)/3 + c)

def cardano(m,n):
    c = (m**3)/27
    h = n/2.0
    r = 0.3333333333

    rad1 = ( h + sqrt(h**2 + c))
    rad2 = (-h + sqrt(h**2 + c))
    return round(rad1**r - rad2**r, 5)
```

```
> python
..
>>> from cubics import *
>>> simplify(3,-1,1)
(-4.0, 4.0)
>>> cardano(6,20)
2.0
>>> cardano(3,2)
```

```
0.59607
```

We get Cardano's result, and confirm the calculation for the second example. This suggests that the error lies in the plotting program.

practical solving

A practical approach to real problems involves first plotting the function so as to know whether there is one real root or three, and get an idea of their values.

For example

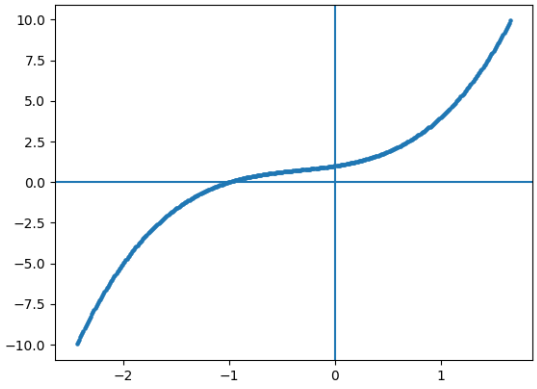
plotter.py

```
from matplotlib import pyplot as plt
import numpy as np

def plot(X,Y):
    plt.scatter(X,Y,s=5)
    plt.axhline()
    plt.axvline()
    #plt.axes().set_aspect('equal')
    plt.savefig('x.png')

def cubic(a,b,c,d):
    def f(x):
        return a*x**3 + b*x**2 + c*x + d
    L = np.linspace(-10,10,2000)
    X = list()
    Y = list()
    for x in L:
        y = f(x)
        if y < -10 or y > 10:
            continue
        X.append(x)
        Y.append(y)
    plot(X,Y)

cubic(1,1,1,1)
```



An actual solver might look something like this:

guess.py

```
import numpy as np

a,b,c = 1, 1, 1

def f(x):
    return x**3 + a*x**2 + b*x + c

def getX(x1,x2):
    N = 1000
    return np.linspace(x1,x2,N)

# assumes we go from f(x) < 0 to f(x) > 0
def guess(x1, x2):
    std_order = f(x1) < f(x2)

    print 'guess'
    print 'x1 = ', str(x1)
    print 'x2 = ', str(x2)
    print 'y1 = ', str(f(x1))
    print 'y2 = ', str(f(x2))
    print
```

```
    X = getX(x1,x2)
    if not std_order:
        X.reverse()
    assert f(X[0]) < 0 and f(X[-1]) > 0

    for i,x1 in enumerate(X):
        x2 = X[i+1]
        # must happen
        if f(x2) > 0:
            if not std_order:
                return x2, x1
            return x1, x2

def close(r):
    e = 1e-12
    return not (r > e or r < -e)

x1 = -2
x2 =  0
i = 0

while i < 100:
    print i+1
    x1, x2 = guess(x1, x2)
    if close(x2 - x1):
        break
    i += 1
```

output

```
> python guess.py
1
guess
x1 =  -2
x2 =  0
y1 =  -5
y2 =  1

2
```

```
guess
x1 =  -1.001001001
x2 =  -0.998998998999
y1 =  -0.00200400701102
y2 =  0.001999998999

3
guess
x1 =  -1.000001002
x2 =  -0.999998997997
y1 =  -2.00400801575e-06
y2 =  2.00400399997e-06

4
guess
x1 =  -1.000000001
x2 =  -0.999999998997
y1 =  -2.00601180111e-09
y2 =  2.00601202316e-09

5
guess
x1 =  -1.0
x2 =  -0.999999999999
y1 =  -2.00772731773e-12
y2 =  2.00817140694e-12

