Part One: Equations

1. Fractions

1.) A fraction remains the same if its numerator and denominator are multiplied with the same number. We call this an expansion of the fraction:

$\frac{2}{3} = \frac{4}{6}$, here we multiplied with 2, that is expanded $\frac{2}{3}$ with 2.

$\frac{2}{3} = \frac{6}{9}$, here we multiplied with 3, that is expanded $\frac{2}{3}$ with 3.

2.) The reverse of expansion is simplification:

$\frac{6}{20} = \frac{3}{10}$, here we simplified $\frac{6}{20}$ with 2.

Sometimes we merely cross out the old numerator and denominator and write the new ones above and under:

$$\frac{\overset{3}{\cancel{6}}}{\underset{10}{\cancel{20}}}$$

3.) If we don't multiply or divide both the numerator and the denominator, only one of them, then of course the fraction value changes, namely:
Multiplying the numerator increases, multiplying the denominator decreases the value.
Dividing the numerator decreases, dividing the denominator increases the value:

$\frac{2}{3}$ becomes double by changing it to $\frac{4}{3}$. $\frac{2}{3}$ becomes half by changing it to $\frac{2}{6}$.

$\frac{4}{5}$ becomes half by changing it to $\frac{2}{5}$. $\frac{5}{4}$ becomes double by changing it to $\frac{5}{2}$.

This can give many variations of how to increase or decrease a fraction. Even more when increase and decrease are done at the same time. Luckily, all these can be combined into the following five rules of fraction multiplications and divisions:

4.)

a.) Multiplying fractions can be done by top with top, bottom with bottom:

$\frac{2}{3} \cdot \frac{4}{9} = \frac{8}{27}$.

b.) Dividing fractions can be replaced by multiplying, if the second is turned upside down:

$\frac{2}{3} : \frac{4}{9} = \frac{2}{3} \cdot \frac{9}{4} = \frac{18}{12}$

c.) Before we perform the top with top, bottom with bottom multiplications, we should check for simplifying but including all numerators and denominators:

$\frac{2}{3} \cdot \frac{4}{9} = \frac{8}{27}$ was correct because the top 2 and 4 can not be simplified with any of the 3 or 9. But:

$\frac{\cancel{2}}{\cancel{3}} \cdot \frac{\overset{3}{\cancel{9}}}{\underset{2}{\cancel{4}}} = \frac{3}{2}$ As we see, we didn't write the 1 above the 2 and under the 3.

This is an accepted abbreviation. By the way, if one misses the simplification, the result is still good. Indeed, $\frac{18}{12} = \frac{3}{2}$, the left can be simplified with 6.

d.) Whole numbers are merely fractions with 1 denominators: $\frac{2}{3} \cdot 10 = \frac{2}{3} \cdot \frac{10}{1} = \frac{20}{3}$.

This of course, should be done at once without writing it out and rather remembering:

$\frac{2}{3} \cdot 10 = \frac{20}{3}$.

e.) The fraction "of" an amount is merely an other way of saying the multiplication:

"Two thirds of ten" means $\frac{2}{3} \cdot 10 = \frac{20}{3}$.

5.) Adding or subtracting fractions can not be done as mechanically as multiplications.
The only obvious situation is when the denominators are the same:
$\frac{2}{3} + \frac{1}{3} = \frac{2+1}{3} = \frac{3}{3} = 1$ and $\frac{2}{3} - \frac{1}{3} = \frac{2-1}{3} = \frac{1}{3}$.
All other additions or subtractions must be done by this rule, that is we have to achieve common denominators. Luckily, this is easy with proper expansions!
To find the common denominator we should start with the largest denominator.
If all others divide this, then it can be used:
$\frac{2}{3} + \frac{3}{4} + \frac{5}{12} = ?$ The largest denominator is 12.
3 and 4 both divide it, so 12 can be used as common denominator:
$\frac{2}{3} + \frac{3}{4} + \frac{5}{12} = \frac{?}{12} + \frac{?}{12} + \frac{5}{12}$.
The ? values can be obtained by seeing how much expansion was done.
From 3 to 12, the expansion was 4 times, so 2 also must be multiplied with this.
From 4 to 12, the expansion was 3 times, so 3 also must be multiplied with this. So:
$\frac{2}{3} + \frac{3}{4} + \frac{5}{12} = \frac{8}{12} + \frac{9}{12} + \frac{5}{12} = \frac{22}{12} = \frac{11}{6}$.
If the largest denominator is not "good", that is the others don't divide it, then we have to try the double, triple, and so on. Sooner or later, we'll succeed.
$\frac{8}{9} - \frac{5}{6} = ?$ 9 is the bigger one, but 6 doesn't divide it. $2 \cdot 9 = 18$, already works, because 6 divides it. So: $\frac{8}{9} - \frac{5}{6} = \frac{?}{18} - \frac{?}{18}$.
The ? values can be obtained again as the expanded numerators.
9 → 18 was doubling, so 8 must be doubled too.
6 → 18 was tripling, so 5 must be tripled too. So, finally:
$\frac{8}{9} - \frac{5}{6} = \frac{16}{18} - \frac{15}{18} = \frac{1}{18}$.

6.) Whole numbers can be added and subtracted easily as fractions with 1 denominator:
$\frac{1}{2} + \frac{2}{3} + 2 = \frac{1}{2} + \frac{2}{3} + \frac{2}{1} = \frac{?}{6} + \frac{?}{6} + \frac{?}{6} = \frac{3}{6} + \frac{4}{6} + \frac{12}{6} = \frac{19}{6}$.
If only two members are, and one is a whole, then the situation is always just using the denominator and multiplying the whole with that:
$\frac{2}{3} + 2 = \frac{2}{3} + \frac{2}{1} = \frac{2}{3} + \frac{?}{3} = \frac{2}{3} + \frac{6}{3} = \frac{8}{3}$.
An old fashioned notation of adding a whole number is the so called, mix number.
This contains a bigger written whole number and a fraction part:
$2\frac{2}{3} = 2 + \frac{2}{3} = \frac{6}{3} + \frac{2}{3} = \frac{8}{3}$. As we see, we can get it at once by remembering:
$2\frac{2}{3}$, here the line means that they must be multiplied and then added to the top.
If a fraction has bigger numerator than denominator, it can be changed to such mix number easily, by checking how many times the denominator fits into the numerator and what remains: $\frac{37}{12} = 3\frac{1}{12}$, because 12 went into 37, 3 times and $3 \cdot 12 = 36$, so 1 remained.
Mix numbers are only used for giving initial or final values.
For calculations we always must use fraction form!

2. Brackets

1.) The agreed order of calculations is that times and division are carried out before additions: $5 + 3 \cdot 2 = 5 + 6 = 11$. Some calculators obey this rule, but some can't. So with those, we have to start entering the multiplication.
If we want to specify our own order of calculations, we can use brackets:
$(2 + 3) + [5\,(3 + 5)] = 5 + [5 \cdot 8] = 5 + 40 = 45$.
As we see, the $(2 + 3)$ bracket and the [ ] bracket was unnecessary. Sometimes we still use such brackets just to express the groups. The $(3 + 5)$ bracketing was vital though! Without it, $5 \cdot 3 + 5 = 20$ had been in [ ]. As we also see, 5 ( ) was used without the multiplication dot, because it was obvious what to do.

2.) Multiplying a bracket sum can be done member by member too:
$5\,(3 + 5) = 5 \cdot 3 + 5 \cdot 5 = 15 + 25 = 40$. Which indeed, is the same as $5 \cdot 8 = 40$.
If there are letters, the same happens: $5\,(x + 5) = 5 \cdot x + 5 \cdot 5 = 5x + 25$.
The signs must be multiplied first:

$$\begin{array}{c} -+ \quad -- \\ \downarrow \quad\;\; \downarrow \end{array}$$

$-5\,(x - 5) = -5x + 25$. As we see, the x after the opening bracket was regarded as +.
Indeed, in the beginning of lines or after opening brackets or equation signs, we omit +.
If there are more letters, we should multiply them in alphabetical order!
Multiplying a number by itself can be abbreviated as squares and cubes, and so on, with other exponent: $x\,x = x^2$, $x\,x\,x = x^3$, . . .

$$\begin{array}{ccc} -- & -+ & -- \\ \downarrow & \downarrow & \downarrow \end{array}$$

$3 - 5ax\,(-2x + ax - b) = 3 + 10ax^2 - 5a^2x^2 + 5abx$.

3.) When two bracket sums are multiplied, we have to multiply every member of one bracket with every member of the other:
$(3x - 5 + ab)\,(-2ax + 1) = -6ax^2 + 3x + 10ax - 5 - 2a^2bx + ab$

The two lines showed how we multiplied 3x with both members of the second bracket.
Similarly, we went through with -5 and then with ab.
After a lot of multiplications we have to combine the numbers and sum letter products to shorten the result:
$3 + (3ax - a - 1)\,(x - 3x^2 + 1) = 3 + 3ax^2 - 9ax^3 + 3ax - ax + 3ax^2 - a - x + 3x^2 - 1 =$

$= 2 + 6ax^2 - 9ax^3 + 2ax - a - x + 3x^2$.
$3 - 1$ gave 2, and $3ax^2 + 3ax^2$ gave $6ax^2$, and finally $3ax - ax$ gave 2ax.

4.) The reverse of multiplication is called <u>factorization</u> and it is much harder.
For example, above giving the final result, nobody would be able to figure out the original form. A special case is quite simple though, namely when we only take out a single non bracketed factor:
$6a^2x^2b + 9a^2x - 3a^3xyb = ?$.
For the numbers, the common divider is 3.
a appears in all members, namely twice or rather as square or second power, $a^2 = aa$.
x also appears in all, but in the last two members only as itself, that is first power.
So, $3a^2x$ can be taken out and thus $? = 3a^2x\,(2bx + 3 - aby)$.
As we see, the example was not given in alphabetical orders, but we strived for this.

5.) Three special bracket products are very important:
$(a + b)^2 = (a + b)\,(a + b) = a^2 + ab + ab + b^2 = a^2 + 2ab + b^2$.
$(a - b)^2 = (a - b)\,(a - b) = a^2 - ab - ab + b^2 = a^2 - 2ab + b^2$.
$(a + b)\,(a - b) = a^2 - ab + ab - b^2 = a^2 - b^2$.

6.) A special application of the previous three special rules themselves, is to change an $x^2 \pm \frac{n}{d}x$ expression into one with a single appearance of x. Here, $\frac{n}{d}$ is a fraction.

For example, $x^2 + \frac{2}{3}x = ?$ Obviously, x^2 and x can not be combined, so to find an expression with a single x, would be quite surprising and very useful later for equations. The idea is that x is regarded as "a" and the half of $\frac{2}{3}$, that is $\frac{1}{3}$, as "b" in the $(a+b)^2 = a^2 + 2ab + b^2$ formula. Indeed, then:

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

Then just subtracting $\left(\frac{1}{3}\right)^2$ from both sides, we indeed obtained $x^2 + \frac{2}{3}x$ as:

$\left(x+\frac{1}{3}\right)^2 - \left(\frac{1}{3}\right)^2$, and so with one appearance of the x letter.

If the fraction is negative, we simply use negative in the bracket too, but the fraction square is always negative. Sometimes the halving of the fraction must be done by the denominator:

$x^2 - \frac{3}{4}x = ?$ The half of $\frac{3}{4}$ is $\frac{3}{8}$, so: $x^2 - \frac{3}{4}x = \left(x-\frac{3}{8}\right)^2 - \left(\frac{3}{8}\right)^2$.

Interestingly, our final results can be verified by using the $a^2 - b^2 = (a+b)(a-b)$ third rule form above too. Indeed,

$$\left(x+\frac{1}{3}\right)^2 - \left(\frac{1}{3}\right)^2 = \left(x+\frac{1}{3}+\frac{1}{3}\right)\left(x+\frac{1}{3}-\frac{1}{3}\right) = \left(x+\frac{2}{3}\right)x = x^2 + \frac{2}{3}x \text{ or}$$

$$\left(x-\frac{3}{8}\right)^2 - \left(\frac{3}{8}\right)^2 = \left(x-\frac{3}{8}+\frac{3}{8}\right)\left(x-\frac{3}{8}-\frac{3}{8}\right) = x\left(x-\frac{6}{8}\right) = x\left(x-\frac{3}{4}\right) = x^2 - \frac{3}{4}x.$$

3. One Variable First Order Equation

One variable means that only one letter appears and we will use x.
First order means that x will only be multiplied with numbers, but not with itself and thus, x^2, x^3 or higher powers can not appear.
The basic rule is to keep an equation in balance!
If we just change a side to an other form, that doesn't affect the balance. A special case of such change is when we combine different members on the sides.
If we add, subtract, multiply or do anything to one side, then the other must be altered the same way. What balancings we use, is up to us! Our goal is to express x with numbers!
The first step of course, is to change all mix numbers to fractions.
The second, is to get rid of all the numbers from the left, and all the x-s from the right.
This can be achieved by subtracting them or adding, if they were negatives:

$$2\frac{1}{2} - \frac{1}{2}x - 1 = \frac{2}{3}x - 1\frac{1}{3} \qquad / \text{ change}$$

$$\frac{5}{2} - \frac{1}{2}x - 1 = \frac{2}{3}x - \frac{4}{3} \qquad / -\frac{5}{2} + 1 - \frac{2}{3}x$$

These three changes must be done to both sides!
On the left, $\frac{5}{2}$ disappears, $-\frac{1}{2}x$ remains, -1 disappears and $-\frac{2}{3}x$ will appear.
Instead of writing the $-\frac{1}{2}x$ and $-\frac{2}{3}x$ separately, we'll take out the common x and write it as, $\left(-\frac{1}{2} - \frac{2}{3}\right)x$.
On the right side, the $\frac{2}{3}x$ disappears, $-\frac{4}{3}$ remains and $-\frac{5}{2}$ and $+1$ will appear. So:

$$\left(-\frac{1}{2} - \frac{2}{3}\right)x = -\frac{4}{3} - \frac{5}{2} + 1$$

We should combine the numbers, but for this, first we have to use common denominators:

$$\left(-\frac{1}{2} - \frac{2}{3}\right)x = -\frac{4}{3} - \frac{5}{2} + 1 \qquad / \text{ c.d.}$$

$$\left(-\frac{3}{6} - \frac{4}{6}\right)x = -\frac{8}{6} - \frac{15}{6} + \frac{6}{6} \qquad / \text{ combine}$$

$$-\frac{7}{6}x = -\frac{17}{6}$$

We are almost finished to get x, except $-\frac{7}{6}$ multiplies it.

To get rid of it, we'll have to divide both sides with $-\frac{7}{6}$

$$-\frac{7}{6}x = -\frac{17}{6} \qquad / : -\frac{7}{6}$$

$$x = -\frac{17}{6} : -\frac{7}{6} = +\frac{17}{6} \cdot \frac{6}{7} = \frac{17}{\cancel{6}} \cdot \frac{\cancel{6}}{7} = \frac{17}{7} = 2\frac{3}{7}$$

Now lets use the five steps again through an other example:

$$\frac{2}{3} - x - 1\frac{3}{4} = 2\frac{1}{2}x - 1 \qquad / \text{ change}$$

$$\frac{2}{3} - x - \frac{7}{4} = \frac{5}{2}x - 1 \qquad / -\frac{2}{3} + \frac{7}{4} - \frac{5}{2}x$$

$$\left(-1 - \frac{5}{2}\right)x = -1 - \frac{2}{3} + \frac{7}{4} \qquad / \text{ c.d.}$$

$$\left(-\frac{2}{2} - \frac{5}{2}\right)x = -\frac{12}{12} - \frac{8}{12} + \frac{21}{12} \qquad / \text{ combine}$$

$$-\frac{7}{2}x = \frac{1}{12} \qquad / : -\frac{7}{2}$$

$$x = \frac{1}{12} : -\frac{7}{2} = -\frac{1}{12} \cdot \frac{2}{7} = -\frac{1}{\underset{6}{\cancel{12}}} \cdot \frac{\cancel{2}}{7} = -\frac{1}{42}$$

We can verify our result by writing it into the changed original equation:

$$\frac{2}{3} - -\frac{1}{42} - \frac{7}{4} = \frac{5}{2} \cdot -\frac{1}{42} - 1$$

$$\frac{2}{3} + \frac{1}{42} - \frac{7}{4} = -\frac{5}{84} - 1 \qquad / \text{ c.d.}$$

$$\frac{56}{84} + \frac{2}{84} - \frac{147}{84} = -\frac{5}{84} - \frac{84}{84} \qquad / \text{ combine}$$

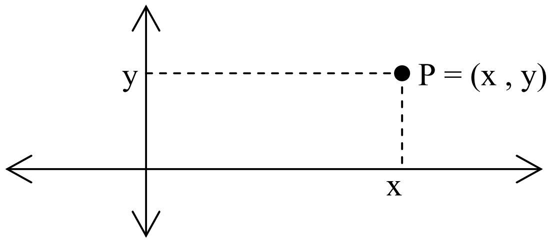
$$-\frac{89}{84} = -\frac{89}{84}$$

Thus, our solution was correct.

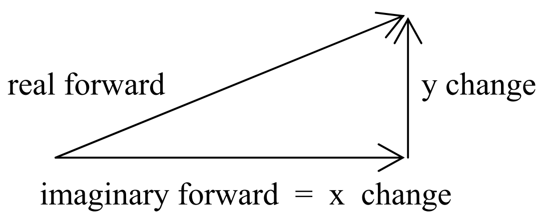
Of course, we could aim for expressing x in the opposite way, that is having it on the right side and all the numbers on the left. Sometimes the original equation has x only on the right side, and then we save a few steps by proceeding this way.

4. The "Hyper Drill"

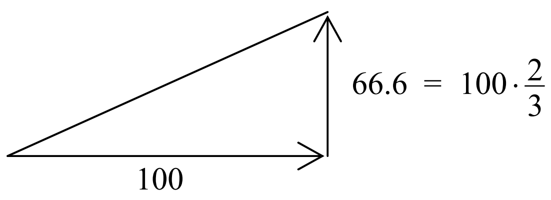
This is not a mining equipment, rather a type of exercise that has infinite many variations and perfects the usage of the Descartes coordinate system and the solving of equations at the same time. In the Descartes system, the points of the plane are located by the x and y coordinates. These can be obtained by simply drawing perpendiculars to the x , y number lines:



The agreement is that on x, the plus values are on the right, while on y, they are upwards.
The real goal of Descartes was to combine geometry with algebra. And indeed, instead of just single points, we can regard lines, circle, and so on. These of course, contain infinite many points. So, they could only be given as sets of points. But how to give a whole set?
If we can find an equation containing x and y, so that it is only true for the points of our geometrical shape, then the equation itself is actually a set of all those x , y values, that satisfy it. But even better is the fact, that if two shapes like a line and a circle cross each other, then the crossing points are satisfying both equations. So calculating the common solutions is actually giving the crossing points. Thus, algebra can solve geometrical problems. And also in reverse, algebraic solutions can be looked and checked by the pictures.
The "hyper drill" calculates the crossing of two lines!
The equations of lines can be best given as $y = s x + h$.
Here, s is the slope and h is the crossing of the y line.
The slope means what we use in street signs too. Instead of angles, it gives the dangerously steep road's slope as a percentage, like 13%. This means that the elevation, that is the increasing of y, is 13% of the travel forward. This is still, a little bit ambiguous, because the "forward" could mean the actual upwards travel on the road or the really straight forward distance in an imaginary line inside the road:

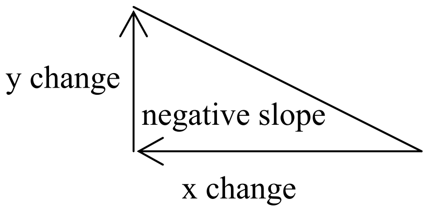


As we see, the imaginary change is easier for us, because it is exactly the x change.
So, slope $= \dfrac{\text{y change}}{\text{x change}}$ or y change = x change · slope . For example:
If the slope is $\frac{2}{3}$, and we go forward 100 in x, then y will change: $100 \cdot \frac{2}{3} = \frac{200}{3} = 66.6$

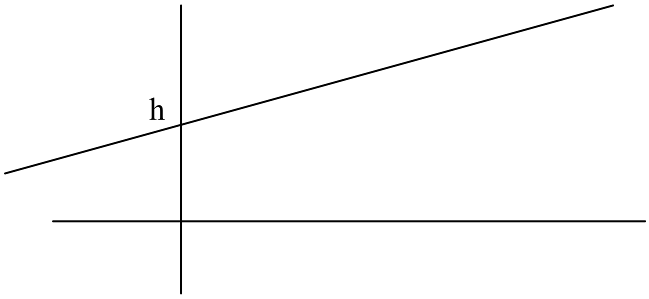


The real beauty of this concept is that it can be easily generalized for negative values.
Negative slope means dropping, that is decreasing y, when x goes forward to the right.
But even if we use negative x change, that is going backwards, the rule remains the same:
y change = x change · slope .

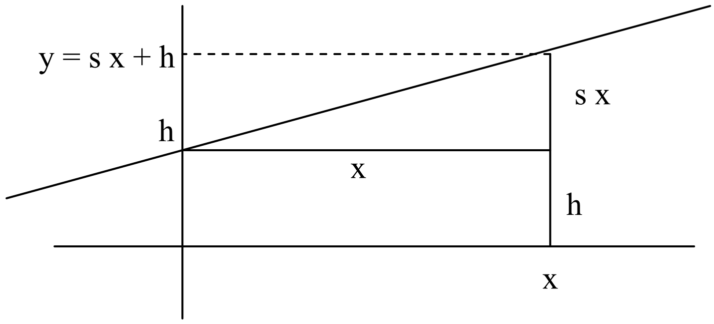
If for example, the x change is negative, and the slope was also negative, then this will give a positive y change, which indeed it should be:



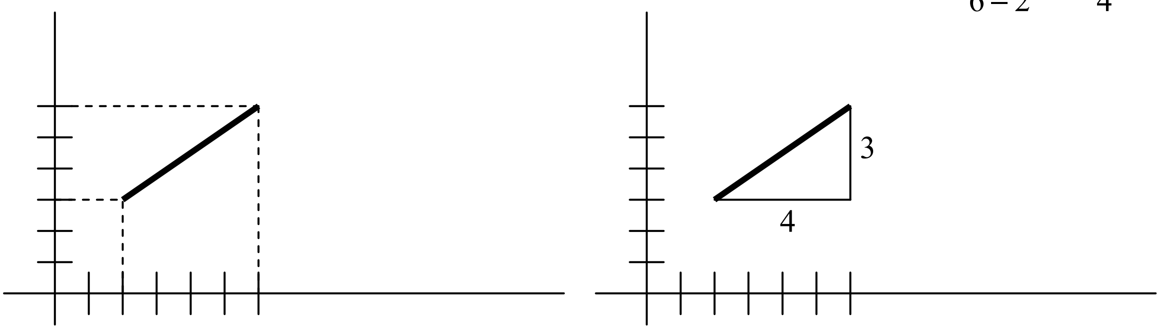
So we don't have to worry about the signs, they will always come out correct by the multiplications.
Since x is measured from the center y line, it's logical to start with h height at here.
So the h initial height of a line is actually the y-crossing of it:



Then the y value, that is the height, at an x distance can be calculated from this initial height and the y change, which is the $x \text{ change} \cdot \text{slope}$:

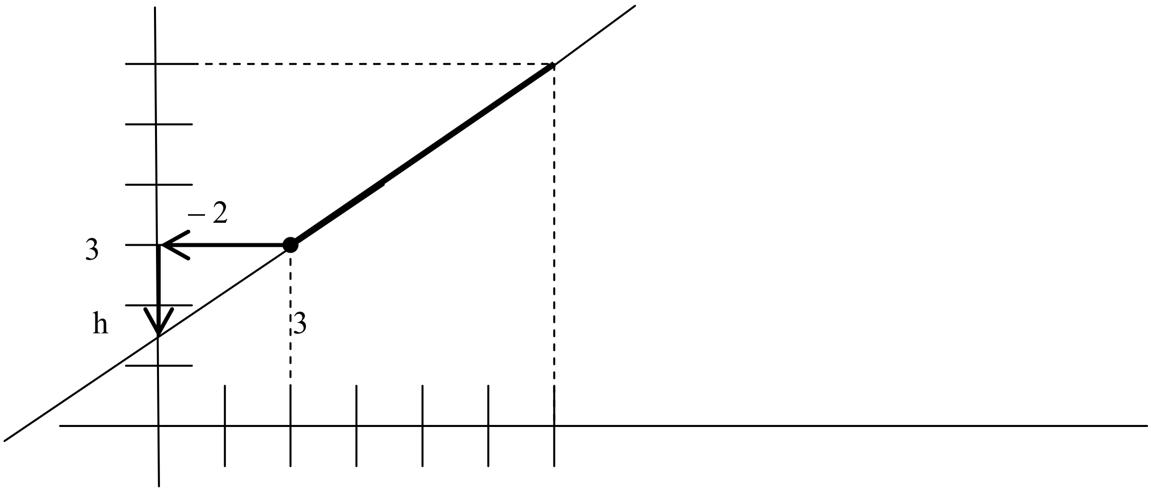


Again, if h is negative, everything still works out!
Now we only have to give two lines, and then find their common x , y values.
But we won't give the two lines as equations directly, rather give them geometrically.
The simplest way is by two points, for each line. From these points, we'll write the equations ourselves. All we have to do is find out the s slopes and the h y-crossings.
The first is easy because the differences of the y and x coordinates of the points give at once the slope: For example, if two points are (2 , 3) and (6 , 6), then the slope is $\frac{6-3}{6-2} = \frac{3}{4}$:



By the way, the coordinate differences are much easier to see if we draw a triangle as above.

For the y crossing we have to choose one of the points, namely the closer one to the y line! Then move towards the y line, and simply add to the y coordinate of the chosen point the $move \cdot slope$ value, with the already obtained slope:



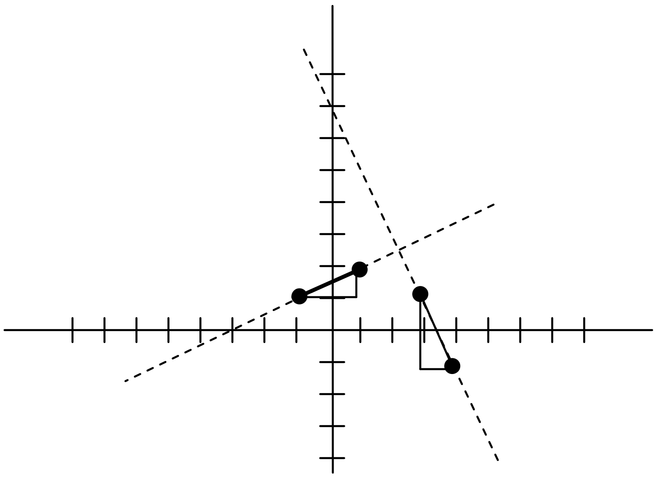
$$h = 3 - 2\cdot\frac{3}{4} = 3 - \frac{3}{2} = \frac{6}{2} - \frac{3}{2} = \frac{3}{2}.$$

3 was the y coordinate, that is the height of the closer point to the y-line.
From this point, a -2 move took us onto the y-line exactly to 3.
This move is always the opposite of the x coordinate of the closer point.
Indeed, if it were on the other side of the y-line, it were negative, but the move is positive.
The change from the height 3 of the closer point can be obtained as the $x\ change \cdot slope$.

In our case it was: $-\ 2\cdot\frac{3}{4}$.

Now the equation is easy: $y = \frac{3}{4}x + \frac{3}{2}$.

If we repeat the same for an other line, then we get two:
y = . . . and y = . . . equations.
Then of course, the two right sides must be equal for common (x , y) points, so we get a single equation for x. That can be solved and then get y too, from any of the above equations.
By the way, it is better to put the more complicated right side equations on the left.
The following pages each contain a full example:



Lets start with he “left” line! On the triangle the y side is 1, the x side is 2.
The line itself goes up, so the slope will be positive.

slope $= \frac{1}{2}$

Both points are 1 distanced from the y-line so we could choose either of them, but we choose the left, because it's easier since then the move is positive. This point's height is 1, so:

y-cross $= 1 + 1 \cdot \frac{1}{2} = \frac{2}{2} + \frac{1}{2} = \frac{3}{2} = 1\frac{1}{2}$

equation: $y = \frac{1}{2}\, x + \frac{3}{2}$

The other line goes downward, so its slope is negative:

slope $= -\frac{2}{1} = -2$

y-cross $= 1 - 3 \cdot -2 = 1 + 3 \cdot 2 = 7$

equation: $y = -2\,x + 7$

Making an equation from the right sides:

$$\frac{1}{2}\, x + \frac{3}{2} = -2\,x + 7 \qquad / -\frac{3}{2} + 2\,x$$

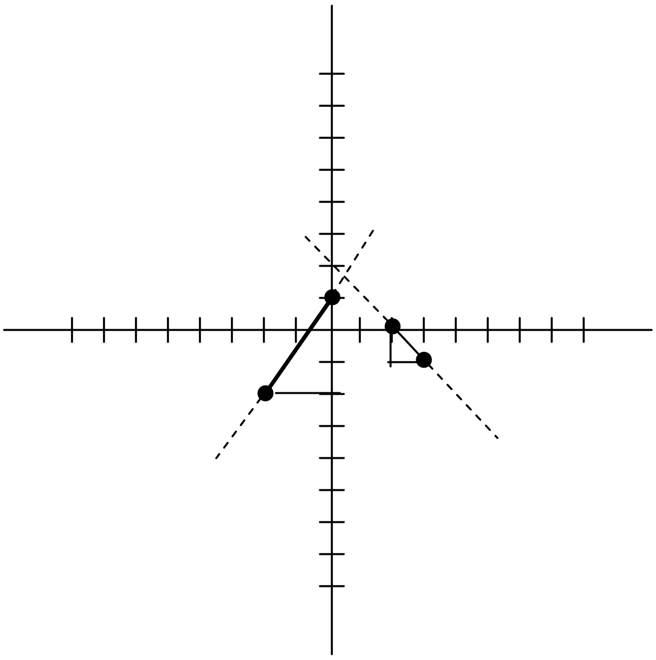
$$\left(\frac{1}{2} + 2\right) x = 7 - \frac{3}{2} \qquad / \text{ c.d.}$$

$$\left(\frac{1}{2} + \frac{4}{2}\right) x = \frac{14}{2} - \frac{3}{2} \qquad / \text{ combine}$$

$$\frac{5}{2}\, x = \frac{11}{2} \qquad / : \frac{5}{2}$$

$$x = \frac{11}{2} : \frac{5}{2} = \frac{11}{2} \cdot \frac{2}{5} = \frac{11}{5} = 2\frac{1}{5}$$

$$y = -2 \cdot \frac{11}{5} + 7 = -\frac{22}{5} + \frac{35}{5} = \frac{13}{5} = 2\,\frac{3}{5}$$



slope $= \frac{3}{2}$

y-cross $= 1 - 0 \cdot \frac{3}{2} = 1$ The point is on the y line already.

equation: $y = \frac{3}{2}x + 1$

---

slope $= -\frac{1}{1} = -1$

y-cross $= 0 - 2 \cdot -1 = 2$

equation: $y = -x + 2$

$\frac{3}{2}x + 1 = -x + 2$ / $-1 + x$

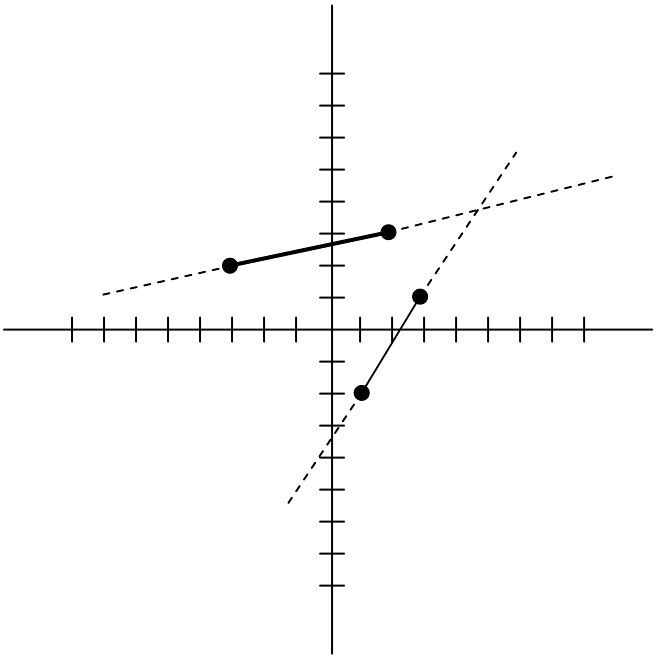
$\left(\frac{3}{2} + 1\right)x = 2 - 1$ / c.d.

