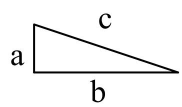
Part One: "Sum Square" Theorems

R

This first part is a beautiful combination of algebra and geometry, without bogging down into the nuances, rather getting results fast and visually.
The more detailed, separate investigations of the two fields will be in Part Two and Three.

R

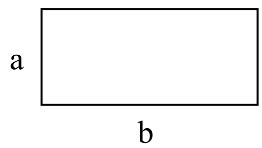
For two numbers a and b, the most important ones that follow from them, come through the three basic operations: addition, subtraction and multiplication. In other words, we can look at their sum difference and product: $a + b$, $b - a$, $a \cdot b$.
The most important geometrically constructed distance from a and b is achieved by placing these two in right angle, that is in 90^o and then regarding the third c side in this right angle triangle:



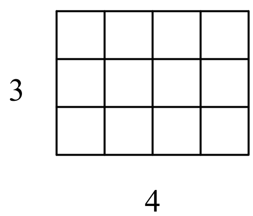
The followings will examine how the a , b values and these four derived values, that is the sum, difference, product and c third side are relating to each other.
The obtained formulas are the basics both in algebra and geometry. Interestingly, the tricks to derive all these for positive a , b and also to easily visualize and thus, remember them, can be seen from the different possibilities of looking at a square with side $a + b$.
This explains the title.

D

a and b will be any two given numbers, so that b is the bigger one.
s will denote $a + b$, their sum.
d will denote $b - a$, their difference.
p will denote $a \cdot b$, their product.
If a , b are both positive then they can be regarded as simple distances and in this case, they determine a rectangle with b base and a height side:

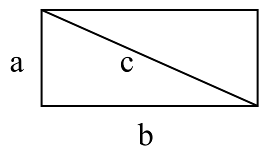


The p product is then the area of the rectangle as an example shows it:



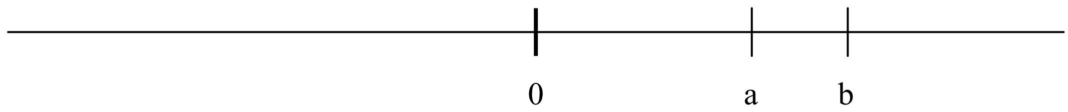
area = $3 \cdot 4 = 12$.

The diagonal of the a , b rectangle is denoted as c.

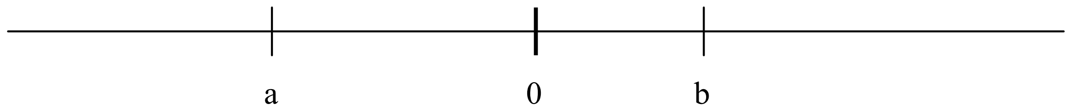


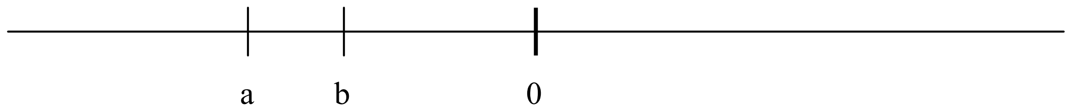
The area of the a, b, c triangle is obviously half of the rectangle and so is: $\dfrac{a \cdot b}{2} = \dfrac{p}{2}$

For the general case of a , b numbers, it’s better to visualize them as points on a number line.
If both are positive, then they are on the positive right side of the line:

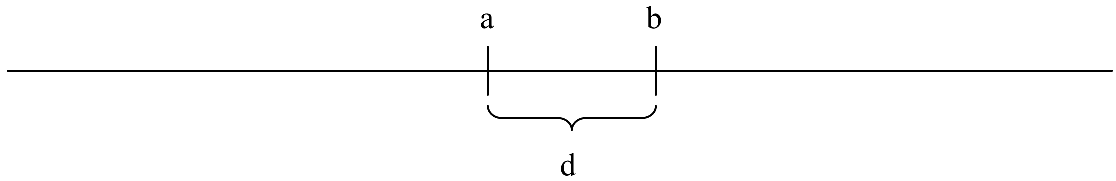


If one or both are negative, then those are on the left side:

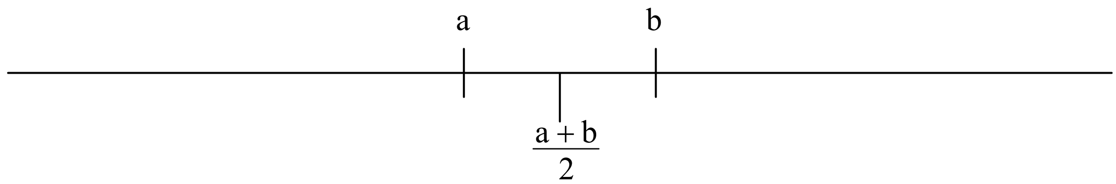




The agreement that b is the bigger one means that b must be always to the right from a.
So if they are both negative as in the last picture, then the length from 0, which is the so called absolute value, must be smaller for b, in spite of b being the bigger as a number!
The d difference is always positive and it is merely the distance between the a , b points, regardless of where they are:



The product and sum for two a , b points can not be as simply visualized as the d difference!
Luckily, the $\frac{a+b}{2} = \frac{s}{2}$ half sum is also very simple because it is the average of the two values a , b and thus, is at their middle points:

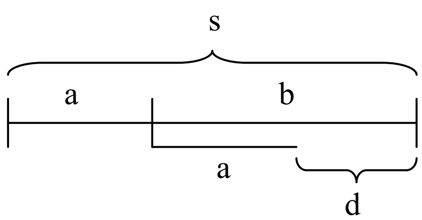


T

How to get a , b back from their sum and difference: $a = \frac{s-d}{2}$, $b = \frac{s+d}{2}$

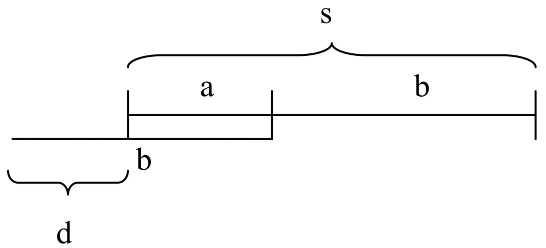
P

For positive a , b they can be regarded as distances.
Measuring a and b next to each other and also a into b shows both s and d:



$\longrightarrow$ $s = 2a + d$ or $2a = s - d$

Instead of measuring a into b, we can measure b over a, and thus we get the other rule:



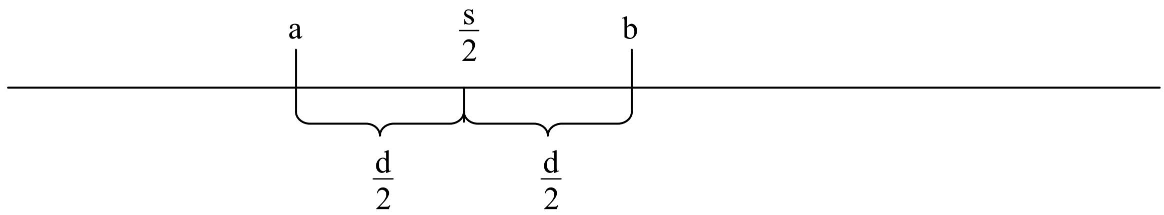
$\longrightarrow \quad 2b = s + d$

For arbitrary numbers a , b we can go algebraically, by simply using the meanings of s and d:

$s + d = a + b + b - a = 2b$

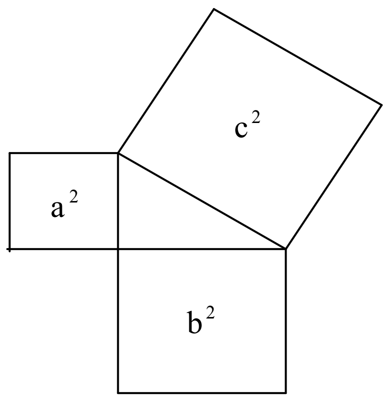
$s - d = a + b - (b - a) = a + b - b + a = 2a$