>
```

It's pretty clear that $x = -1$ is the real root.

$$x^3 + x^2 + x + 1 = (x + 1)(x^2 + 1)$$

The product of the two other roots is $x^2 + 1$, that is, they are $\pm\, i$.

Recall that $d = -pqr$ so

$$d = -\ [\ (-1) \times i \times -i\ ] = -\ [\ (-1) \times 1\ ] = 1$$

Chapter 19

Continuum of numbers

In a previous chapter we showed that given any two rational numbers one can find a rational number which lies between them.

Three related statements are also true. We will show that

◦ for any two rational numbers one can find a real number which lies between them

◦ for any two real numbers one can find a rational number which lies between them

◦ for any two real numbers one can find a real number which lies between them

continuum

◦ Between any two *real* numbers it is always possible to find a rational number.

Proof: pick

$$N \in \mathbb{N} \text{ such that } N > \frac{1}{b-a}$$

Then

$$\frac{1}{N} < b - a$$

Define the set $\mathbf{A}$ as follows:

$$\mathbf{A} = \{ \frac{m}{N} : \ m \in \mathbb{N}\}, \quad \text{a subset of } \mathbb{Q}$$

The claim is that

$$\mathbf{A} \cap (a, b) \neq \emptyset$$

There do exist numbers within the open interval (a, b) that are in the set $\mathbb{Q}$.

The proof is by contradiction. Assume on the contrary that the set $\mathbf{A}$ does not contain a rational number lying inside this interval. In other words:

$$\mathbf{A} \cap (a, b) = \emptyset$$

Now, find the largest integer m_1 such that $m_1/N < a$ (it is OK if m_1 is equal to 0). Then the next rational number in $\mathbf{A}$ must be larger than b since the set intersection is empty:

$$\frac{m_1 + 1}{N} > b$$

But this implies that

$$\frac{m_1 + 1}{N} - \frac{m_1}{N} > b - a$$

$$\frac{1}{N} > b - a$$

which contradicts our condition on N above. Hence the assumption is false and so

$$\mathbf{A} \cap (a, b) \neq \emptyset$$

Thus there must exist a rational number r in $\mathbf{A}$ such that $a < r < b$.

example

Consider the open interval: $(\sqrt{2}, \sqrt{3})$.

$$a = \sqrt{2} \approx 1.414$$

$$b = \sqrt{3} \approx 1.732$$

$$b - a \approx 0.3178$$

$$\frac{1}{b - a} \approx 3.1462$$

Pick $N \geq 4$, for example

$$N = 4: \quad 1.414 < \frac{6}{4} = 1.5 < 1.732$$

$$N = 5: \quad 1.414 < \frac{8}{5} = 1.6 < 1.732$$

$$N = 6: \quad 1.414 < \frac{9}{6} = 1.5 < 1.732$$

(In this case $N = 2$ and $N = 3$ happen to work as well).

◦ Between any two rational numbers it is always possible to find a real number.

One proof consists of finding a *particular* irrational in the interval (a, b), where a and b are rational. For $a < b$, we simply add to the number a the following

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

c is smaller than $b - a$ (because $\sqrt{2}/2 < 1$) so the result $a + c$ lies between a and b. We also know that c is irrational, because $\sqrt{2}$ times any rational number is irrational. Finally, $a+c$ is irrational because adding $\sqrt{2}$ times a rational number to any rational number produces an irrational number.

Proof of the first preliminary requirement: $\sqrt{2}$ times a rational is irrational. Suppose for integer p, q, r, s we have

$$\sqrt{2}\ \frac{p}{q} = \frac{r}{s}$$

then

$$\sqrt{2} = \frac{rq}{ps}$$

But the right-hand side is rational, so this is a contradiction.

For the second requirement, again by contradiction suppose

$$\sqrt{2}\ \frac{p}{q} + \frac{s}{t} = \frac{u}{v}$$

for integer p, q, r, s, u, v. But the right-hand side of

$$\sqrt{2} = \frac{q}{p}(\frac{u}{v} - \frac{s}{t})$$

is rational, so this is a contradiction.

Note in passing that powers are different. What do you think about

$$r = \sqrt{2}^{\sqrt{2}}$$

You may think r is "likely" to be irrational. Just a mess. But how about

$$r^{\sqrt{2}}$$

Whether r is rational or irrational

$$r^{\sqrt{2}} = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^2 = 2$$

!!

∘ Between any two real numbers it is always possible to find another real number. This one is subtle. Suppose the two real numbers are "really, really close."

We suppose that they are not equal, so they must be different, say $a < b$.

Since they are different, at some stage in the decimal expansions of a and b, there must be a first position at which a and b differ. If b does not have a 0 at the next position, terminate there and that will be c.

For example:

$$a = 1.23456789129..$$

$$b = 1.23456789133..$$

$$c = 1.23456789130..$$