$\left(\frac{3}{2} + \frac{2}{2}\right)x = 2 - 1$ / combine

$\frac{5}{2}x = 1$ / $: \frac{5}{2}$

$x = 1 : \frac{5}{2} = 1 \cdot \frac{2}{5} = \frac{2}{5}$

$y = -\frac{2}{5} + 2 = -\frac{2}{5} + \frac{10}{5} = \frac{8}{5} = 1\frac{3}{5}$



slope $= \frac{1}{5}$

y-cross $= 3 - 2 \cdot \frac{1}{5} = \frac{15}{5} - \frac{2}{5} = \frac{13}{5} = 2\frac{3}{5}$

equation: $y = \frac{1}{5}x + \frac{13}{5}$

---

slope $= \frac{3}{2}$

y-cross $= -2 - 1 \cdot \frac{3}{2} = -\frac{4}{2} - \frac{3}{2} = -\frac{7}{2} = -3\frac{1}{2}$

equation: $y = \frac{3}{2}x - \frac{7}{2}$

$$\frac{1}{5}x + \frac{13}{5} = \frac{3}{2}x - \frac{7}{2} \quad / -\frac{13}{5} + \frac{3}{2}x$$

$$\left(\frac{1}{5} - \frac{3}{2}\right)x = -\frac{7}{2} - \frac{13}{5} \quad / \text{ c.d.}$$

$$\left(\frac{2}{10} - \frac{15}{10}\right)x = -\frac{35}{10} - \frac{26}{10} \quad / \text{ combine}$$

$$-\frac{13}{10}x = -\frac{61}{10} \quad / : -\frac{13}{10}$$

$$x = -\frac{61}{10} : -\frac{13}{10} = +\frac{61}{10} \cdot \frac{10}{13} = \frac{61}{13} = 4\frac{9}{13}$$

$$y = \frac{3}{2} \cdot \frac{61}{13} - \frac{7}{2} = \frac{183}{26} - \frac{7}{2} = \frac{183}{26} - \frac{91}{26} = \frac{92}{26} = \frac{46}{13} = 3\frac{7}{13}$$

5. Word Problems

1.) Two numbers are 99 together. What are they if:

a.) One is 7 bigger than the other.

b.) One is twice of the other.

c.) One is 15 smaller than the other.

d.) One is one tenth of the other.

e.) Adding 1 to the smaller, it becomes exactly one ninth of the other.

f.) Adding 1 to the smaller, and subtracting 1 from the bigger, the bigger becomes double of the smaller.

g.) Adding 2 to the smaller and subtracting 2 from the bigger, one becomes double of the other.

h.) One of them is odd and is almost half of the other. The difference of the two is almost 40.

Solutions:

a.) Let the smaller number be x. Then the bigger is x + 7.
Their sum is 99, so:

$$x + x + 7 = 99 \quad / \text{ combine}, -7$$
$$2x = 92 \quad / :2$$
$$x = 46$$

So this is the smaller number, and the bigger is 46 + 7 = 53.
They are correct because their sum 46 + 53 is indeed 99.

b.) Let the smaller number be x. Then the bigger is 2x.
Their sum is 99, so:

$$x + 2x = 99 \quad / \text{ combine}$$
$$3x = 99 \quad / :3$$
$$x = 33$$

So this is the smaller number, and the bigger is $2 \cdot 33 = 66$.
They are correct because their sum 33 + 66 is indeed 99.

c.) Let the smaller number be x. Then the bigger is x + 15.
Their sum is 99, so:

$$x + x + 15 = 99 \quad / \text{ combine}, -15$$
$$2x = 84 \quad / :2$$
$$x = 42$$

So this is the smaller number, and the bigger is 42 + 15 = 57.
They are correct because their sum 42 + 57 is indeed 99.

d.) Let the smaller number be x. Then the bigger is 10x.
Their sum is 99, so:

$$x + 10x = 99 \quad / \text{ combine}$$
$$11x = 99 \quad / :11$$
$$x = 9$$

So this is the smaller number, and the bigger is $10 \cdot 9 = 90$
They are correct because their sum $9 + 90$ is indeed 99.

e.) Let the smaller number be x. Then the bigger is $9(x + 1)$
Their sum is 99, so:

$$x + 9(x + 1) = 99 \quad / \text{ change}$$
$$x + 9x + 9 = 99 \quad / \text{ combine}, -9$$
$$10x = 90 \quad / :10$$
$$x = 9$$

So this is the smaller number, and the bigger is $9(9 + 1) = 9 \cdot 10 = 90$.
They are correct because their sum $9 + 90$ is indeed 99.

f.) Let the smaller number be x. Then the bigger is $99 - x$
Adding 1 to the smaller makes it $x + 1$.
Subtracting 1 from the bigger makes it $99 - x - 1$.
This is the double of the first, so:

$$2(x + 1) = 99 - x - 1 \quad / \text{ change, combine}$$
$$2x + 2 = 98 - x \quad / -2 \; +x$$
$$3x = 96 \quad / :3$$
$$x = 32$$

So this is the smaller number, and the bigger is $99 - 32 = 67$.
They are correct because $67 - 1 = 66$ is indeed, double of $32 + 1 = 33$.

g.) Let the smaller number be x. Then the bigger is $99 - x$.
Adding 2 to the smaller makes it $x + 2$.
Subtracting 2 from the bigger makes it $99 - x - 2$.
We only know that one is double of the other, so either:

$$2(x + 2) = 99 - x - 2 \quad \text{or}$$
$$2(99 - x - 2) = x + 2 \quad / \text{ change}$$
$$198 - 2x - 4 = x + 2 \quad / \text{ combine } +2x$$
$$194 = x + 2 + 2x \quad / \text{ combine } -2$$
$$192 = 3x \quad / :3$$
$$64 = x$$

This should be the smaller number. Then the bigger number should be $99 - 64 = 35$. But it became smaller, so this solution is false.
Thus, using the other possibility:

$2(x+2)$	$=$	$99 - x - 2$	/ change, combine
$2x + 4$	$=$	$97 - x$	/ $-4 + x$
$3x$	$=$	93	/ $:3$
x	$=$	31	

This is the smaller number. Then the bigger is $99 - 31 = 68$.
They are correct because $68 - 2 = 66$ is indeed double of $31 + 2 = 33$.

h.) Let the smaller number be x. Then the bigger is $99 - x$.
The "almost half" is not quite exact, but definitely means less than half.
Obviously, only the smaller can be less than half of the bigger, so:

x	$<$	$\frac{99 - x}{2}$	/ $\cdot 2$
$2x$	$<$	$99 - x$	/ $+ x$
$3x$	$<$	99	/ $:3$
x	$<$	33	

The difference of the two numbers can obviously be obtained by subtracting the smaller, that is x from the bigger, that is from $99 - x$. So it is: $99 - x - x$.
This being "almost" 40, can again be used as smaller than 40, so:

$99 - x \quad - x$	$<$	40	/ combine -99
$-2x$	$<$	-59	/ $: -2$
x	$>$	$-59 : -2 = +29.5$	
x	$\geq$	30	

At the division with -2 the inequality had to be turned around! Indeed, any sign change causes this! For example: $1 < 2$ but, $-1 > -2$
The $\geq$ symbol means larger or equal and among whole numbers $x > 29.5$ clearly means that $x \geq 30$.
This and the previously obtained $x < 33$ together mean that $x = 30$ or 31 or 32.
Since x was odd, it must be 31. Then the bigger number is $99 - 31 = 68$.
31 is indeed, "almost" half of 68, and their difference 37, is "almost" 40.

2.) A family consists of a father, a mother, a son and a daughter. How old are they if:

a.) The sum of all their ages is 100 years.
The father is three times as old as the son.
The mother is three times as old as the daughter.
20 years ago, the mother was the same age as the daughter is now.

b.) The sum of the parent's ages is four times as the sum of the children's.
When the daughter was born, it was ten times.
18 years from now, the father will be twice as the son will be, and the mother will be twice as the daughter will be.

c.) The sum of the parent's ages is 100 years.
The father is three times as old as the son.
The difference of the children's age is 10 years.
The mother is as old as the children together.

d.) 20 years from now, the son will be the same age as the mother is today and the sum of their ages will be 100.
The father is older than the mother, but less than six times the daughter.
8 years ago the family consisted of only three members.

Solutions:

a.) Let the age of the daughter be x. Then the mother is 3x.
But since 20 years ago she was x, today she is also x + 20. So:

$$3x = x + 20 \quad / \; -x$$
$$2x = 20 \quad / \; :2$$
$$x = 10$$

This is the daughter's age and so the mother is $3 \cdot 10$ or $10 + 20$ both = 30.
The mother and the daughter together are $30 + 10 = 40$.
So the father and the son together are $100 - 40 = 60$.
Let now the son's age be x. Then the father is 3x, so:

$$x + 3x = 60 \quad / \; \text{combine}$$
$$4x = 60 \quad / \; :4$$
$$x = 15$$

This is the son's age and the father is $3 \cdot 15 = 45$.
This is correct because $45 + 15 = 60$.

b.) Let the sum of the children's ages be x. Then the sum of the parent's is $4x$.
18 years from now, everybody's age will be 18 more, so both the sum of the children and the parents will be 36 more, that is $x + 36$ and $4x + 36$.
Since the parents will be each twice as one of the children, their sum will be also twice as the sum of the children's:

$$
\begin{aligned}
4x + 36 &= 2(x + 36) && / \text{ change} \\
4x + 36 &= 2x + 72 && / -2x - 36 \\
2x &= 36 && / :2 \\
x &= 18
\end{aligned}
$$

This is the children's sum and the parent's is $4 \cdot 18 = 72$.
Now let the age of the daughter be x. Then the son is $18 - x$.

When the daughter was born the "children" only consisted the son.
Since the daughter is x today, this was x years ago and so, the son was x years younger, that is $18 - x - x = 18 - 2x$. This was the sum of the "children".
The parents were also x years less each, so their sum was $72 - 2x$.
This was ten times as the sum of the "children". So:

$$
\begin{aligned}
72 - 2x &= 10(18 - 2x) && / \text{ change} \\
72 - 2x &= 180 - 20x && / -72 + 20x \\
18x &= 108 && / :18 \\
x &= 6
\end{aligned}
$$

This is the daughter and so the son is $18 - 6 = 12$.
18 years from now the son will be $12 + 18 = 30$, the father twice, that is 60.
And so he is today, $60 - 18 = 42$.
18 years from now the daughter will be $6 + 18 = 24$, the mother twice, that is 48.
And so she is today, $48 - 18 = 30$.

c.) Let the age of the son be x. Then the father is $3x$ and the mother $100 - 3x$.
Since the mother's age is the sum of the children's, thus the daughter's age is the mother's minus the son's, that is $100 - 3x - x = 100 - 4x$.
Since the children have 10 years difference in their age, thus either:

$$
\begin{aligned}
100 - 4x - x &= 10 && \text{or} \\
x - (100 - 4x) &= 10 && / \text{ change} \\
x - 100 + 4x &= 10 && / \text{ combine } + 100 \\
5x &= 110 && / :5 \\
x &= 22
\end{aligned}
$$

This is the son and then the daughter is $100 - 4 \cdot 22 = 12$.
The father is $3 \cdot 22 = 66$ and the mother $100 - 66 = 34$.
This is a pretty big age difference, but not that impossible.
The mother's age is indeed, the sum of the children $12 + 22 = 24$.
But this would mean that when the son was born, she was 12.
That's more than unusual, so lets try the first possibility.

$$100 - 4x \;\; - x \quad = \quad 10 \quad / \text{ combine } - 100$$

$$-5x \quad = \quad -90 \quad / \;: -5$$

$$x \quad = \quad 18$$

This is the son and then the daughter is $100 - \; 4 \cdot 18 \; = 28$.
The father is $3 \cdot 18 \; = 54$ and the mother $100 - 54 = 46$.
The mother's age is indeed, the sum of the children $18 + 28 = 46$.
We must regard this as the solution.

d.) Let the age of the son be x. Then 20 years from now, he will be $x + 20$.
This is the mother now and she will be 20 more, that is $x + 20 + 20 = x + 40$.
Their total will be 100, so:

$$x + 20 \;\; + \;\; x + 40 \quad = \quad 100 \quad / \text{ combine } -20 - 40$$

$$2x \quad = \quad 40 \quad / \;: 2$$

$$x \quad = \quad 20$$

This is the son and so the mother is $20 + 20 = 40$ and indeed, in 20 years they will be together $40 + 60 = 100$.
Now let the daughter be x.
Eight years ago the son was $20 - 8 = 12$ years old, so well alive and thus, the daughter had to be the missing member. In other words, she is less than eight.

$$x \quad < \quad 8$$

Since the father is older than the mother, but less than six times the daughter, thus the mother is also less than six times the daughter. So:

$$40 \; < \; 6x \quad / \;: 6$$

$$6.6 \; < \; x$$

$x \; < \; 8$ and $x \; > \; 6.6$ together mean that $x = 7$. Thus, the daughter is 7.
The father must be more than 40, but less than $6 \cdot 7 \; = 42$. So he is 41.

6. More Variables

R

As we saw, in our previous word problems, even when they were asking for more unknowns, we could succeed with using only one variable equations successively.
Sometimes however, even if there is one unknown only, we might have to use more variables.
For example:
Two cities are 126 km apart on the bank of a straight river. A steam boat travels down the river in 7 hours from one city to the other, while it needs 9 hours to travel upstream.
What is the speed of the river?

Solution:

Let the speed of the river be x and the speed of the steam boat in still water be y.
Then the boat's speed down is y + x, while upstream is y – x.
Thus, under 7 hours down stream, the traveled distance is 7 (y + x).
While under 9 hours upstream, the traveled distance is 9 (y – x).
Thus:

$$7(y + x) = 126$$

$$9(y - x) = 126$$

Dividing the first equation with 7 and the second with 9, we'll get:

$$y + x = 18$$

$$y - x = 14$$

Subtracting the second equation from the first:

$$(y + x) - (y - x) = 18 - 14 \quad / \text{ Comb.}$$

$$2x = 4 \quad / :2$$

$$x = 2$$

We could have argued as follows too:
The speed of the boat down stream is $\frac{126}{7} = 18$ km/h, up stream $\frac{126}{9} = 14$ km/h.
The difference in speed is 18 – 14 = 4 km/h, which is twice the river's speed.
But this argument is a bit over complicated and it was much simpler to get rid of the y variable.
Such elimination of variables can always easily lead us to the solutions.
There are two ways to this. Either by multiplying one of the equations with a number and then adding or subtracting with the other or, by expressing the variable by the others.
For example in:

$$2x - 5y = 3$$
$$x + y = 7$$

If we multiply the second by 2, it becomes, 2x + 2y = 14, then subtracting the first from this:

$$2y + 5y = 14 - 3$$
$$7y = 11$$
$$y = \frac{11}{7} = 1\frac{4}{7}$$

Then to get x, we can put y's value into one of the original equations and solve it for x.

With the other method of "expressing", the solution could be as follows:
From the second equation, subtracting x from both sides gives: $y = 7 - x$
Writing this into the first:

$$\begin{aligned} 2x - 5(7 - x) &= 3 \\ 2x - 35 + 5x &= 3 \\ 7x &= 38 \\ x &= \frac{38}{7} = 5\frac{3}{7} \end{aligned}$$

Then, $y = 7 - x = 7 - 5\frac{3}{7} = 1\frac{4}{7}$.

As we see, "expressing" is a bit more complicated at the beginning, but it gives the other unknowns successively backwards. So in average, both elimination methods require the same amount of calculations.

Sometimes however, we can succeed with some unexpected tricks much easier:

1.) $x + y = a$
$y + z = b$
$x + z = c$

Here a, b, c are any given numbers but we can find an easy solution for all such possible numbers as follows: Lets add together <u>all three</u> equations:

$$2x + 2y + 2z = a + b + c \quad / :2$$

$$x + y + z = \frac{a+b+c}{2}$$

Subtracting each equation from this we get at once each unknown.

$$z = \frac{a+b+c}{2} - a = \frac{a+b+c}{2} - \frac{2a}{2} = \frac{b+c-a}{2}$$

$$x = \frac{a+b+c}{2} - b = \frac{a+b+c}{2} - \frac{2b}{2} = \frac{a+c-b}{2}$$

$$y = \frac{a+b+c}{2} - c = \frac{a+b+c}{2} - \frac{2c}{2} = \frac{a+b-c}{2}$$

2.) $x y = a$
$y z = b$
$x z = c$

Here we should multiply all three equations together to get:

$$x^2 y^2 z^2 = a b c \quad / \sqrt{\ }$$

$x y z = \sqrt{a b c}$ Dividing this with each equation, we get each unknown:

$$z = \frac{\sqrt{abc}}{a} = \sqrt{\frac{abc}{a^2}} = \sqrt{\frac{bc}{a}}$$

$$x = \frac{\sqrt{abc}}{b} = \sqrt{\frac{abc}{b^2}} = \sqrt{\frac{ac}{b}}$$

$$y = \frac{\sqrt{abc}}{c} = \sqrt{\frac{abc}{c^2}} = \sqrt{\frac{ab}{c}}$$

3.) $x(y+z) = a \rightarrow xy + xz = a$
$y(x+z) = b \rightarrow yx + yz = b$
$z(x+y) = c \rightarrow zx + zy = c$

Regarding xy, xz and yz as new unknowns we can use 1.) to find them as:
$xy = A$
$yz = B$
$xz = C$ Then we can use 2.) to find x, y, z.

4.) $x + y + xy = a$
$y + z + yz = b$
$x + z + xz = c$

Lets add 1 to each equation.

$$x + y + xy + 1 = (x+1)(y+1) = a+1$$

$$y + z + yz + 1 = (y+1)(z+1) = b+1$$

$$x + z + xz + 1 = (x+1)(z+1) = c+1$$

Multiplying them all: $(x+1)^2 \ (y+1)^2 \ (z+1)^2 = (a+1)(b+1)(c+1)$

Thus, $(x+1)(y+1)(z+1) = \sqrt{(a+1)(b+1)(c+1)}$ dividing this with each:

$$z + 1 = \frac{\sqrt{(a+1)(b+1)(c+1)}}{a+1} = \sqrt{\frac{(b+1)(c+1)}{a+1}}$$

$$y + 1 = \frac{\sqrt{(a+1)(b+1)(c+1)}}{b+1} = \sqrt{\frac{(a+1)(c+1)}{b+1}}$$

$$x + 1 = \frac{\sqrt{(a+1)(b+1)(c+1)}}{c+1} = \sqrt{\frac{(a+1)(b+1)}{c+1}}$$

7. Linear Equation System and Determinants

R

Linear means the same as first order, that is that the variables are only multiplied with numbers, but not with each other or themselves.
So if x^2, y^2 or xy appear in an equation, then it is not linear or first order.

$$2x = 3 + 5y$$
$$y - x = 7$$

is a linear equation system, because both x and y are only multiplied with numbers.
A more organized form of the above system is:

$$2x - 5y = 3$$
$$-x + y = 7$$

because the same variables all appear under each other, and the numbers without variables are all appearing on the right side. Even if a variable doesn't appear in an equation, we can use 0 to keep the general form. For example:

$$\begin{aligned} -x &+ 3y + 0z = -2 \\ 2x &- y + 5z = 3 \\ 0x &+ y + 2z = 0 \end{aligned}$$

So the multipliers of the variables can be given as the table:

$$\begin{pmatrix} -1 & 3 & 0 \\ 2 & -1 & 5 \\ 0 & 1 & 2 \end{pmatrix}$$

Such table is called a matrix.
This uniform writing of the linear equation system suggests the question, whether there is an instant way to calculate all unknowns without the complicated eliminating methods of the variables.
There is such method, what's more for any number of unknowns. To introduce this we have to use $a, b, c, \ldots$ variables for the normally given data numbers too.
So, $x, y, \ldots, z$ are our alphabetical variables for the unknowns, while $a, b, \ldots, e$ for the given numbers. In our English alphabet, z comes after y, but here, we assume that z is the last unknown. We can imagine other letters between y and z, for example, $y_1, y_2, \ldots$
Also, e is the last alphabetical variable for our data. Think of e = "end".
Between b and e, there can be as many variables as we want, but usually we'll use c for any of these. This will make sense for a c = "column" meaning too.
On the other hand, for the numbers that stand on the right side of the equations, we'll always use the r letters which will make sense, not only as r = "right", but as r = "replacement".

D

1.) An <u>n-square matrix</u> is an n by n table of numbers.
We'll usually place them in round brackets and denote them column by column with alphabetical letters $a, b, \ldots, c, \ldots, e$ and row by row with subscripts.

$$\begin{pmatrix} a_1 & b_1 & . & . & e_1 \\ a_2 & b_2 & . & . & e_2 \\ . & . & . & . & . \\ . & . & . & . & . \\ a_n & b_n & . & . & e_n \end{pmatrix}$$

2.) An <u>n-order</u> is an ordering of the $1, 2, 3, \ldots, n$ numbers.
For example, a 4-order is 3, 1, 2, 4.

3.) The <u>pick</u> by an n-order i, j, . . . from a matrix, is $a_i, b_j, \dots$
For example, the pick by the 4-order 3, 1, 2, 4 is a_3, b_1, c_2, d_4.

4.) The <u>pick product</u> is the product of the pick.

5.) An <u>assignment</u> is assigning + or – to each n-order.

6.) The <u>determinant</u> of an n-square matrix with a chosen assignment is the sum of all pick products, with the assigned signs.
We denote this sum by changing the round bracket of the matrix to straight.

T

1.) Decomposition

a.) By the i-th row:

$$\begin{vmatrix} a_1 & b_1 & . & e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & e_n \end{vmatrix} = \begin{vmatrix} \emptyset & . & . & . \\ a_i & 0 & 0 & 0 \\ \emptyset & . & . & . \\ \emptyset & . & . & . \end{vmatrix} + \begin{vmatrix} . & \emptyset & . & . \\ 0 & b_i & 0 & 0 \\ . & \emptyset & . & . \\ . & \emptyset & . & . \end{vmatrix} + \dots + \begin{vmatrix} . & . & . & \emptyset \\ 0 & 0 & 0 & e_i \\ . & . & . & \emptyset \\ . & . & . & \emptyset \end{vmatrix}$$

b.) By the c column:

$$\begin{vmatrix} a_1 & b_1 & . & e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & e_n \end{vmatrix} = \begin{vmatrix} \emptyset & \emptyset & c_1 & \emptyset \\ . & . & 0 & . \\ . & . & 0 & . \\ . & . & 0 & . \end{vmatrix} + \begin{vmatrix} . & . & 0 & . \\ \emptyset & \emptyset & c_2 & \emptyset \\ . & . & 0 & . \\ . & . & 0 & . \end{vmatrix} + \dots + \begin{vmatrix} . & . & 0 & . \\ . & . & 0 & . \\ . & . & 0 & . \\ \emptyset & \emptyset & c_n & \emptyset \end{vmatrix}$$

The $\emptyset$ places can contain any numbers. The dots are the unchanged members.

2.) Multiplying

a.) By the i-th row:

$$x \begin{vmatrix} a_1 & b_1 & . & e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & e_n \end{vmatrix} = \begin{vmatrix} . & . & . & . \\ x\,a_i & x\,b_i & . & x\,e_i \\ . & . & . & . \\ . & . & . & . \end{vmatrix}$$

b.) By the c column:

$$x \begin{vmatrix} a_1 & b_1 & . & e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & e_n \end{vmatrix} = \begin{vmatrix} . & . & x\,c_1 & . \\ . & . & x\,c_2 & . \\ . & . & . & . \\ . & . & x\,c_m & . \end{vmatrix}$$

3.) Splitting

a.) By the i-th row:

$$\begin{vmatrix} . & . & . & . \\ a_i + \bar{a}_i & b_i + \bar{b}_i & . & e_i + \bar{e}_i \\ . & . & . & . \\ . & . & . & . \end{vmatrix} = \begin{vmatrix} . & . & . & . \\ a_i & b_i & . & e_i \\ . & . & . & . \\ . & . & . & . \end{vmatrix} + \begin{vmatrix} . & . & . & . \\ \bar{a}_i & \bar{b}_i & . & \bar{e}_i \\ . & . & . & . \\ . & . & . & . \end{vmatrix}$$

b.) By the c column:

$$\begin{vmatrix} . & . & c_1 + \bar{c}_1 & . \\ . & . & c_2 + \bar{c}_2 & . \\ . & . & . & . \\ . & . & c_n + \bar{c}_n & . \end{vmatrix} = \begin{vmatrix} . & . & c_1 & . \\ . & . & c_2 & . \\ . & . & . & . \\ . & . & c_n & . \end{vmatrix} + \begin{vmatrix} . & . & \bar{c}_1 & . \\ . & . & \bar{c}_2 & . \\ . & . & . & . \\ . & . & \bar{c}_n & . \end{vmatrix}$$

P

1.) a.) The first member gives all the pick products containing a_i.
The second member, the pick products containing b_i. And so on.
b.) The first member gives all the pick products containing c_1.
The second member, the pick products containing c_2. And so on.
2.) , 3.) Go similarly as 1.)

D

The alternating assignment is obtained as follows:
In each n-order we'll give a sign to each member and the total product of these signs will be the one assigned to the whole order.
The first member has the sign by: $+1$, -2 , $+3$, -4 , $+5$, . . .
For example, if the first member is the second, its sign is –.
Then we take out this member and again alternate the rest: $+1$, -3 , $+4$, -5 , . . .
This will tell the sign of the second member. Then we take that out again and re-alternate the remaining numbers. And so on. Of course, the last remaining number is always +, so we don't have to bother about that, when we multiply all the signs together.
For example, the alternating assignment for 2 , 4 , 1 , 5 , 3 = – + + – = +
From now on we use this alternating assignment for all orders and determinants.

T

1.) Changing two neighbouring members in an n-order, the sign changes to opposite.
2.) Changing any two members is also changes the sign to opposite.
3.) Exchanging two columns or rows makes the determinant change sign.
(Unless it was 0, and thus didn't have a sign at all.)
4.) If two columns or rows are the same, then the determinant is 0.
5.) Adding a column to an other doesn't change the value of the determinant.
(Similarly for rows.)
6.) If a column is the same as the sum of other columns, or their multiplied variants, then the value of the determinant is 0. (Similarly for rows.)

P

1.) The signs given to all other members than the two exchanged remain the same.
The smaller of the two neighbourings remain the same too after the exchange, but the bigger one changes. Thus, all together only one sign changes.
2.) Every change of two members can be obtained by successive neighbouring changes, as follows: First, we move one member next to the other. Then, use one single neighbouring exchange. And thirdly, we move the exchanged member back to the position of the other. The back and forth movements are the same many neighbouring exchanges, so together are even many. Plus the single exchange makes the total odd. And thus, the sign opposite.
3.) By 2.) all pick products become opposite and thus the total too.
4.) Suppose that the determinant were not 0, but had a sign.
Exchanging the identical columns or rows keeps the determinant identical too.
By 3.), if it had a sign, it couldn't be identical, but were opposite in sign.
5.) By first T 3.) we can split the new determinant into the original and one with repeated columns. But then this second member is 0 by 4.).
6.) By 5.), the sum of any number of columns can be replaced into one of the added columns. (With keeping the same value of the determinant.)
If one column is a sum of others, then the sum can be replaced into one of the members and thus, obtain a same value determinant with two equal columns. But by 4.) this is 0.
If one column is the sum of not others, but only some multiple variants of them, then first we can multiply the columns, which changes the value of the determinant with the multipliers. But still this leads to a 0 value, and thus the original had to be 0 too.

T Lets use an $\begin{pmatrix} r_1 \\ . \\ . \\ r_n \end{pmatrix}$ "replacement" column, that we exchange with each column of an original matrix which had a D determinant value. The obtained new determinants are denoted as D_a , D_b , ... , D_e according to which column is replaced. Then:

$$a_i\ D_a + b_i\ D_b + \dots + e_i\ D_e = r_i\ D$$

P We'll only show it for $i = 1$. The general case can be seen similarly or we can replace the i-th row with the first in all determinants and thus, change the sign on both sides.

$$a_1 \begin{vmatrix} r_1 & b_1 & . & e_1 \\ . & . & . & . \\ . & . & . & . \\ r_n & b_n & . & e_n \end{vmatrix} + b_1 \begin{vmatrix} a_1 & r_1 & . & e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & r_n & . & e_n \end{vmatrix} + \dots + e_1 \begin{vmatrix} a_1 & b_1 & . & r_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & r_n \end{vmatrix} = r_1 \begin{vmatrix} a_1 & b_1 & . & e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & e_n \end{vmatrix}$$

$$\|\qquad\qquad\|\qquad\qquad\|\qquad\qquad\|$$

$$\begin{vmatrix} a_1r_1 & a_1b_1 & . & a_1e_1 \\ . & . & . & . \\ . & . & . & . \\ r_n & b_n & . & e_n \end{vmatrix} + \begin{vmatrix} b_1a_1 & b_1r_1 & . & b_1e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & r_n & . & e_n \end{vmatrix} + . + \begin{vmatrix} e_1a_1 & e_1b_1 & . & e_1r_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & r_n \end{vmatrix} = \begin{vmatrix} r_1a_1 & r_1b_1 & . & r_1e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & e_n \end{vmatrix}$$

$$\|\qquad\qquad\|\qquad\qquad\|\qquad\qquad\|$$

$$\begin{vmatrix} a_1r_1 & 0 & 0 & 0 \\ \emptyset & . & . & . \\ \emptyset & . & . & . \\ \emptyset & . & . & . \end{vmatrix} + \begin{vmatrix} b_1a_1 & 0 & 0 & 0 \\ \emptyset & . & . & . \\ \emptyset & . & . & . \\ \emptyset & . & . & . \end{vmatrix} + \dots + \begin{vmatrix} e_1a_1 & 0 & 0 & 0 \\ \emptyset & . & . & . \\ \emptyset & . & . & . \\ \emptyset & . & . & . \end{vmatrix} = \begin{vmatrix} r_1a_1 & 0 & 0 & 0 \\ \emptyset & . & . & . \\ \emptyset & . & . & . \\ \emptyset & . & . & . \end{vmatrix}$$

$$+\qquad\qquad+\qquad\qquad+\qquad\qquad+$$

$$\begin{vmatrix} 0 & a_1b_1 & 0 & 0 \\ . & \emptyset & . & . \\ . & \emptyset & . & . \\ . & \emptyset & . & . \end{vmatrix} + \begin{vmatrix} 0 & b_1r_1 & 0 & 0 \\ . & \emptyset & . & . \\ . & \emptyset & . & . \\ . & \emptyset & . & . \end{vmatrix} + \dots + \begin{vmatrix} 0 & e_1b_1 & 0 & 0 \\ . & \emptyset & . & . \\ . & \emptyset & . & . \\ . & \emptyset & . & . \end{vmatrix} = \begin{vmatrix} 0 & r_1b_1 & 0 & 0 \\ . & \emptyset & . & . \\ . & \emptyset & . & . \\ . & \emptyset & . & . \end{vmatrix}$$

$$+\qquad\qquad+\qquad\qquad+\qquad\qquad+$$
$$\vdots\qquad\qquad\vdots\qquad\qquad\vdots\qquad\qquad\vdots$$
$$+\qquad\qquad+\qquad\qquad+\qquad\qquad+$$

$$\begin{vmatrix} 0 & 0 & 0 & a_1e_1 \\ . & . & . & \emptyset. \\ . & . & . & \emptyset \\ . & . & . & \emptyset \end{vmatrix} + \begin{vmatrix} 0 & 0 & 0 & b_1e_1 \\ . & . & . & \emptyset. \\ . & . & . & \emptyset \\ . & . & . & \emptyset \end{vmatrix} + \dots + \begin{vmatrix} 0 & 0 & 0 & e_1r_1 \\ . & . & . & \emptyset. \\ . & . & . & \emptyset \\ . & . & . & \emptyset \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 & r_1e_1 \\ . & . & . & \emptyset. \\ . & . & . & \emptyset \\ . & . & . & \emptyset \end{vmatrix}$$

The diagonally positioned determinants on the left are the same as the right.
The other members on the left cancel each other if paired by the mirrored position to the diagonal ones. Indeed, each pair can be obtained with a column exchange. Thus, are opposite.