But we can go visually too by remembering that $\frac{a+b}{2} = \frac{s}{2}$ is the middle point of a , b:

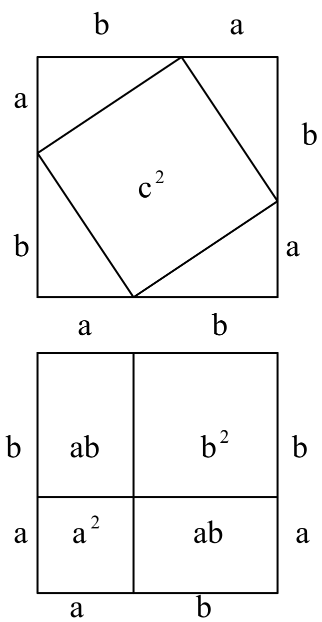


$$a = \frac{s}{2} - \frac{d}{2} = \frac{s-d}{2} \quad , \quad b = \frac{s}{2} + \frac{d}{2} = \frac{s+d}{2}$$

T Pythagoras Theorem $c^2 = a^2 + b^2$



P Lets regard the $a + b = s$ sided square in two different ways:



$$s^2 = c^2 + 4\,\frac{a \cdot b}{2} = c^2 + 2ab.$$

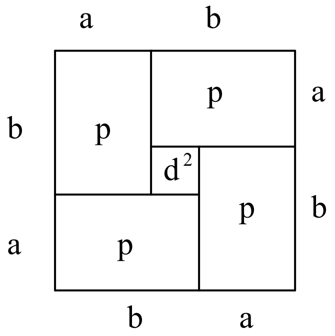
$$s^2 = a^2 + b^2 + 2ab$$

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

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

T How to get the difference from the sum and product: $d = \sqrt{s^2 - 4p}$

P For positive a , b both the proof and the meaning of the formula at once follows from a new third way of looking at the $a + b = s$ sided square:



$s^2 = d^2 + 4p$ or $d^2 = s^2 - 4p$

taking square root of both sides gives d.

For general a , b numbers we can again use the algebraic meanings:

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

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

Thus, $s^2 - 4p = a^2 + 2ab + b^2 - 4ab = a^2 - 2ab + b^2 = d^2$

T How to get back a , b from their sum and product: $a = \dfrac{s - \sqrt{s^2 - 4p}}{2}$, $b = \dfrac{s + \sqrt{s^2 - 4p}}{2}$.

Or in a combined way: $a , b = \dfrac{s \mp \sqrt{s^2 - 4p}}{2}$

P From our first theorem: $a = \dfrac{s - d}{2}$, $b = \dfrac{s + d}{2}$.

From this last theorem, $d = \sqrt{s^2 - 4p}$.

Placing this into the aboves we get a and b with s and p.

T Product of two simple first order expressions of x:

$(x + a)(x + b) = x^2 + xb + ax + ab = x^2 + (a + b)x + ab = x^2 + sx + p$

R “First order” means that x is only with power 1. But $x^1 = x$.

“Simple” means that x is not even multiplied by a number except 1. But $1 \cdot x = x$.

As we see, the product of the two simple first order expressions of x lead to a simple second order, because the x^2 was only multiplied by 1. The multiplier of x became s and the number member became p.

The obvious question is that if we start with a simple second order expression of x, can we always replace it with a product of two first order ones?

Since the members of a product are also called factors, this problem can also be stated as whether a simple second order expression can always be factorized.

The answer is no!

Seemingly, the problem is very simple because only the $s = a + b$ and $p = ab$ appear in our above formula for the result product and earlier we just found a formula how to get a and b from s and p. So we might think that:

$$x^2 + sx + p = \left(x + \frac{s - \sqrt{s^2 - 4p}}{2}\right)\left(x + \frac{s + \sqrt{s^2 - 4p}}{2}\right)$$

But this is only true if s is a sum and p is the product of two numbers.
Unfortunately, we can start with s , p values that simply can't be the sums and products of the same two numbers. Our formula even tells when this happens! Indeed, we only run into trouble of calculating a and b if the square root is not calculable. That of course only happens if the number under the square root, namely $s^2 - 4p$ is negative. For example:

$$x^2 + 3x + 5 = \left(x + \frac{3 - \sqrt{9 - 20}}{2}\right)\left(x + \frac{3 + \sqrt{9 - 20}}{2}\right) \text{ is useless because}$$

$\sqrt{9 - 20} = \sqrt{-16}$ is meaningless. There is no number whose square is -16.
In those cases, where the square root is calculable, the formula gives the factorization:

$$x^2 + 3x - 5 = \left(x + \frac{3 - \sqrt{9 + 20}}{2}\right)\left(x + \frac{3 + \sqrt{9 + 20}}{2}\right) = \left(x + \frac{3 - \sqrt{29}}{2}\right)\left(x + \frac{3 + \sqrt{29}}{2}\right)$$

T Factorizing a general second order expression of x:
First we take out the multiplier of x^2 for the whole expression.

Example: $6x^2 + x - 2 = 6\left[x^2 + \frac{1}{6}x - \frac{2}{6}\right] = 6\left[x^2 + \frac{1}{6}x - \frac{1}{3}\right]$

Then, $\frac{1}{6} = s$ and $-\frac{1}{3} = p$. So, a , b can be obtained if these s , p values are suitable and the square root will be calculable. a , b =

$$\frac{s \mp \sqrt{s^2 - 4p}}{2} = \frac{\frac{1}{6} \mp \sqrt{\frac{1}{36} + \frac{4}{3}}}{2} = \frac{\frac{1}{6} \mp \sqrt{\frac{1}{36} + \frac{48}{36}}}{2} = \frac{\frac{1}{6} \mp \sqrt{\frac{49}{36}}}{2} = \frac{\frac{1}{6} \mp \frac{7}{6}}{2} = \begin{cases} -\frac{1}{2} \\ \frac{8}{12} = \frac{2}{3} \end{cases}$$

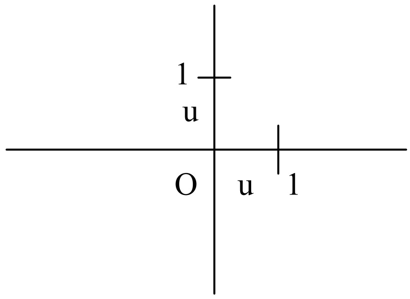
Thus, $6x^2 + x - 2 = 6\,(x - \frac{1}{2})\,(x + \frac{2}{3})$.

T Solution of a second order equation:
A second order equation is merely a second order expression required to be equal to 0.
A product can only be 0 if some of its factors are 0. Thus, the factorization of the second order expression at once gives the solutions. For example,
$6x^2 + x - 2 = 6\,(x - \frac{1}{2})\,(x + \frac{2}{3}) = 0$, can only be if $x - \frac{1}{2} = 0$ or $x + \frac{2}{3} = 0$.
So the solutions or so called roots are: $x_1 = \frac{1}{2}$ and $x_2 = -\frac{2}{3}$.
As we see, the roots are simply the opposite signed values of a and b.
Thus, we can even give a general formula for the roots from the a , b formula by simply changing the signs:

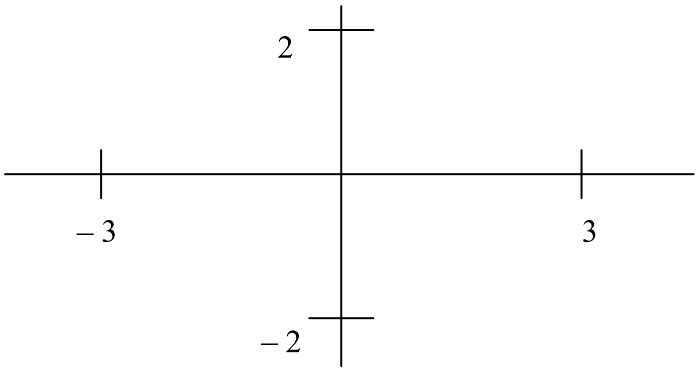
$$x_{1,2} = \frac{-s \pm \sqrt{s^2 - 4p}}{2}$$