b must have some digit following this first position where it does not match a, and which is also not equal to zero (otherwise it would be a terminating decimal and thus a rational number). So we can always find a place to terminate to form c.

Eternity is a very long time, especially towards the end.

(credited to Woody Allen)

variations of infinity

In other words there is *no least number* x such that $x > 0$, for example, and no greatest number x such that $x < 1$.

Proof: Assume that m is the smallest number > 0. The rational number $m/2 < m$ is also greater than zero, but smaller than m. Thus, m is not the smallest positive number.

In general, there is no number that is the closest number to another number.

That is actually OK. Here's what's really weird. Cantor proved that the set $\mathbb{Q}$ is *countably finite*. Each element in $\mathbb{Q}$ can be paired in order with a member of $\mathbb{N}$.

The idea of the proof is to show that one can set up a correspondence between $\mathbb{N}$ and $\mathbb{Q}$, assigning each number $r \in \mathbb{Q}$ in a particular order to $1, 2, 3, \ldots$. Here is the figure from Courant and John:

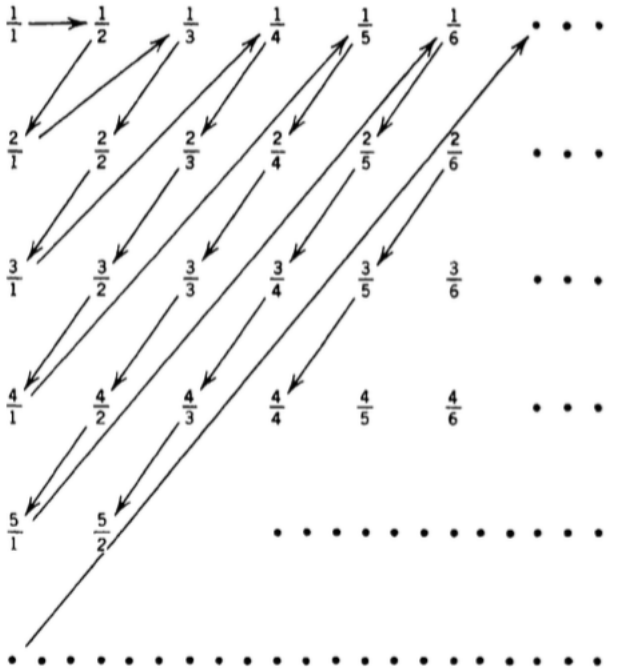


Figure 1.S.1 Denumerability of the positive rationals.

Basically, each row contains all the rational numbers with a particular numerator, and each column contains all the numbers with a particular denominator, arranged in strict increasing order.

Next, set up the sequence indicated by the arrows:

$$\frac{1}{1} \quad \frac{1}{2}, \frac{2}{1} \quad \frac{1}{3}, \frac{2}{2}, \frac{3}{1} \quad \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1} \quad \frac{1}{5}, \frac{2}{4}, \frac{3}{3}, \frac{4}{2}, \frac{5}{1} \quad \frac{1}{6}, \frac{2}{5}, \frac{3}{4}, \frac{4}{3}, \frac{5}{2}, \frac{6}{1} \ldots$$

Then remove all fractions that are duplicates because they are not in lowest terms.

$$\frac{1}{1} \quad \frac{1}{2}, \frac{2}{1} \quad \frac{1}{3}, \frac{3}{1} \quad \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1} \quad \frac{1}{5}, \frac{5}{1} \quad \frac{1}{6}, \frac{2}{5}, \frac{3}{4}, \frac{4}{3}, \frac{5}{2}, \frac{6}{1} \ldots$$

Finally, each r in this sequence is assigned to a natural number (in the sequence $\mathbb{N}$), establishing the denumerability property. $1/3$ is paired with 4 and $3/1$ is paired with 5, and so on.

Cantor showed that such a correspondence (which we just established for $\mathbb{Q}$), is impossible for $\mathbb{R}$. The proof of this is not hard, but we will skip it here. You can check out the chapters on Georg Cantor in Dunham's *Journey Through Genius*.

Thus, the rational numbers are said to be "countably infinite", while the real numbers are not countable. (There is also a proof that the transcendental numbers are much more numerous than the non-transcendental ones).

We say that the set of numbers greater than 0 has *no least element*. We can test this by picking the smallest rational member imaginable, but subsequently, we can always find a smaller rational element (say, by halving that number).

And once we get really close with the small rational element, there are infinitely more irrational than rational ones waiting beyond. And yet, given any such very close irrational number, we can always find a smaller rational number, still larger than the bound.

I told you it was weird.

This property of the real numbers, that there is no closest number to any given number, accounts for virtually all of the theoretical difficulties in calculus which are solved by the use of limits and the apparatus of δ and ϵ or alternatively, neighborhoods. We will get to that in a bit.

Part V

Addendum

Chapter 20

References

- Acheson. *The Calculus Story.*
- Alcock. *Mathematics Rebooted.*
- Courant, Robbins, Stewart. *What is Mathematics?.*
- Dunham. *Euler: The Master of Us All.*
- Dunham. *Journey Through Genius.*
- Dunham. *The Mathematical Universe.*
- Hamming. *Methods of Mathematics Applied to Calculus, Probability, and Statistics.*
- Kline. *Calculus.*
- Lockhart. *Measurement.*
- Maor. *e, the Story of a Number.*
- Maor. *To Infinity and Beyond.*
- Nahin. *An Imaginary Tale. The Story of* $\sqrt{-1}$.
- Nelsen. *Proofs Without Words.*
- Silverman. *A Friendly Introduction to Number Theory.*
- Simmons. *Precalculus mathematics in a nutshell.*
- Spivak. *The Hitchhiker's Guide to Calculus.*

- Stewart. *Significant Figures.*
- Strogatz. *The Joy of x.*
- Thompson. *Calculus Made Easy.*