D

$$\begin{array}{ccccccccc} a_1 x & + & b_1 y & + & \dots & + & e_1 z & = & r_1 \\ a_2 x & + & b_2 y & + & \dots & + & e_2 z & = & r_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_n x & + & b_n y & + & \dots & + & e_n z & = & r_n \end{array}$$

Is an n variable linear equation system where $x, y, \dots, z$ are the n unknowns.

T

If the D determinant of the left multipliers of the unknowns is not 0, and
$D_a, D_b, \dots, D_e$ denote again the replaced ones with the r right side data, then:
$x = \frac{D_a}{D}$, $y = \frac{D_b}{D}$, . . . , $z = \frac{D_e}{D}$ are solutions and the only ones.

P

By previous theorem, $a_i D_a + b_i D_b + \dots + e_i D_e = r_i D$.
Dividing both sides with D, we can see that the claimed ones are indeed solutions.
Now enough to show that if there are two set of solutions, then $D = 0$.
Let $x_1, y_1, \dots, z_1$ and $x_2, y_2, \dots, z_2$ be two set of solutions.
(The two set can have common members, but not all.)
The $\bar{x} = x_1 - x_2$, $\bar{y} = y_1 - y_2$, . . . , $\bar{z} = z_1 - z_2$ differences will satisfy:
$a_i \bar{x} + b_i \bar{y} + \dots + e_i \bar{z} = 0$
All $\bar{x}, \bar{y}, \dots, \bar{z}$ can't be 0, because we had two sets. Dividing with a non zero, say $\bar{y}$, all b_i can be expressed from the others. Thus, this column will be a sum of multiples of others. Then by 6.), $D = 0$.

T

Calculating Determinants

1.) Expansion by row (column similarly.)

$$\begin{vmatrix} a_1 & b_1 & . & e_1 \\ . & . & . & . \\ . & . & . & . \\ a_n & b_n & . & e_n \end{vmatrix} = a_i \begin{vmatrix} x & . & . & . \\ x & x & x & x \\ x & . & . & . \\ x & . & . & . \end{vmatrix} + b_i \begin{vmatrix} . & x & . & . \\ x & x & x & x \\ . & x & . & . \\ . & x & . & . \end{vmatrix} + \dots + e_i \begin{vmatrix} . & . & . & x \\ x & x & x & x \\ . & . & . & x \\ . & . & . & x \end{vmatrix}$$

The x-s mean omitting the members and thus, obtaining one smaller sized determinant.

2.) "Criss Cross" Rule: $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$

3.) Cramer Rule:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{matrix} a_1 & b_1 & c_1 & a_1 & b_1 \\ a_2 & b_2 & c_2 & a_2 & b_2 \\ a_3 & b_3 & c_3 & a_3 & b_3 \end{matrix} = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3$$

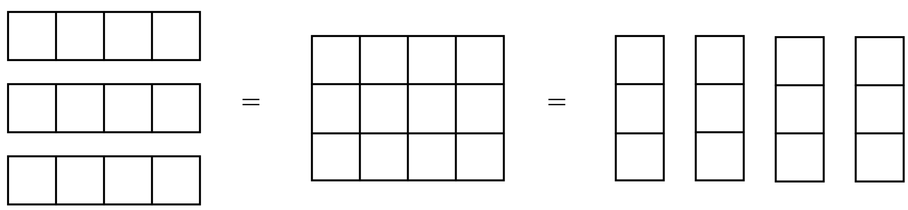
8. From Variants Of Fractions To Prime Factorization

R

The “strange” feature of fractions is that different ones can still be equal. For example, $\frac{2}{3} = \frac{4}{6}$. We called this expansion and simplification, but not all equal fractions are expanded or simplified from each other. Indeed, $\frac{4}{6} = \frac{10}{15}$ and the 10 is not multiple of 4. Yet, both sides are expansions of $\frac{2}{3}$. Thus, the obvious question is whether all equal fractions are merely expansions of each other or of a common third one. The answer is yes, and we’ll show this in the followings by a very simple geometrical way.
An other hidden problem was left at the simplifications themselves. We assumed that crossing out the common factors from the numerators and denominators leads to the unexpanded forms of which we spoke above. For example, $\frac{4}{6} = \frac{2}{3}$ and $\frac{10}{15} = \frac{2}{3}$. Here we only simplified with one factor in both cases, but in more complicated fractions or products of fractions, we can do many simplifications. So the question is whether the final simplified form does depend on what order we do these simplifications, or not. We’ll show that the order is immaterial. Amazingly, the solution of this second problem follows directly from the solution of the first, that is from the expansion of any two equal fractions from a common one.
In the followings, we start from scratch and won’t rely on any earlier naïve concepts.

R

From counting, the addition of numbers follows by natural intuitions. Indeed, $4 + 3$ can be achieved as continuing the counting from 4 with 3 steps more. The fact that $4 + 3 = 3 + 4$ does not follow from this procedure, but gradually we learn that addition is also the combining of sets. Then, 4 apples plus 3 apples being the same as 3 apples plus 4 apples, is obvious.
Multiplication is the repeated additions of identical members.
For example, $3 \cdot 4 = 4 + 4 + 4$, on the other hand, $4 \cdot 3 = 3 + 3 + 3 + 3$.
These being the same is not obvious at all and doesn’t follow from this meaning.
On the other hand, the geometrical meaning of $3 \cdot 4$ can be the number of tiles in a rectangle with sides 3 and 4. Then, this area can be added by rows or columns:

$3 \cdot 4 =$  $= 4 \cdot 3$

As we learn the times table, we get a sense of “natural evidence” for multiplications. This becomes the most controversial at the breaking up of numbers into products of smallest units, so called primes. For example, $100 = 2 \cdot 50 = 2 \cdot 2 \cdot 25 = 2 \cdot 2 \cdot 5 \cdot 5$. Here we proceeded in the order to always find the next smallest possible prime. But we can go differently too!
For example: $100 = 5 \cdot 20 = 5 \cdot 2 \cdot 10 = 5 \cdot 2 \cdot 2 \cdot 5$. The above mentioned independence of a product from its two members easily generalizes to more members, so it’s not surprising that the number 100 is the value in both cases. However, it is far from obvious that we ended up with the same primes at all! To feel this, we should imagine a huge number instead of 100.
And yet, it is true for any number: No matter what next possible prime numbers we choose. In the end, the same set of primes will appear at every sequence of choices, only in different order. This so called “unique prime factorization theorem” is without doubt the most important fact of the natural numbers. Our goal is a crystal clear proof of it:

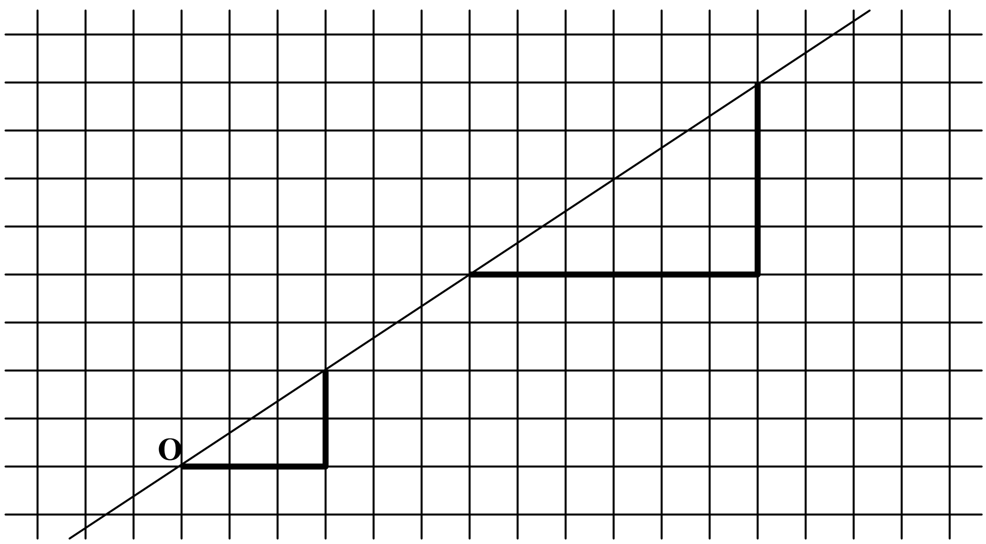
D

In the followings, every letter stands for natural numbers: 1 , 2 , 3 , . . .

1.) x <u>divides</u> y or is a <u>divider</u> of y, if $y = m\,x$, that is if y is a multiple of x.
We include the $m = 1$, that is $y = x$ case too.
1 divides every number because $y = y\,1$.

2.) x and y are <u>a simple pair</u> if the only number that divides both of them is 1.
For example: 8 and 15 are a simple pair because:
The dividers of 8 are: 1 , 2 , 4 , 8 and the dividers of 15 are: 1 , 3 , 5 , 15.
(Another name used for being simple pair is being relative primes.)

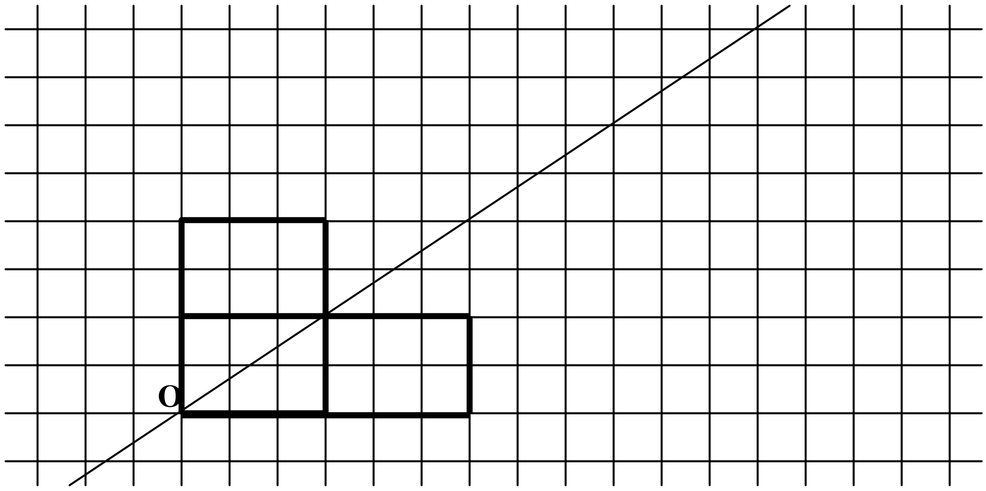
3.) For any two x , y we define the <u>fraction</u>: $\frac{y}{x}$. The y is <u>numerator</u> and x <u>denominator</u>.
The value of a fraction can be defined as the partial section of a distance, area or any other geometrical size. So, we cut the full size in x equal parts and take y many of these. Then, different fractions may be equal, for example $\frac{2}{3} = \frac{4}{6}$, because taking "two thirds" of something is the same as taking "four sixth" of it.
A completely different geometrical definition of fractions could be the slope of lines that connect two P , Q grid points in a grid system. The horizontal difference of P , Q is x units, while the vertical is y. Then, $\frac{y}{x}$ is the ratio of elevation compared to the advance.
Declining lines could even be interpreted as negative fractions, but we ignore this now. The equality of different fractions would then mean the parallelity of the lines. Or, if we only regard lines through a fixed O origin, then the equal fractions are only between pairs of points on one line:



The line in this picture is the $\frac{2}{3} = \frac{4}{6}$ line.

Without such geometrical definitions, the equality of two $\frac{y}{x}$ and $\frac{Y}{X}$ fractions can be also defined as $X\,y = x\,Y$.

This equality of two products can also be seen in the grid system as the equal area of two rectangles. In our above example, $6 \cdot 2 = 3 \cdot 4$
Also, if these two triangles are moved into each other's corner, then the continuation of the other sides will cross on the previously used line with $\frac{2}{3} = \frac{4}{6}$ slope:



4.) The equal fractions are also called as <u>variants</u> of each other.

For any two variants $\frac{y}{x} = \frac{Y}{X}$ if $x < X$ then also $y < Y$,

so we can speak about <u>smaller or bigger variants</u>.

5.) $\frac{Y}{X}$ is an <u>expansion</u> of $\frac{y}{x}$, if $Y = m\,y$ and $X = m\,x$.

The expansion is a variant, because $\frac{m\,y}{m\,x} = \frac{y}{x}$.

If $\frac{Y}{X}$ is an <u>expansion</u> of $\frac{y}{x}$, then we can also say that:

$\frac{y}{x}$ is a <u>simplification</u> of $\frac{Y}{X}$.

6.) $\frac{y}{x}$ is a <u>simple fraction</u> if x and y are a simple pair.

This is logical with the previous simplification name because:

A $\frac{y}{x}$ fraction is simple exactly if it can't be simplified.

7.) $\frac{y}{x}$ is a <u>minimal fraction</u> if there is no smaller variant of it.

For example: $\frac{2}{3}$ is a minimal fraction because none of $\frac{1}{2}$ or $\frac{1}{1}$ are equal to it.

If y or x is 1, then we obviously have a minimal fraction.

The $\frac{y}{1}$ are called wholes and the $\frac{1}{x}$ are called reciprocals.

All minimal fractions must be simple. Indeed, otherwise, that is if X and Y are not a simple pair, then they have a c common divider and so the fraction can be simplified:

$\frac{Y}{X} = \frac{c\,y}{c\,x} = \frac{y}{x}$.

The reverse, that is that all simple fractions are minimal is far from obvious.

The opposite would simply mean that besides the minimal fraction, there are other simple variants too. In short, two simple fractions could be equal. The impossibility of this is not evident if we regard fractions with huge numerators and denominators.

8.) A number is a prime, if it is not 1, but it can only be divided by 1 and itself.
$2, 3, 5, 7, 11, 13, 17, \ldots$ are primes.
$4, 6, 8, 9, 10, 12, 14, 15, 16, \ldots$ are not primes, rather so called composites.
The exclusion of 1 from the primes was logical because the composites can all be "composed" from primes.
For example: $4 = 2\cdot2$, $6 = 2\cdot3$, $9 = 3\cdot3$, $12 = 2\cdot2\cdot3$, $\ldots$
This prime composition or prime factorization is unique except of the order of the appearing primes. If 1 were allowed as a prime, then it could be repeated as many times as we wish, thus making the factorizations not unique.
This uniqueness of prime factorization follows from the above mentioned identity of minimal and simple fractions. So, in the end we obtained results about the products of natural numbers, that can be easier proved by looking at the divisions, that is fractions. This is typical in mathematics to widen the scope of a field, to get an easier proof. If we wanted to restrict our attention, to naturals and products only, then the same proofs would become much more artificial and concealed.

T

1.) If $\frac{Y}{X}$ is a bigger variant of $\frac{y}{x}$, then $\frac{Y-y}{X-x}$ is a variant of them too.

2.) If $\frac{y}{x}$ is minimal and is a variant of $\frac{Y}{X}$, then $\frac{Y}{X}$ is a multiple variant of $\frac{y}{x}$.

3.) If $\frac{y}{x}$ is simple, then it is minimal.

4.) If x divides a $y\,z$ product, but x and y are a simple pair, then x divides z.

5.) If a p prime divides a $q_1\, q_2$ product, then p divides at least one of them.

6.) If a p prime divides a $q_1\, q_2 \ldots q_n$ product, then p divides at least one of them.

7.) If a p prime divides a $q_1\, q_2 \ldots q_n$ product of primes, then p is one of them.

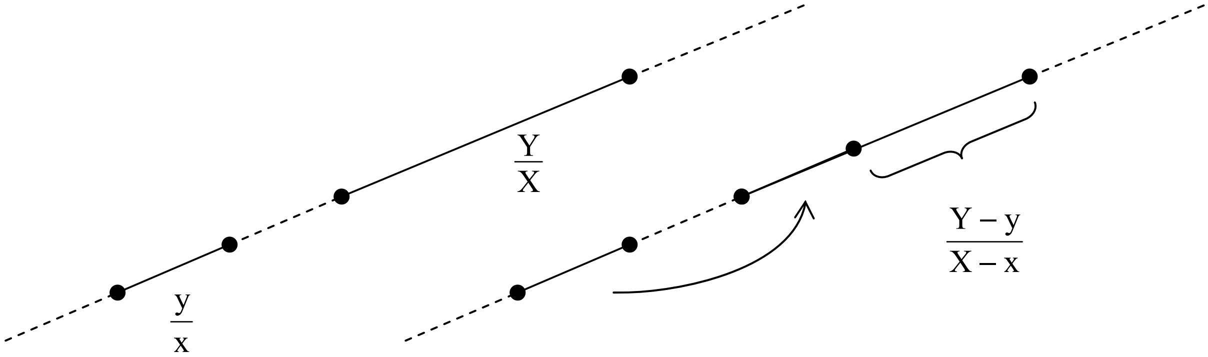
8.) If $p_1\, p_2 \ldots p_m = q_1\, q_2 \ldots q_n$ are equal prime products, then the p and q primes are the same except maybe in different order.

P

1.) $\frac{Y}{X} = \frac{y}{x} \rightarrow$

$$Y x = X y \rightarrow Y x - y x = X y - x y \rightarrow x(Y-y) = y(X-x) \rightarrow \frac{Y-y}{X-x} = \frac{y}{x}$$

This fact can be seen from the line representation of fractions by simply sliding the $\frac{y}{x}$ fraction into the $\frac{Y}{X}$ and thus, the $\frac{Y-y}{X-x}$ appears at once:

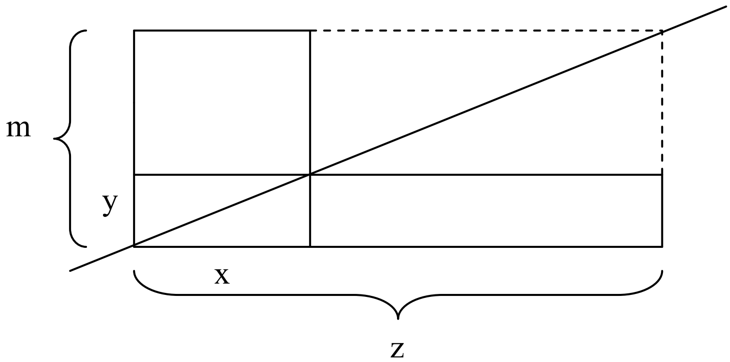


2.) $\frac{Y}{X}$ must be a bigger variant, since $\frac{y}{x}$ was a minimal. Thus, we can repeatedly subtract y from Y and x from X and get new variants. If there were a final remainder of y in Y, and x in X, then this last variant would be smaller than $\frac{y}{x}$, contradicting that it was minimal. Thus, there is no remainder and so $\frac{Y}{X}$ was indeed multiple variant of $\frac{y}{x}$.

3.) If it were not minimal, then there were an other minimal among its variants, of which it were an $m \neq 1$ multiple variant by 2.). Thus, m would be a common divider of x and y.

4.) $m\,x = y\,z \rightarrow \frac{m}{z} = \frac{y}{x}$ but, $\frac{y}{x}$ is a simple fraction, and thus by 3.) minimal.

Then, by 2.) $\frac{m}{z}$ is multiple variant of it, and so z is multiple of x.



5.) If p doesn't divide q_1, then p and q are a simple pair, and thus by 4.), p divides q_2.

6.) If p doesn't divide q_1, then by 5.) it divides $q_2\ q_3 \ldots q_n$.
If it doesn't divide q_2, then again by 5.) it divides $q_3 \ldots q_n$ and so on.
Finally, p must divide q_n.

7.) By 6.) p divides a q, but then it must be equal to it too, because q is prime now.

8.) p_1 is one of the q-s by 7.), so we can divide with these.
Then p_2 is also one of the q-s, so we can divide again, and so on.

9. Infinite Decimals, Irrational Numbers

T

The decimal system makes the basic operations of whole numbers easily calculable, digit by digit:

Addition:

```
  3 5 7
+   7 9
  -----
  1 1
  4 3 6
```

Fast way:

```
  3 5 7
+   7 9
  -----
  4 3 6
   (1 1  remainders only in head)
```

Subtraction:

```
  3 5 7
-   7 9
  -----
  1 1
  2 7 8
```

Fast way:

```
  3 5 7
-   7 9
  -----
  2 7 8
   (1 1  borrowed only in head)
```

Multiplication:

```
3 5 7  •  79
2 3 4
  1 5 9
  2 4 6
    7 5 3
---------
  2 2
2 8 2 0 3
```

Fast way:

```
   3 5 7  • 79
 2 4 9 9
   3 2 1 3
 ---------
 2 8 2 0 3
```

Division:

```
1 9 9 8  :  5  =  3 9 9
1 5
  4 9
  4 5
    4 8
    4 5
      3  remainder
```

Fast way:

```
1 9 9 8  :  5  =  3 9 9
  4 9
    4 8
      3
```

We have to be careful for 0-s:

```
2 4 5 9  :  6  =  4 0 9
  0 5 9
      5
```

We can continue the division process by bringing down newer and newer 0-s and thus, get an infinite decimal form of the result:

$$2459 : 6 = 409.8333\ldots = 409.8\overline{3}$$

```
2 4 5 9  :  6  =  4 0 9 . 8 3 3 3 . . .  =  4 0 9 . 8 3̅
  0 5 9
      5 0
        2 0
          2 0
```

$$469 : 75 = 6.2533\ldots = 6.25\overline{3}$$

```
469 : 75 = 6.2533... = 6.253̄
 190
   400
     250
       250
```

In general:

$$\frac{m}{n} = .B_1 B_2 \ldots B_b\ P_1 P_2 \ldots P_p\ P_1 P_2 \ldots P_p \ldots = .\underbrace{B_1 B_2 \ldots B_b}_{\text{beginning}}\ \underbrace{\overline{P_1 P_2 \ldots P_p}}_{\text{period}}$$

Interestingly, the reverse problem, that is how to find the fraction for a infinite periodical decimal, is also very simple, namely:

$$.B_1 B_2 \ldots B_b \overline{P_1 P_2 \ldots P_p} = \frac{B_1 \ldots B_b}{1\underbrace{0\ldots0}_{b}} + \frac{P_1 \ldots P_p}{\underbrace{9\ldots9}_{p}\ \underbrace{0\ldots0}_{b}}$$

Example: $.2305757\ldots = \frac{230}{1000} + \frac{57}{99000}$.

$$6.2533\ldots = 6 + \frac{25}{100} + \frac{3}{900} = 6 + \frac{1}{4} + \frac{1}{300} = \frac{7200}{1200} + \frac{300}{1200} + \frac{4}{1200} = \frac{7504}{1200} = \frac{469}{75}$$

This fact that all periodical decimals are actually fractions, prove it at once that there must be numbers that are not fractions, namely the infinite decimals that are non periodical. If we create an infinite decimal with randomly picked digits, it should be obviously such, but we can even use rules to generate the digits and yet not have a repeating period. For example:
$0.1234567891011121314151617181920 2122\ldots =$ not periodical.

R The fractions are also called rationals, while the numbers that are non fractions as irrationals.

The infinite decimal system made it obvious that there are irrational numbers and it even suggests that there are more irrationals than rationals. But this “obviousness” is a false formalism, that jumps through the original problem of what are numbers at all.
If we start with distances then the fractions or rational numbers are merely the exact whole divisions of a fixed unit interval. Then the problem of irrationality is to create a distance that is not obtainable from exact division of the unit. This was investigated by the greeks already.
The infinite decimal solution of course can be translated back to distances as adding together the smaller and smaller distances that correspond to the infinite many digits. Thus, the infinite decimal system also shows at once that infinite many distances can add up to a single distance. This second problem was also investigated by the greeks, but the two problems were not combined.
Today, when we look at an infinite decimal like $0.12345678910111213\ldots$ we don’t actually visualize how it is a distance made up as: $0.1 + 0.02 + 0.003 + \ldots$
In order to appreciate the fact that infinite many small distances can add up to a single one, we should start with the simplest case of: $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots = 1$

Indeed, $\frac{1}{2} + \frac{1}{4} = \frac{2+1}{4} = \frac{3}{4} = 1 - \frac{1}{4}$

$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{4+2+1}{8} = \frac{7}{8} = 1 - \frac{1}{8}$, and so on.

The general rule that leads to the numerators is $2^n + 2^{n-1} + \ldots + 2 + 1 = 2^{n+1} - 1$.
This itself can be proved easily step by step showing it for higher values of n:

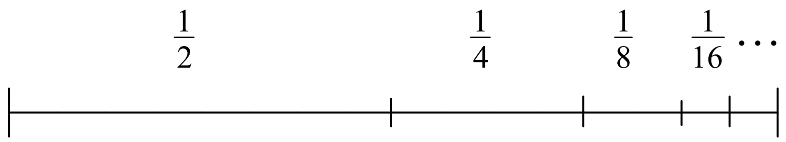
$$1 = 2 - 1 \qquad / \bullet 2 , + 1$$

$$2 + 1 = 4 - 2 + 1 = 4 - 1 \qquad / \bullet 2 , + 1$$

$$4 + 2 + 1 = 8 - 2 + 1 = 8 - 1 \qquad / \bullet 2 , + 1$$

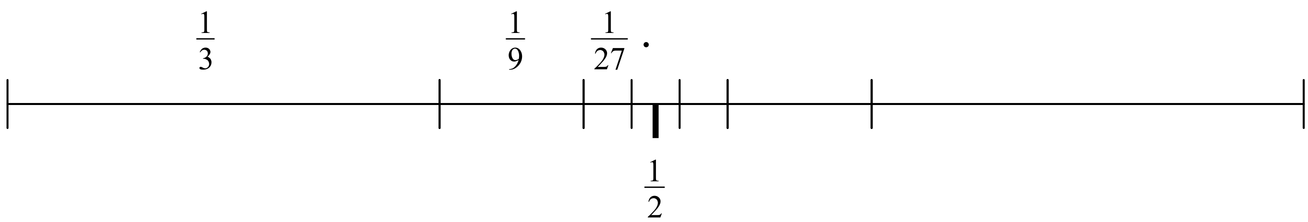
$$\vdots$$

On a unit distance the $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots = 1$ equality can be seen directly too:



The next simplest case would be: $\frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \frac{1}{81} + \ldots = ?$

Here a drawing with distances shows quite convincingly that the sum should be $\frac{1}{2}$:



To prove this formally is quite easy with some tricks used for the sum as an equation:

$$x = \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \ldots \qquad / \bullet \frac{2}{3} = 1 - \frac{1}{3}$$

$$\frac{2}{3} x = \left(\frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \ldots\right)\left(1 - \frac{1}{3}\right) = \frac{1}{3} - \frac{1}{9} + \frac{1}{9} - \frac{1}{27} + \frac{1}{27} - \frac{1}{81} + \ldots = \frac{1}{3} \qquad / : \frac{2}{3}$$

$$x = \frac{1}{3} \bullet \frac{3}{2} = \frac{1}{2}$$

In fact, the same trick works for the general case with an s starting value and q multiplier or quotient as called:

$$x = s + sq + sq^2 + sq^3 + \ldots \qquad / \bullet (1 - q)$$

$$x\,(1 - q) = (s + sq + sq^2 + sq^3 + \ldots)(1 - q) = s - sq + sq - sq^2 + \ldots = s \qquad / : (1 - q)$$

$$x = \frac{s}{1 - q}$$

R

Achilles Paradox

How deeply disturbed the greek thinkers were by the fact that infinite many values can add up to a single finite one, can be seen from the famous paradox of Achilles and the turtle.
They claimed that even if the runner Achilles is 100 times faster than a turtle, he shouldn't be able to catch up with the turtle that starts with an s advantage.
Indeed, they argued that by the time Achilles reaches the point where the turtle started, that is s distance, the turtle will be $\frac{s}{100}$ further ahead. When Achilles reaches this point then the turtle will be again $\frac{s}{10000}$ away. And so on, the turtle is "always" ahead.
The error is the false application of the "always". Just because something happens infinite many times, it doesn't mean that it will be forever. Indeed, if the rain starts now, then there were infinite many times just before when it didn't rain, namely a minute ago, half a minute ago, a third minute ago, and so on.

$s + \frac{s}{100} + \frac{s}{10000} + \ldots = \frac{s}{1-\frac{1}{100}} = \frac{100\,s}{99} = s \bullet 1.1111\ldots$ is exactly the distance where Achilles reaches the turtle.

R

Anti Achilles Paradox

The solution to the Achilles paradox, that is the acceptance of the fact that infinite many smaller and smaller distances added up to a single finite value might make us jump to the wrong conclusion that smaller and smaller amounts <u>always</u> add up to a finite value.
This is not so and we can easily create smaller and smaller values that in the end add up to infinity. The easiest way is to start with adding up a fix value, say 1, infinite many times, which is obviously adding up to infinity: $1 + 1 + 1 + 1 + \ldots = \infty$
Then, we can distribute each member into more and more pieces, and thus getting smaller and smaller members. For example, with equal distributions:

$$1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \ldots = \infty$$

Quite surprisingly, but not as surprisingly as without this introduction, it's also true that:

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots = \infty$$

The general question of when the smaller and smaller amounts add up to infinity, is quite hard.
The square reciprocals for example are not enough to produce infinity:

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \ldots = \frac{\pi^2}{6} \qquad (\pi = 3.14\ldots)$$

On the other hand, the prime reciprocals will give infinity:

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \ldots = \infty$$

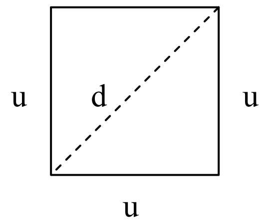
This suggests that there are more primes than squares, and indeed between all consecutive square numbers there are more and more primes:

Between 1 and 4, are 2 and 3.
Between 4 and 9, are 5 and 7.
Between 9 and 16, are 11 and 13.
Between 16 and 25, are 17, 19, 23, and so on.

Amazingly, it is still not proven that between every two consecutive squares, there is at least one prime.

R Irrational Surds

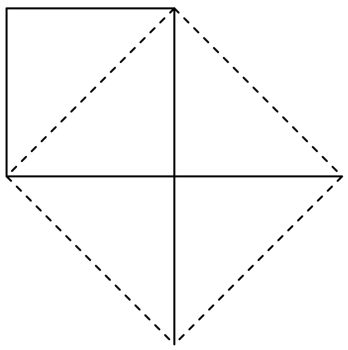
As I mentioned the greeks looked for the irrational numbers as actual distances that can not be obtained from the exact, whole divisions of a unit. The first concrete example they found was the diagonal of a square with the unit sides:



$$d \neq \frac{m}{n} u$$

in other words if $u = 1$, d is irrational.

From the Pythagoras Theorem $d^2 = u^2 + u^2$ so with $u = 1$, $d^2 = 2$ or $d = \sqrt{2}$.
By the way, this can be seen easily without the Pythagoras Theorem too from:



The irrationality of $\sqrt{2}$ means quite simply that there is no $\frac{m}{n}$ fraction, so that $\left(\frac{m}{n}\right)^2 = 2$.