Part Two: Coordinates and Numbers

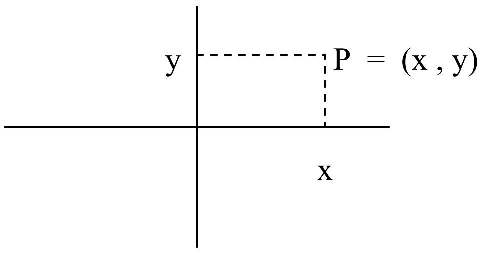
D 1.) The Descartes coordinate system is two perpendicular number lines with common units. Usually one of the lines is placed horizontally, which is called the x coordinate line and the other y is vertical:



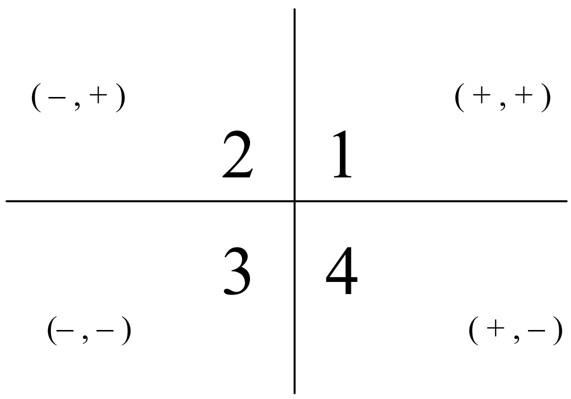
Also, the unit on x is placed on the right and on y above from the O crossing point or origin. Thus, the positive values are on the right on x and on the top on y:



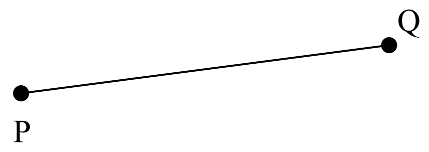
Any P point in the plane now can be identified by its coordinates if we draw perpendiculars to the coordinate lines from P. Thus, P = (x , y):



As we see, in the first quarter or "quadrant", both x and y are positive. On the next, x is negative, then both are negative, finally x positive and y negative:

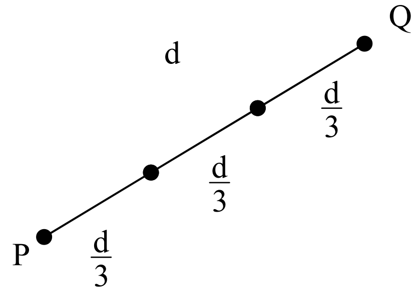


2.) If P and Q are two points in space, (or a plane), then the PQ segment or interval is the line going through P and Q, but only regarded between P and Q:

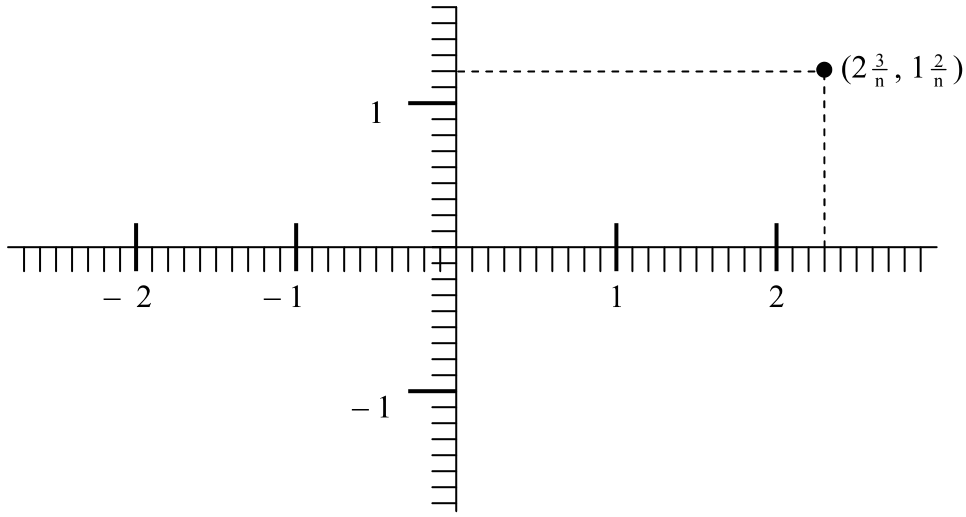


If the P , Q end points are included, we denote the segment as [P , Q]. But if we want to leave them out, we write (P , Q). If we want only one end to be included, we can even use [,) or (,] brackets for such half "closed" and half "opened" segments.

The n-dividers of a segment are the points that are equally distanced, namely $\frac{d}{n}$ away from each other. Where d is the original full distance.



As we see, there are two 3-dividers and in general there are $n - 1$ many n-dividers.
The 2-divider is simply the middle point of P and Q.
If we place the n-dividers between every consecutive whole numbers on the x and y coordinate lines, then we get all fraction coordinates with n denominators:



3.) If we place all possible dividers between the P , Q points then this leads to infinite many points. Some different dividers can be identical points. For example, the third 4-divider is again the 2-divider middle point. To prove that in the total, we still have infinite many dividers, is easy by realizing the more general fact, that between any two dividers, there are other dividers. This follows from the even more general law that:

If D_1 and D_2 are dividers of P and Q, then any D divider of D_1 and D_2 is already a divider of P and Q.

Indeed, if D_1 was an m-divider, D_2 an n-divider and D is a k-divider of D_1 and D_2, then D is also an m n k divider of P and Q. This follows from the simple fact, that an n-divider is also an (m n)-divider for any n.

This amazing "density" of the dividers might even suggest that the dividers cover the whole segment. The question of doubting this is the following:

Is there a point between two P , Q that is not a divider of them?
As we saw on the number lines, the dividers are simply the fractions, so the same question then is whether there are numbers that are not fractions. The fractions are also called rational numbers from the word ratio, meaning division and then in this case, division of two whole numbers. So, an other way to ask the question is: Are there non rational or “irrational” numbers?
By renaming the problem of course, we don’t achieve anything, so the original non divider point is the heart of the matter.
The real reason of these renamings of the original problem and the real reason of overlooking it today, is that the usage of numbers instead of actual distances became combined with the infinite decimal system. On the calculators, we only see a certain amount of digits, but they still suggest the feeling that for example, between 0 and 1, every number is an infinite decimal in the form of:

$$0 . d_1 \; d_2 \; d_3 \; . . .$$

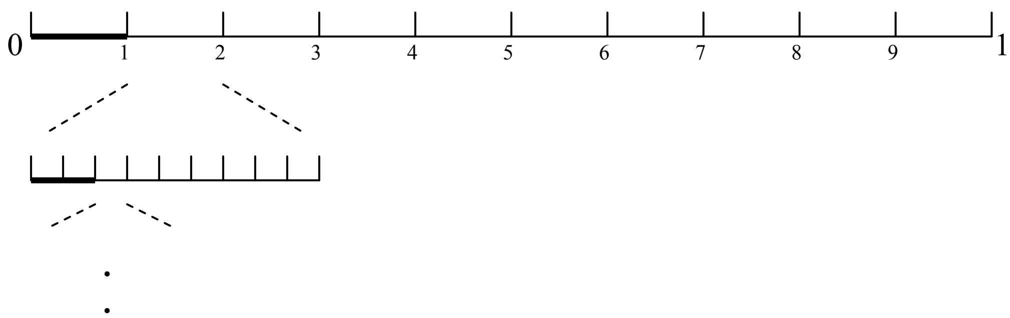
A fraction when put into this infinite decimal form, is either “finite”, meaning that it has all zeros after a point, or it has a repeating group like 85 in:

$$0 . 2\;3\;0\;1\;85\;85\;85 \; . . .$$

This can be proved easily if we actually do the division of a fraction, $\frac{m}{n}$ digit by digit.
Indeed, the remainders must repeat, because they can all come from $1 , 2 , . . . , n - 1$.
So then, the fractions obviously can’t be all the numbers, because we can easily create an infinite decimal that has no repetition. For example, we can just use all numbers in increasing order:

$$0 . 1\;2\;3\;4\;5\;6\;7\;8\;9\;1\;0\;1\;1\;1\;2\;1\;3\;1\;4\;1\;5 \; . . .$$

The real missing point today, of course lies behind this naïve belief, that every infinite decimal is indeed a distance on the number line. Nobody actually checks where this above number lies.
Indeed, it would be very hard to locate it, but at least in general we should look and see that every time we continue the decimal with a new digit, then we actually make new 10-dividers of the last smallest divisions. For example, at $0 . 1\;2\;3 \; . . .$ first we cut $[0 , 1]$ in ten and go up to the first divider. Then we cut the next again in ten, and go up to 2, then 3 and so on:



After the total infinite decimal, we had to go infinite many forward moves, but always smaller and smaller ones. Thus, the total can and in our case indeed is just a finite length. The reached point can still be just a simple divider, but now expressed in 10-dividers infinitely, or it can be a point that is not a divider at all. So the existence of the infinite many moves forward as total, combined with the law of dividers or fractions being repetitive moves, yields the seemingly natural or plausible feeling that there must be irrationals. In fact, it even suggests that irrationals are more than rationals. Indeed, for an infinite decimal to be repetitive is much less likely than being without repetition or maybe even being without any rules or to be random. This also shows that our modern day learning by getting visions on a silver plate is not faulty, in fact it can even raise deeper problems, like the randomness above.

The real problem is that the visions obtained on silver plates are not broken down to their original meanings. So then, in spite of getting the right plausibilities about numbers, when one is asked whether the dividers can cover a segment or not, he is clueless. He can't see it being the same problem, that he thought to understand with numbers.
So, we'll proceed without the usage of the infinite decimal system and show the deeper details with strictly distances. The most important feature of our investigations is what already shows from the earlier parts of these definitions, namely that to see how the numbers, that is points on a line behave, one has to go out into the plane. Amazingly, this goes even for quite simple facts about the fractions themselves.

4.)
A A P point in the Descartes plane is
a.) Grid point, if both coordinates are whole numbers
b.) Rational, if both coordinates are fractions
c.) Irrational, if none of the two coordinates is fraction.
d.) Mixed, if one coordinate is fraction, the other is not.

B A half line going from the O origin is a "direction", namely:
a.) Grid direction, if it contains grid point besides O
Gridless direction, if it contains no grid point except O.
b.) Rational, if it contains rational point besides O.
Rationalless, if it contains no rational point except O.
c.) Irrational, if it contains irrational point.
Irrationalless, if it contains no irrational point.
d.) Mixed, if it contains mixed point.
Totally mixed if it contains only mixed points except O.
Mixless if it contains no mixed points.

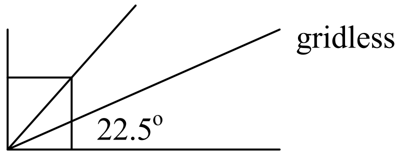
Totally grid, rational, or irrational direction is clearly impossible because we can pick one of the coordinates freely and as we'll see in the next theorem, there are irrationals.

T 1.) Fractions
a.) If $P = (x , y)$ is a grid point, then $2P = (2x , 2y)$, $3P = (3x , 3y)$, . . . are also grid points on the same direction.
b.) If P , Q are grid points on a same direction, P being closer to O, and $Q \neq mP$, then there is other grid point on the direction, even closer to O than P.
c.) If C is the closest grid point on a direction to O, then all the grid points are: $C , 2C , 3C$, . . .
d.) If two fractions $\frac{a}{b}$ and $\frac{A}{B}$ are equal in value, but $\frac{A}{B}$ having bigger numerator and denominator then either $A = m\,a$ and $B = m\,b$ or there is an even "smaller" $\frac{a_o}{b_o}$ version of the same value, so that $a = n\,a_o$, $b = n\,b_o$, $A = m\,a_o$, $B = m\,b_o$.
In short, all equal fractions are merely the expansions of a simplified smallest version.

2.) Irrationals
a.) All dividers of an $[A , B]$ interval can be listed as a sequence.
b.) All fractions can be listed as a sequence .
c.) A sequence of P_1 , P_2 , P_3 , . . . points can not cover any $[A , B]$ interval.
d.) There are non divider points on any $[A , B]$ interval.
There are irrationals between any two numbers.

3.) Directions
a.) A rational direction is grid direction too. A gridless direction is rationalless.
b.) The coordinate lines are rational and mixed, but irrationalless directions.
c.) Any non coordinate, rationalless direction is mixless. (Thus, irrational too.)
Any non coordinate, mixed direction is rationalless.
So by picking any mixed point, we get a gridless direction too.

d.) Any non coordinate direction between any of its two A , B points has irrational point.
This at once shows that there are no totally mixed directions, but more importantly, how much more irrational points are than rational. Each of the three sets: the rational, the irrational and the mixed points are dense in the plane, but they are very different!
The rational and the mixed points can be avoided completely with full directions by simply picking one rational or mixed point to be on the direction. But, the irrationals are "so dense", that they must appear on every non coordinate direction between any two points.



4.) Concrete irrationals

a.) The $\frac{45^{\circ}}{2} = 22.5^{\circ}$ direction is gridless.

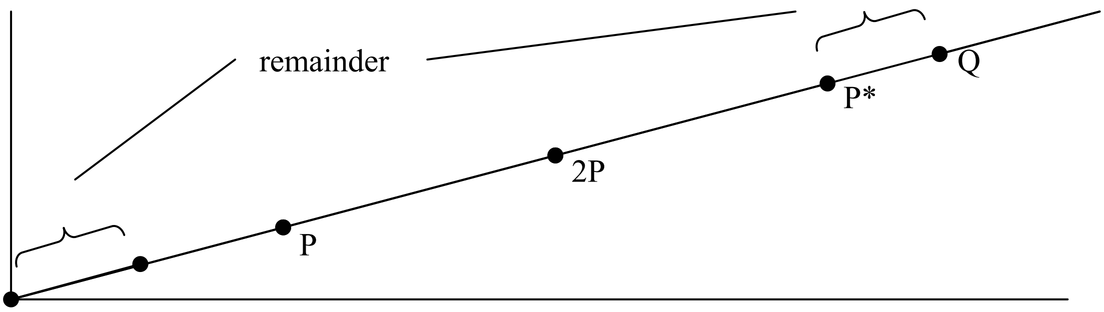
b.) If x is rational, and (x , y) is on the 22.5° direction, then y is irrational.

c.) If (x , x) is any rational point on the 45° direction, then its d distance from O is irrational.

d.) A d distance is the distance of the (x , x) point from O, if and only if, $d^2 = 2\ x^2$.
So by c.) this is impossible for x and d, both rational.
In particular, $d^2 = 2\ x^2$ is impossible for whole numbers either.

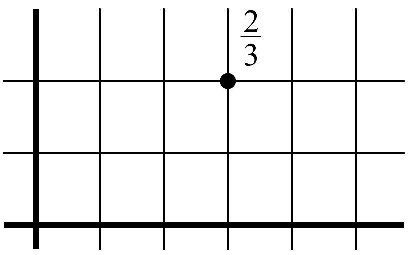
P 1.) a.) Trivial

b.) Measure OP repeatedly as many times as possible before passing Q. If this last point is P*, then shifting back the P* Q remainder distance to O, we get a closer grid point than P:



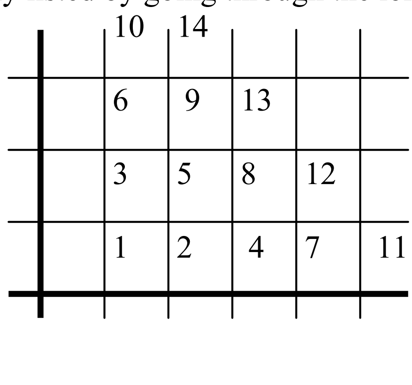
c.) Trivial by b.)

d.) Regard the $\frac{y}{x}$ fraction, at the (x , y) grid point, then follows from a.) b.) c.)