Or to put it even more concretely $m^2 = 2\,n^2$ is impossible, that is the double of a square number can't be a square itself. Our earlier result of the unique prime factorization of numbers, proves this at once. Indeed, a square number must have every prime factor even many times, so as a special consequence n^2 has either no or even many 2 factors. Then, $2\,n^2$ must have odd many 2 factors, unlike the left side m^2. This argument shows that not only $\sqrt{2}$ is irrational, but in general any so called "surd" of a whole number is either a whole or irrational:
The surd $\sqrt[k]{x}$ is the y value for which $y^k = x$.
For a whole number w, $\sqrt[k]{w}$ can be whole, for example $\sqrt[3]{8} = 2$.
In general, if all prime factors of w are in multiples of k, then $\sqrt[k]{w}$ will be a whole.
However, if $\sqrt[k]{w}$ is not whole, then it can't be an $\frac{m}{n}$ fraction either.

Indeed, $\left(\frac{m}{n}\right)^k = w$ means $m^k = w\,n^k$.

Now if w has a p prime factor not with a multiplicity of k, then p would appear with multiplicity of k in m^k, but with not this multiplicity in the right side $w\,n^k$.

Part Two: Geometry
1. Euclidian Construction

R

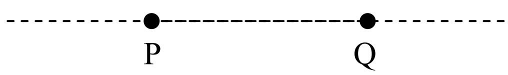
Euclidian constructions use ruler and compass. But even the usage of these are restricted.

A ruler can only be placed on two points to draw an infinite line through them. The compass can be placed on any two already obtained points and this distance is kept if we remove the compass. So, it can be used to draw a circle around any other point with this radius. Points are obtained as crossings of lines or circles. The main restriction is that the ruler can not be used as a measuring device and it can not be moved in wanted positions. For example, we might feel that after drawing two circles, we could find their common tangent by simply placing our ruler to touch them both. If we look a bit closer we can understand why this is not allowed. When we place our ruler on two points, we might have to move the ruler too, but this motion can be made totally exact in the following way: We stick a pin into one of the points and then resting the ruler against it, we can turn the ruler until it will exactly go through the other point. With two circles, the situation would be very different! We can easily slide the ruler until it touches one of the circles, but then it might double cross or not cross at all the other circle. If we pin this touching point and turn the ruler to touch the other circle, then the touching point will not remain touching point anymore. So to obtain a perfect touch on both circles, we would have to make infinite many corrections.
Euclid not only devised his constructions, but also listed the axioms that rule the points and lines of a plane. These axioms are self evident by our intuitions and when we proceed with constructions we use them without even noticing. Of course, to prove more complicated theorems, it's useful to see the chain of assumptions that were used. To find tricky sequence of constructions don't always require such theorems and then the used axioms are unimportant. Rather we need an exact way of telling our sequence of constructions.
I will present such method:

D

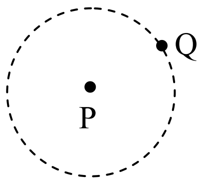
Elementary construction steps:

1.) Line across two points: < PQ >

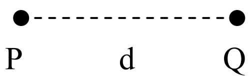


A line across a single P point can be picked as < P >.
< P > can also abbreviate an earlier obtained line.

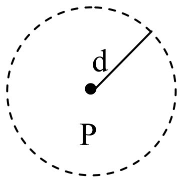
2.) Circle around a point and passing through an other: (PQ)



3.) Placing a distance: d = PQ



4.) Circle around a point with a d radius: (Pd)



(PQR) is thus the circle around P with d = QR radius.
Later we could use (P,Q,R) for the circle through the points P,Q,R.
Or (P,Q) for a circle through two given points.

5.) Crossings:

two lines:

<. . .> }
<. . .> }

one line one circle:

<. . .> }
(. . .) } < P, Q

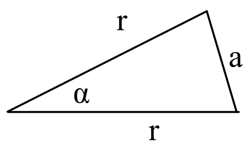
two circles

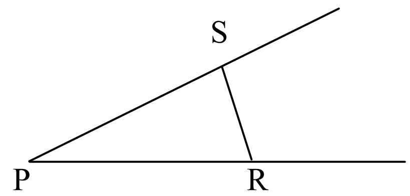
(. . .) }
(. . .) } < P, Q

If we just want to pick an arbitrary P point from a line or a circle, we'll also use this notation as: <. . .> } P (. . .) } P.

6.) Angle, copied: < < P > α > α is given as an a chord in a circle with r radius.

< P >, (Pr) } R   (Ra), (Pr) } S   < PS >



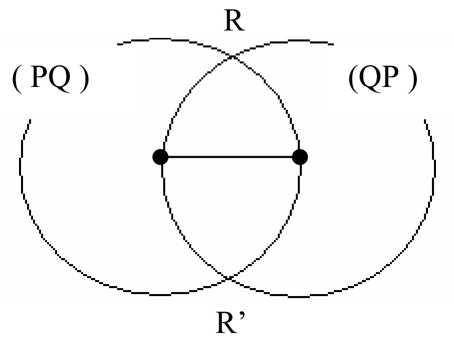


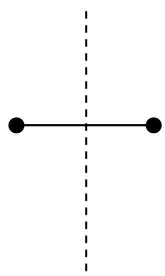
If < P > is < PQ >, then < < P > α > is abbreviated as < PQ α >.

T

Basic constructions:

1.) Symmetry line: < P ⊥ Q > = (PQ), (QP) } < R, R'   < RR' >





2.) Perpendicular line: < < . . > ├ R > = (Rd), < . . > } < S, T   < S ⊥ T >

If < . . > is < PQ >, then < < . . > ├ R > is abbreviated as < PQ ├ R >.

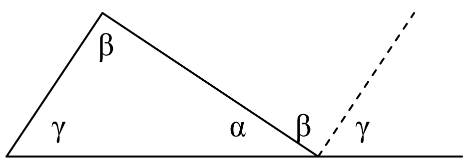
3.) Parallels through PQ: < P || Q > = < PQ ├ P > , < PQ ├ Q >

4.) Parallel line: < < . . > || R > = < < < . . > ├ R > ├ R >

T

Outer "supplementing" angle of one angle of a triangle is equal to the other two. That is:
Three angles of a triangle is 180°.

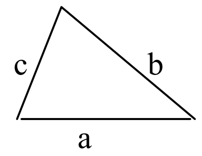
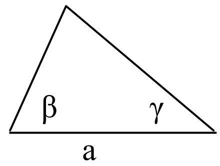
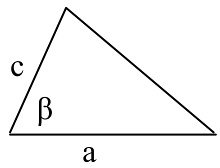
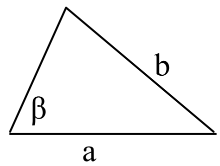
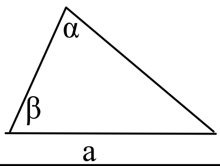
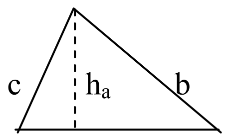
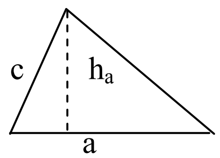
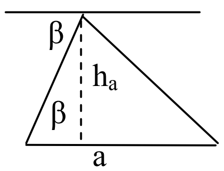
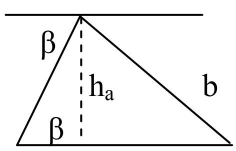
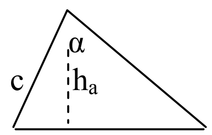
P



T

Basic triangle constructions:

From the given data (distances or angles), we have to find some A, B, C points, so that the ABC triangle possess the given data. Apart from the sides and angles, we can also use as data the so called "height lines", that go from a corner perpendicularly to the opposite side.

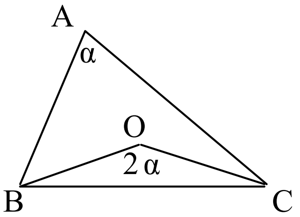
Data	Sketch	Construction
a, b, c		$a = BC$ (Bc) (Cb) } A
a, β, γ		$a = BC$ $< BC\beta >$ $< CB - \gamma >$ } A
a, c, β		$a = BC$ $< BC\beta >$ (Bc) } A
a, b, β		$a = BC$ $< BC\beta >$ (Cb) } < A, A’
a, α, β		$a = BC$ $\gamma = 180 - \alpha - \beta$ we can apply the second case
b, c, h_a		$h_a = AH$ $<A \parallel H> = <A>, <H>$ (Ac), $<H>$ } B (Ab), $<H>$ } C
a, c, h_a		$h_a = AH$ $<A \parallel H> = <A>, <H>$ (Ac), $<H>$ } B (Ba), $<H>$ }
a, β, h_a		$h_a = AH$ $<A \parallel H> = <A>, <H>$ $<<A> \beta>$, $<H>$ } B (Ba), $<H>$ } C
b, β, h_a		$h_a = AH$ $<A \parallel H> = <A>, <H>$ $<<A> \beta>$, $<H>$ } B (Ab), $<H>$ } C
α, c, h_a		$h_a = AH$ $<A \parallel H> = <A>, <H>$ (Ac), $<H>$ } B $< AB\alpha >$, $<H>$ } C

2. Symmetry Lines and Circle

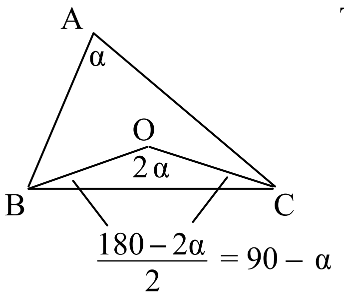
T

If A , B , C are not on one line, that is they form a triangle, then:

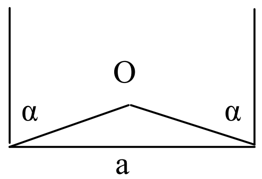
1.) The symmetry lines of the three sides go through one point.
2.) There is and there is only one circle going through A , B , C.
3.) Any side looks twice the angle from the center of the circle than from the corner.



4.) One side and the angle across determines already the circle around a triangle.



To construct O from a and α :

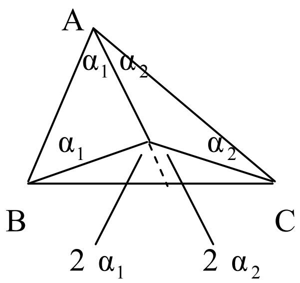


P

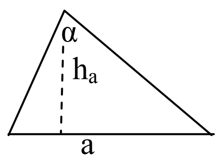
1.) < A ⊥ B > and < A ⊥ C > can not be parallel and thus, they cross in an O point. AO = BO because O is on < A ⊥ B >. AO = CO because O is on < A ⊥ C >. Thus, BO = CO and so O is on < B ⊥ C > too.

2.) AO = BO = CO, that is all the three points are equal distanced from O and so the circle around O with this radius goes through all three points. Any other circle would have to have its center on the same symmetry lines, so it can only be the same.

3.)



T

<table>
<tr><td>a , α , h_a</td><td></td><td>$h_a = A_0H_0$ <A_0 ∥ H_0> = < A_0 > , < H_0 > } B (Ba) } C
< H_0 >

< BC 90 – α > } O
< CB – (90 – α) >

(OB) } < A
< A_0 > A’</td></tr>
<tr><td colspan="3">Alternate construction including the 90 – α angles:
a = BC , < B ∥ C > = < B > , < C >

(B h_a) }
< B >

(C h_a) }
< C >

< < C > α > }
< < B > – α >

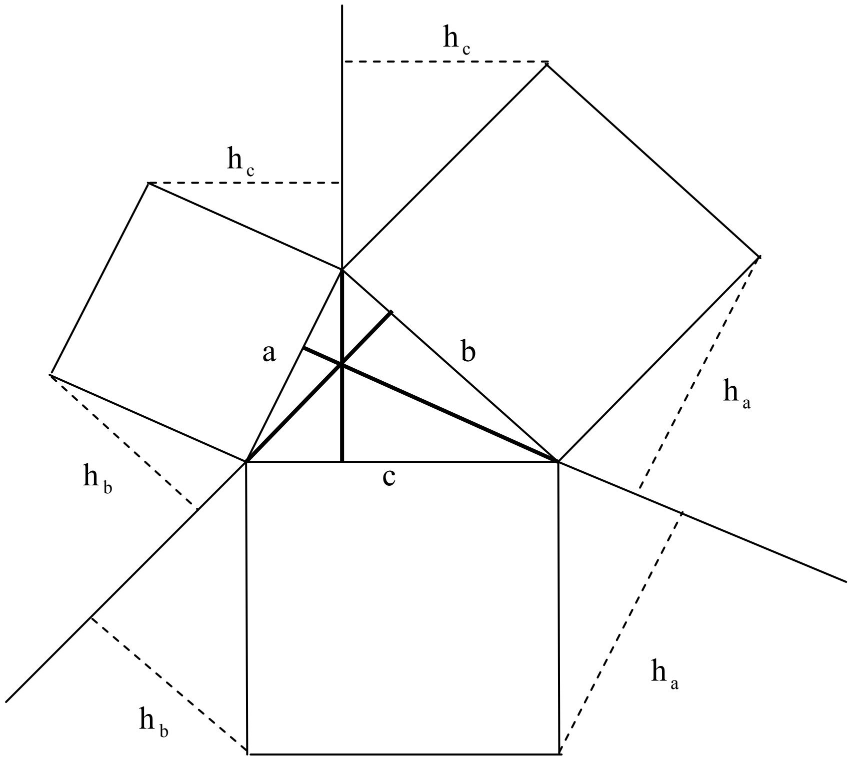
(OB) } < A
< B_0C_0 > A’</td></tr>
</table>

3. Height Lines and Side Squares

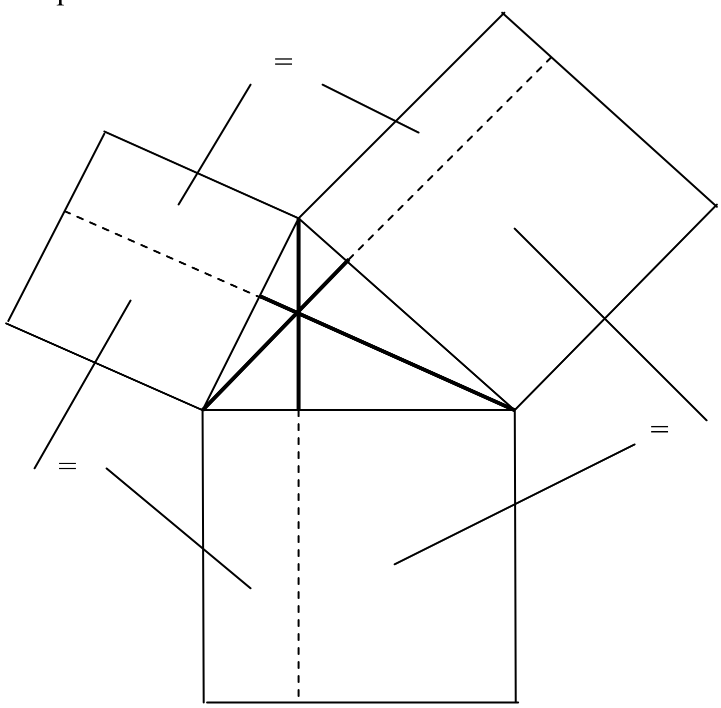
T

1.) The height lines of a triangle go through one point.

2.) If we draw squares on each side, then the new points obtained next to a corner are equal distanced from the height line of that corner, namely their distance is that height.



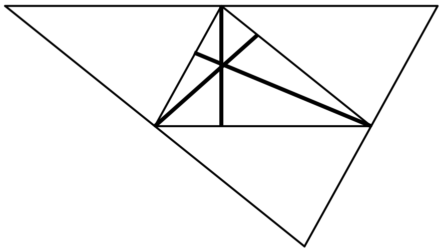
3.) If we cut the side squares in two by the height lines, then the parts that meet at a corner are equal in area.



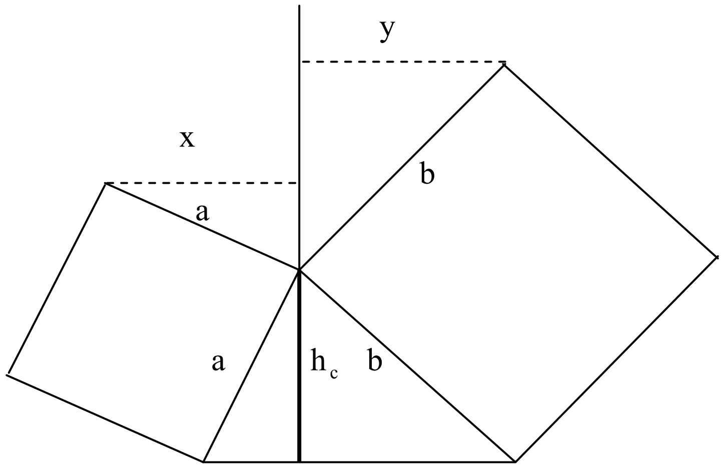
4.) Pythagoras Theorem: If $\gamma = 90^{\circ}$ then $a^2 + b^2 = c^2$

P

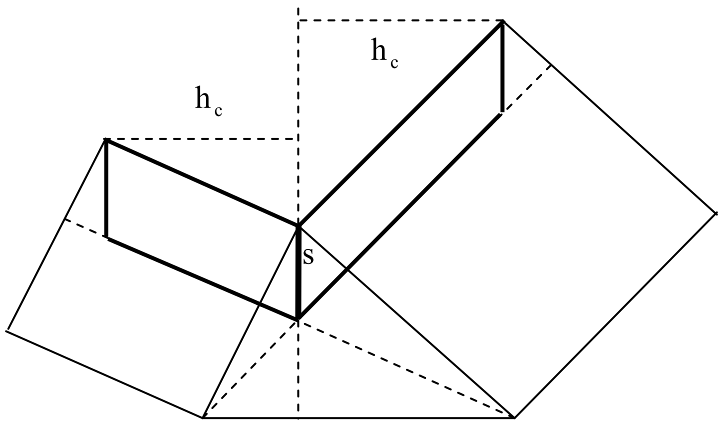
1.) Lets draw parallels with each side across the opposite corners. Thus, we obtain a bigger triangle in which the old heights become the symmetry lines.
(Thus follows from 1.) of the first theorem in the previous section.)



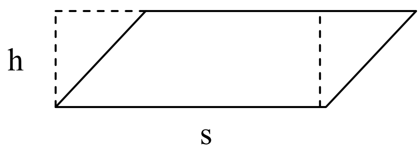
2.) The angle between a and h_c is the same as between a and x.
Thus, the two triangles are identical and so, $x = h_c$. Similarly, $y = h_c$.



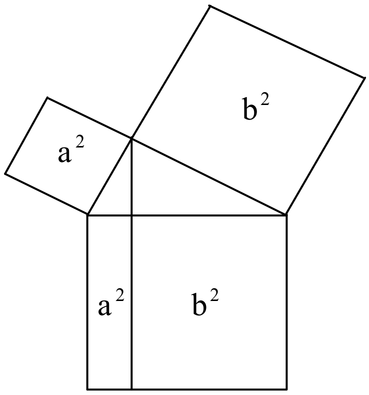
3.)



By drawing parallels with the height, we can change the square parts into parallelograms with a common s side on the height. The height of both parallelograms is h_c by 2.).
Thus, they both have the same $s\,h_c$ areas.
Here we used the fact that the area of a parallelogram is calculated as side multiplied by the height. This follows from the same being true for a rectangle:

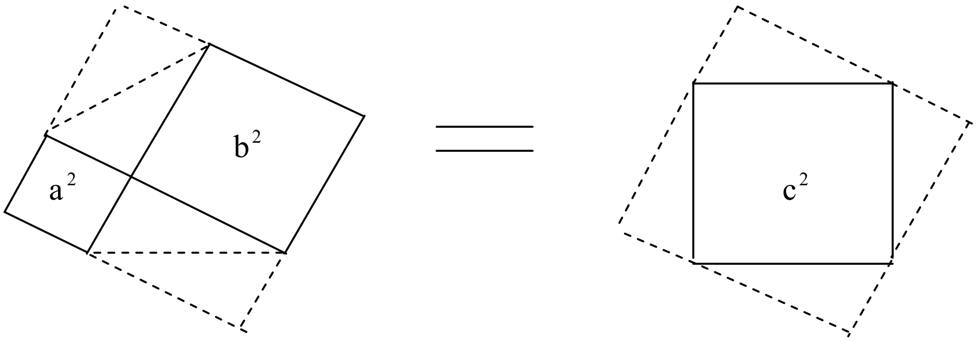


4.)

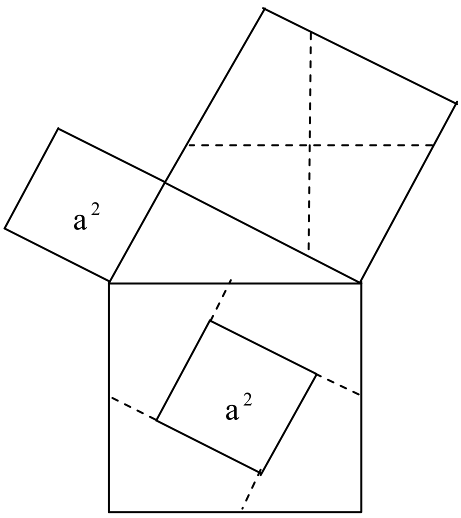


The a^2 and b^2 are the full parts and thus, are equal to the two parts of c^2.

R There are direct proofs for the Pythagoras Theorem that don't show the two square parts of c^2. One way to show that $a^2 + b^2$ is the same as c^2, is by adding the same areas to both and then obtaining identical objects. The oldest proof adds four of the a, b, c triangles and achieves identical squares with $a + b$ sides:



The other more direct way is to cut the squares themselves into sections and re-arrange them so that a^2 and b^2 becomes c^2. The simplest of these is cutting only the bigger b^2 into four identical pieces through its center, and arranging these around the whole a^2 shifted into the center of c^2:

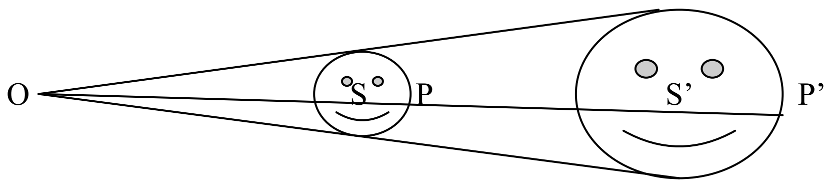


4. Similarity

D

1.) Two point sets are λ-proportional or similar if there is a $P \leftrightarrow P'$ one to one correspondence between their points, so that $P'Q' = \lambda PQ$ for all P', Q'.
If $\lambda = 1$ the two sets are isometric (iso = equal , metric = size).

2.) The S' set is a λ-projection of S from an O point if every P' point of S' is obtained by connecting a P point of S with O and changing OP to $OP' = \lambda OP$.



T

1.) Any triangle that has equal sides with an other can be moved over it with two of the followings: shifting, mirroring or turning.

2.) If {A', B', C'} is λ-proportional to {A, B, C}, then it can be obtained as a moved copy of a λ-projection.

3.) Two triangles are similar if and only if they have the same angles.

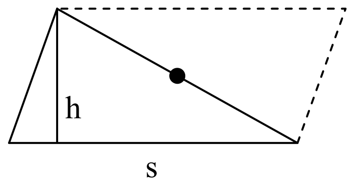
4.) Two sets are similar if and only if they have the same angles.

T

Area of triangle $= \dfrac{a\,h_a}{2} = \dfrac{b\,h_b}{2} = \dfrac{c\,h_c}{2}$

P

A triangle mirrored to the middle point of a side makes a parallelogram:

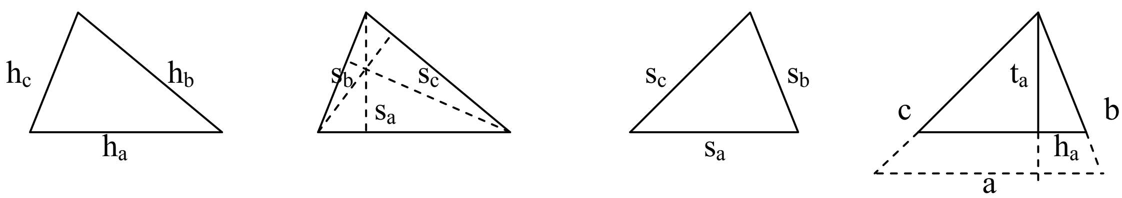


Parallelogram's area = s h

Triangle's area $= \dfrac{s\,h}{2}$

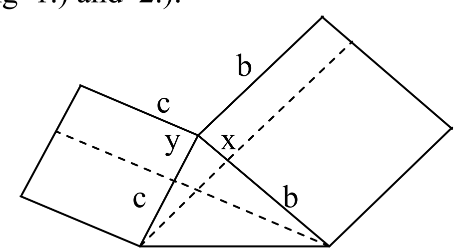
T

Triangle construction from the three heights. Twice the area = $a\,h_a = b\,h_b = c\,h_c$.
If the triangle constructed from h_a, h_b, h_c, has s_a, s_b, s_c, "secondary" heights, then twice the area of this triangle = $h_a\,s_a = h_b\,s_b = h_c\,s_c$. Thus, dividing the two equations:
$\dfrac{a}{s_a} = \dfrac{b}{s_b} = \dfrac{c}{s_c}$ and so the a , b , c triangle is similar to the s_a , s_b , s_c. This one can be constructed from the three sides and then increased until its t_a tertiary height becomes h_a.



R

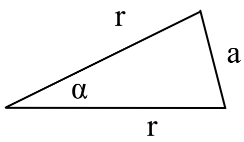
We can give a proof for 3.) of the theorem in the previous section with similarity, without using 1.) and 2.):

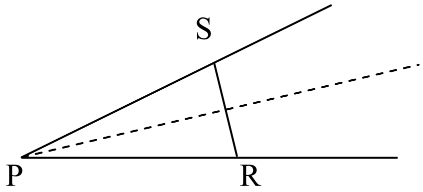


$\dfrac{x}{c} = \dfrac{y}{b} \rightarrow x\,b = y\,c$

5. Dividing Distances and Angles

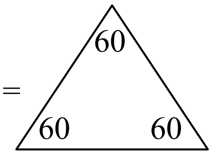
T Halving an angle $= < S \perp R >$





T Constructible special angles:

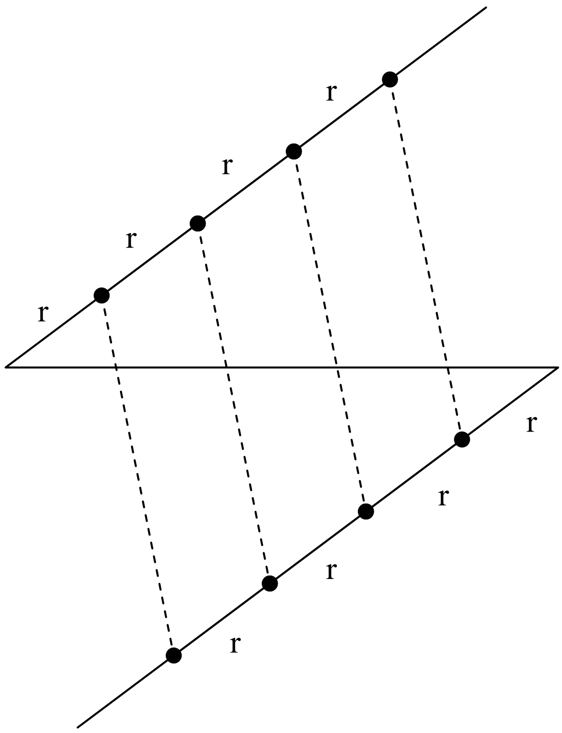
90° = perpendicular line , 60° =



45° , 30° can be obtained by halving these.

T n section of a d distance:

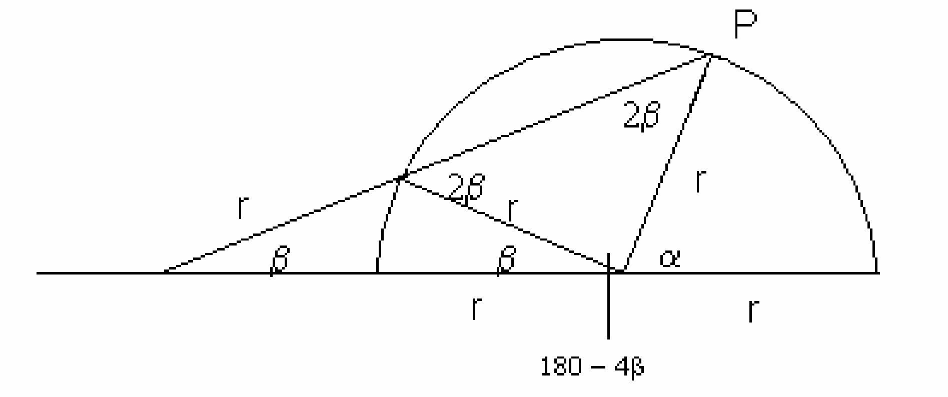
Measure arbitrary r distance, $(n - 1)$ times on any angle.
Do the same on the opposite end and opposite side.
Connect the $(n - 1)$ many pair of points with parallel lines:



R The n section of an angle is not constructible for arbitrary n and arbitrary angle.

With repeated halvings, any angle can be cut into $n = 2^k$ many equal parts. But:
The trisection of an angle is not constructible for an arbitrary angle. In particular:
The trisection of 60° is not constructible! In other words, 20° is not constructible!
Since 60° is the third of 180° and it is constructible, thus obviously:
For particular angles, we can construct the n section, even if n is not 2^k.
The most important special question was, what sections of the full 360°, are constructible?
In other words, how many sided symmetrical polygons can be constructed in a circle?
Gauss solved this problem when he was eighteen years old and this made him to become a mathematician, rather than a philologist. His statue stands on a seventeen sided polygon, because that was the smallest sided, that was not known to be constructible before him.

R Archimedes devised the following “construction” to get the third of an α angle:
Measure α up in a half circle and from the obtained P point of the circle, draw a line that determines a distance equal to the radius between the other crossing with the circle and the base line of the half circle! Then the angle between this line and the base line is $\frac{\alpha}{3}$.



Indeed, $\beta + 180 - 4\beta + \alpha = 180 \rightarrow \beta = \frac{\alpha}{3}$.

The problem is that we used our ruler as a measurer, which is not allowed.
We can measure distances only with the compass.

6. Non Commeasurable Distances

R

This is the old Greek way of looking at the irrational distances.

Indeed, if b is a unit then an a distance being irrational means that $a \neq \frac{m}{n} b$.

This is the same as $\frac{a}{m} \neq \frac{b}{n}$ for any m , n natural numbers.
In even simpler way, there is no u distance that would be common unit for a and b, that is $a = m u$ and $b = n u$ is impossible.
This was a better way of looking at irrationality, namely for finding actual examples a , b non commeasurable distances.

Among the $\frac{m}{n}$ fractions, the vital relation is the expansion and simplification.

$\frac{2}{3} = \frac{4}{6}$ and here $\frac{4}{6}$ is an expansion of $\frac{2}{3}$ by 2, while $\frac{2}{3}$ is a simplification of $\frac{4}{6}$.

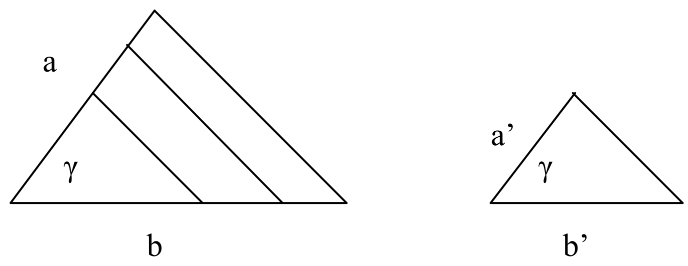
The $\frac{2}{3}$ can't be simplified anymore, so it is a simple fraction.

There is only one simple fraction among the equal fractions, or to put it another way, two simple fractions can't be equal. This fact might seem obvious from the practices of fraction simplifications, but it's far from obvious logically. Indeed, looking at two equal fractions with huge numerators and denominators, nothing guarantees that when simplified, they end up to be identical. The crucial concept was of course, the "minimality".
Among the equal fractions, there has to be a singular one that is minimal, in the sense of having the smallest possible numerator and denominator. Then, it was quite easy to show that all other fractions have to be extensions of this minimal. Thus, the minimal is also the simple.

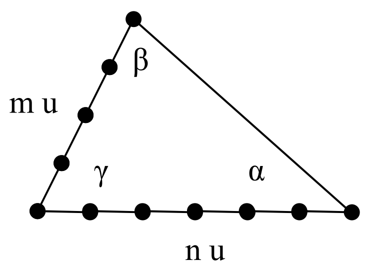
This "simple" versus "minimal" idea turns out to be still lingering among the general $\frac{a}{b}$ distance ratios. And they will provide the two methods to create incommeasurable distances.

Of course, if a , b are distances then there is no simplest or minimal among the $\frac{a}{b}$ ratios.

Already with a fix γ angle between them, we can continually create arbitrary small same ratios:



So the whole fractional simplicity and minimality seems meaningless directly.
But, a new question can be asked about the a , b distances. Namely, whether a u unit could be measured into both of them. That is whether $a = m u$, $b = n u$ is possible:



As we said, the Greek mathematicians realized that not all a , b distances can have common u unit. The basic proportionality of Euclidean geometry means that if an $(a , b , \gamma , \alpha , \beta)$ triangle leads to a , b that have common unit, then any $(a' , b' , \gamma , \alpha , \beta)$ is such too. Indeed, if (a' , b') is a p proportional change of (a , b) then $\frac{a}{b} = \frac{a'}{b'} = \frac{p\,a}{p\,b}$ and more importantly, if $a = m\,u$ and $b = n\,u$ then also $a' = m\,p\,u = m\,u'$ and $b' = n\,p\,u = n\,u'$.
So, the $u' = p\,u$ is a common unit for a' , b'.
In reverse, if an a , b have no common unit, then the proportional a' , b' have neither.
So the key to find distances without common units, is essentially finding γ , α , β angles where this happens. Of course, two of them already determines the third.
This theoretical idea of going for similar triangles is also the practical way to find distances without common unit. Namely, if $(a = m\,u , b = n\,u , \gamma , \alpha , \beta)$ implies another triangle $(a' = m'\,u , b' = n'\,u , \gamma , \alpha , \beta)$ so that $a' < a$ then the original triangle is impossible.
Observe that besides the obvious fix angle, that is similarity of the triangles, the crucial condition is that the new sides a' , b' are made of the same old u unit. Plus, a decrease: $a' < a$.
We can easily guess the argument that makes the original triangle impossible:
Exactly due to the proportionality of the Euclidean geometry, we can create from $(a' , b' , \gamma , \alpha , \beta)$ a new $(a'' , b'' , \gamma , \alpha , \beta)$ again and then again repeatedly.
But then, $a = m\,u > a' = m'\,u > a'' = m''\,u > \ldots$ is a contradiction.
Namely, how could multiples of a fix u decrease infinitely?
If we require a bit more precision, then it turns out that this impossibility is a bit trickier then seemed. Indeed, a first version could be that the infinite decrease gets arbitrary small, yet it can't be less than u. First of all, decreasing numbers don't necessarily have to decrease to 0. Secondly, a 0 multiple could reach 0 at once. So a better argument would make sure that 0 is not reachable. Then, we must have infinite many $a > a' > a'' > \ldots$ but, $a = m\,u$ means that only maximum m possible multiples could at all be, namely $u , 2\,u , 3\,u , \ldots , m\,u$.
Unfortunately, this:

$$(a = m\,u , b = n\,u , \gamma , \alpha , \beta) \quad \rightarrow \quad (a' = m'\,u , b' = n'\,u , \gamma , \alpha , \beta)$$

method is still very vague. How could we guarantee the new m' , n' without the particular knowledge of m , n ? In other words, a' , b' should be obtained directly from a , b .
The solution is very simple. All we have to require is that the new a' , b' are made from multiples of the old a , b , that is: $a' = p\,a \pm q\,b$, $b' = r\,a \pm s\,b$
Here, p , q , r , s are natural numbers and the $\pm$ means adding or subtracting distances.
Then, if $a = m\,u$ and $b = n\,u$, it guarantees at once that $a' = m'\,u$ and $b' = n'\,u$, simply because $p\,a , q\,b , r\,a , s\,b$, are all u multiples again and their sums or differences too.
So our method for impossibility of $a = m\,u$, $b = n\,u$ is now:

$$(a , b , \gamma , \alpha , \beta) \quad \rightarrow \quad (a' = p\,a \pm q\,b \ , \ b' = r\,a \pm s\,b \ , \ \gamma , \alpha , \beta)$$

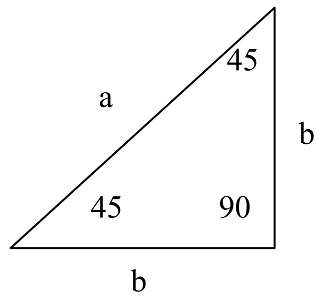
with the added requirements: $a' < a$, $a' , b' \neq 0$
The $a' < a$ goal of course especially emphasizes the $-$ choices in $\pm$.
This method is an "opposite" of the minimality among fractions.
Another method is the opposite of simplicity. Namely, if:

$$(a = m\,u , b = n\,u , \gamma , \alpha , \beta) \quad \rightarrow \quad m , n \text{ have common } f \text{ factor}$$

Then, $(a' = \frac{m}{f}\,u , b' = \frac{n}{f}\,u , \gamma , \alpha , \beta)$ is a similar triangle, so again we obtain an infinite decrease: $a > a' > a'' > \ldots$

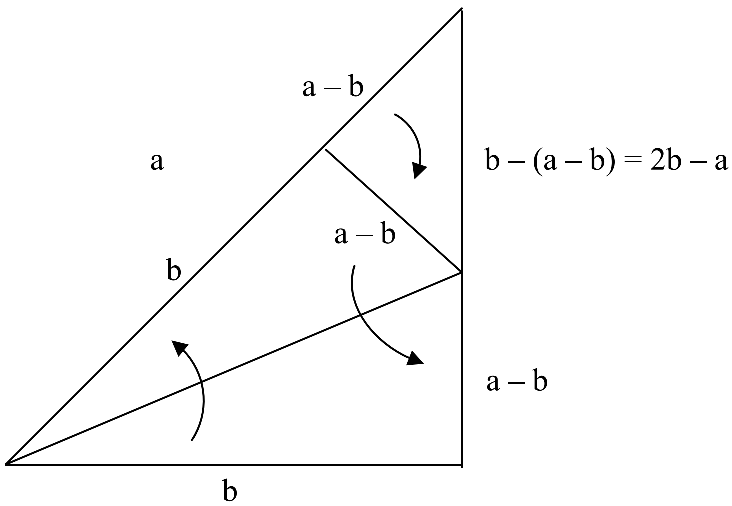
Such f common factor of m , n could be guaranteed by some special relation of a , b coming from the γ , α , β .
For example, $a^q = p\, b^q$ with a p prime number guarantees f = p.
Indeed, $(m\, u)^q = p\,(n\, u)^q$ → $m^q = p\, n^q$ so, m^q is dividable by p. Since p is prime, thus m must be dividable by p too. Using m = p r , then
$(p\, r)^q = p^q\, r^q = p\, n^q$, so $n^q = p^{q-1}\, r^q$ so n^q is dividable by p too and again, this implies that n is dividable too. Thus, indeed, p is common factor of m and n.

T The simplest (a , b , 45 , 90 , 45) triangle:



can be used for both methods:

P For the "± reduction" method:

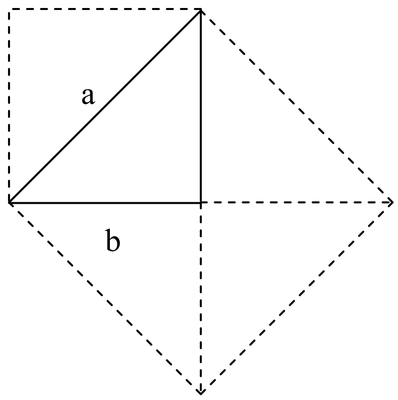


We mirrored the b side onto the a side through the angle halfer.
Thus, we obtained the leftover a – b , which determined the same in the two triangles.
This determined the leftover from the other b side as b – (a – b) = 2b – a .
The small 2b – a , a – b , sided triangle is similar to the original, because one of its angle was already 45 and an other is the mirroring of 90. So indeed:

(a , b , 45 , 90 , 45) → (a' = 2b – a , b' = a – b , 45 , 90 , 45)

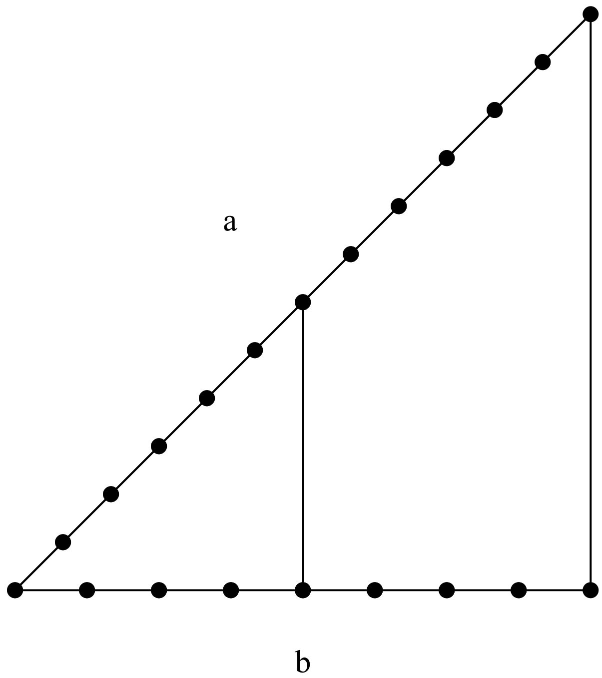
It's also obvious that a' = 2b – a < a .
Thus, our first method, that is similarity and ± combination from the old sides, proves that the original triangle can't have common units. a = m u , b = n u is impossible.
Just to repeat the argument: If a = m u , b = n u were, then a' = m' u , b' = n' u were too.
Then, (a'' , b'') could be created similarly and so on, giving infinite many smaller and smaller versions. But with a starting a = m u we can only have finite many smaller multiples.

The second “common f factor” method applies too, because:



$a^2 = 2\,b^2$

So $a = m\,u$, $b = n\,u$ implies that m , n are both even. Thus, the ($a = m\,u$, $b = n\,u$, 45 , 90 , 45) triangle can be halved exactly on units:



But this half triangle is similar too, so with the repeated argument, it must have even many units again. So we can half it again, and so on, getting decreasing infinite many u multiples. This is impossible, thus, our original assumption of $a = m\,u$ and $b = n\,u$ is false.
To see the 2 common factor, that is the evenness of m and n directly:

$$a^2 = 2\,b^2 \rightarrow (m\,u)^2 = 2\,(n\,u)^2 \rightarrow m^2 = 2\,n^2 \rightarrow m^2 \text{ is even} \rightarrow m \text{ is even}$$

$$m = 2\,r \rightarrow (2\,r)^2 = 4\,r^2 = 2\,n^2 \rightarrow n^2 = 2\,r^2 \rightarrow n^2 \text{ is even} \rightarrow n \text{ is even}$$

T

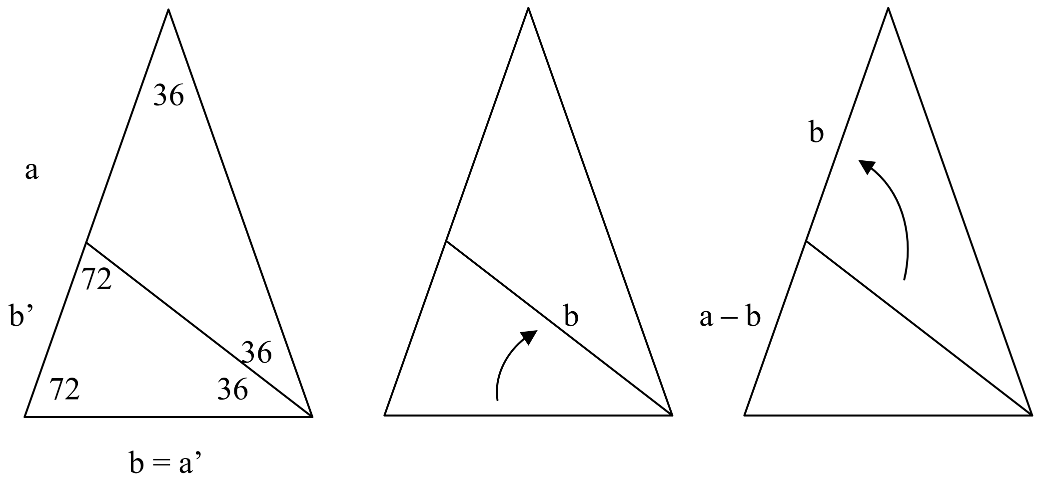
The “± reduction” method can be applied to:

1.) (a , b , 72 , 72 , 36) with $a' = b$, $b' = a - b$

2.) (a , b , 90 , 67.5 , 22.5) with $a' = b$, $b' = a - 2\,b$

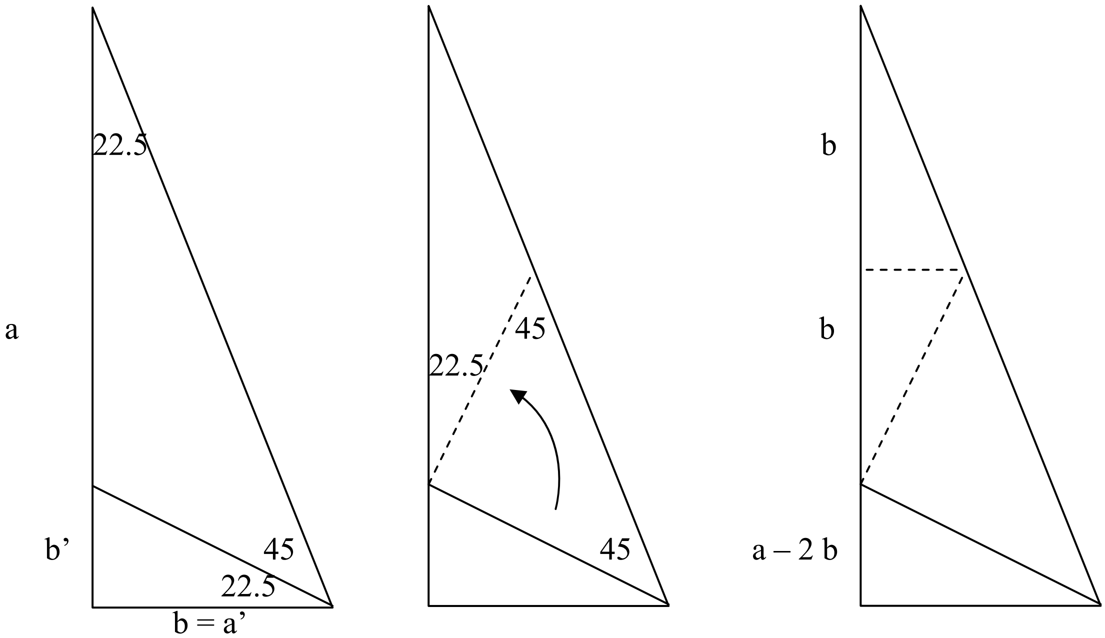
P

1.)



$$\frac{a}{b} = \frac{b}{a-b} = \frac{a'}{b'}$$

2.)



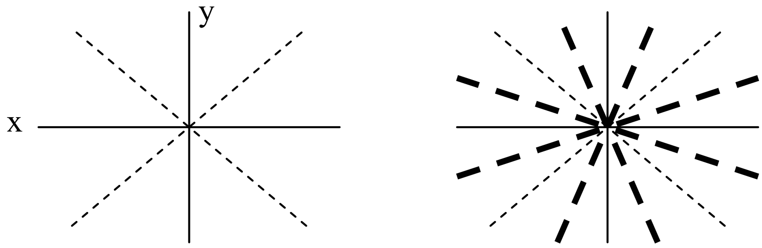
$$\frac{a}{b} = \frac{b}{a-2b} = \frac{a'}{b'}$$

R

The non existence of common units for a , b also means the non existence of common multiples. Indeed, $a = m\,u$, $b = n\,u$ → $u = \dfrac{a}{m} = \dfrac{b}{n}$ → $n\,a = m\,b$.

With $\gamma = 90°$, this can be represented as the existence of lines through the origin of a Descartes coordinate system that never go through a grid point again.

Indeed, the $\beta = 22.5°$ or $67.5°$ angle lines from the origin must be such:



R

Same $\frac{a}{b}$ ratio can come out from different γ , α , β angles.
In fact, we can use other ratios to specify the angles for an a , b.
This is exactly how the old trick works to construct 72° and 36°.
We derived the $\frac{a}{b} = \frac{b}{a-2b}$ from the special relation of 72° and 36°.
But the same $\frac{a}{b} = \frac{b}{a-2b}$ can be obtained by creating new angles that we are not even interested in. Rather, it will help to create the (a , b , 72 , 72 , 36) triangle.

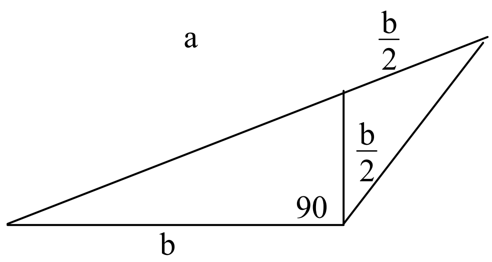
T

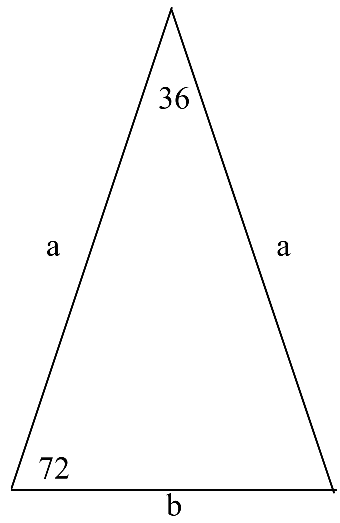
Construction of 72° and 36°:

Start with arbitrary b distance. Erect 90 and measure $\frac{b}{2}$ onto it.

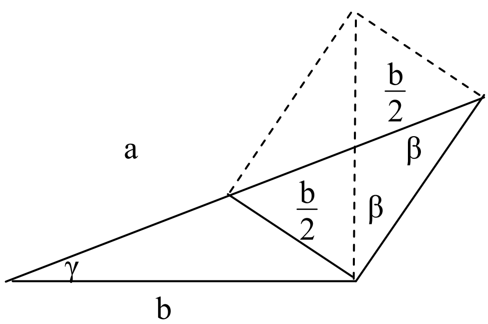
Connect it with the other end of b and continue it with $\frac{b}{2}$ in the opposite direction.

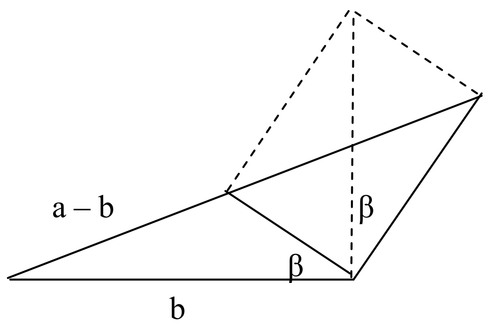
This gives our a . Using this on b for both sides gives the 72 , 72 , 36 triangle





P





Thus, $\frac{a}{b} = \frac{b}{a-b}$ just as it was in the $\gamma = 72^\circ$, $\beta = 36^\circ$ case.

R

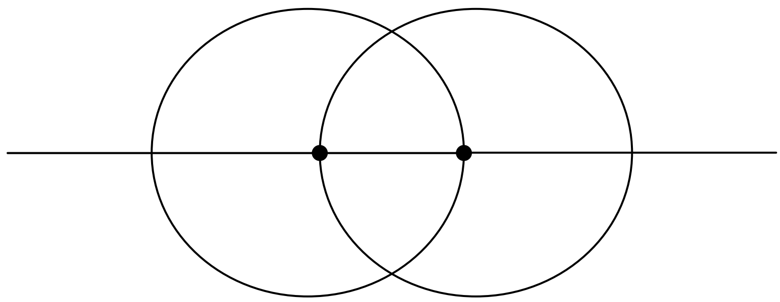
The obtained construction of 72° is also a method of the construction of a pentagon, because:
$\frac{360}{5} = 72.$

7. Constructibility

D

Constructible points and sets

Let two points be fix in a plane as the given unit. All other constructible points will be obtained as crossings of constructible lines and circles.
The constructible lines are connectors of constructed points.
The distance between two constructed points is also regarded as constructible.
The constructible circles are ones with centers as constructed points and radius as constructed distance.
Finally, the angles between constructed lines are also regarded as constructible.
The first constructible set is the line of the unit.
The second constructible sets are the circles with centers of the unit ends and radius as the unit:

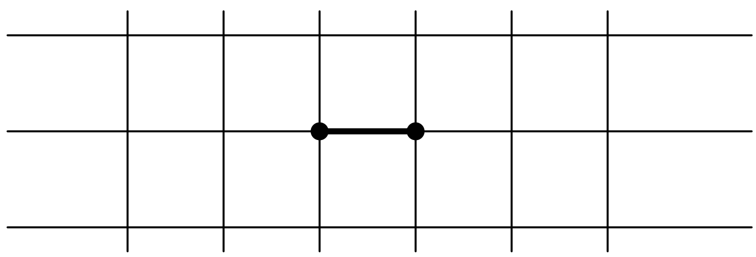


Thus, we obtain four new points, namely two on the unit line and two crossings of the circles.
These two crossing points of the circles can be connected with each other and also with all four points on the unit line. So, nine new lines can be obtained.
These ten lines will have a lot of crossing with each other and the two circles.
Then new lines and circles can be constructed. And so on.
There is no fix sequence to obtain all constructible sets. Above we merely started in a seemingly logical way. The obtainable sets are of course still well determined.
Indeed, a set is constructible if there is a particular finite sequence that obtains it.

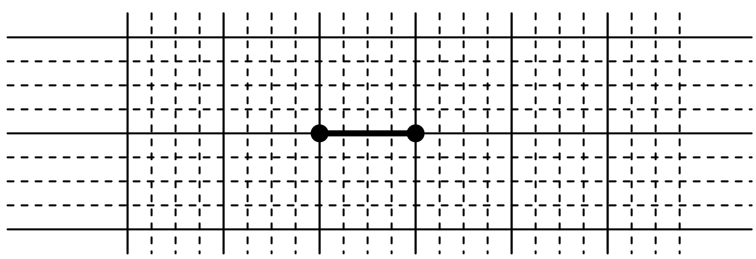
T

1.) 60° , 30° , 90°, are constructible.

2.) The infinite unit grid is constructible:



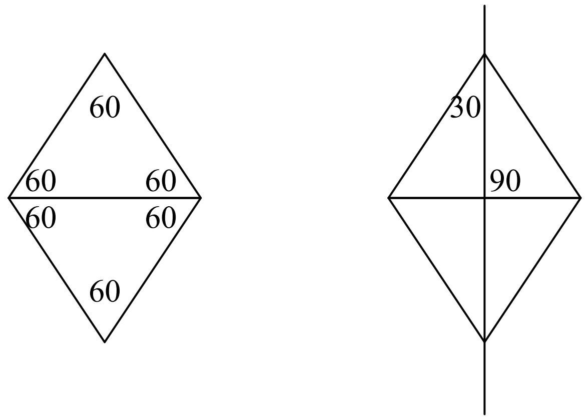
3.) The infinite fractional grid is constructible:



4.) Connecting any two points on the fractional grid, they cross in a fractional grid already there.

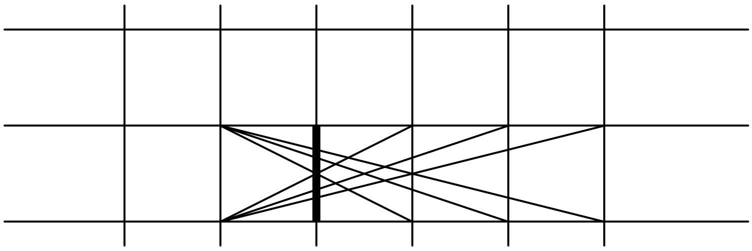
P

1.)

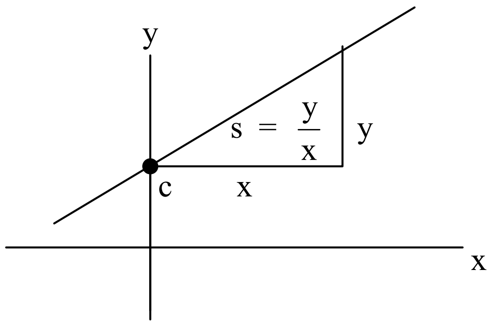


2.) We can repeat the unit and use 90°.

3.) Crossing already the points of two grid line will cut all fractions.



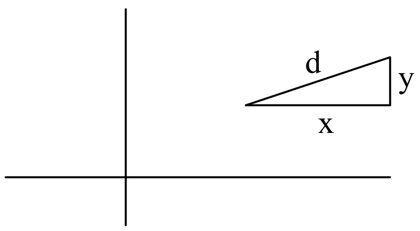
4.) Two grid lines can be regarded as x , y coordinates and then any fractional s sloped and c crossing of y has the equation: $y = s\,x + c$



The crossing of two such lines is a linear equation system and is easy to see that it leads to fractional solutions.

R

This "criss cross completeness" of the fractional grids is interesting, but seems as a detour from constructability. The real reason the whole coordinate view is going to be useful is the Pythagoras Theorem. Usually it is expressed as $a^2 + b^2 = c^2$, but actually it also means that any d diagonal distance in a coordinate system can be calculated from the x , y coordinate distances:



$$d = \sqrt{x^2 + y^2}$$

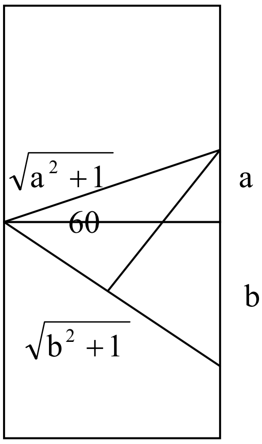
Before we even go to our main object, which is proving that 20° is not constructible, I prepare that with an easier, but just as big surprise:

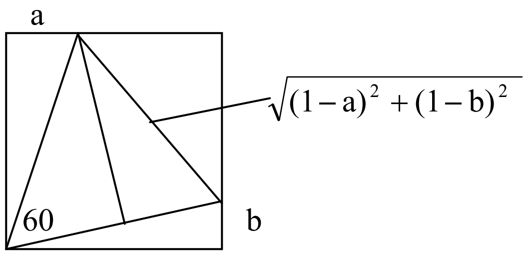
T

There are no fractional lines that would be in 60^o.

P

Such two lines could be in two possible situations relative to the unit grids measured from the crossing of the lines:





Both lead to the same impossibility but we only follow the first:

$$(a+b)^2 = \cancel{a^2} + \cancel{b^2} + 2ab = \left(\sqrt{a^2+1}\,\frac{\sqrt{3}}{2}\right)^2 + \left(\sqrt{b^2+1} - \frac{\sqrt{a^2+1}}{2}\right)^2 =$$

$$\frac{3}{4}(a^2+1) + b^2 + 1 + \frac{1}{4}(a^2+1) - \sqrt{a^2+1}\,\sqrt{b^2+1} =$$

$$\cancel{a^2} + \cancel{b^2} + 2 - \sqrt{a^2+1}\,\sqrt{b^2+1}$$

$$\sqrt{a^2+1}\,\sqrt{b^2+1}\,\sqrt{3} = 2 \bullet \text{area} = a+b$$

So $\sqrt{3} = \dfrac{a+b}{\sqrt{a^2+1}\,\sqrt{b^2+1}} = \dfrac{a+b}{2 - 2ab}$ would be rational.

R

The fractional grids could be constructed with a limited use of the compass, as only a measurer on a line, plus a 90^o ruler. But this would lead to non fractional distances too. For example, the diagonal of a unit grid is $\sqrt{2}$ and it could be measured onto other lines.
I think 60^o could still not be obtained this way. If anybody knows how to prove this, I'd like to hear about it.
Now we turn to our goal to prove that 20^o is not constructible at all. This of course means at once that we can not trisect angles in general, because a 60^o is constructible.
The amazing thing about this proof of the inconstructibility of 20^o is that it first translates the whole problem into algebra:

T

Algebraization Theorems

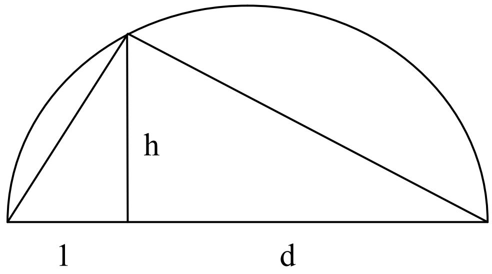
1.) Distance constructability condition.
A distance is constructible if and only if it is expressible by $+, -, \bullet, \div, \sqrt{\ }$ from 1.
In short, if it is square root expressible.

2.) 20^o constructability condition.
20^o is constructible if and only if an x distance is constructible for which $x^3 = 3x + 1$.

3.) 20^o non constructability condition.
If $x^3 = 3x + 1$ has no square root expressible solution then 20^o is not constructible.

P

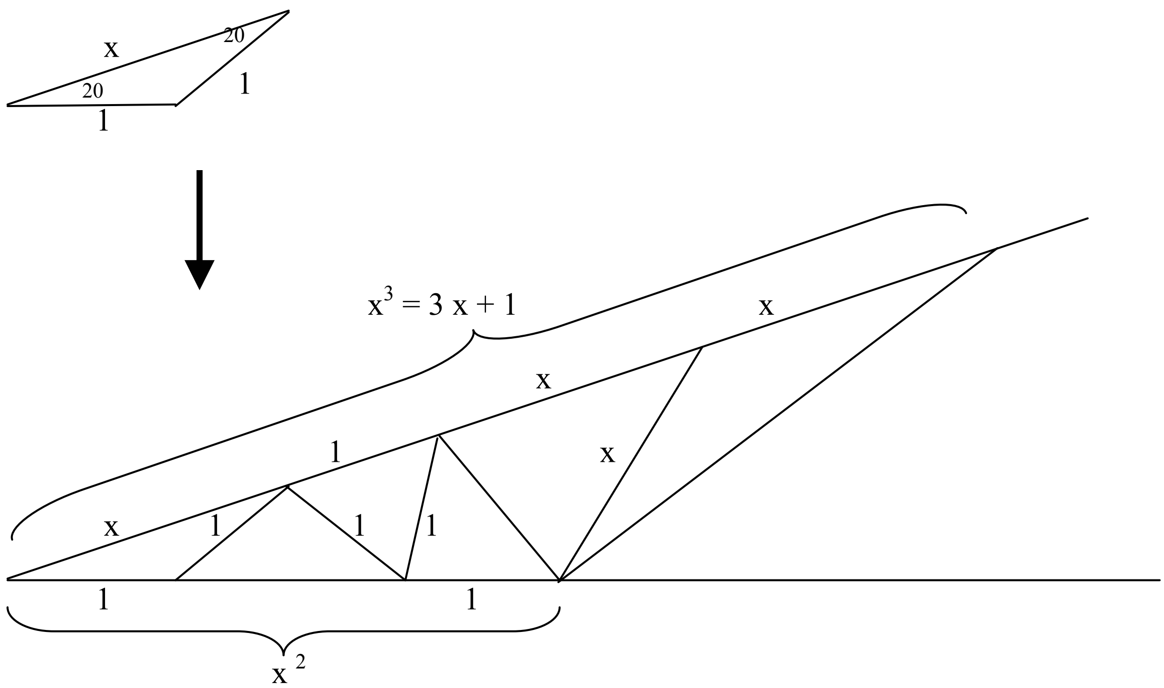
1.) First of all, it’s easy to construct the square root of a d distance as follows:



$$\frac{h}{1} = h = \frac{d}{h} \rightarrow h^2 = d \rightarrow h = \sqrt{d}$$

The reverse is that any constructible distance is square root expressible from the distances that are used in the construction. Indeed, then the used ones are again square root expressible from what they used, and so on, everything is from 1.
It would be better to go backwards through the directly used distances, but this is a bit ambiguous when we use earlier crossing points. So here, the coordinate representation helps again. Indeed, this way we can talk about square root expressible points too. Namely meaning the coordinates. In fact, square root expressible lines are connections of such points and square root expressible circles are ones using such center and radius.
Then all we have to show is that crossing of lines and circles keeps square root expressibility. This is not too difficult.