2.) a.) Start with the one 2-divider. Then the two 3-dividers, then the 4-dividers and so on.
If an earlier divider is repeated (like the second 4-divider), we can simply skip it.

b.) Assign the positive fractions to the grids in the first quadrant as in 1.) d.)
These can be easily listed by going through the longer and longer finite diagonals:



This is the same as listing the fractions by the increasing total of their numerators and denominators:

$\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{2}{2}, \frac{3}{1}, \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1}, \frac{1}{5}, \frac{2}{4}, \frac{3}{3}, \frac{4}{2}, \frac{5}{1}, \ldots$

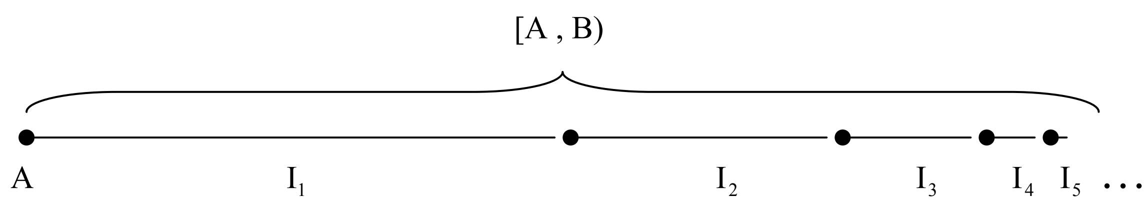
To take care of the negative fractions as well, we can put them in an other quadrant and go through the diagonals alternatingly. In fact, all four quadrants' grids including the coordinates can be listed by a simple spiraling order, starting from the origin:

17	16	15	14	13
18	5	4	3	12
19	6	1	2	11
20	7	8	9	10

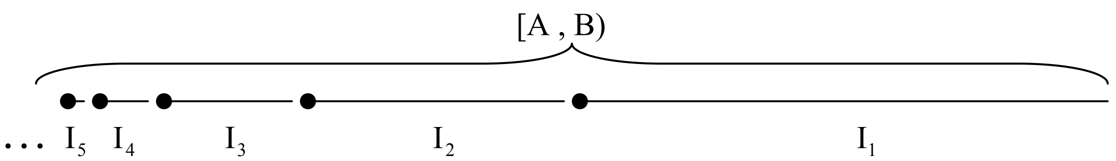
c.) In order to prove this we need a vital new axiom about how points are located on a line. Before introducing this, I will create an illusion, a straight forward argument, that seems to contradict either what we proved in a.) , b.) or some very natural new assumptions. So let in [A , B], P_1 = halving point , P_2 = first 3-divider , P_3 = second 3-divider , P_4 = first 4-divider , P_5 = third 4-divider , P_6 = first 5-divider , and so on.

Now let I be any (arbitrary small) interval in [A , B]! Lets cut a piece, say half off from I, and call it I_1. This can be placed over P_1, for example, so that P_1 becomes the middle point of it. Then lets cut off again the half of the leftover of I, call it I_2 and place it over P_2 ! And so on, we have cut I into I_1 , I_2 , . . . pieces and place them over all the dividers, so that they cover all of them. But the dividers are dense in [A , B], so these covering intervals must overlap and thus cover the full [A , B]. I can even go further and increase the illusion by showing the following concrete example of how points can disappear or appear by merely shifting intervals:

Lets regard [A , B) and cover it with repeatedly halved left closed, right open intervals:



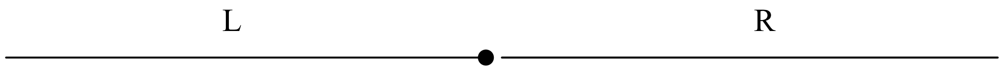
Observe that all missing right ends are covered by the next left end, so indeed the full [A,B) is covered. Now lets rearrange the I_n intervals to reverse order, that is, I_1 over the second half of [A , B), then I_2 before it, and so on:



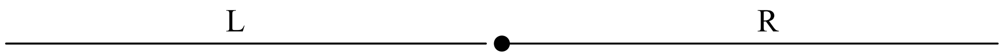
Seemingly everything is the same. But no! "A" has disappeared! Yes, it is not covered anymore. Similarly, we can lose or gain more, even infinite many points. But as I said, all this was just an other "special effect" to make us believe that a small I part of [A , B] can cut to cover the whole [A , B]. So what's happening, what is true and what is not?

The loss of points is true. But we can't increase the total lengths! So, a smaller I can't cover [A , B]. We'll show this soon. But then our doing exactly that with covering the dividers was faulty too! Yes, so then the dividers are not sequencable? Yes they are. The fault was where we didn't expect, namely by simply saying that the little pieces overlap. Some surely will overlap, but they don't form a properly overlapping cover system. There will be holes between them, in fact lots of holes. That's very hard to see. Now, that I revealed the error in our intuitions about "density", we might shout, eureka, because then we can go quite in reverse to prove c.). Indeed, any P_1 , P_2 , P_3 , . . . can be covered by pieces of an I. But if they can't cover [A , B] then the P_1 , P_2 , P_3 , . . . sequence can't do that either, to start with, otherwise the pieces were obviously doing the cover too. Maybe this uncoverability of [A , B] by a smaller I is actually the new axiom I promised. In light of the lost and gained points, this would be a very bad axiom, because only the shorter length makes it true, but our intuitions don't regard that factor at all. We have to find something much simpler and indeed, Dedekind found a simplest possible assumption we need. In fact, we already used this tacidly earlier when covered [A , B) with half opened intervals. This axiom of Dedekind could be called the Line Cutting axiom and it says that if a line is cut in two halves, then either the left has a right end point or the right, a left end one:

either:



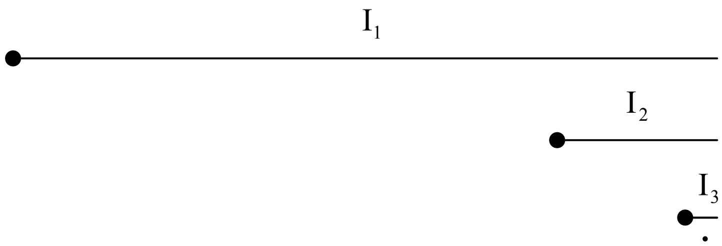
or:



A consequence of this axiom for an S set of points is the following:
If S doesn't go to infinity towards the right, that is there are points to its right, then either S has a right most point or the ones to the right have a left most. Indeed, all we have to do is regard the points that are completely right from S as an R set and the ones that are not as L. Then if R has no left end, then L has right end and it must belong to S, otherwise it would be in R. Similarly for the left case.
These points that are either ends of S or ends of the left or right points from S are called the infinum and supremum of S, in symbol, inf S and sup S. These are the generalizations of the everyday notions of minimum and maximum. If an S set doesn't reach towards infinity in the left or right, it doesn't necessarily mean that it has minimum or maximum, because the set can have smaller and smaller or bigger and bigger points. The infinum and supremum in these cases don't belong to S, rather to the points that are left or right from S.
From this result, one follows for intervals, which is also called Cantor's Common Point theorem:
If $[P_1 Q_1]$, $[P_2 Q_2]$, . . . are closed intervals placed in each other, then there has to be point that is in all of the intervals. In fact, these "common" points are on a [P , Q] interval unless P = Q is the single common point.
The order of points is obviously:

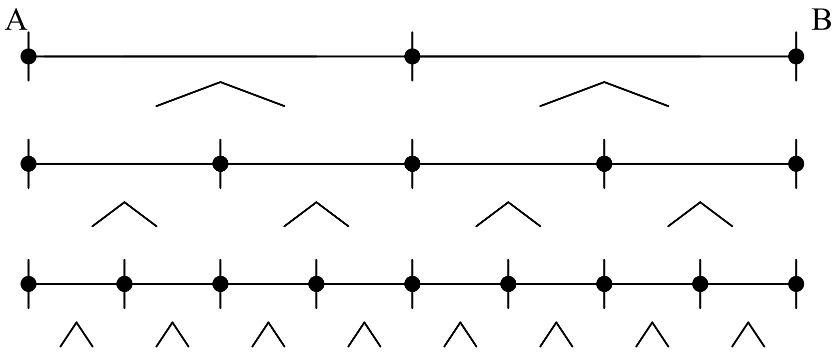
$$P_1 \le P_2 \le P_3 \le . . . \quad . . . \le Q_3 \le Q_2 \le Q_1$$

From this clearly $P = \sup \{P_1 , P_2 , . . .\}$ and $Q = \inf \{Q_1 , Q_2 , . . .\}$
Indeed, this P , Q and the points between them are the only ones (or one if P = Q) that are right to all P_n and left to all Q_n
Our proof clearly relied on the P_n , Q_n end points of the intervals, and indeed, with even just half closed intervals, placed in each other, we can easily narrow down to the missing end point and thus, have no common point:



no common point

Now a good test of the Common Point theorem will be to use it to prove the uncoverabilty of [A , B] by a smaller I. First of all, lets observe that if we only cut I into finite many pieces, then the uncoverability of [A, B] with those, follows from the addition of intervals through geometry. It's only when we have infinite many pieces, that the total length is not a simple sum. In higher math, we learn that a sum of infinite many numbers can't be changed by rearranging the members, if they are all positive, but that result is based on limits, which use the very theorems we introduce right now.
The jump to infinite many cuts, would also follow if we could show that, if in an infinite cut and arrangement, no finite many can cover, then the total neither.
This is true, but only for all open interval members:
If a collection and position of open intervals is such that no finite many of them covers [A,B] then their total is not covering [A , B] either.
To show this, lets cut [A , B] in half repeatedly! In other words, lets regard the 2 , 4 , 8 , . . 2^n–dividers:

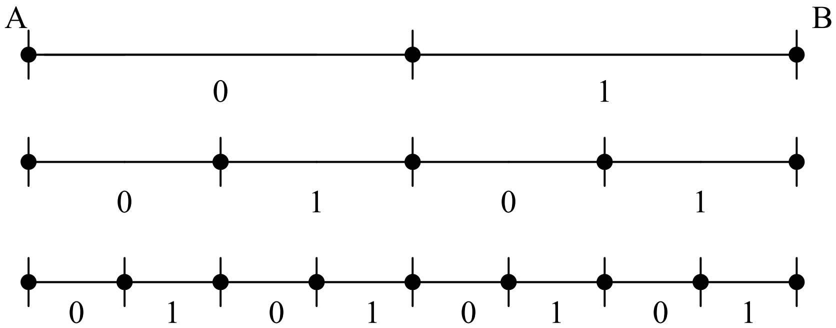


By the Common Point theorem, every narrowing sequence of halves will determine a point of [A,B]. In other words, every path in our "tree" of branchings, will determine an end point. If an end point is in an open interval, then it must be inside of it. So then not only the point itself, but a small enough full interval in our above list containing it, must be in the open interval too. Then of course, all the halvings of that will stay in the open interval, so in other words, the continuing tree remains inside. If we cut off these continuations, then our tree will be having lots of dead ends, namely those that would have lead to points in the open intervals. The remaining tree is still infinite long! Indeed, if there were a stage in the halvings, where every branch is cut off, then regarding the collection of all those open intervals that contained the last branches, would be a finite set covering the full [A,B]. Does this infinity of our remaining tree mean that there is an actual infinite path in it? If yes, then we are finished, because such path determines a point and it can't be in any open interval, so the open intervals indeed, don't cover the full [A , B].
To prove the existence of an infinite path, we would have to tell how to make the left or right choices. We can't give an effective way to choose, but we <u>can</u> give a theoretically correct choice, which guarantees infinite path. Lets choose always such branch, that is infinite in its continuation! For example, at the start, it can't be that both branches lead to finite continuing tree, because then the total tree were finite too, namely the longer of the two branches plus one. So at least one of the branchings is infinite and choose one such. Then again, at the two branchings of that, at least one half leads to infinite tree, so choose that. And so on, we never get into dead end and thus, achieve an infinite path.

Now to show that a smaller interval can't cover a bigger in general, is fairly easy. Indeed, if we use closed or half closed pieces of an I, and we want to show that even (A , B) is not coverable, then all we have to do is get a smaller I_0 in (A , B) outside I, which is always possible if I is less than (A , B). Then, cut from I_0 infinite many open pieces and cover A , B and the end points of the I pieces with these! Then a cover of (A , B) from I would also mean a cover of [A,B] and with open pieces of $I + I_0$, contradicting our result.

This hair splitting was not necessary to show that a sequence P_1 , P_2 , P_3 , . . . can't cover [A , B], because we could use all open intervals to cover the P-s, so the theorem for open intervals is enough. Even so, this proof is too long, and it was merely an exercise to introduce the Common Point theorem. Now, we show how this theorem proves at once that P_1 , P_2 , P_3 , . . . can't cover [A , B]. Indeed, let's go in the P_1 , P_2 , P_3 , . . . sequence until we encounter two points that are in (A , B). If there weren't such, or only one, then they obviously couldn't cover the full [A , B]. So suppose these are P_i and P_j. Then, [P_i , P_j] is a nested interval in [A , B]. Now we can continue and find in the sequence two points that are in (P_i , P_j) . And so on, we'll clearly get a sequence of nested intervals, which therefore must have a common P point. This P can't be any of P_1 , P_2 , P_3 , . . . Indeed, all points became eventually outside of some newly chosen inner interval.

The halvings used in our over complicated proof show that the Common Point theorem can be generalized, not just to guarantee the existence of points, but to actually identify every point of [A , B]. In fact, instead of halvings, we can use any repeated n-dividers. For example, the repeated 10-dividers give our usual decimal system. Still, the halvings are the simplest and if the first halves are denoted as 0 and the second as 1, we get the binary "decimal" forms of points.



Every infinite 0 , 1 sequence will determine a point of [A , B].
We had to use closed intervals, because the Common Point theorem requires this!
Unfortunately, this means that the halves have common end points and thus, different left right sequences, can locate same points. Indeed, for example:
0 1 1 1 1 1 1 1 . . . and 1 0 0 0 0 0 0 0 . . . both locate the middle point of [A , B].
The same discrepancy appears in the decimal system, when we use infinite many 9-s.
For example 0 . 1 2 9 9 9 9 9 . . . is the same as 0 . 1 3 0 0 0 0 0 . . . = 0 . 1 3
It's easy to see that ambiguity can only be for these sequences containing identical digits from a point and they all locate the dividers, in our case the halving points.
The most famous proof of that a P_1 , P_2 , P_3 , . . . can't cover [A , B] is using this generalized theorem of Cantor, in other words, the binary forms of the points in [A , B].
Unfortunately, usually these proofs are incomplete, because they only show that the binary forms can not be sequenced. The error comes from the fact that as we saw above, the binary forms are not unique for all points of [A , B]. So what we really have to show is that the set of all possible 0 , 1 sequences, except the ones having all 1-s from a point are not sequencable. Indeed, if we avoid the ambiguous forms by always choosing the zero ending versions, then every point of [A , B] has a unique 0 , 1 sequence, and all sequences appear except the ones with 1 from a point.

These sequences that are all 1 from a point, of course locate the halving points. They are part of all dividers, so a sequence by a.). To be specific:
$s_1 = 0\,1\,1\,1\,.\,.\,.$ = the middle point, $s_2 = 0\,0\,1\,1\,1\,.\,.\,.$ = the first quarter ,
$s_3 = 1\,0\,1\,1\,1\,.\,.\,.$ = the third quarter, $s_4 = 0\,0\,1\,1\,1\,.\,.\,.$ = the first 8-divider,
$s_5 = 0\,1\,0\,1\,1\,1\,.\,.\,.$ = the third 8-divider, $s_6 = 1\,0\,0\,1\,1\,1\,.\,.\,.$ = the fifth 8-divider,
$s_7 = 1\,1\,0\,1\,1\,1\,.\,.\,.$ = the seventh 8-divider. Then come the 16-dividers and so on.
Now if the 0 , 1 sequences that are not 1 from a point could also be sequenced as an other $S_1, S_2, S_3, \dots$ sequence, then: $s_1, S_1, s_2, S_2, s_3, S_3, \dots$ would be a sequencing of all possible 0 , 1 sequences. Or in reverse, if all possible 0 , 1 sequences are not sequencable, then the ones that are not 1 from a point are not sequencable either.
So now it's indeed enough to show that the total set of all possible 0 , 1 sequences is not sequencable! Suppose there were such sequencing as:

First =	0	1	0	1	1	1	0	1	0	0	1
Second =	0	0	1	1	0	0	1	1	1	1	0
Third =	1	1	1	0	0	1	0	0	1	0	1
Fourth =	0	1	0	1	1	1	0	1	0	0	1

.
.
Lets list the diagonal sequence: D = 0 0 1 1 . . .
This can very well be one of the sequences on our list, but now lets turn every digit to its opposite! This so called anti-diagonal sequence:
$\overline{D}$ = 1 1 0 0 . . . can not be on our list! Indeed, every digit of it is differing from the ones on the list, in at least one digit, namely at the diagonal.

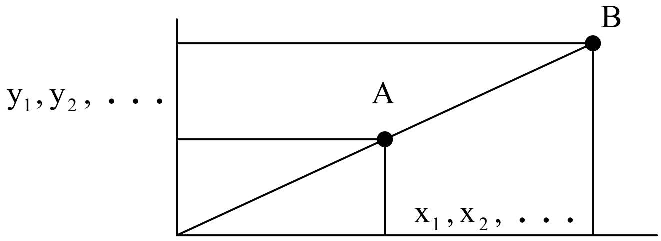
d.) The first follows at once from a.) and c.). Indeed the dividers are merely a sequence, while the total is not. The second similarly from b.) and c.).
A more specific way would be to find actual non divider point and it is done in 4.) of this theorem, and the later theorems.
An in between way would be not to give single actual examples, but still show somehow effectively that there are non divider points. This is done by realizing the fact, we already mentioned in D 3.) that the dividers must have periodic decimal or in general any fixed n-divider sequence forms. Thus, all non periodic sequences are non dividers. This also shows how many more non dividers are than dividers.

3.) a.) If $(\frac{m}{n}, \frac{p}{q})$ is on the direction, then $n\,q\,(\frac{m}{n}, \frac{p}{q}) = (m\,q, n\,p)$ is on too.

b.) Trivial by the existence of irrationals.

c.) If the rational (a, b) point is on the direction, then for any other (x, y) on it: $\frac{y}{x} = \frac{b}{a}$
Thus, $y = x\,\frac{b}{a}$ is rational or irrational exactly when x is so.

d.) List the rational x coordinates between the A , B points as $x_1, x_2, \dots$
and the rational y coordinates as $y_1, y_2, \dots$

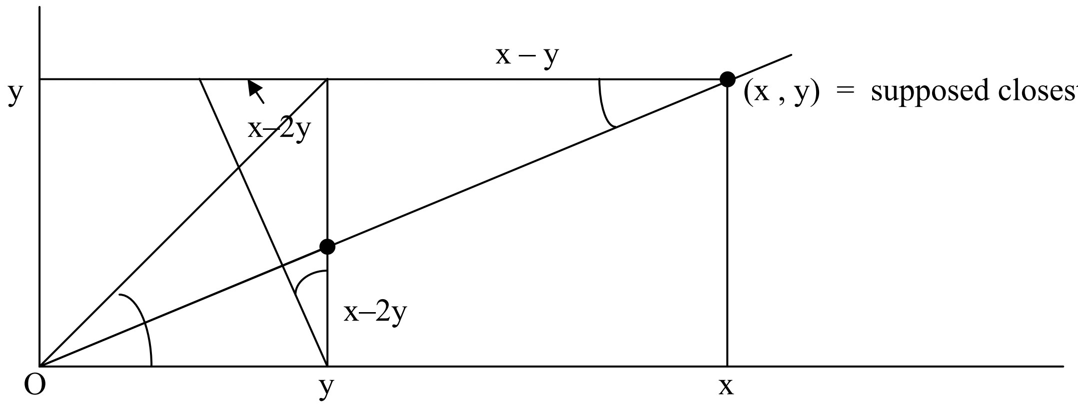


Lets regard the following ordering of points between A , B: $x_1, y_1, x_2, y_2, \dots$
$P_1, P_2, P_3, P_4, \dots$

R

Here P_n is the point between A and B that has the coordinate above it.
The "the" was correct, because there is only one P for every x_i or y_i.
Rational points of course will appear twice under their x and y coordinates as well.
But most importantly, if there were no irrational points between A , B then every point would appear and thus were listed, contradicting 2.) c.) .

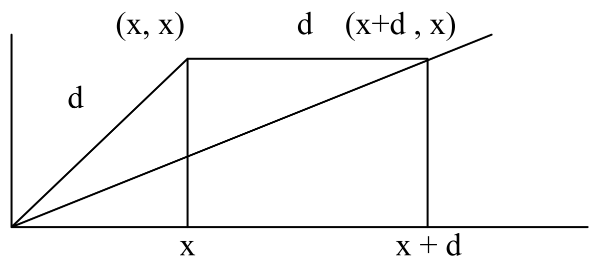
4.) a.)



The diagonal of the square is $x - y$ and the reflected y, subtracted yields the $x - 2y$ value. This in the equal sided triangle, gives the other same side on the top, and then this proves by the equal angles, that the value at y must be also $x - 2y$. This of course, contradicts that (x , y) was the closest grid to O.

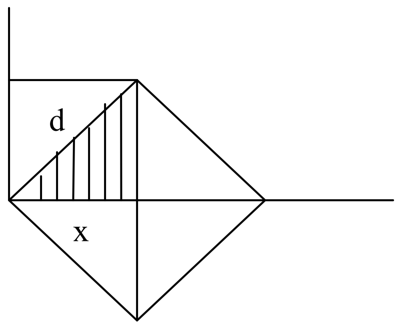
b.) If $(\frac{m}{n}, \frac{p}{q})$ were on the 22.5° direction, then $n\,q\,(\frac{m}{n}, \frac{p}{q}) = (m\,q, n\,p)$ were a grid on it too, contradicting a.).

c.)



If d were rational, then the (x+d , x) rational point were on the 22.5° direction too, contradicting b.).

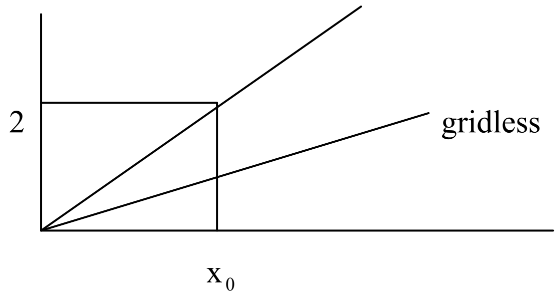
d.)



x^2 is twice of the shaded triangle, while d^2 is four times.
Thus, d^2 is twice of x^2.

R The halving of the 45° to get gridless direction can be generalized to any grid direction that contains grid with $y = 2$ value:

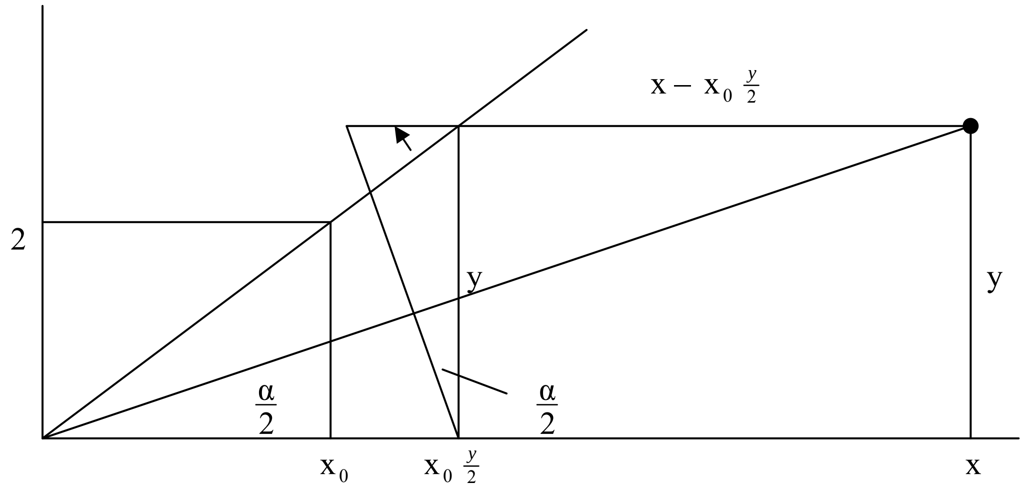
T If a $(x_0 , 2)$ grid point's direction's angle is α, then the $\frac{\alpha}{2}$ direction is gridless:



P Suppose (x , y) were the closest grid to O on the $\frac{\alpha}{2}$ direction.

First of all, $x > x_0$ because before x_0 or at x_0 the $\frac{\alpha}{2}$ direction can't reach up to 1.

Secondly, $x > y$ because $\frac{\alpha}{2} < 45^o$. Thus, the situation of (x , y) is as:



From the figure, we can see that actually $x > x_0 y$. Indeed, at x, the continuation of y would go up to the α angle line as $\frac{x}{x_0} 2$ height and y is smaller than half of this.

Now we'll show that $(y , x - x_0 y)$ would be a closer grid point on the $\frac{\alpha}{2}$ angle line than (x , y).

As we said, $y < x$, so all we have to show is that $x - x_0 y$ is on the $\frac{\alpha}{2}$ angle line.

The little side at the bottom of the arrow is, $x - x_0 \frac{y}{2} - x_0 \frac{y}{2} = x - x_0 y$.

Then, the other side at the arrow is the same. Turning the y-sided and $\frac{\alpha}{2}$ angle triangle onto the x line, we can see that indeed $(y , x - x_0 y)$ is on the $\frac{\alpha}{2}$ direction.

R The impossibility of $d^2 = 2 x^2$ for whole numbers obviously implies the same for rationals, because $(\frac{m}{n})^2 = 2 (\frac{p}{q})^2$ would imply $(m q)^2 = 2 (n p)^2$.

So the irrationality of $(1 , 1)$'s distance from O also follows from $d^2 = 2 x^2$ being impossible for wholes. It's interesting to see direct proofs for this. We'll give one algebraic and one geometric:

T $d^2 = 2 x^2$ is impossible for whole numbers.

P_1

Lets call the h (n) halvability of an n number, how many times it can be halved.
If n is odd, then of course h (n) = 0.

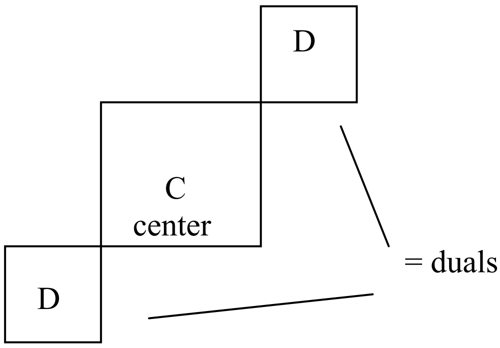
For example, h (8) = 3, because $\frac{8}{2} = 4$, $\frac{4}{2} = 2$, $\frac{2}{2} = 1 =$ odd.

h (100) = 2 because $\frac{100}{2} = 50$, $\frac{50}{2} = 25 =$ odd

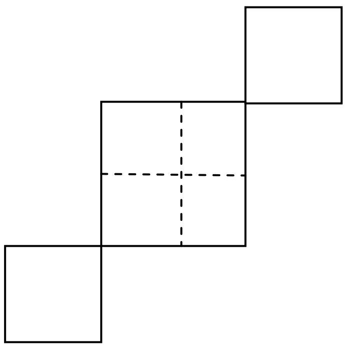
We claim that h (m n) = h (m) + h (n).
Clearly, m n can be halved at least h (m) + h (n) times, by simply halving m, h (m) many times and then n, h (n) many times. But also that's all the halvability, because after these m becomes an odd m_0 and n an odd n_0. Then, m_0 n_0 is odd again, so is not halvable anymore.
To see that odd times odd is odd, observe: (2a + 1) (2b + 1) = 2a (2b + 1) + 2b + 1 = even + 1.
Then, a d^2's halvability is h (d^2) = h (d) + h (d) = 2 h (d) = even. Similarly for x^2.
But, h (2 x^2) = h (2) + h (x^2) = 1 + even = odd.
So, $d^2 = 2\ x^2$ is impossible because the left side has even halvability and the right has odd.

P_2

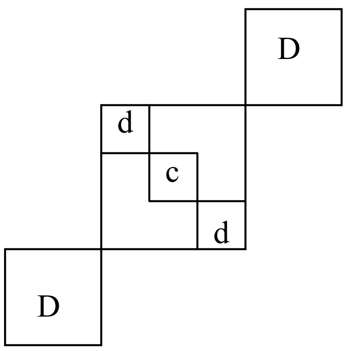
1.) Lets call two identical squares placed at opposite corners of a third square outside continuing the sides, a square triplet.



2.) A D – C – D square triplet is continuable if C > D, but D > $\frac{C}{4}$.



3.) The continuation of a continuable D – C – D triplet is obtained by pushing D-s into C, and then regarding the overlapping c center square and the missing d squares from C:



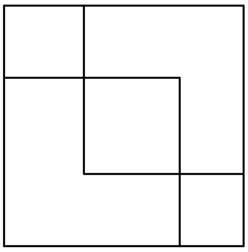
4.) A triplet is infinitely continuable if the continuation is again and again continuable.
5.) Lets regard a fix grid system, and the triplets that fit into this call grid triplets.

The followings are true:

1.) If $D - C - D$ is continuable and its continuation is $d - c - d$, then $C - 2D = 2d - c$.
2.) If $D - C - D$ is such that $C = 2D$, then it is continuable
3.) If $D - C - D$ is continuable grid triplet, then the $d - c - d$ continuation is grid too.
4.) If $D - C - D$ is grid triplet, it can't be infinitely continuable.
5.) If $D - C - D$ is grid triplet, then $C = 2D$ is impossible.

Indeed:

1.)



$C - 2D$ is the two missing corners minus the center, because it is covered twice.

2.) If $C = 2D$, then obviously $C > D$ and also $D = \frac{C}{2} > \frac{C}{4}$.
3.) Trivial
4.) The continuations are always smaller.
5.) Lets start with a $D - C - D$ in a grid system and continue it as far as possible!
By 4.) it can only happen finite many times.
By 3.) they are all in grids.
By 2.) the final non continuable triplet has a $c - 2d \neq 0$.
By 1.) the original $2C - D$ is the same or negative of it.

R In the last second proof, claim 2 is much sharper, namely $C = 2D$, if and only if the triplet is infinitely continuable.