2.)



3.) Trivial by 1.) and 2.).

D

1.) A set of real numbers is operation complete if the + , – , • , ÷ of members is member too. At ÷ the divider is assumed to be non zero.

2.) An S set of real numbers is a square root base for an x real number if S is operation complete and there are some u , v , w in it, so that $x = u + v\sqrt{w}$.

3.) $E_0, E_1, \ldots, E_N$ is a square root extension sequence if:

a.) E_0 = the fractions

b.) There are $w_0, w_1, \ldots, w_N$ in each $E_0, E_1, \ldots, E_N$, so that:

c.) $E_n = \{x \,;\, x = u + v\sqrt{w_{n-1}}$ so that $u, v \in E_{n-1}\}$

T

1.) Every square root expressible number is element in a square root extension sequence.

2.) In a square root extension sequence $E_0 \subset E_1 \subset \ldots \subset E_N$ and E_{n-1} is a square root base for any $x \in E_n$.

3.) If $x^3 + a\,x^2 + b\,x + c$ has an x_1 root and an S square root base for x_1 contains a, b, c, then S contains a root too.

4.) If $a, b, c \in E_0$ but $x^3 + a\,x^2 + b\,x + c$ has no root in E_0 then it has no root in any square root extension sequence.
Thus, it has no square root expressible root either.

5.) $x^3 - 3\,x - 1$ has no root in E_0.
Thus, it has no square root expressible root.
20^o is not constructible.

P

1.) Start from any square root of fractions and then widen outwards.

Example: For $\sqrt{\sqrt{\frac{2}{3}} + \sqrt{5}}$:

$\sqrt{\frac{2}{3}} \in E_1$, $\sqrt{\frac{2}{3}} + \sqrt{5} \in E_2$, $\sqrt{\sqrt{\frac{2}{3}} + \sqrt{5}} \in E_3$

2.) The widening is obvious because $0 \in$ every E_n and we can use $u + v\sqrt{w}$ with $v = 0$.
For the square root base "extension" we only have to show that every E_n is operation complete. $+, -, \bullet$ are obvious and for $\div$ observe:

$$\frac{u + v\sqrt{w}}{p + q\sqrt{w}} = \frac{(u + v\sqrt{w})(p - q\sqrt{w})}{(p + q\sqrt{w})(p - q\sqrt{w})} = \frac{u\,p - v\,q\,w + (v\,p - u\,q)\sqrt{w}}{p^2 - q^2\,w} = r + s\sqrt{w}$$

3.) Suppose $x_1 = u + v\sqrt{w}$ with $u, v, w \in S$.
$(u + v\sqrt{w})^3 + a\,(u + v\sqrt{w})^2 + b\,(u + v\sqrt{w}) + c = 0$ that is:
$u^3 + 3\,u^2 v\sqrt{w} + 3\,u\,v^2 w + v^3\,w\sqrt{w} + a\,u^2 + 2\,a\,u\,v\sqrt{w} + a\,v^2\,w +$
$b\,u + b\,v\sqrt{w} + c = 0$ re-arranged as:
$(u^3 + 3\,u\,v^2 w + a\,u^2 + a\,v^2\,w + b\,u + c) + [3\,u^2 v + v^3\,w + 2\,a\,u\,v + b\,v]\sqrt{w} = 0$
If $[\quad] \neq 0$ then $\sqrt{w} = -\frac{(\quad)}{[\quad]}$, so $u + v\sqrt{w} \in S$.
If $[\quad] = 0$ then $(\quad) = 0$ too and putting $x_2 - v\sqrt{w}$ into $x^3 + a\,x^2 + b\,x + c$
we get $(u - v\sqrt{w})^3 + a\,(u - v\sqrt{w})^2 + b\,(u - v\sqrt{w}) + c =$
$(\quad) - [\quad]\sqrt{w} = 0 - 0\sqrt{w} = 0$. So x_2 is root too.
Since $x_1 \neq x_2$, thus $-a - x_1 - x_2$ is root too. (See Lemma)
But $-a - x_1 - x_2 = -a - (u + v\sqrt{w}) - (u - v\sqrt{w}) = -a - 2\,u \in S$.

Lemma:
If $x_1 \neq x_2$ are roots of $x^3 + a\,x^2 + b\,x + c$ then $-a - x_1 - x_2$ is root too.
Proof:

1.) For any $x^3 + a\,x^2 + b\,x + c$ and x_1 there are p, q, r that
$x^3 + a\,x^2 + b\,x + c = (x - x_1)\,(x^2 + p\,x + q) + r$

2.) If x_1 is root of $x^3 + a\,x^2 + b\,x + c$ then $r = 0$.

3.) If $x_1 \neq x_2$ are roots of $x^3 + a\,x^2 + b\,x + c$ then x_2 is root of $x^2 + p\,x + q$ and there is x_3 so that $x^2 + p\,x + q = (x - x_2)(x - x_3)$.
So $x^3 + a\,x^2 + b\,x + c = (x - x_1)\,(x - x_2)\,(x - x_3) =$
$x^3 - (x_1 + x_2 + x_3)\,x^2 + (x_1\,x_2 + x_1\,x_3 + x_2\,x_3)\,x - x_1\,x_2\,x_3$
So, $x_3 = -a - x_1 - x_2$ is root too.

4.) By the widening $E_0 \subset E_1, E_2, \ldots, E_N$. Thus, by 3.)
a root in E_n implies root E_{n-1}, then in E_{n-2} and so on, finally in E_0 too.
But this was assumed to be false.

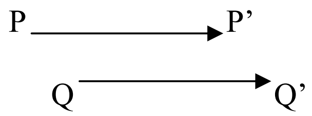
5.) Suppose a simple $\frac{a}{b}$ fraction were root: $(\frac{a}{b})^3 - 3\,\frac{a}{b} - 1 = 0$.
Then, $a^3 - 3\,a\,b^2 - b^3 = 0$.
Every prime factor of a is also of $a^3 - 3\,a\,b^2$, so of b^3 and b too.
Every prime factor of b is also of $-3\,a\,b^2 - b^3$, so of a^3 and a too.
Thus, a, b can't have prime factors, so they are 1. But $a = b = 1$ is not a solution.

8. Isometries Of Space

D PQ denotes the interval between P and Q.
PQ = RT denotes that the two intervals have the same length.
PQ || RT denotes that the two intervals are parallel.
PQ $\vee$ RT denotes that the two intervals are not parallel.

D
1.) A transformation is any P' assigned points to all P.
2.) S' = {P' ; P $\in$ S}.
3.) P is fixpoint is P' = P.
4.) S is a fix set if S' = S.
5.) S is conserved set if S' $\subseteq$ S.
6.) P* = middle point of PP'.
7.) $P_\perp$ = middle perpendicular plane of PP' = perpendicular to PP' through P*.
8.) The identity is the P' = P transformation.

D A P' non identity is:
1.) Isometry (iso = equal, metry = length) if P'Q' = PQ for all P , Q.
2.) Shift if all PP' , QQ' are:
   a.) parallel
   b.) same length
   c.) same directional



3.) Mirroring to a Π plane, if for all P, $\Pi = P_\perp$.
4.) Turn with an α angle around an L line with one of its directions chosen,
   if P_L denotes the perpendicular projection of P to L,
   then $P'P_L$ is the α turn of PP_L in the plane perpendicular to L,
   looking from the chosen direction.
   Of course, looking from the other direction, the angle would be $-\alpha = 360 - \alpha$.

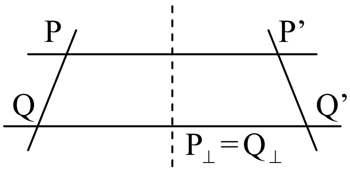
T Any P' transformation must obey exactly one of the followings:
1.) All PP' are parallel and all P'Q' are parallel to PQ.
2.) All PP' are parallel, but not all P'Q' are parallel to PQ.
3.) Not all PP' are parallel, but all $P_\perp$ go through a fix L line.
4.) Not all PP' are parallel, and not all $P_\perp$ go through a fix L line.

P The four exclude each other and one must be true.

T The four cases of the previous theorem for an isometry are:
1.) Shift
2.) Mirroring
3.) Turn
4.) Turn then mirroring or turn then turn.

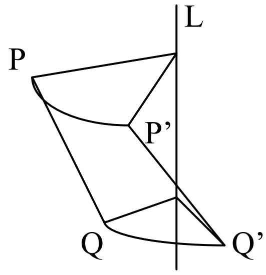
P
1.) All PP' are parallel by definition and P'Q' = PQ → PP' = QQ'.

2.) Let P'Q' be one that is not parallel with PQ. Then, P'Q' = PQ → $P_\perp = Q_\perp$.



For any third R point, R'P' $\vee$ RP or R'Q' $\vee$ RQ must be true, so $R_\perp = P_\perp$ or $Q_\perp$

3.) If PP' $\vee$ QQ', then $P_\perp \vee Q_\perp$, so they cross in an L line.
We claim that P and P' and Q and Q' are not only equal distanced from this L line, but are turned with the same angle. That is, if P_L and Q_L denote their drop to L, then $P P_L P' \angle = Q Q_L Q' \angle$:



If all $R_\perp$ go through L, then of course, this same turn works for all R to get R'.

4.) Let PP' $\vee$ QQ' and turn S around the L line, that is the crossing of $P_\perp$ and $Q_\perp$ with the α angle determined by P' and Q'. If this turned set is S_α, then it is isometric to S' and this S_α → S' isometry has at least two fix points, namely P and Q.
If all $R_\alpha R'$ are parallel, then by 1.) and 2.), we have a shift or mirroring, but since we have fix point, we can't have a shift, so we have a mirroring. If not all $R_\alpha R'$ are parallel, then by 3.), we have a turn around the PQ line, because these are fix, so all $R_\perp$ go through them.

D Special isometries are:
The three basic: shift, mirroring, turn, and the following three combinations:
1.) A turned mirroring is a turn around an L and a mirroring to a plane perpendicular to L.
2.) A shifted mirroring is a shift and a mirroring to a plane parallel to the shift.
3.) A screw around L is a turn around L and a shift parallel to L.

T In the special combinations, using "and" instead of "then" was justified because their order was immaterial. For example, the screw is a turn then shift or shift then turn.

T All special isometries have conserved line.

P Trivial one by one.

T If an L line is conserved, then the isometry is special, namely:
1.) L is either fix or mirrored to a Π plane or shifted in itself.
2.) If L is fix then, P' is either a mirroring to a Π, containing L or a turn around L.
3.) If L is mirrored to Π, then P' is either the same mirroring in the whole space or a turned mirroring to Π around L.
4.) If L is shifted, then P' is either the same shift in the whole space or the same shifted mirroring to a Π containing L, or a screw with the same shift.

P

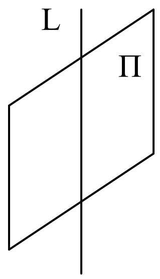
1.) If L has two fix points, then the whole L is fix.
   If L has only one fix point O, then any other P must be mirrored to O to have PO = P'O' = P'O.
   If L has no fix points, then for any P , Q on L we have PP' = QQ'.
2.) If all RR' are parallel, then since we have fix points, we can't have a shift, so it must be a mirroring to a Π and it must contain all fix points, including L.
   If not all RR' are parallel, then we must have a turn around L because all $R_\perp$ must contain all fix points, including L.
3.) Π is conserved too and either it is fix or for all R point of it, $R_\perp$ contains L, so Π is turned in itself. If Π is fix, then in the space we have the same mirroring as in L, or if Π is turned, we have a turned mirroring.
4.) If R_L is the drop of R to L, then the RR_L distances are preserved. In other words, looking perpendicularly to L, we have an Θ plane in which P' is still an isometry.
   But here, L is a fix point, so we have the following possibilities:
   a.) The whole Θ plane is fix, and thus, P' is the same shift in the space as in L.
   b.) There is only a line fix in Θ, through L, which in space is a Π plane. Then, in Θ, we have a mirroring to Π and thus, a shifted mirroring to it in space.
   c.) Only L is fix in Θ and then we have a turn in Θ and thus, a screw in space.

T

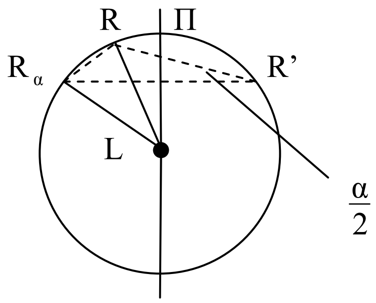
"Turn then mirroring" replaced by special isometries:
1.) A turn then mirroring to a plane going through the turn line is a single mirroring.
2.) A turn then mirroring to a plane parallel to the turn line is a shifted mirroring.
3.) A turn the mirroring to a plane crossing the turn line is a turned mirroring.

P

1.)

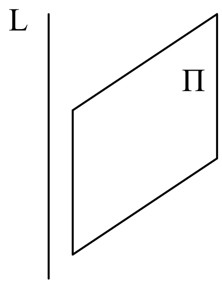


from top:

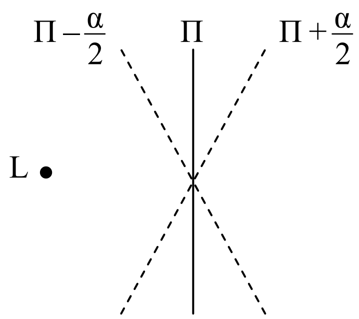


Thus, all RR' will be parallel and L is fix, so we have a mirroring.

2.)



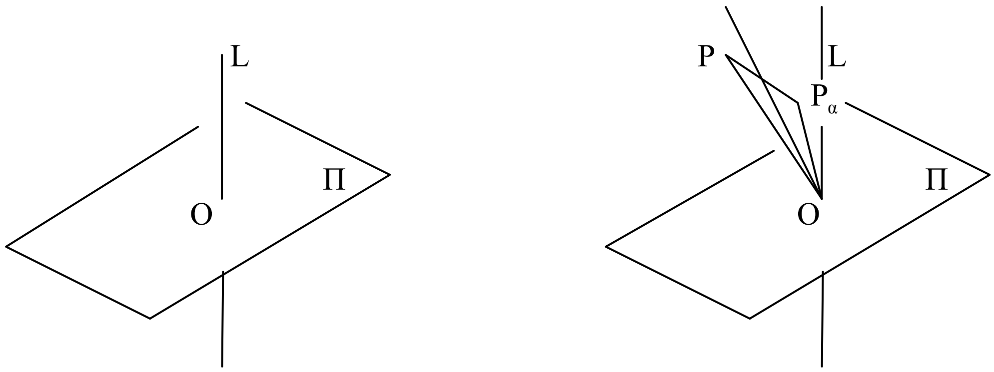
from top:



We turned Π with $\pm\frac{\alpha}{2}$. These are mirror of each other to Π and also, the α turn of $\Pi-\frac{\alpha}{2}$ is $\Pi+\frac{\alpha}{2}$. Thus, $\Pi-\frac{\alpha}{2}$ is conserved by the turn then mirroring.
If a line is conserved, it either has a fix point and is a mirroring or hasn't and is a shift.
Here $\Pi-\frac{\alpha}{2}$ can't have fix points, so it's shifted in itself, and thus the whole plane is a shifted mirroring to this same line.

Then, layer by layer, the whole space is a shifted mirroring to $\Pi - \frac{\alpha}{2}$.

3.)



By 3.) of previous theorem, enough to show that there is an L_0 line, that is mirrored in itself. The P point of the above picture defines such L_0, if P and P_α are symmetrical to the perpendicular to Π. Indeed, P_α mirrored to Π will fall on the continuation of the $PO = L_0$ line.

T "Turn then turn" replaced by special isometries:

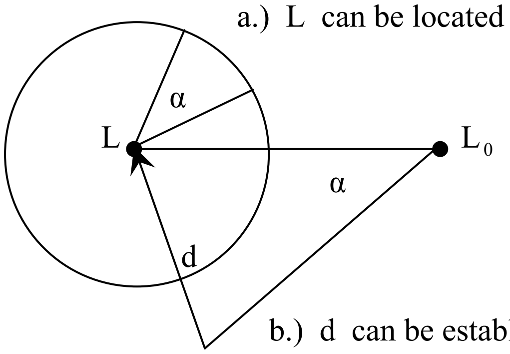
1.) a.) Every α turn around an L_0 line followed by a d shift perpendicular to L_0, is just an α turn around an L parallel to L_0.

b.) For every α turn around an L line, and any L_0 line parallel to L, the turn can be replaced by same α turn around L_0 followed by a shift perpendicular to L_0.

2.) A turn followed by a turn around an axis that is parallel to or crosses the first, can be replaced by a single turn.

3.) A turn followed by a turn around an axis, that is not in the same plane as the first, can be replaced by a screw.

P 1.) a.) L can be located as the point in the figure and it is fix:



b.) d can be established from the figure and L becomes fix.

2.) There are two lines so that the first turn turns the first line in the second, while the second turn turns the second line into the first. Thus, the first line will be a fix line.

3.) Let L_1 be the axis of the first and L_2 of the second! Let L_0 be the parallel with L_2 that crosses L_1. Then by 1.) b.), the turn around L_2 can be replaced one around L_0 followed by a shift perpendicular to them. Then by 2.), the turns around L and L_0 can be replaced by one around an L.
The followed shift can be decomposed into a perpendicular and parallel component to L. The perpendicular shift melts into giving a new turn by 1.) a.).

T 1.) Every isometry is a special one.

2.) Every isometry has a conserved line.

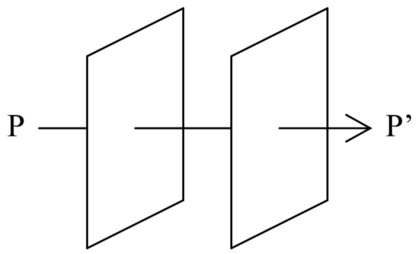
T

Every isometry can be replaced by a sequence of mirrorings, namely:

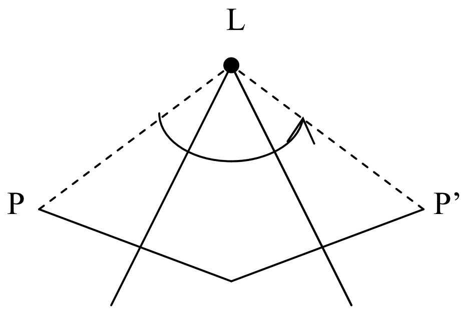
1.) A shift can be replaced by two mirrorings, both with perpendicular plane to the shift.
2.) A turn around L can be replaced by two mirrorings with planes crossing in L.
3.) Every special isometry is:
    a.) one mirroring or
    b.) two mirrorings or
    c.) three mirrorings, where the last is perpendicular to the first two or
    d.) four mirrorings, where the second two are perpendicular to the first two.

P

1.)



2.)



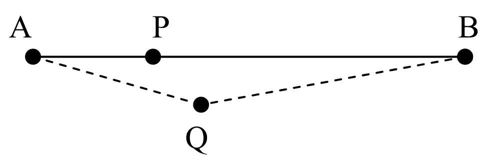
3.) a.) A mirroring itself is basic.
    b.) A shift is two parallel mirrorings by 1.).
    A turn is two mirrorings with crossing planes by 2.)
    c.) A shifted mirroring is two parallel mirrorings followed by a perpendicular one.
    A turned mirroring is two crossing mirrorings followed by a perpendicular one.
    d.) A screw is a turn then shift or shift then turn and thus is:
    two crossing mirrorings followed by two perpendicular ones that are parallel or
    two parallel mirrorings followed by two perpendicular ones that are crossing.

9. Parallelity

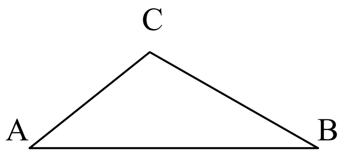
R

The most fundamental two concepts of geometry are the points and distances.

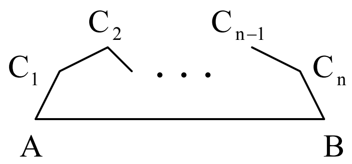
In the modern view, the points are simply the elements of the whole space and lines , circles, planes and other geometrical objects are set of points in the space. Or to put it an other way, these are special subsets of the space. The connecting interval between two points is also a set of points, in fact it is the simplest special set by which all the others will be defined.
Sometimes we simply call this connecting interval as a distance, but it is not quite correct, because the proper meaning of the distance is its length. To measure a length, we need a unit length, so we first need a comparing of distances. The simplest relationship is the equality of two distances. Or to say it properly, two connecting intervals having the same length. This is also a basic concept.
So as we see, actually we have three basic concepts: Points, connecting intervals and equality of such connecting intervals. A, B, C, D, E, F, G, H and M, N, O, P, Q, R letters will be used for points, while I, J, K, L, S, T, U, . . . for sets of points. For the connecting interval, we simply put the two points next to each other. So for example, AB is the set of points on the connecting interval of A and B.
The intended meaning of AB is the shortest connection from A to B. This also means that if we pick a P and Q point from AB, then not only P and Q, but all the shortest connecting points from P to Q must be in AB. Indeed, otherwise we could shorten AB.
So in exact form: $P , Q \in AB \rightarrow PQ \subseteq AB$.
This is the most fundamental axiom of connecting intervals. But it doesn't guarantee yet, the full meaning of intervals as shortest connections. It didn't even use the concept of lengths.
The equality of lengths will be simply denoted as AB = CD meaning that the AB and CD connecting intervals are equal long.
The physical meaning of this equality is not as simple as it seems. If we take a measuring tape or a rigid object that is equal to AB, then we have to move it to place it over CD. Before Relativity, nobody would have worried about this, but now we all heard of the change of length caused by motion. Of course, that effect of Special Relativity only causes problem, while the object is in motion, so moving and then slowing down should be okay. On the other hand, General Relativity claims that gravitation changes the length too, so that causes a bigger problem. Anyway, we can ignore these problems and rely on our intuitions.
The connecting intervals and their equality of length offers an obvious road to define smaller and bigger lengths too. Indeed, all we need is an axiom that claims that for any two intervals, exactly one of them will be equal to a beginning interval of the other, and then this can be defined as the smaller. Amazingly, we could have gone the opposite way too, that is start with a smaller, bigger comparison of lengths as basic concept and define the sets of connecting intervals through this. Indeed, a P point is on the AB interval, if and only if AP and PB are minimal. In other words, for any Q point that is not on the interval, AQ or QB (or maybe both) must be bigger than AP or PB:



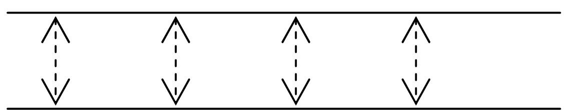
With this approach, AB = CD or AB < CD would only mean the comparison of lengths assigned to pairs of points, not to their connecting interval and the actual interval would only become meaningful later. I think that our above approach is more natural.
The idea of measuring one interval onto the other to define the smaller and bigger length, can be modified to define addition and subtraction of lengths too. Then, AB < CD can be replaced by AB = CD – PQ or CD = AB + PQ.
The most plausible claim about lengths is that for any three A, B, C points, AC + CB > AB.



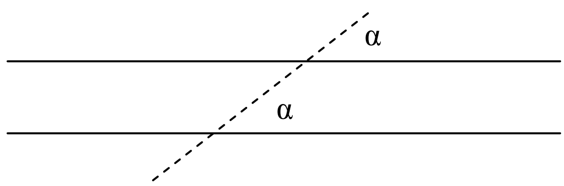
This so called triangle inequality expresses that from “A to B”, the connecting interval is shorter than going through an “unnecessary” C.
By repeated application this also implies that AB is shorter than any
$AC_1 + C_1C_2 + \ldots + C_{n-1}C_n + C_nB$ broken connection:



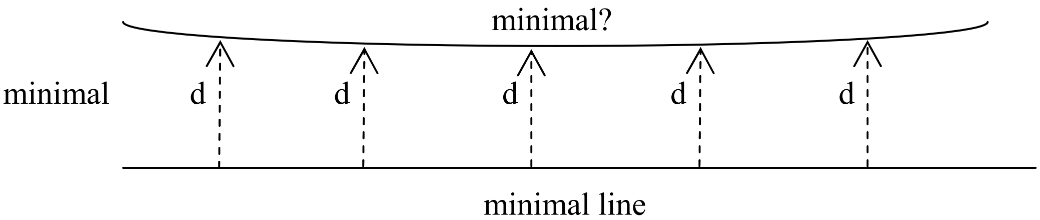
The triangle inequality doesn’t follow just from the comparability of lengths. It requires some axioms. Once however it is proved, we have the original meaning of connecting intervals as shortest connections indeed established.
Lines are usually regarded as basic concepts but with our approach they can be defined as simply the combined sets of wider and wider intervals. So $< AB >$ can denote the combined set of all intervals that contain AB. Of course, it is still far from obvious that if P and Q are points from $< AB >$, then this same line would be obtained starting from PQ.
In short: $P, Q \in < AB > \rightarrow < AB > = < PQ >$. To prove this, we need again new axioms.
Those who regard lines as basic concepts only seemingly avoid this complication because for any two points they have to order a line. And thus, for one line the different pairs of points must determine the same. In addition at the end, they still have to regard lines as sets of points.
Above, in $< AB > = < PQ >$ we used the equal sign for the equality of two sets meaning that they have the same points. This is a little bit contradictory to our previous usage of AB = CD, which only meant the equal length. But, otherwise we would need two equal signs. One for sets and one for lengths. The actual equality of the AB and CD intervals is practically never needed, so we can use the other meaning instead of this. So, in short, the = sign means equality of sets except for intervals where it means the equal length.
By the fact, that any two points of a line determine it, it’s easy to show that two lines can only have one common point (or none). Indeed, if they had two A and B, then these would determine the same line. Having a common point is also called as crossing for lines.
After the lines, usually the planes are introduced, but if we restrict our attention to a single plane, then the real problems are easier to see, so in the followings we only deal with this.
Just above we mentioned that two lines can have no common point at all, which in space is obvious because two randomly chosen lines simply avoid each other. If we are in one fix plane, then it is still possible that two lines don’t cross each other, indeed this is what we see as parallelity. There are two alternative approaches to parallelity. The first is to claim that the parallel lines keep the same distance from each other.



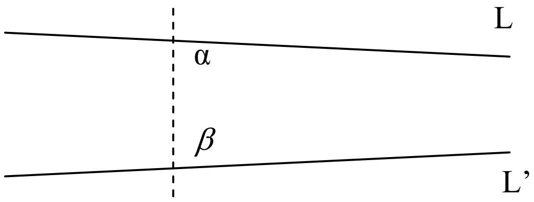
The other is to say that they are going in the same direction. If we ask more about what we should mean by same direction, then we soon realize that it can only be specified relative to a third crossing line. Indeed, the two parallels must have the same angle to the third one.



Everyone would guess that the first approach, that is the kept distance is easier to define precisely than the second, because that requires the concept of angles as opposed to distances.
As it will turn out, the equidirectionality is simpler than fix distance. But already now we can see an advantage of the equidirectionality as follows:
By drawing two equal angled lines to a third, at least we know that we are dealing with two lines. On the other hand, if we start with one line, then measure up the same lengthed intervals from every point, then it is not so obvious that the obtained points form a line at all. Indeed, intervals and lines are minimal connecting "curves", so a horizontal minimality moved minimally vertically could very well be non minimal anymore.



It became obvious already for Euclid that parallelity is the crucial problem of geometry.
He firmly believed that the three concepts, non crossing, having fix distance or having same direction are identical, but when he tried to prove their identity, he went into strange logical circles. Finally, he chose as axiom the assumption that if two lines are both having less than a right angle to a third towards each other, then these two lines must cross on this side of the connecting line:

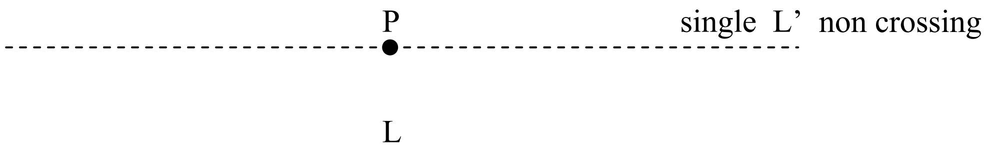


$\alpha\ ,\ \beta\ <\ 90^{\circ}\ \rightarrow$ L , L' cross

Later, it turned out, that a much simpler assumption could still prove the identity of the three form of parallelity, namely John Playfair and Legendre stated about the same time:

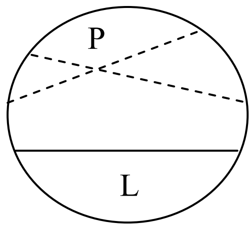
Axiom Of Parallelity: In a fix plane:

For an L line, and an outside P point, there is only one non crossing L' line through P:



In spite of its beauty it was still regarded as too complicated and so, up until the 19-th century, many mathematicians wanted to avoid this axiom by somehow deriving it from the simpler other axioms. Some of them attempted to find such derivation by the following way: They assumed that the Axiom Of Parallelity is false, that is assumed more non crossing L' lines and then tried to reach a contradiction from this. This in effect would indeed make a proof for the unique non crossing line. To their biggest surprise instead of a contradiction weird, but beautiful possibilities followed. The first person who went into this jungle and became convinced that it must be a meaningful reality was Janos Bolyai, the son of an old friend of Gauss. When the father showed his son's results to Gauss, he replied by saying that he can't praise the young man, otherwise he would have to praise himself, because he already realized all that long ago. This was of course, a terrible blow to Janos. Gauss indeed, explored the

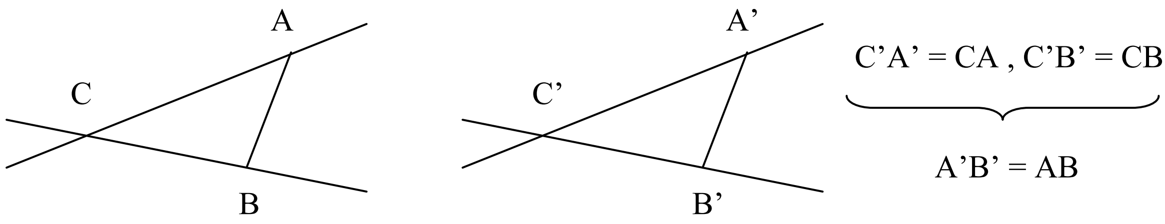
possibilities of many non crossing lines, but left his notes in his drawers because he thought the world is not ready yet. Later, when he learnt that a third person, Lobachevsky, also realized that something lies behind the non euclidian parallels, Gauss finally openly praised Lobachevsky, but still didn't tell him about the young Bolyai.
These three people, Gauss, Bolyai and Lobachevsky still remained within classical mathematics. To see why, we have to jump to just a few decades forward, when Beltrami finally made the crucial step not only in the history of mathematics, but probably in the whole history of human intellect. He realized that the points of a euclidian plane itself can be used to show that a non euclidian plane can exist.
If we shrink a plane towards a central point so that the whole plane fits within a disc, then obviously the lines will bend into curves. This itself is an interesting idea because it shows that the same truths can be kept by replacing the lines with other special curves. But Beltrami went further and realized that properly changing the lengths, we can even change some truths while keeping others. Amazingly with this proper change, the infinite lines of the plane will not bend, rather become the chords in the disc. Of course, the circle around the disc does not belong to the disc and the chords are actually infinitely long, because towards the edges, smaller and smaller real distances of the disc would mean bigger and bigger imaginary distances. Most importantly, all the simple axioms about the plane remain true, but obviously we'll have more non crossing lines to an L through a P outside:



But this tricky idea is still not the point! To see why this meant such a big leap in the history of human thinking, we have to go back a bit and remember where all the non euclidian investigations started from. They wanted to prove the parallelity axiom from the other simple axioms. Then, as an alternate strategy for this, they assumed the parallelity axiom to be false and tried to reach a contradiction. Finally, Gauss, Bolyai, Lobachevsky realized that such contradiction will not be obtained, rather an amazing new geometry can be developed. This personal conviction of them is very admirable, but still didn't prove that the parallelity axiom can not be derived from the simpler ones. Beltrami not only created the above mentioned strange model of plane geometry, but realized that it proves it without a shadow of a doubt, that the parallelity axiom is not derivable logically from the others. But how can this be if at that time the logic of mathematics was not even worked out yet. And that's the whole point!
Beltrami's realization was the actual seed that spawned the whole new mathematics, Set Theory and Logic. Indeed, we don't have to know about sets or mathematical logic, and still obtain something fundamental about them. Namely, even if we don't know what "logical" means exactly, one thing is sure: If from some assumptions, others should follow logically, then in every reality where those assumptions are true, all those that "logically follow", should be true too. Now, in the above disc model of the plane, all the simple axioms of normal plane are true. But then everything that logically follows from those simple axioms must be true there too. And of course, the single non crossing through an outside point is not true and so it can't be a logical consequence of the simpler axioms.
In the followings, we go back from this giant leap of Beltrami, and simply investigate the possibilities without the assumption of single non crossing lines. Unlike Gauss, Bolyai and Lobachevsky who investigated the new possibilities just to see what a wider world may look like, we will obtain something very positive and amazing for the parallelity axiom itself. Namely, the first seven theorems will prove that we don't have to assume a single non crossing line for all lines and outside points, because one such will automatically imply it for all. Finally, only the eighth theorem will be a description how the world without single non crossing lines can look like.

D

The crucial assumption for our following theorems is the concept of angle.
Our intuitive conviction that two crossing lines determine a certain angle must be reduced to the concept of lengths. And indeed, what we really mean by equal angles is the fact that measuring the same lengths on the sides, we get identical connecting distances.



This is actually the definition of equal angles. Then the axiom we need is that once two angles are equal, the same equality of connecting distances to equal side distances is true for all possible side distances.

T_1

U is unique non crossing with L through P

↓

L and U are equidirectional to all connectors.

1.) If two lines are equidirectional to a third, crossing them in P , Q points, then the lines are mirrored of each other to the M middle of P and Q.

2.) Such two lines can only cross in a C point that is mirror of itself to M.

3.) a.) If such two lines exist, then all lines going through P and Q will cross.
b.) If there are non crossing lines, one through P and one through Q, then any equidirectional lines to PQ will not cross.

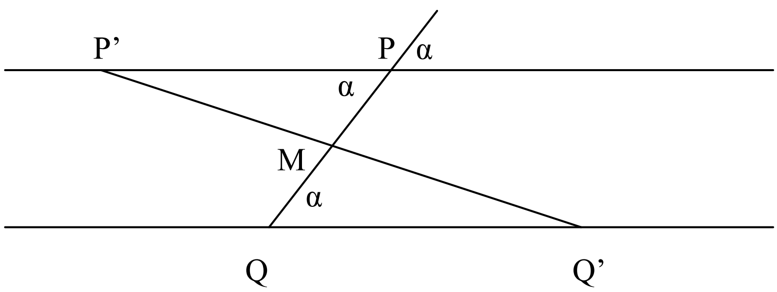
4.) If there is a non crossing line with L through P, then any line equidirectional with L to any connector from P is non crossing.

5.) If there is a U unique non crossing line with L through P, then U is equidirectional with L to any connector from P.

6.) This U is equidirectional with L to any connector.

P

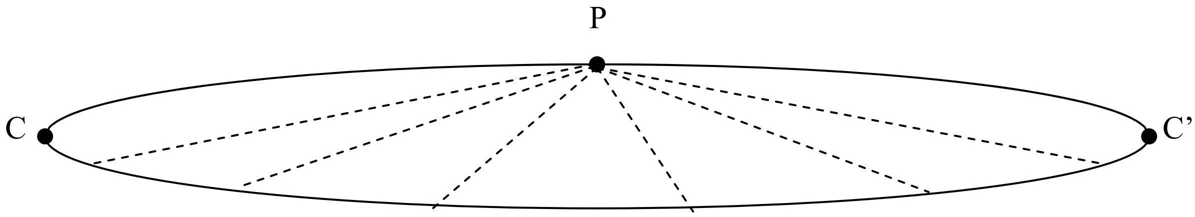
1.)



The same α angle appears at P on the other side of the PQ connector.
Then this means that measuring any distance from P and Q on the two lines in opposite directions and connecting these P' and Q' with the M middle point of PQ will give identical triangles. Thus, P'M and MQ' are on one line and equal, so in other words, indeed, P' and Q' are mirrored to M.

2.) First it seems, that such two lines that are mirrored to an M point can't cross at all. Indeed, we can argue that if they cross in a C on the right side, then mirroring C to M, we get a C' that again must be on both lines. Thus, the two lines would connect the C and C' points contradicting that two points determine only one line.
This argument however, is faulty because it assumed that C' will be a different point from C. If C' is the same as C, then the two lines doesn't make a contradiction.
Of course, then C is a mirrored point of itself, which seems just as contradictory.
Indeed, then C has a distance from itself. Such C could be called a weird point and can be easily excluded if we assume that points only have 0 lengths. Yet, we don't make this assumption right now, and play with the idea of weird points a little bit longer.

3.) a.)



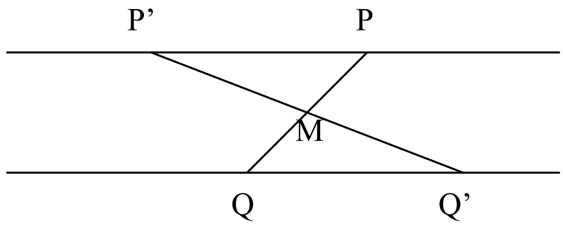
Sweeping through all possible connectors from a P point, we obtain all lines through P.

b.) It is merely an other way of saying a.).

4.) Special case of 3.) b.).

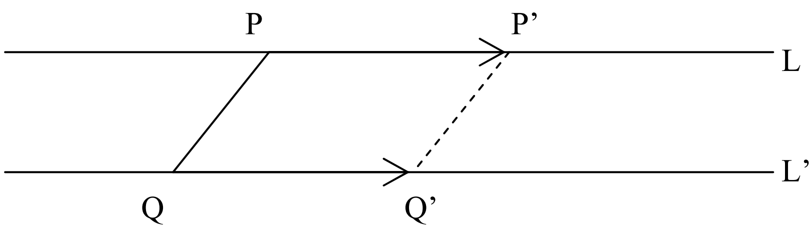
5.) If U weren't equidirectional with L for a PQ connector, then we could draw an equidirectional with L through P, which were a different and non crossing by 4.).
This would contradict the assumption of having only one non crossing through P.

6.) Let P'Q' be a connector, so that P' $\neq$ P. Let M be the middle of P'Q'.
Lets connect P with M to obtain a Q crossing on L. For this PQ connector the equidirectionality was proved in 5.). But this means being mirrored to M.
And this means being equidirectional to any line going through M, including P'Q'.



R We already mentioned that the concept of being fix distanced from an L line is controversial.
Now, we still generalize that concept, avoiding the perpendicularity and rather use any connector as the starting "distance" between the two lines:

D L' is fix distanced with L from a PQ connector if:
P $\in$ L', Q $\in$ L and measuring any common distance on L' and L from P and Q in the same direction, will lead to P' and Q' points that are the same distanced as was PQ.



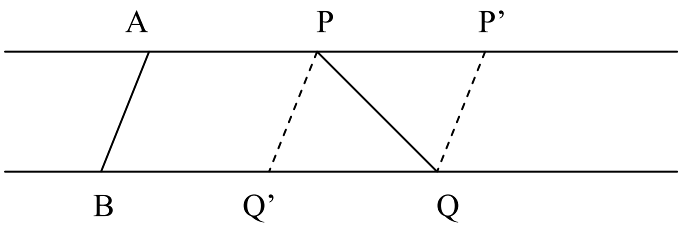
PP' = QQ' → P'Q' = PQ.

T_2 Equidirectionality to two non halving connectors avoids crossing.

L' is fix distanced with L from one connector.
↓
L' is equidirectional with L to two non halving connectors.
↓
L' is equidirectional with L to all connectors.
↓
L' is fix distanced with L from all connectors.
↓
L' doesn't cross L.

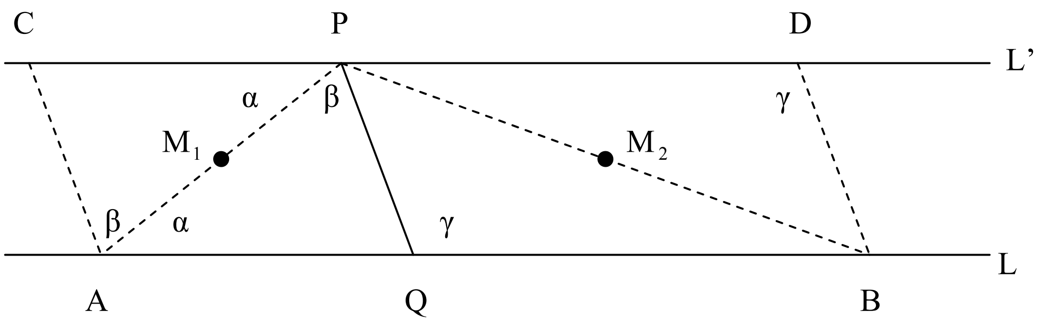
P

1.) Let the one connector be AB from which L and L' are fix distanced and let PQ be an arbitrarily chosen new connector:

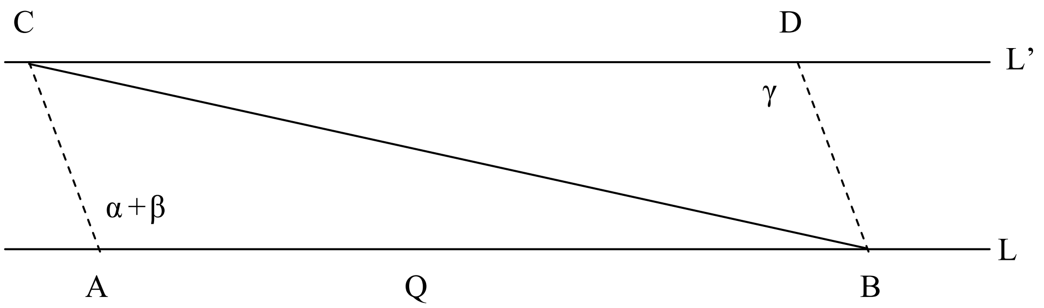


We measured AP from B to get Q' and BQ from A to get P'.
Then PQ' = QP' because of the fix distancedness. Then since
PP' = AP' – AP = BQ – AP and QQ' = BQ – BQ' = BQ – AP thus, PP' = QQ'.
These are in opposite directions, so the lines are indeed equidirectional to PQ.

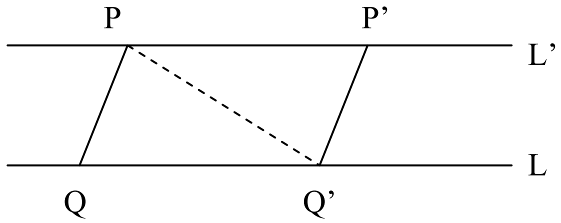
2.) Let M_1 and M_2 be the two middle of the non halving connectors and let PQ be a third connector. Connect P with M_1 and M_2 to obtain A , B on L and then the same distanced C , D on L':



The pairs of α , β , γ angles in the picture are equal because we used connectors through M_1 and M_2. All we have to show is that $\alpha + \beta = \gamma$. And indeed:



3.) Let PQ be a connector from which we want to show the fix distancedness. Lets measure any same distances from P and Q in the same direction to obtain P' and Q'.



PQ' used as connector, the equidirectionality implies P'Q' = PQ.

4.) Suppose L and L' would cross in a C point. Measuring back any same distance from C on L and L', we would get a connector to which the fix distancedness would imply that C is this fix distanced from itself. This seems like again just a weird crossing point of the two lines, but actually this is a deeper impossibility, because we obtained it by measuring back arbitrary distances from C. For example, we can measure back a d distance and obtain a c length connector. Then approaching C arbitrarily close, we would have points at c distance. This contradicts the triangle inequality.

T_3 Triangles if there are no weird points.

1.) Outer angle at a corner can't be equal to the inner angle at an other corner.

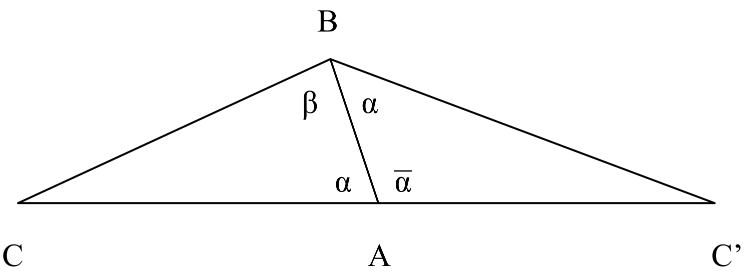
2.) Outer angle at a corner is bigger than the inner angle at an other corner.

3.) Two inner angles together are less than 180°.

4.) The three inner angles together are less or equal to 180°.

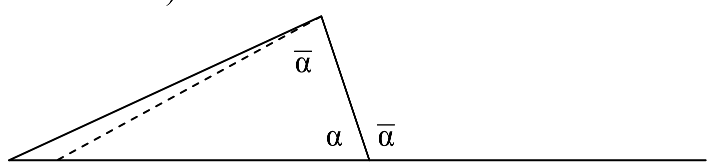
P

1.)



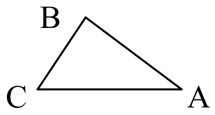
We measured the BC length from A on the continuation of CA to obtain C'.
If $\overline{\alpha}$ were equal to β, then the ABC triangle were identical with ABC'.
But then the angle at B, in the ABC' triangle would be also α.
Thus, the full angle at B were $\alpha + \beta = \alpha + \overline{\alpha} = 180^\circ$ and so, C , B , C' were on a line contradicting that A,B,C is a triangle that is, B is not on the < CA > line.
Of course again, this contradiction would be avoided if C and C' were the same points.
If there are no such weird points, then the contradiction proves that $\overline{\alpha} \neq \beta$.

2.) By 1.) we only have to show that $\overline{\alpha} < \beta$ is impossible. If this were the case, then measuring $\overline{\alpha}$ on AB at B, we would get a new triangle leading to the same contradiction as in 1.):

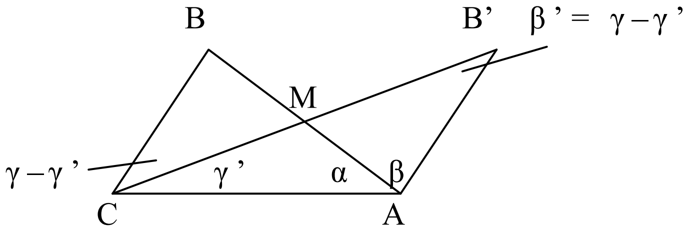


3.) By 2.), $\overline{\alpha} > \beta$, thus adding α to both sides, $\alpha + \overline{\alpha} = 180^\circ > \alpha + \beta$.

4.) Let $\alpha \le \beta$, that is $BC \le AC$.



Lets mirror C to the M middle of AB to get a new B'.



$AB' = BC \le AC \rightarrow \gamma' \le \beta' = \gamma - \gamma' \rightarrow \gamma' \le \frac{\gamma}{2}$.

The angle sum of the A, B', C triangle is: $\alpha + \beta + \gamma - \gamma' + \gamma' = \alpha + \beta + \gamma$.

So, if the angle sum of the A, B, C triangle were $180^\circ + \delta$, then so would be of the A, B', C triangle but with having an angle half or less than γ.

Repeating such replacements, we would reach a triangle with $180^\circ + \delta$ angle sum, but with one angle being less than δ.

Then, the other two angles would be still more than 180°, contradicting 3.).

T_4 Triangles for fix base line and corner above.

A triangle based on L and cornered across at P has 180° angle sum.

↓

An L' line through P is equidirectional with L to two connectors from P.

↓

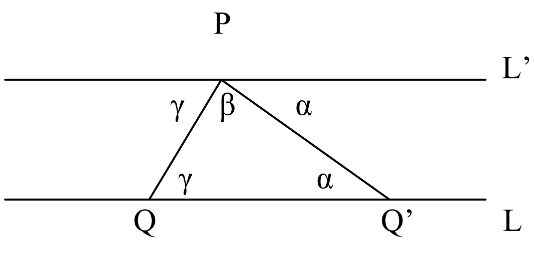
L' is equidirectional to all connectors.

↓

Any triangle based on L and cornered at P has 180° angle sum.

P

1.)



2.) Follows from 1.) and T_2 1.) because PQ and PQ' are non halving.

3.) The picture above in 1.) proves it again.

T_5 Triangles for changing lines.

There is an L' so that they are equidirectional to two connectors from $P \in L'$

$\updownarrow$

There is a triangle with 180° angle sum.

$\updownarrow$

Any triangle has 180° angle sum.

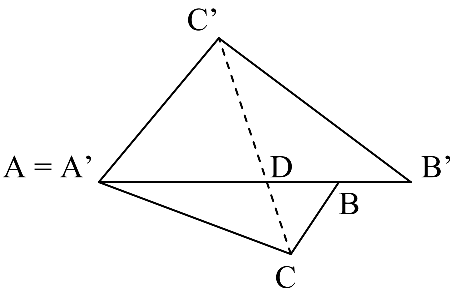
$\updownarrow$

All L and L' that are equidirectional to one connector, are equidirectional to all connectors.

P

1.) Trivial again by picture in proof of T_4 1.).

2.) By previous theorem, if an A, B, C triangle has 180° angle sum, then all triangles with a common corner and base line across will be such again. With this trick, from A, B, C we can reach any A', B' C'. Indeed, lets place A over A' and B onto < A'B' >:



ABC = 180° → ACD = 180° → ADC = 180° → A'B'C' = 180°.

3.) Follows from 1.) and T_2 1.).

T_6

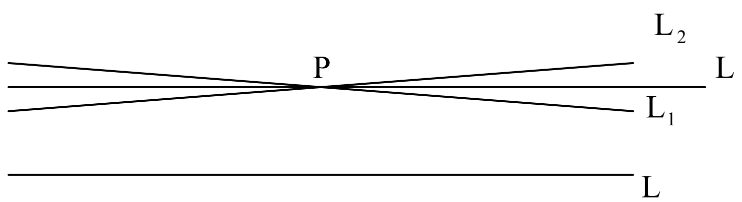
U is equidirectional with L to two connectors from $P \in U$

↓

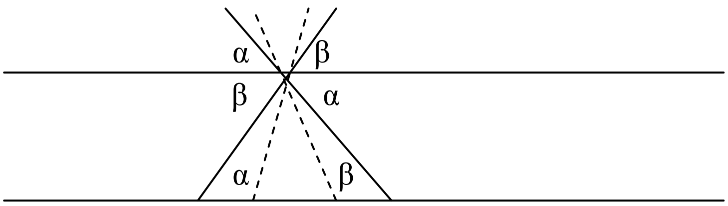
U is unique non crossing with L through P.

P

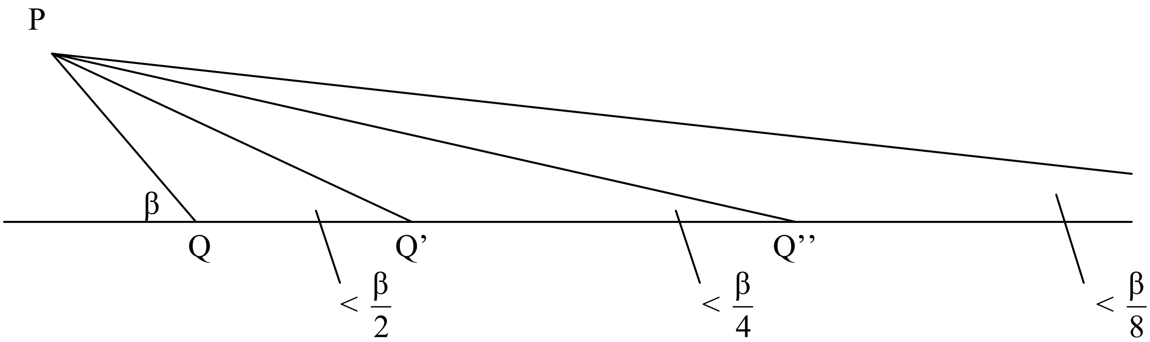
1.) If two L_1 , L_2 lines are non crossing with L, through P, then all L' lines in between them are not crossing either. Thus, the non crossings form a "bundle".



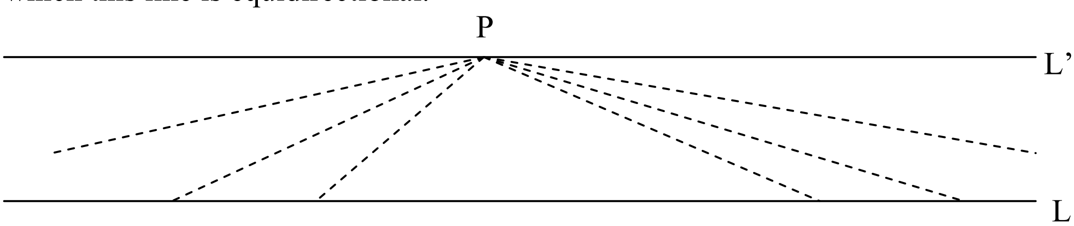
2.) If two connectors from P go in α and β to L and there is a line through P equidirectional to both connectors, then the δ angle of the non crossing bundle is smaller than α + β .



3.) There are arbitrary small angled connectors to a line.
Lets start with any β angled PQ connector and measure the connector on L to get a new Q'. The PQ' connector must have less than $\frac{\beta}{2}$ angle by T_3 4.).
And so on, we always get less than half of the previous angles.



4.) If there is an L' line through P that is equidirectional to two connectors from P, then the bundle is a single line.
Indeed, by T_2 1.) this line is equidirectional to all connectors from P.
Then we can apply 3.) in both directions to obtain pairs of arbitrary flat connectors to which this line is equidirectional.



Thus, by 2.), the non crossing bundle must have 0 angle, that is L' is the single non crossing.

T_7 **There is a unique non crossing U line with a fix L through a fix P**

↓

There is a unique non crossing U line with any L through any P outside.

P By T_1 5.), and T_5 all L lines and P outside have an L' equidirectional to all connectors.
Thus, by T_6 this L' is the U unique non crossing through P.

T_8 If there are more non crossings with L through a P point then:

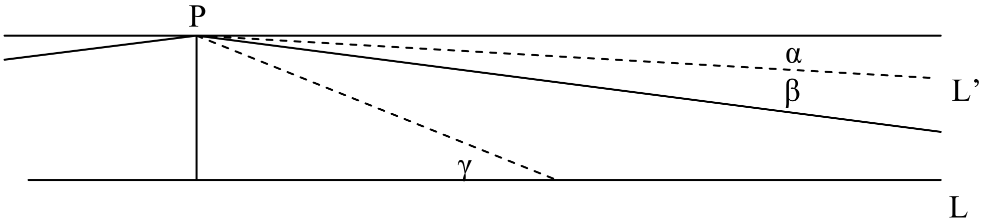
1.) The edge lines of the non crossing bundle through P are non crossing either.
2.) All non edge non crossings are equidirectional to exactly one connector.
3.) The connectors and the equidirectional lines to them turn towards each other.

If there are no non crossings with L through a P point then:

4.) All lines through P are equidirectional to exactly one connector.
5.) The connectors and the equidirectional lines to them turn in the same direction.

P

1.) A crossing line has other crossing ones arbitrary close in both directions.
The edge of the non crossings has only close ones in one direction.
2.) Lets draw the perpendicular connector from P and draw a perpendicular line to this at P. This will be the middle line of the bundle which is equidirectional with L to this perpendicular connector. Any other L' line inside the bundle will have an α angle to this middle one and β to the edge. It will also lean toward L on one side, say the right.
Lets change the perpendicular connector toward the right to flatter and flatter positions.
It's γ angle to L will go from 90° to 0°. On the other hand, its angle to L' will go from $90^{\circ} - \alpha$ to β. Thus, the two angles will be equal at exactly one position of the connector:



R

We already mentioned Beltrami's model for more non crossing lines through a point.
Strangely, for the "weirder" no non crossing lines, there was a model already known for two thousand years. Indeed, the spherical geometry was investigated by the greeks.
The farthest, opposite two points of a sphere called antipodal. On the earth, such are the north and south pole. Of course, every P point has an antipodal pair, because we can simply mirror P to the center of the sphere. The circle's going through antipodal points, that is having their center at the center of the sphere are called main circles. On the earth, such are the equator and the time zone circles. Every two points determine a main circle going through them, because the center of the sphere and the two points are on a plane that crosses the sphere in this circle. Also, between any two points, this main circle is the shortest path on the sphere. So as we see, the main circles are behaving as lines. Unfortunately, we were wrong above because not any two points will determine a unique main circle. The exceptions are the antipodal pairs through which infinite many main circle go, like the time zones through the north and south pole. An other consequence of this "error" is that if the main circles are the lines, then two lines will cross not in one point but rather in the two antipodal ones. With a simple trick, we can get over this problem, and obtain a "perfect" model. The antipodal pairs should be regarded as single points. The strangeness of these new points is then easy to see, because going on a main circle we arrive back to our points, in fact not in a full circle, but already at halfway.
So the lengths are limited!

Part Three: Complex Numbers
1. Definitions

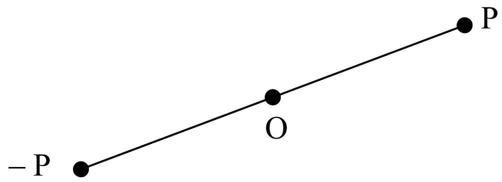
R

The real numbers are the comparing of distances or if we choose a unit length on a line, then the points of the line themselves can be regarded as the real numbers.
It is amazingly simple to generalize this and regard all points of a plane as numbers.
This has consequences in two directions: In algebra, equations will have wider solutions and in geometry the coordinate systems become simpler.
And yet, this whole approach is avoided in high schools. This is solely due to the reluctance of teachers who themselves are unfamiliar with complex numbers. But complex numbers can not be avoided! Engineering and science must use them! So in tertiary education they are introduced but seem strange to students and so the cycle repeats, complex numbers stay out of basic mathematics. It is interesting to compare the complex numbers with calculus. The question whether calculus should be taught in high schools always reoccurs, but complex numbers are consistently avoided, even though they are much more basic and simpler too.
I hope the followings will convince the readers to give a chance to complex numbers in the wider education.

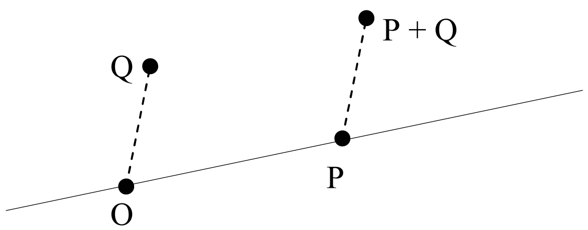
D

The points of a plane can be regarded as positions if we choose a fix O origin.

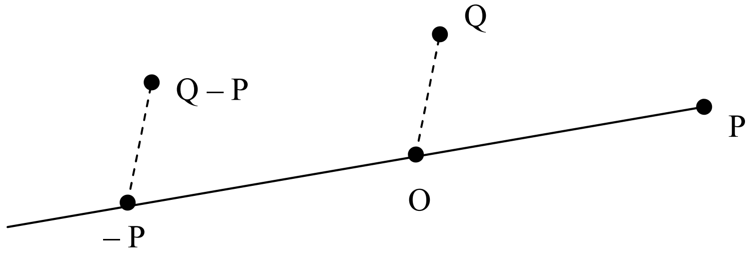
1.) The simplest operation of a point is the opposite of it, that is – P from P.
This means the mirroring of P to the fix O:



2.) The addition of two points, that is P + Q means that the first, that is P, is regarded as a new origin and Q is positioned from there. Thus, the OQ distance is shifted along the OP line until O goes to P.



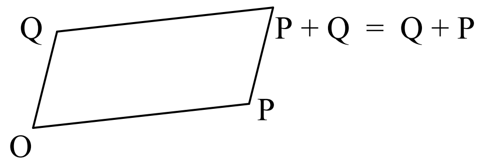
3.) Q – P is simply defined as – P + Q.



T

1.) P + Q = Q + P
2.) Q – P + P = Q

P

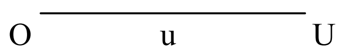
1.) 

If OP is shifted to Q, to get P + Q then O, P, Q, P + Q is a parallelogram, so OQ is shifted to the same point.

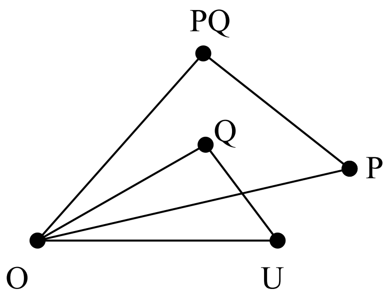
2.) Q – P is shifting Q towards – P and then + P is shifting it back to Q.

D

To define the multiplication of points, we need more than a fix O point of origin, namely we need a fix U unit point too. The u distance between O and U is the unit length.



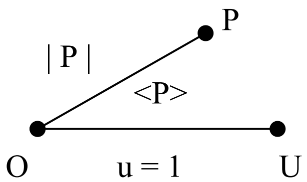
The same basic idea is used to define P • Q = PQ as was for P + Q, that is Q will be repositioned by using P. But now not a shifting is applied, rather a turning and changing of unit length. So, P will be regarded as the new U.



The "repositioning" means that PQ must be relative to OP the same as Q was to OU.
In other words, the O , U , Q triangle must be similar to the O , P , PQ.
Thus, the Q , O , U angle is the same as the PQ , O, P. And O , PQ length is $OQ\,\dfrac{OP}{OU}$.

D

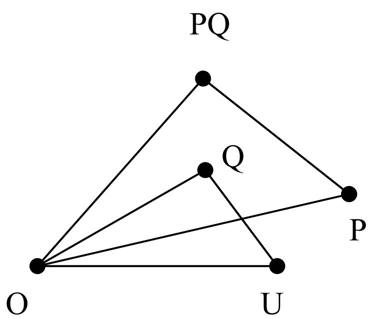
| P | : = distance of P from O, using u as unit.
<P> : = angle of OP from OU.



T

1.) | PQ | = | P | | Q |
2.) <PQ> = <P> + <Q>

P



1.) $| PQ | = | Q | \dfrac{|P|}{u} = | P | | Q |$

2.) <PQ> = Angle PQ , O , U =
Angle P , O , U + Angle PQ , O , P =
< P > + Angle Q , O , U = <P> + <Q>

T $PQ = QP$

P Both $|\ |$ and $<\ >$ are the same for the two sides.

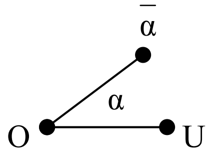
D $\dfrac{P}{Q}$:= the point R so that $RQ = P$.

T $\left|\dfrac{P}{Q}\right| = \dfrac{|P|}{|Q|}$ $\quad \left\langle\dfrac{P}{Q}\right\rangle = \langle P\rangle - \langle Q\rangle$

T $(P \pm Q)\, R = PR \pm QR$

P The turning and increase of a parallelogram keeps it a parallelogram.

D 1.) All P points of the plane are determined by their $|P|$ distance and $\langle P\rangle$ angle. The $u = 1$ distanced points from O, that is the points on the unit circle are called the unit points. These are determined by merely their α angle and can be denoted as $\overline{\alpha}$.

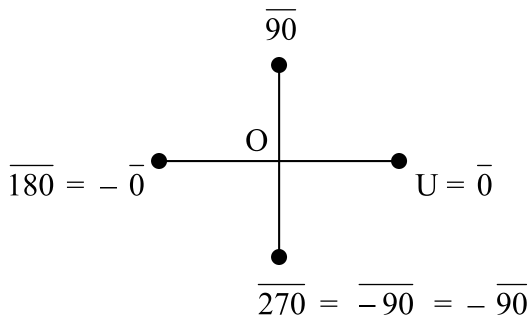


The unit vector of a P point is the unit vector on OP and is denoted as P^{o}.
Thus, $P^{o} = \overline{\langle P\rangle}$ and $P = |P|\, P^{o} = |P|\, \overline{\langle P\rangle}$.

2.) Our earlier rules for the multiplication of points can now be expressed as:

$$PQ = |P|\,\overline{\langle P\rangle}\ \ |Q|\,\overline{\langle Q\rangle} = p\,\overline{\alpha}\ q\,\overline{\beta} = p\,q\,\overline{\alpha}\ \overline{\beta} = p\,q\ \overline{\alpha+\beta}$$

3.) The bar notation is especially useful for concrete angles. In fact, U itself can be denoted as the zero angled unit vector, that is $\overline{0}$. The four perpendicular unit vectors are:
$\overline{0}$, $\overline{90}$, $\overline{180} = -\,\overline{0}$, $\overline{270} = \overline{-90} = -\,\overline{90}$



Usually $\overline{0}$, $\overline{90}$, $-\,\overline{0}$, $-\,\overline{90}$ are abbreviated as 1 , i , -1 , $-i$.
In other words, the unit vector is identified with the number unit 1.

T 1.) $-1\,P = -P$

2.) $i^2 = -1$

3.) $\dfrac{1}{i} = -i$

P

1.) $-1\,P = \overline{180}\;\; p\overline{\alpha} = p\;\overline{\alpha+180} = p\;\overline{-\alpha} = -p\overline{\alpha} = -P$

2.) $i^2 = i\,i = \overline{90}\;\overline{90} = \overline{180} = -1$

3.) $\dfrac{1}{i} = \dfrac{\overline{0}}{\overline{90}} = \overline{-90} = -\,\overline{90} = -i$

2. Exponentiation

D

$(p\overline{\alpha})^n = p\overline{\alpha}\;\; p\overline{\alpha}\;\ldots\; p\overline{\alpha} = p\,p\ldots p\;\; \overline{\alpha+\alpha+\ldots+\alpha} = (p^n)\;\overline{\alpha n}$

$(p\overline{\alpha})^x = (p^x)\,\overline{\alpha x}$

T

1.) $(p\overline{\alpha})^x\;\;(p\overline{\alpha})^y = (p\overline{\alpha})^{x+y} = p^{x+y}\;\;\overline{\alpha(x+y)}$

2.) $(p\overline{\alpha})^x\;\;(q\overline{\beta})^x = (p\overline{\alpha}\;q\overline{\beta})^x = p^x\,q^x\;\;\overline{(\alpha+\beta)\,x}$

3.) $\dfrac{(p\overline{\alpha})^x}{(q\overline{\beta})^x} = \left(\dfrac{p\overline{\alpha}}{q\overline{\beta}}\right)^x = \left(\dfrac{p}{q}\right)^x\;\overline{(\alpha-\beta)\,x}$

4.) $\left[(p\overline{\alpha})^x\right]^y = (p\overline{\alpha})^{xy} = p^{xy}\;\;\overline{\alpha\,xy}$

D

$\sqrt[x]{P}$:= set of all Q points for which $Q^x = P$.
For $x = 2$ we omit this 2 from the root.

T

$P^{\frac{1}{x}} = (p\,\overline{\alpha})^{\frac{1}{x}} = p^{\frac{1}{x}}\;\overline{\frac{\alpha}{x}} \in \sqrt[x]{P}$

$\sqrt[x]{p\overline{\alpha}} = \left\{ p^{\frac{1}{x}}\;(\overline{\tfrac{\alpha}{x}}) \quad , \quad p^{\frac{1}{x}}\,(\overline{\tfrac{\alpha}{x}+\tfrac{360}{x}}) \quad , \quad p^{\frac{1}{x}}\;(\overline{\tfrac{\alpha}{x}+2\,\tfrac{360}{x}}) \quad , \;\ldots\ldots\ldots\ldots \right\}$

T

For $x = n$ natural, there are exactly n roots. For example:

$\sqrt{1} = \sqrt{\overline{0}} = \left\{ 1\,\overline{\tfrac{0}{2}}\;,\; 1(\overline{\tfrac{0}{2}+\tfrac{360}{2}}) \right\} = \left\{\overline{0}, \overline{180}\right\} = \{1, -1\}$,

$\sqrt{-1} = \sqrt{\overline{180}} = \left\{\overline{90}, \overline{270}\right\} = \{i, -i\}$

$\sqrt[3]{1} = \sqrt[3]{\overline{0}} = \left\{\overline{0}\;,\;\overline{120}\;,\;\overline{240}\right\}$

$\sqrt[3]{-1} = \sqrt[3]{\overline{180}} = \left\{\overline{60}\;,\;\overline{180}\;,\;\overline{300}\right\} = \left\{\overline{60}\;,\;-1\;,\;\overline{300}\right\}$

$\sqrt[4]{1} = \sqrt[4]{\overline{0}} = \left\{\overline{0}\;,\;\overline{90}\;,\;\overline{180}\;,\;\overline{270}\right\} = \{1\;,\;i\;,\;-1\;,\;-i\}$

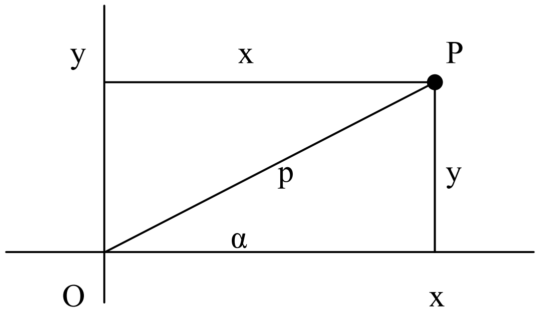
$\sqrt[4]{-1} = \sqrt[4]{\overline{180}} = \left\{\overline{45}\;,\;\overline{135}\;,\;\overline{225}\;,\;\overline{315}\right\}$

3. Complex Form

R Giving the points of a plane as $p\overline{\alpha}$ is usually called the polar coordinate form.
We already mentioned in Part 2 Section 4, the Descartes coordinate system.
With the use of the Pythagoras theorem and the sin , cos , tan , so called trigonometric functions, we can calculate from polar to Descartes and back:

$$p\overline{\alpha} = (x , y) = (p \cos \alpha , p \sin \alpha)$$

$$(x , y) = p\overline{\alpha} = \sqrt{x^2 + y^2}\ \overline{(\text{arc} \tan \frac{y}{x})}$$



Only the Descartes coordinate system is usually taught in high school, in spite of the fact that the polar system is just as useful. For example, Newton's major result, the derivation of Kepler's First Law can much easier be proven with polar coordinates than with Descartes'.
The Descartes system can be dramatically improved if instead of two separate coordinates, the P point is actually represented as the sum of the coordinates. Of course, x and y can not be simply added because that would be just an other number on the single real number line.
We need separate notation for the points of the perpendicular y-axis. And this is easy if we use the $\overline{90} = i$ unit on it. Indeed then, y is actually yi, so $P = x + yi$.
By the way, the letter i comes from the word "imaginary", though in the way we introduced it there was nothing imaginary about it. But when it was first used it simply meant a number so that its square is -1. There is such number on the real number line, so looking from there, it seems imaginary.
The huge advantage of the $x + yi$ so called complex form is that all the coordinate calculations come out "by themselves":

D Complex form: $P = x + iy$

T

$$P \pm Q = (x + iy) \pm (v + iw) = (x \pm v) + i (y \pm w)$$

$$P Q = (x + iy) (v + iw) = xv + ixw + iyv + i^2 yw = (xv - yw) + i (xw + yv)$$

$$\frac{P}{Q} = \frac{x+iy}{v+iw} = \frac{x+iy}{v+iw}\,\frac{v-iw}{v-iw} = \frac{xv-ixw+iyv-i^2yw}{v^2+w^2} = \frac{(xv+yw)+i\,(yv-xw)}{v^2+w^2} =$$
$$\frac{xv+yw}{v^2+w^2} + i\,\frac{yv-xw}{v^2+w^2}$$

R

The complex multiplication when calculated from the polar form, even leads to the sums of sin and cos "by itself". Repeating it with unit length:

$$\overline{\alpha+\beta} \quad = \quad \overline{\alpha}\ \overline{\beta}$$

$$\cos(\alpha+\beta)+i\sin(\alpha+\beta) = (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) =$$

$$\cos\alpha\cos\beta + i\cos\alpha\sin\beta + i\sin\alpha\cos\beta + i^2\sin\alpha\sin\beta =$$

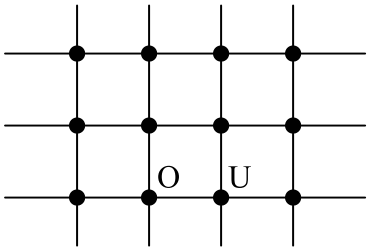
$$\cos\alpha\cos\beta - \sin\alpha\sin\beta + i(\cos\alpha\sin\beta + \sin\alpha\cos\beta)$$ Thus,

$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

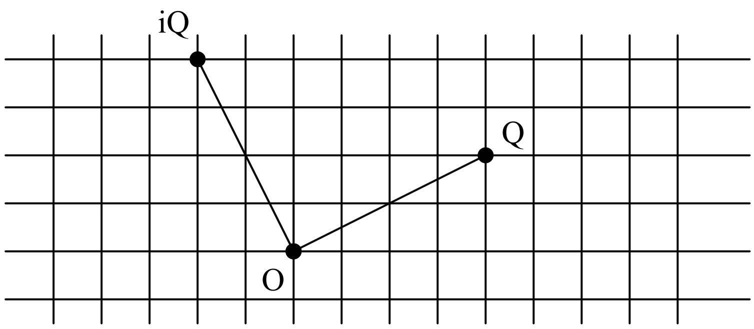
$$\sin(\alpha+\beta) = \cos\alpha\sin\beta + \sin\alpha\cos\beta$$

D

A chosen OU unit interval determines the "grid" points by measuring the unit repeatedly, horizontally and vertically.



These grid points are the same as the (x , y) coordinate points with x and y whole numbers. An even better way to regard them is the x + iy complex numbers with x and y wholes. Then the above listed coordinate calculations of + , – and • show at once that these three operations of whole complex numbers lead again to wholes.
For + , – this is natural from the original meaning too, because shifting a grid point with a grid value goes again to a grid. For a $PQ = p\overline{\alpha}\ \ q\overline{\beta}$ multiplication, the original meaning of using Q as the new unit doesn't explain at once why PQ will be a grid.
The following argument helps:
1.) The 90° turned version of Q, that is iQ is obviously a grid:



2.) The repetitions of Q or iQ in their own directions are again leading to grids.

Then PQ can be visualized as P = (x , y) = x + iy but with new Q and iQ units, instead of the 1 and i. In other words, Q is repeated x times, while iQ is y times. These are grids by 2.), and so their sum is grid again.

The division of complex wholes of course leads out of them, in other words, the ratio of two grid points doesn't have to be a grid. Just as among whole numbers, here too the dividabilities are the most crucial.

While among naturals, every number is dividable by 1, and among positive and negative wholes, or integers, every number is dividable by 1 or – 1, here among the grids, we have four units, 1, – 1, i , – i and every grid is dividable by these. So now we should call a grid composite if it is the product of two grids, both different from the units. The non composite grids could be also called as primes, just as among the natural numbers.
Gauss was the first who realized this amazing generalization of number theory to two dimensional grids or complex wholes, and so they are also called Gaussian integers.
The most fundamental theorem about primes among the naturals is that every number can be uniquely decomposed into prime factors, except of course the order of the members.
This is the reason why 1 is not regarded as a prime! Indeed, even though it is not a composite, so it should be a prime, if we allowed it to be a factor then it could be repeated as many times as we wish, so the decomposition weren't unique. Similarly, among complex wholes the 1 , – 1, i , – i units are not regarded as primes, and then indeed every complex whole or Gaussian integer can be uniquely decomposed into primes.
Gauss completely answered the question, what grids are the primes. Most amazingly, this shed new light on the already known fact that half of the natural primes can not be written as square sums, while the other half can be uniquely. Namely, the 4k – 1 primes can not be, the 4k + 1 can be. Indeed, $3 \neq a^2 + b^2$, $5 = 1^2 + 2^2$, $7 \neq a^2 + b^2$, $11 \neq a^2 + b^2$, $13 = 2^2 + 3^2$, $17 = 1^2 + 4^2$, $19 \neq a^2 + b^2$, . . . The "reason" for this is that among the complex wholes the old 4k + 1 primes are not primes anymore.
Indeed, $(1 + i2)(1 - i2) = 1^2 - i^2 2^2 = 1^2 + 2^2 = 5$

$(2 + i3)(2 - i3) = 2^2 + 3^2 = 13$, . . . and so on.

This of course doesn't explain why exactly the 4k + 1 ones are such and why the decomposition is unique. The uniqueness follows from the fact that these two factors are always primes. And this follows from a basic law proved by Gauss, namely that for grids not on the x , y axis, the primes are merely the ones with $x^2 + y^2$ being a natural prime.

4. Unique Prime Factorization Through The Grid Fractions

R I already showed among naturals this same root, so now only repeat the main idea and then redo the details for gird numbers. The most practical appearance of prime factors comes through the simplification of fractions. We cross out the common factors of numerators and denominators. Slowly we get the impression that these omittable factors are inherent to a number. If we ask people to list all the factors of a number, then they usually proceed by dividing the number with gradually increasing primes. For example, for 120, they would go as: $120 = 2\cdot 60 = 2\cdot 2\cdot 30 = 2\cdot 2\cdot 2\cdot 15 = 2\cdot 2\cdot 2\cdot 3\cdot 5$. Of course, we can go in other order too, for example: $120 = 3\cdot 40 = 3\cdot 2\cdot 20 = 3\cdot 2\cdot 5\cdot 4 = 3\cdot 2\cdot 5\cdot 2\cdot 2$.
Somehow it feels obvious that we ended up with the same factors. Yet, we can show that this is far from obvious. Indeed, if we assume the possibility of different breakdowns, leading to different prime factorizations, then omitting the common prime factors, it would simply mean that: $p_1\cdot p_2\cdot\ .\ .\ .\ .\ \cdot p_m = q_1\cdot q_2\cdot\ .\ .\ .\ .\ \cdot q_n$ with different primes on the left and on the right. What makes this so impossible? Why couldn't be that $3\cdot 11\cdot 13 = 17\cdot 19$.
Using bigger numbers, it soon becomes clear that nothing makes this impossible, in fact the two sides can be very close. Yet never equal! The real reason for this is a wider truth about numbers that don't even have to be primes, only relative primes. Two numbers are called such relative primes if they have no common factor, except the obvious 1. If one number is multiple of the other, that is $a = mb$ then b is still regarded as common factor so a , b are not relative primes. The simple general fact that makes the different prime products impossible to be equal, is that if a number divides a product, say a divides bc, but a and b are relative primes, then a must divide c. Indeed, if this is true, then $p_1\cdot p_2\cdot\ .\ .\ .\ .\ \cdot p_m = q_1\cdot q_2\cdot\ .\ .\ .\ .\ \cdot q_n$ is impossible because all p-s are relative primes to the q-s and on the other hand, none of the p-s can divide any of the q-s. But our new general fact can be put in an even better form and amazingly it also relates to our elementary school experiences with fractions. Indeed, the a dividing bc means $ma = bc$, so $\frac{a}{b} = \frac{c}{m}$. Then, a and b being relative primes means that the $\frac{a}{b}$ fraction is simplified completely and the claim that a must divide c means that the $\frac{c}{m}$ fraction is merely an expansion of the $\frac{a}{b}$ fraction. So what we claim is that a simplified fraction can only be equal to its expansions. Expansions of course, are not simplified! If we simplify them, we get back the fractions that we expanded. So in other words, we claim that two simplified fractions can not be equal. This again seems natural from our experiences with small fractions, but if we imagine simplified fractions with bigger and bigger numerators and denominators, then we see the non obviousness of our claim. Luckily here at fractions a simple and heuristic way of proving our claim also emerges. The fundamental idea is to forget about the simplified fractions and rather regard the "minimal" ones. For two equal fractions, if one of them has a smaller numerator, then it obviously has a smaller denominator too, because the ratios are the same. Thus, we can simply talk about "smaller" and "bigger" versions of equal fractions. Clearly, there has to be a smallest version among the equal fractions. If we could prove that all other are merely expansions of this minimal, we were finished, because then the minimal one were the single non expanded one and thus it were the simplified one too.
The proof of why the non minimal fractions are expansions of the minimal again uses a heuristic idea. All we need is that if $\frac{a}{b} = \frac{A}{B}$ and $a < A$, that is $\frac{A}{B}$ is a "bigger" version of $\frac{a}{b}$, then $\frac{A-a}{B-b}$ is again a version of the same fraction value, that is $\frac{a}{b} = \frac{A-a}{B-b}$.
But this is obvious, because:

$$\frac{a}{b} = \frac{A}{B} \quad / \cdot b \ \cdot B$$

$$aB = Ab \quad / -ab$$

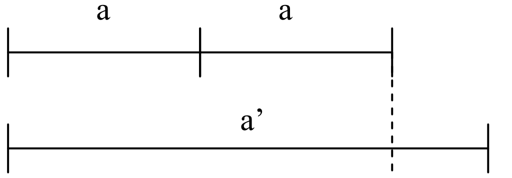
$$a(B-b) = b(A-a) \quad / : b \ : (B-b)$$

$$\frac{a}{b} = \frac{A-a}{B-b}$$

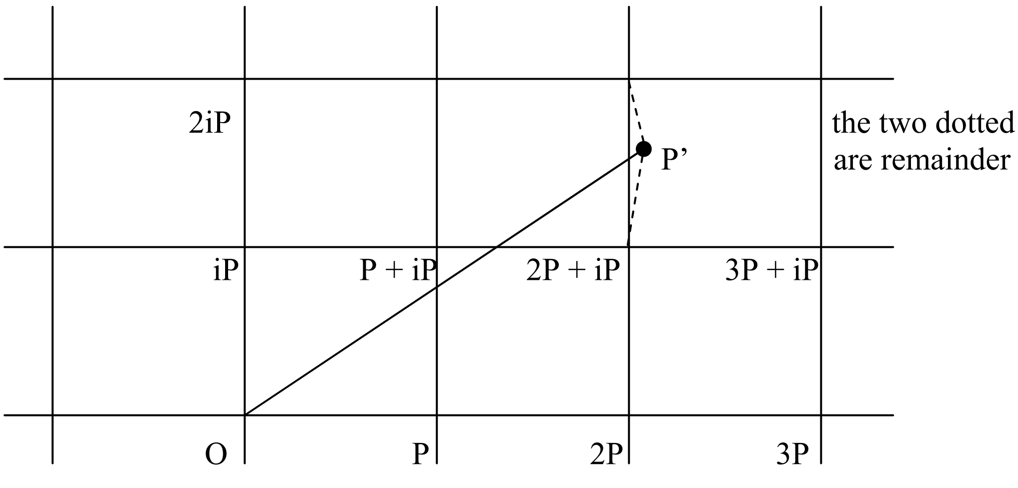
So now, that we see that $\frac{a}{b} = \frac{A}{B} = \frac{A-a}{B-b}$, we can continue subtracting an other a and b from A and B. So, $\frac{a}{b} = \frac{A}{B} = \frac{A-a}{B-b} = \frac{A-2a}{B-2b} = \frac{A-3a}{B-3b} = \ldots$

Sooner or later, we must end up with either $\frac{a}{b}$ again if A , B were multiples of a , b, or if not then with an $\frac{A-ka}{B-kb}$ smaller version than $\frac{a}{b}$. Thus, if $\frac{a}{b}$ was the smallest minimal version, then this second case is impossible so $\frac{A}{B}$ had to be an expansion.

As we see, the heart of the argument was that $A - ka$ becomes a remainder smaller than a.
If we try to use the same argument for grid points P , P', then a simple repeated sequence of subtractions of P doesn't always get closer to P' than | P |! But that's okay, because a whole multiple of P is now two dimensional! Then, such multiple of P does get closer to P' than P. So, the concept of remainder survives! The generalization to two dimension is even better understood if we use already in one dimension not $a < A$ whole numbers rather $a < a'$ arbitrary distances:



Just as the remainder here appeared as we repeated a and placed a' over it, similarly, a remainder of a P' point relative to a P, appears if we "repeat" P. But now repeating means using it as a unit for a grid system. Luckily, from the four corners of the grid, where P' is inside, at least two of them will be closer than | P | and they can be regarded as remainder:



T

1.) For any point in a plane with a grid system, there is a grid point closer than the unit.

2.) For every P , P’ there is a W grid point, so that: $| P' - WP | < | P |$

3.) If $\frac{P}{Q} = \frac{P'}{Q'}$ and $P' - MP \neq 0$ then $\frac{P}{Q} = \frac{P' - MP}{Q' - MQ}$

4.) If $\frac{P}{Q}$ is a minimal among some grid ratios, that is,
there are no $\frac{P'}{Q'}$ grid ratios, so that $\frac{P'}{Q'} = \frac{P}{Q}$ and $| P' | < | P |$,
then all other $\frac{P'}{Q'}$ equal grid ratios are expansions of $\frac{P}{Q}$, that is
$P' = WP$ and $Q' = WQ$

5.) If $\frac{P}{Q}$ is a simplified grid ratio, that is,
P and Q have no common dividers except the units 1, – 1 , i , – i ,
then all other equal grid ratios are expansions of $\frac{P}{Q}$.

$\frac{P}{Q}$ being simplified is also called as P and Q being relative primes.

6.) If P and Q are relative primes, but P divides QP’, then P divides P’.

7.) A grid point is prime if nothing divides it except the units and it is not a unit itself.
If $P_1, P_2, \ldots, P_m$ are some primes different from the $Q_1, Q_2, \ldots, Q_m$ primes,
then $P_1 \cdot P_2 \cdot \ldots \cdot P_m = Q_1 \cdot Q_2 \cdot \ldots \cdot Q_m$ is impossible.

8.) Every W grid point can be written as a unique product of primes except their order and changes by multiplying the members with units.

P

1.) , 2.) Trivial by the picture before the theorem.

3.)

$$\frac{P}{Q} = \frac{P'}{Q'} \qquad / \cdot Q \,,\ \cdot Q'$$

$$PQ' = P'Q \qquad / - PMQ$$

$$P\,(Q' - MQ) = Q\,(P' - MP) \qquad / : (Q' - MQ) \,,\ : Q$$

$$\frac{P}{Q} = \frac{P' - MP}{Q' - MQ}$$

4.) By 2.) there is W so that $| P' - WP | < | P |$.
If P’ were not WP, then $P' - WP \neq 0$ and so by 3.), $\frac{P' - MP}{Q' - MQ}$ were a smaller version of $\frac{P}{Q}$, contradicting that it was minimal.

5.) By 4.), all grid ratios are either minimal or expansion of a minimal.
The simplified ones can not be expanded versions, so they are minimal.
Thus, 4.) applies to them.

6.) P dividing QP’ means that PQ’ = QP’, that is $\frac{P}{Q} = \frac{P'}{Q'}$.
Thus, by 5.) if P and Q are relative primes, then P’ = WP.

7.) All Q-s are relative primes to all P-s but none of the Q-s can divide any P-s.

8.) If a W have two different forms, then crossing out the identical prime factors, we would still end up with the impossible situation of 7.).

5. Exponentiation With i

R

We might hope that after the rationals and reals, both of where exponentiation failed to be a perfect operation, finally among complex numbers, we can succeed and thus, any base to any exponent, can be meaningful. We are wrong! Here too, only special cases can be defined.
First of all, the general exponentiation is easily reduced to the i exponent of a real number and a turned unit:

$$(p\,\overline{\alpha})^{x+iy} = (p\,\overline{\alpha})^{x}\,(p\,\overline{\alpha})^{iy} \text{ and then } (p\,\overline{\alpha})^{iy} = \left[(p\,\overline{\alpha})^{y}\right]^{i} = \left[p^{y}\right]^{i}\left[\overline{\alpha\, y}\right]^{i}$$

Now, we’ll show that these two cases, that is the i exponent of positive reals and turned units contradict each other.
First of all, the turned units could only be defined in a trivial way by simply remaining the same. Indeed, even for the simplest turn, that is $\overline{180} = -1$:

$$\left[(-1)^{i}\right]^{2} = \left[(-1)^{2}\right]^{i} = 1^{i} = 1 \quad \text{so} \quad (-1)^{i} \in \sqrt{1} = \{1\,,\,-1\,\}$$

$(-1)^{i} = 1$ is impossible, because raising both sides to i would give:

$$(-1)^{i\,i} = (-1)^{i^{2}} = (-1)^{-1} = -1 = 1^{i} = 1 \text{ a contradiction.}$$

So indeed, $(-1)^{i} = -1$ can only be. But then this inherits to all turns too:

$$(\,\overline{\alpha}\,)^{i} = \left[(\,\overline{180}\,)^{\frac{\alpha}{180}}\right]^{i} = \left[(-1)^{i}\right]^{\frac{\alpha}{180}} = \left[-1\right]^{\frac{\alpha}{180}} = \left[\overline{180}\right]^{\frac{\alpha}{180}} = \overline{\alpha}$$

Even if we accepted this trivial way of defining the i exponentiation of turned units, it would make the other half, that is the i exponentiation of positive real numbers impossible for anything else than the real unit 1. Indeed, let $x^{i} = p\overline{\alpha} = x^{\beta}\,\overline{\alpha}$. Then:

$$(x^{i})^{i} = x^{i^{2}} = x^{-1} = \frac{1}{x} = \left[x^{\beta}\,\overline{\alpha}\right]^{i} = \left[x^{\beta}\right]^{i}\overline{\alpha} = \left[x^{i}\right]^{\beta}\overline{\alpha} = \left[x^{\beta}\,\overline{\alpha}\right]^{\beta}\overline{\alpha} =$$

$x^{\beta^{2}}\,(\overline{\alpha\beta + \alpha})$. Thus, $\alpha\beta + \alpha = 0$ and so, $\beta = -1$ because $\alpha \neq 0$.

Then, $\frac{1}{x} = x^{\beta^{2}} = x$ and so, $x = 1$.

Thus, the choice is quite logical! We drop the idea of defining i exponents for turned units and rather define it for all positive real numbers.

Amazingly, the turned units will be still remaining with us, because if the i exponent of one a positive real number is unit length, that is $\overline{\alpha}$ then all other b-s will become so too.
Indeed, if $a^i = \overline{\alpha}$ and $b = a^m$ then, $b^i = (a^m)^i = (a^i)^m = (\overline{\alpha})^m = \overline{m\,\alpha}$.
So we only have to define this α for a concrete a positive, not 1, real number and then, the i exponents of all others are determined at once. Or an even better idea is simply to choose the a real number for which a^i is the simplest turned unit, namely $\overline{180} = -1$.
And then, for any $b = a^m$, $b^i = \overline{m\ 180}$.
Euler chose this a number to be e^{π}, so he defined $(e^{\pi})^i = -1$.
The reason for this choice is explained in the following section:

6. Euler's Formula

R

Euler discovered many infinite sums, including the following three:

$$e^a = 1 + \frac{a}{1} + \frac{a^2}{2!} + \frac{a^3}{3!} + \ldots$$

$$\cos a = 1 - \frac{a^2}{2!} + \frac{a^4}{4!} - \frac{a^6}{6!} + \frac{a^8}{8!} - \ldots$$

$$\sin a = a - \frac{a^3}{3!} + \frac{a^5}{5!} - \frac{a^7}{7!} + \frac{a^9}{9!} - \ldots$$

(Elementary derivation of these can be found in the book Infinite Sums)

He also knew that:

$$i^2 = -1 \;,\; i^3 = -i \;,\; i^4 = 1 \;,\; i^5 = i \;,\; i^6 = -1 \;,\; i^7 = -i \;,\; i^8 = 1 \;,\; i^9 = i \;, \ldots$$

Thus, if we write in the e^a sum $i\,a$ in place of a, then we obtain:

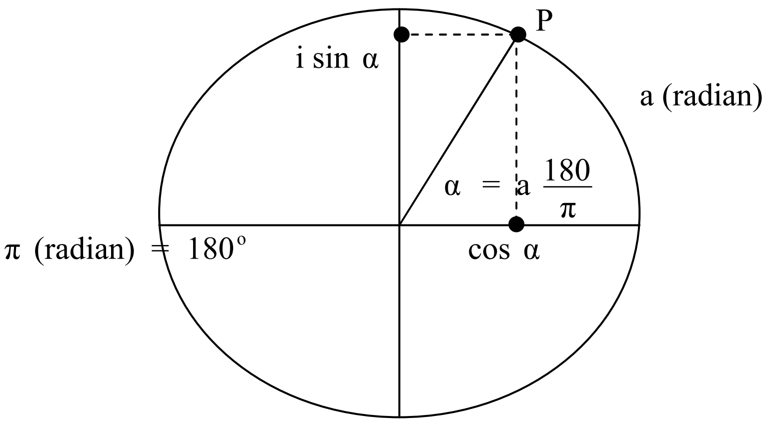
$$e^a = 1 + \frac{i\,a}{1} + \frac{(i\,a)^2}{2!} + \frac{(i\,a)^3}{3!} + \frac{(i\,a)^4}{4!} + \ldots =$$

$$\underbrace{1} + i\frac{a}{1} \underbrace{- \frac{a^2}{2!}} - i\frac{a^3}{3!} \underbrace{+ \frac{a^4}{4!}} + i\frac{a^5}{5!} \underbrace{- \frac{a^6}{6!}} - i\frac{a^7}{7!} \underbrace{+ \frac{a^8}{8!}} + i\frac{a^9}{9!} \ldots$$

As we see, the odd members (marked) give exactly cos a, while the rest of them i multiplied with the members of sin a. Thus:

$$e^{i\,a} = \cos a + i \sin a$$

In the sums for cos and sin, the a argument had to be measured not as angles, but as the circumference of the unit circle. This is also called as the radian. For example, $\alpha = 180^{\circ}$ corresponds to $a = \pi$ because that's the length of the half circle.
$\cos a + i \sin a$ is itself a P point, namely a unit with $\alpha = a\,\frac{180}{\pi}$ turn.



Thus, $e^{i\,a} = (e^{a})^{i} = \cos a + i \sin a = \overline{a\frac{180}{\pi}}$.

In particular, with $a = \pi$: $e^{i\,\pi} = (e^{\pi})^{i} = \overline{180} = -1$.

From this e^{i} can be calculated as $e^{i} = (e^{i\,\pi})^{\frac{1}{\pi}} = (\overline{180})^{\frac{1}{\pi}} = \overline{\frac{180}{\pi}}$

Then, for any $b = e^{m}$ base, we can easily calculate the b^{i} power too:

$$b^{i} = (e^{m})^{i} = (e^{i})^{m} = \left(\overline{\frac{180}{\pi}}\right)^{m} = \overline{m\frac{180}{\pi}}.$$

The m exponent here is also called the natural logarithm or ln of b and thus,

$$b^{i} = \overline{\ln b\,\frac{180}{\pi}}.$$

From the i exponentiation we can easily go to any complex exponent:

$$b^{x+iy} = b^{x}\; b^{iy} = b^{x}\;(b^{y})^{i} = b^{x}\left(\overline{\ln(b^{y})\,\frac{180}{\pi}}\right) = b^{x}\left(\overline{y\ln b\,\frac{180}{\pi}}\right) = b^{x}\left(\overline{y\ln b\,\frac{180}{\pi}}\right).$$

So as we see, in the exponentiation with an $x + iy$ complex number, the x real part determines the length b^{x}, while the y imaginary part determines the angle $y \ln b\frac{180}{\pi}$: