Elementary Geometry II

Tom Elliott

August 17, 2024

Contents

Part I

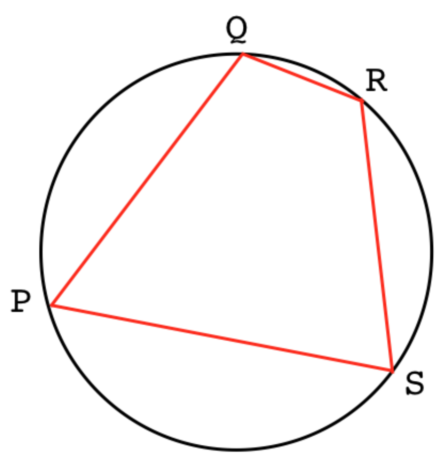
Polygons

Chapter 1

Cyclic quadrilateral

theorem on cyclic quadrilaterals

In this chapter we look at quadrilaterals where the fourth vertex is constrained to lie on the same circumcircle as the other three. We showed this elementary theorem earlier:



- For *any* quadrilateral whose four vertices lie on a circle, the opposing angles are supplementary (they sum to 180°).

Proof.

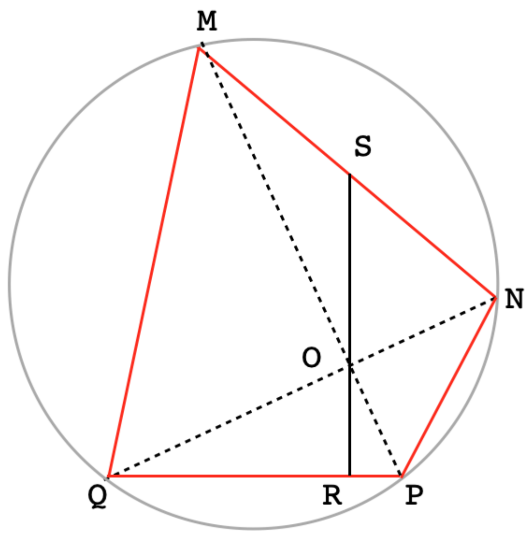
Together, opposing angles exactly sweep out the whole arc of the circle.

□

Brahmagupta's theorem

This is a theorem credited to Brahmagupta.

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

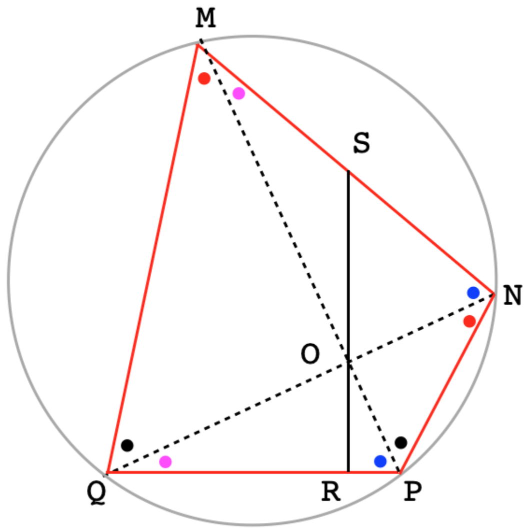


In this example, we have a special cyclic quadrilateral in which the diagonals are perpendicular. Drop the altitude to any side, say QP and extend it to meet the other side MN. Then it bisects that side so $MS = SN$.

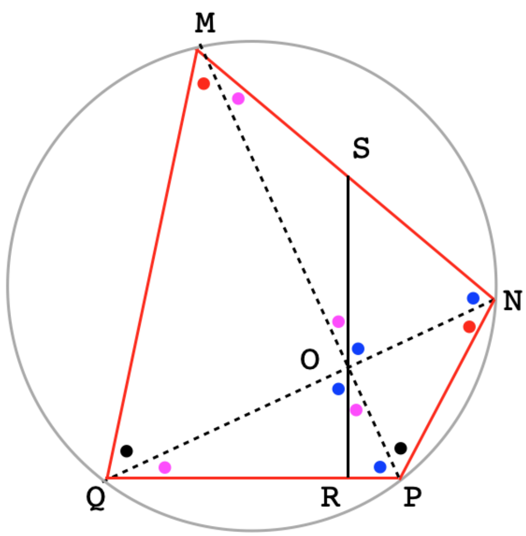
Proof.

This is fairly straightforward. Mark all the angles that are equal because they intersect equal arcs.

We notice that blue and magenta are complementary.



So then, because $SR \perp QP$, we can fill in some more dots and then use vertical angles to show that two triangles are isosceles.

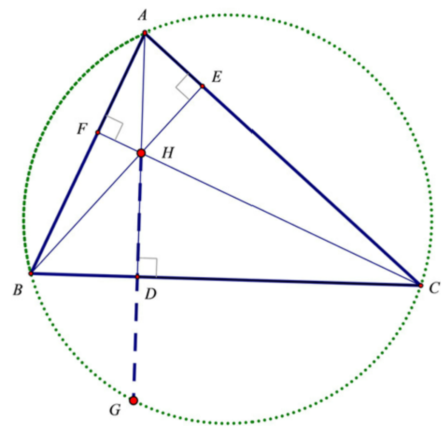


From this, we conclude that $MS = OS = SN$.

$\square$

problem

Arcs of a circle often simplify problems. Draw the altitudes and orthocenter of a triangle and draw its circumcircle.

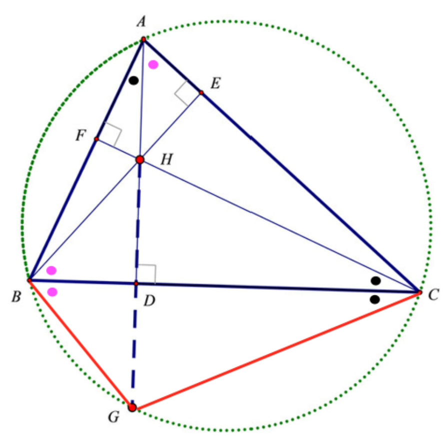


Posamentier gives this relationship. We claim that $HD = DG$.

Proof.

It seems clear that this must be a consequence of a congruence: it looks like $\triangle BCH \cong \triangle BCG$. How to prove that?

By looking at arcs subtended, we can get relationships for the new angles in $\triangle BCG$, and by similar triangles we can get the others shown:



We have congruent triangles by ASA, the component right triangles are congruent for the same reason. Or just say that the altitudes of congruent triangles are also equal.

□

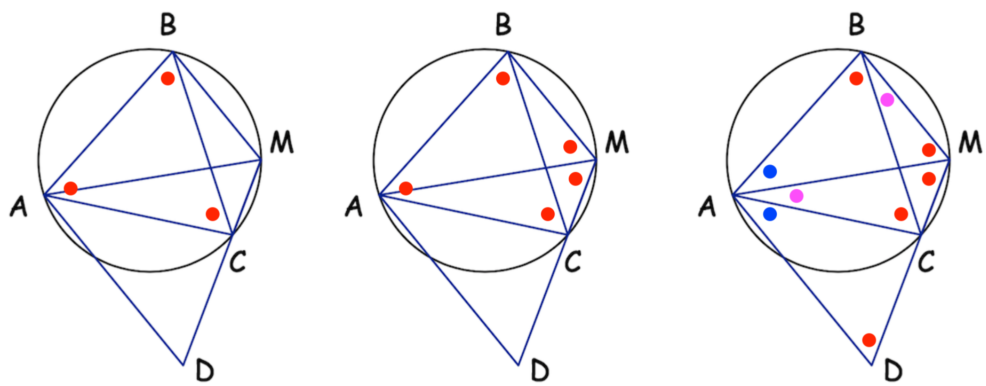
Van Schooten's theorem

This is given as a problem by Surowski (1.3.6).

Given an equilateral triangle ABC draw its circumcircle.

Draw an arbitrary line segment from vertex A through side BC to meet the circle at M. Prove that $AM = BM + MC$.

This is easy to prove as a special case of Ptolemy's theorem, which is coming. Nevertheless, we follow the hint.



Extend MC so that $AM = MD$. Connect AD.

Proof.

Since $\triangle ABC$ is equilateral, we mark each vertex with a red dot.

$\angle AMC$ and $\angle AMB$ are both peripheral and both are subtended by the same arc as an angle of $\triangle ABC$, hence we place two more red dots (middle).

We have drawn $AM = MD$, so $\triangle AMD$ is isosceles. But, since the vertex $\angle AMD$ is equal to an angle that is a vertex of an equilateral triangle, $\triangle AMD$ is equilateral. We place one red dot at D and wait a bit for the other vertex.

The overlapping $\angle BAC$ and $\angle MAD$ are both vertices of equilateral triangles, so they are equal.

Therefore, the non-overlapping parts are also equal (marked with blue dots). We place two magenta dots at angles that are subtended by arc MC.

So now, we know that certainly $\triangle BAM \sim \triangle ACD$.

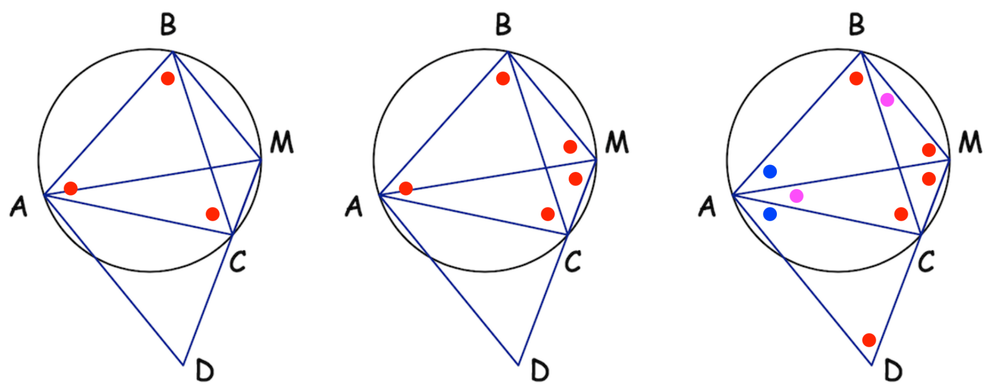
But also, $AM = AD$ because $\triangle AMD$ is equilateral. So we have ASA and $\triangle BAM \cong \triangle ACD$.

By congruent triangles the sides BM and CD are equal. Given $AM = MD$, so

$$AM = MD = MC + CD$$

$$= MC + BM = BM + MC$$

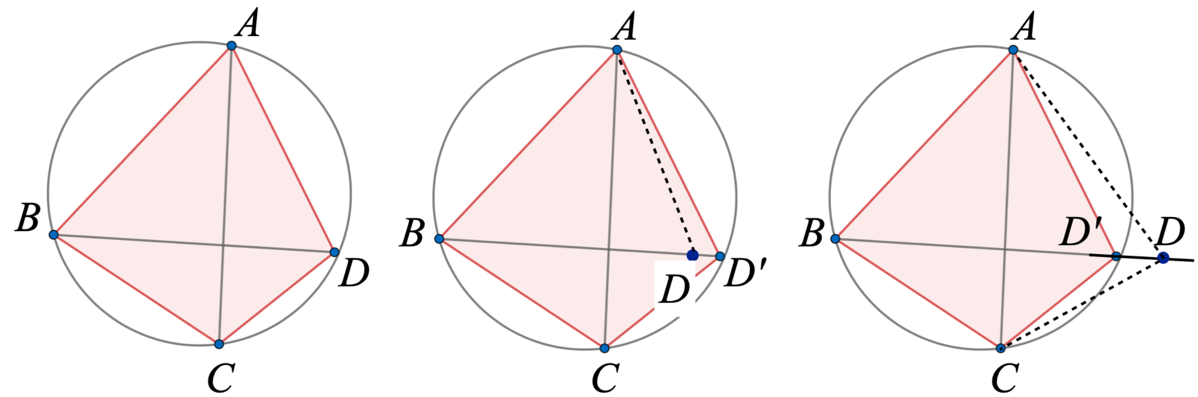
$\square$



converse

Suppose we are given that the angles A and C are supplementary, and that points A, B, and C lie on the circle as drawn.

Then D also lies on the circle.



We proved this previously **here**.

Chapter 2

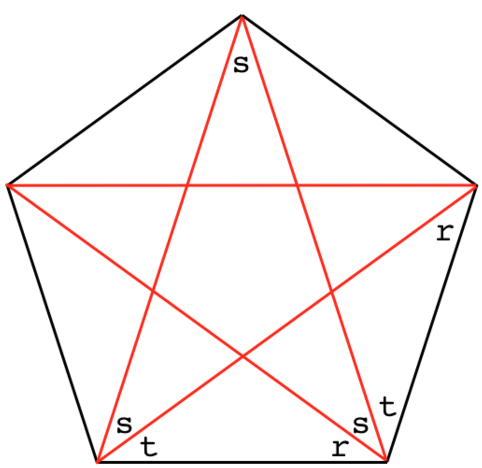
Pentagon

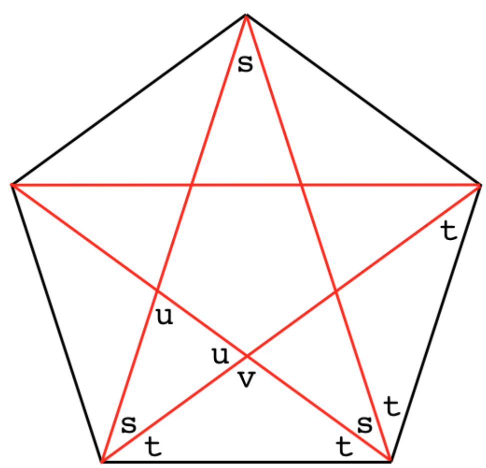
pentagons

In this chapter we explore some properties of a regular pentagon. A pentagon has 5 sides, and a regular polygon has all sides equal.

The regular pentagon has five-fold rotational symmetry. Draw all of the internal chords of the figure and label a few angles.

By rotational symmetry each of the five vertices of the pentagon has the same three components, the central one labeled s, and two flanking ones r and t.





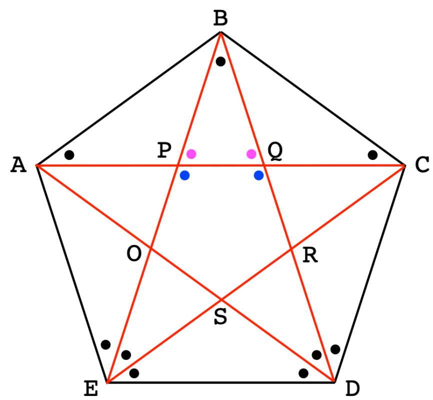
Mirror image symmetry shows that $r = t$, by reflection across the midline. Alternatively, we may invoke the properties of isosceles triangles, using two sides of the pentagon. Hence we relabel the diagram (right panel).

Next, we compute two triangle sums:

$$3s + 2t = 4t + s$$

$$2s = 2t$$

Hence, $s = t$. Relabel the diagram with black dots for the angles previously marked s, r or t.



We observe that five copies of s add up to one triangle (i.e. π).

From here, we might proceed by extending the base ED past point D. The newly formed angle will be supplementary to three copies of s, and therefore we have alternate interior angles when comparing it with $\angle ACD$.

It follows that $AC \parallel ED$.

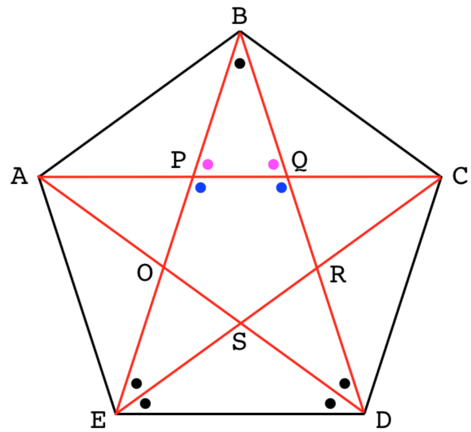
Therefore the angle with the magenta dot is worth $2s$ and the angle with the blue dot is worth $3s$.

Alternatively, compute two triangle sums:

$$5s = v + 2s, \qquad v = 3s$$

$$5s = 2u + s, \qquad u = 2s$$

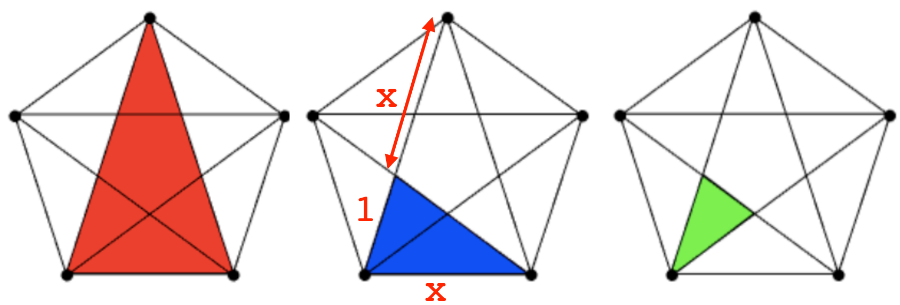
Since $v = 3s$, its measure is the same as the vertex angle of the pentagon. Thus the inner figure is also a regular pentagon.



Our drawing is filled with regular parallelograms, with four sides equal (*proof*: use ASA to show that, e.g. $\triangle CEP \cong \triangle CDE$. So we have two pairs of opposing sides equal).

One can draw two types of isosceles triangles using the chords and sides of the pentagon. One is tall and skinny, with $2s$ at each of the base angles, while the other type is short and fat, with s for the base angles.

Here are three examples of the tall skinny type:

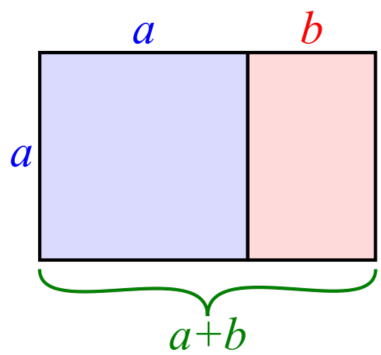


If we look at the medium-sized blue (tall, skinny) triangle, we can scale them so that the base is equal to 1 and then label the long sides as x. The ratio of the long side to the base is $x : 1$ or just x

But these triangles are all similar. In particular, the large red ones have long sides equal to $x + 1$ and base equal to x. The ratio of side lengths is the same for similar triangles, so form the ratio ϕ of the base to the long side (red on the left, blue on the right):

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

Does that look familiar? Here is a picture from wikipedia.



We draw a square with sides a and then extend two parallel sides to make a large rectangle and a small one at the same time. We require the rectangles to have the same ratio of sides.

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

Scale so that $b = 1$ and change notation, using x instead of a:

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

This is our equation from above.

To solve it, rearrange:

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

Substitute into the quadratic equation:

$$x = \frac{1 \pm \sqrt{5}}{2}$$

We are talking about a length. Noticing that one of the solutions is negative, we ignore it for the moment, and take the positive branch of the square root.

x is usually called ϕ, the famous *golden ratio*:

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

Check that ϕ really does solve the equation:

$$\begin{aligned}\phi^2 &= \frac{1+\sqrt{5}}{2} \cdot \frac{1+\sqrt{5}}{2} \\ &= \frac{1}{4}(1+2\sqrt{5}+5) \\ &= 1 + \frac{2+2\sqrt{5}}{4} = 1+\phi\end{aligned}$$

That checks.

Aside on ϕ and the Fibonacci sequence

The Fibonacci sequence is defined as $F_{n+2} = F_{n+1} + F_n$, starting with 1.

The first ten numbers in the sequence are:

```
1 1 2 3 5 8 13 21 34 55 ...
```

Recall that $\phi^2 = 1 + \phi$. The powers of ϕ generate an interesting pattern:

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

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

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

$$\phi^5 = \phi \cdot \phi^4 = 2\phi + 3\phi^2 = 3 + 5\phi$$

Both the first term and the cofactors generate the elements of the Fibonacci sequence from the powers of ϕ

The reason is that $\phi^n + \phi^{n+1} = \phi^{n+2}$, which is the same as the definition for the Fibonacci numbers.

Going back to the solution that we left behind, take the negative branch of the square root, and let us call that other solution ψ.

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

If you look closely, you can easily see that

$$\psi + \phi = 1$$

Since ψ is also a solution of the original equation:

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

Furthermore

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

Now, the left-hand side is just 1, since $\phi + \psi = 1$. Furthermore $\phi^2 = 1 + \phi$ and $\psi^2 = 1 + \psi$ so

$$1 = 1 + \phi + 2\phi \cdot \psi + 1 + \psi$$
$$1 = 3 + 2\phi \cdot \psi$$
$$\phi \cdot \psi = -1$$

Thus, ψ is the negative inverse of ϕ.

ψ comes in handy in the following way. Since ψ solves our original equation, that means the powers of ψ are just like the powers of ϕ

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

$$\psi^3 = 1 + 2\psi$$
$$\psi^4 = 2 + 3\psi$$
$$\psi^5 = 3 + 5\psi$$

So

$$\phi^5 - \psi^5 = 5(\phi - \psi)$$
$$\frac{\phi^5 - \psi^5}{\phi - \psi} = 5$$

5 is the fifth Fibonacci number. If F_n is the nth Fibonacci number

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

This is called Binet's formula. If you work out the denominator you find that it is just $\sqrt{5}$.

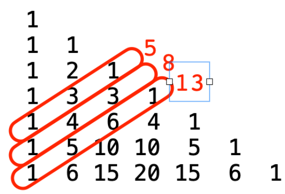
The general equation is

$$F_n = \frac{1}{\sqrt{5}} \cdot (\phi^n - \psi^n)$$

This formula is quite surprising, because the Fibonacci numbers F_n on the left-hand side are *integers*, and yet the first factor on the right-hand side is the inverse of a square root, which is definitely not an integer or even a rational number.

But it turns out that the differences $\phi^n - \psi^n$ contain only odd powers of $\sqrt{5}$. So after multiplying by $1/\sqrt{5}$, we end up only with even powers, which are whole numbers.

In fact, there is a connection between the Fibonacci sequence and Pascal's triangle.



If we let $f = \sqrt{5}$ then Binet's formula says the nth Fibonacci number is equal to

$$(\frac{1}{2})^n \cdot \frac{1}{f} \cdot \ [\ (1+f)^n - (1-f)^n\ ]$$

If you do the two binomial expansions, they are the same except that each term of the second one has a factor of $(-1)^n$. As a result, the odd powers survive, as twice the value. In the case of $n = 5$ we would have

$$(\frac{1}{2})^5 \cdot \frac{1}{f} \cdot \ [\ 10f + 20f^3 + 2f^5\ ]$$

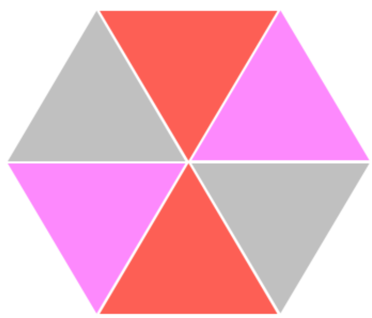
$$(\frac{1}{2})^5 \cdot \ [\ 10 + 20f^2 + 2f^4\ ]$$

The coefficients of the powers of f are twice the alternate coefficients in the binomial expansion for $(1+f)^5$: $5, 10$ and 1. If you work through more examples you'll see there is a cancellation that happens with $1/2^n$ so that this always results in an integer.

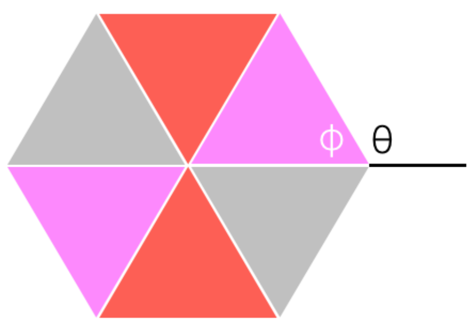
Chapter 3

Polygons

A hexagon can be composed of six equilateral triangles. In an equilateral triangle, the three angles are equal and by the angle sum theorem, their measure is $\pi/3$. This matches the requirement for a full $6 \cdot \pi/3 = 2\pi$ angular measure around a point.



We simply note numerical verification of the exterior angle theorem:



$$\theta = \phi + \phi$$

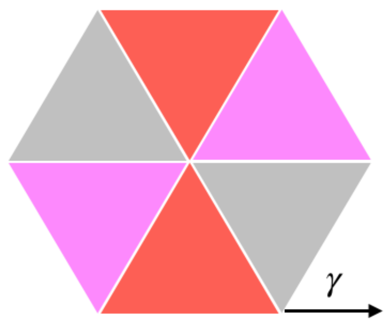
There is a different theorem, why don't we call it the

External angle sum theorem

to distinguish the angle from the exterior angle of a triangle, used above.

Imagine walking all the way around a polygon, to do so we must make n turns. What is the measure of the angle at each turn, and what is the total angle turned?

It's tricky, because the diagram we have drawn above is not the right one to use. We need this:

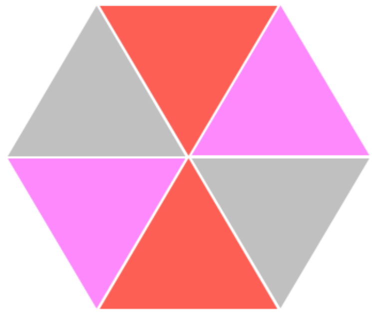


Let's recap what we know for the smaller polygons:

- triangle: $3 \cdot 2\pi/3 = 2\pi$
- rectangle: $4 \cdot \pi/2 = 2\pi$
- pentagon: $5 \cdot 2\pi/5 = 2\pi$
- hexagon: $6 \cdot \pi/3 = 2\pi$

The sum of the external angles is seen to be 2π, which makes sense. We simply turn through $2\pi = 360$ degrees total. Therefore each angle is $2\pi/n$.

approximation for pi



We can use the hexagon to get two approximations, which are lower bounds, for the value of π. Suppose the hexagon is composed of equilateral triangles with sides of unit length.

Then the perimeter is 6 and the diameter is 2 and their ratio is 3.

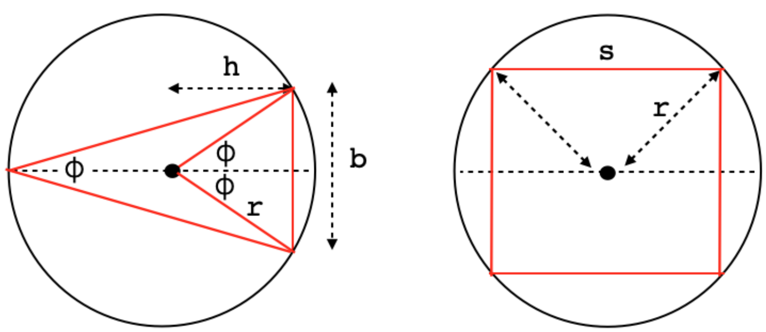
Each triangle has a base of length 1 and a height of length $\sqrt{3}/2$ and an area of $\sqrt{3}/4$. The total area is $6 \cdot \sqrt{3}/4 = 3\sqrt{3}/2 \approx 2.6$, which is not especially close!

area of an n-gon

A standard examination problem from 1900 might be the following: "The radius of a circle is 2. What is the area of a regular inscribed dodecagon?"

We will use some basic trigonometry for this.

Start by making a diagram like the one in the left panel, below.



There are n sides to the polygon, so the central angle for the sector encompassing one side is $\theta = 2\pi/n$. ϕ is one-half that, or π/n.

We will compute the area of the isosceles triangle whose one vertex is at the center of the circle with angle $\theta = 2\phi$. First, half the base is

$$\frac{b}{2} = r \sin \phi$$

so

$$b = 2r \sin \phi$$

and the height is

$$h = r \cos \phi$$

so

$$A_{\triangle} = \frac{1}{2} 2r^2 \sin \phi \cos \phi = r^2 \sin \phi \cos \phi$$

The whole polygon contains n such triangles so its area is

$$A = nr^2 \sin \phi \cos \phi$$

$$= nr^2 \sin \pi/n \cos \pi/n$$

Compute the size of a square inscribed into a unit circle. $n = 4$ so

$$A = 4 \sin \pi/4 \cos \pi/4$$

$$= 4 \frac{\sqrt{2}}{2} \frac{\sqrt{2}}{2} = 2$$

Check this (right panel, above). We have four triangles with height and base both equal to 1 so the total area is $4 \cdot 1/2 = 2$. Alternatively, compute $s = 2 \cdot \sqrt{2}/2 = \sqrt{2}$ so the area is $s^2 = 2$.

The assigned problem had $r = 2$ and a dodecagon ($n = 12$), so $\theta = \pi/6$ and $\phi = \pi/12$). The formula was

$$A = nr^2 \sin \phi \cos \phi$$

We need the sine and cosine of this angle. The way to do that is to use what are called the half-angle formulas with $\theta = \pi/6$, whose sine and cosine are easy. That's real trigonometry, so we won't go through it. However, it makes a nice exam problem because it turns out that for $\phi = \pi/12$, going through some slightly tricky algebra:

$$\sin\phi\cos\phi = \frac{1}{4}$$

and after multiplication by $n = 12$ we get $3r^2$.

A regular hexagon in a unit circle is easy. Then $\phi = \pi/6$ The area is

$$A = 6 \cdot \frac{1}{2}\,\frac{\sqrt{3}}{2} = 3\frac{\sqrt{3}}{2}$$

(We have used the sine and cosine of $\pi/6$ from **here**).

We check this as follows: a regular hexagon is composed of six equilateral triangles with unit sides. The area of each one is $1/2 \cdot 1 \cdot \sqrt{3}/2$. That's correct.

Part II

Revisiting proofs

Chapter 4

Proof by contradiction

We begin to talk more formally about proof. There are three generally recognized types of proof:

- direct proof
- indirect proof, proof by contradiction
- induction

Direct proof is the sort of thing we've been doing. For example, the theorem about the sum of angles in a triangle being equal to two right angles, or the isosceles triangle theorem.

Here is another direct proof, by what could be called enumeration of cases.

odd and even squares

If the perfect square of a number (n^2) is even, then the number n is even also; while if the square is odd, the number must be odd.

To see this, write $n = 2k$ for $k \in 1, 2, 3 \ldots$ as the definition of an even number. Then $n^2 = 4k^2$, which is even.

On the other hand, if n is odd, write $n = 2k + 1$ with $k \in 0, 1, 2 \ldots$, so $n^2 = 4k^2 + 4k + 1$, which is odd.

Since there are only these two cases, we can conclude that the converse is also true: an even square comes from an even number and an odd square from an odd number.

□

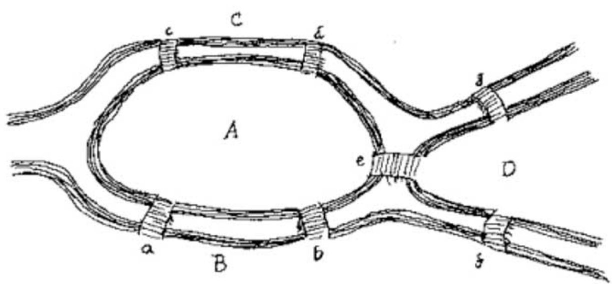
This chapter introduces a new method of proof called proof by contradiction. In Latin it is called *reductio ad absurdum*.

We start with four simple proofs, for problems which are not really related to geometry, but provide insight into the method. They are all very famous.

Bridges of Konigsberg

This problem lies historically at or near the beginning of an important branch of mathematics called graph theory. It was solved by Euler in the early 18th century.

Konigsberg was a Prussian city on the river Pregel (now renamed and part of Russia) sited on two islands, and the challenge is to take a walk through the city, visiting the different islands and both banks of the river, while crossing every bridge exactly *once*.



Euler's Drawing of Konigsberg Bridges

So now, we suppose that it is possible to make this "traversal of the graph".

The key to the proof is to notice that if we pass through an island, or one bank of the river, neither starting nor terminating there, then the number of bridges reaching that area must be even, since we have to cross each bridge once and only once.

There are four regions or "nodes" of the graph, and so there must be at least one region that is neither the beginning nor the end of our journey.

Yet, as the map shows, there is no region with an even number of bridges.

We have reached a contradiction. The original supposition, that the traversal is possible, must not be correct.

□

sum of a rational and irrational number

Let $a = m/n$ be a rational number (m and n are integers) and c an irrational number (c is not the ratio of two integers). We claim that the sum $a + c$ is irrational.

Proof.

Preliminary lemma: by fundamental properties of integer addition and multiplication, the sum and product of two integers are both integers.

Therefore, the sum and difference of two rational numbers

$$\frac{m}{n} \pm \frac{p}{q} = \frac{mq \pm np}{nq}$$

are both rational numbers.

Now, for $a = m/n$ rational and c irrational, assume that $a+c$ is rational. We assume the opposite of what we want to prove.

$$\frac{m}{n} + c = \frac{p}{q}$$

Subtract $a = m/n$ from both sides.

$$c = \frac{p}{q} - \frac{m}{n}$$

The right-hand side is rational, but c is irrational.

This is a contradiction.

Therefore, the sum and difference of a rational and an irrational number are both irrational numbers.

□

A similar approach will show that the product of a rational and irrational is irrational.

square root of 2

Probably the most famous proof by contradiction is the proof by Euclid that $\sqrt{2}$ is rational.

The theorem is that there do not exist two integers p and q such that

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

Proof.

Start by supposing the opposite, that there are two such integers.

If they were both even, that could be easily recognized by looking at the last digit, and they might then be subjected to division by 2 until they no longer share a factor of 2. Suppose that this has been carried out.

Now we need a preliminary result: the square of any even number is even, and the square of any odd number is odd. Write the numbers as $2k$ and $2k+1$ and then

$$(2k)^2 = 4k^2, \qquad (2k+1)^2 = 4k^2 + 4k + 1$$

Clearly, the first is even and the second is odd. Since these are the only two possibilities, an even square implies that the original number is also even.

We had that

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

But this means that p^2 is even, so p is even, so we can rewrite p as $2r$.

$$(2r)^2 = 4r^2 = 2q^2$$
$$2r = q^2$$

Thus, after division by 2, we see that q is also even.

This contradicts our assumption that p and q are not both even.

Therefore, there do not exist two such integers.

□

infinity of primes

The fraction of numbers which is prime (the density of primes) decreases as the integers get larger. There are 25 primes smaller than 100, 168 smaller than 1000, and later on the density is substantially less.

There is a famous theorem which says that the number of primes less than a given number k is approximately:

$$\pi(k) \approx \frac{k}{\ln k}$$

For k equal to one million, $\ln k \approx 13.8$ so $\pi(k)$ is about 7% of k. A quick check with Python gave 9592 for 100,000 and 78,498 for a million.

The question arises, is the number of primes infinite? The answer is yes. Our proof follows Euclid:

Proof.

Suppose, to the contrary, that the number of primes is finite.

Then, there must be a largest prime. Call that number p_n.

Form the product of all the primes and add 1 to it:

$$N = (2 \cdot 3 \cdot 5 \cdots \cdot p_n) + 1$$

Clearly, N is not divisible by any of our known primes, since it leaves a remainder of 1 for any of them. By the definition of prime, N is prime.

This is a contradiction.

Therefore, the number of primes is not infinite.

$\square$.

(Note 1: this proof does not show that N is the next prime greater than p_n. For example, $2 \cdot 3 + 1 = 7$, which is prime, but 7 is not the next prime after 3.)

(Note 2: this proof assumes that we have listed *all* the primes. Another way to set it up is to say "take any finite list of primes..." Then, the resulting N is either prime, or it has a prime factor which is not in the list, since none of those factors divides N.)

Hardy

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

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

Apostol's geometric proof

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

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

Here we will look at a geometric proof from Tom Apostol (see the link). A more elaborate exposition is:

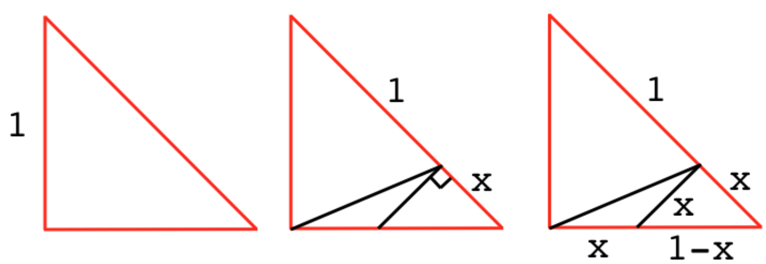
https://jeremykun.com/2011/08/14/the-square-root-of-2-is-irrational-geometric-proof/

Theorem: if there is an isosceles triangle with integer sides, then there is a smaller one with the same property.

Proof.

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

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



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

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

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

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

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

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

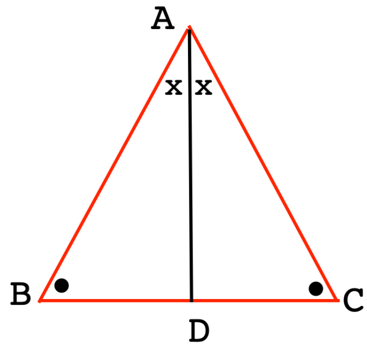
It is clear that this process can continue forever. The sides will never be in "lowest terms" because we can always form a new similar but smaller right triangle, which amounts to evenly dividing both the sides and the hypotenuse by a rational number.

□

isosceles triangles

To prove theorems in geometry, we proceed from known theorems or other accepted knowledge (e.g. axioms), to analyze a particular situation and arrive at a new theorem.

As an example, in thinking about isosceles triangles, we've seen various proofs of the theorem which says that, if two sides of a triangle are equal, then the angles opposite are also equal (**isosceles triangle theorem**).



Proof.

For $\triangle ABC$, if we are given that $AB = AC$, then it follows that the base angles at vertices B and C (marked with black dots) are equal.

One proof (which we gave previously) involves drawing the angle bisector from A, the equal angles being marked with an x. Then, we have SAS, since $AB = AC$, $\angle BAD \cong \angle DAC$ and $AD = AD$ (of course). It follows that AD is not only the angle bisector, but also the altitude and the median of the original triangle.

$\square$

There is a logical difficulty with the last proof which may be missed. It depends on the proof that angle bisection can be done. And *that* proof depends, in turn, on this one. That is the reason why Euclid uses his famously difficult proof of the theorem, which is discussed **here**.

converse theorems

The converse is not necessarily true. As the logicians would say: if $p \rightarrow q$ is a true statement, we do not know whether $q \rightarrow p$ is a true statement. It may be so, or may not.

A famous example. "All men are mortal, Socrates is a man, therefore Socrates is mortal." So if someone is mortal, can we conclude that person is a man? A counterexample will do: remember Cleopatra and the asp?

There is one other deduction that can be made from a statement like $p \rightarrow q$. Suppose it is a true statement that *all cats are black*. We see a four-legged critter that is not black. We may conclude that whatever it is, it is not a cat.

converse of Thales isosceles triangle theorem

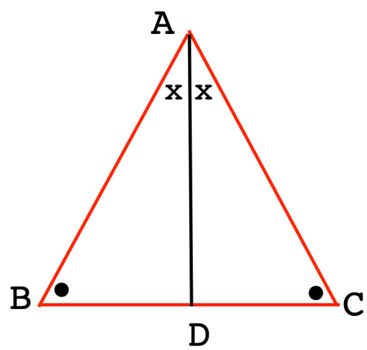
For isosceles triangles, the converse theorem is: given two equal angles in a triangle, the sides opposite are equal.

In this case, there is a direct proof, which we also gave before (**isosceles triangle theorem:** angles $\rightarrow$ sides).

Proof.

We have that the angles marked with black dots are equal.

Once again, draw the angle bisector AD to the base BC. The bisection means that the angles marked with x are equal. Two angles equal in a triangle means all three angles are equal, so we have two equal (and therefore right) angles at D.



We also have the shared side AD, flanked by two known equal angles. Therefore, we have ASA and $\triangle$ congruence. It follows that $AB = AC$ and $BD = DC$.

$\square$

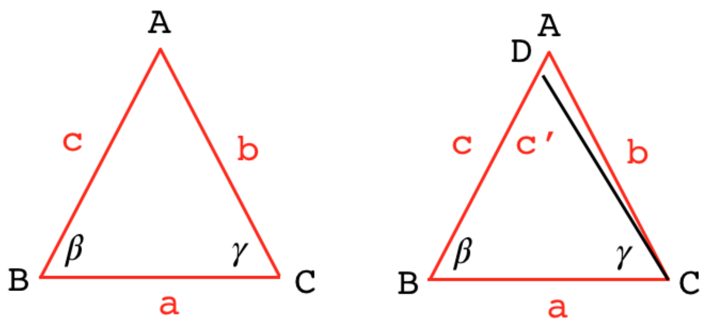
Euclid's proof of the converse (equal angles $\rightarrow$ equal sides) is short and proceeds by the method of contradiction.

Euclid I.6

If in a triangle two angles equal one another, then the sides opposite the equal angles also equal one another.

Proof.

Suppose we have $\triangle ABC$ with equal angles $\beta = \gamma$ at the base (left panel).



To begin with, we assume that the two sides b and c are not equal. We will follow this logic and find a contradiction.

Suppose b and c are not equal, then one of them is greater. Let c be greater, and then cut off b from c at point D such that the new length $c' = b$.

The new triangle has sides c' and a, which flank angle β, while for the original we have side b and side a flanking angle γ. But we constructed $c' = b$, are given that $\beta = \gamma$, and the side a is common.

Therefore the $\triangle DBC \cong \triangle ACB$ by SAS.

But this means that the lesser equals the greater, which is absurd.

Therefore c cannot be unequal to b. It therefore equals it.

Our original assumption that b does not equal c must be false.

$\square$

One subtlety in this proof is that it assumes what is called the law of trichotomy. Suppose we have two line segments AB and AC. Then only one of three possibilities can be true: (i) $AB > AC$, (ii) $AB < AC$, or (iii) $AB = AC$.

This is never explicitly assumed by Euclid.

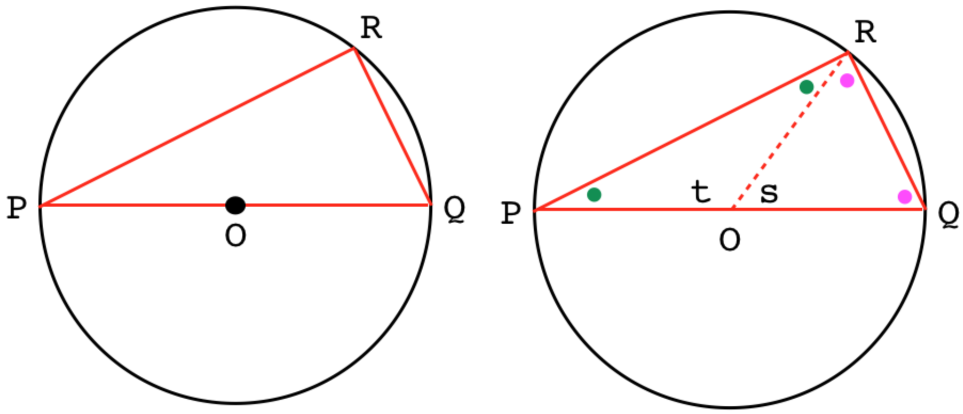
circle theorem from Thales

Here is our beautiful theorem about circles.

- Any angle inscribed in a semicircle is a right angle.

Think of three points on the circumference of a circle, forming a triangle. If two of the points form a diameter of the circle (the line joining them passes through the center), then the angle formed at an arbitrary but distinct third point is always a right angle.

In this figure, $\angle PRQ$ is a right angle (left panel).



Proof.

Draw the radius OR (right panel).

The two smaller triangles produced ($\triangle OPR$ and $\triangle OQR$) are both isosceles (two sides equal), since two of their sides are radii of the circle.

We have for the whole triangle two green dots and two magenta ones, and for the angle at R one of each. Hence the angle at R is one-half the total measure of the triangle, namely, one right angle.

□

Thales circle theorem converse

We have proved that given a semicircle and any point on the circle (not on the diameter), the angle formed with the sides drawn from the diameter is a right angle.

What about the converse (right angle → point on the circle)? Given a semicircle and $\triangle APB$, if $\angle APB$ is a right angle, we can prove that P must lie on the circle.

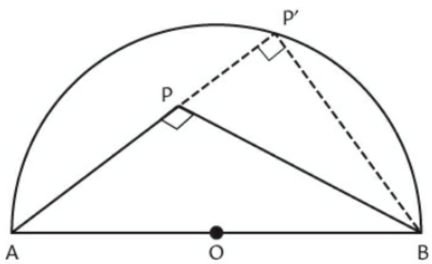


Fig. 58 Proving the converse of Thales' theorem.

Proof.

By contradiction. Assume that $\angle APB$ is a right angle, but P lies inside the circle.

Draw the continuation of AP to form AP' with P' on the circle. Then, by the forward version of the theorem, $\angle AP'B$ is a right angle. But we have assumed that $\angle APB$ is also a right angle.

Therefore, by the parallel postulate, PB is parallel to $P'B$. But these two line segments meet at B.

This is a contradiction.

Therefore APB does not lie inside the circle.

A similar argument shows that P is not outside the circle, either. If it does not lie either outside or inside the circle, it must lie on the circle.

□

converse of the Pythagorean theorem

Let the three sides of a triangle have lengths a, b and c and furthermore, let $a^2 + b^2 = c^2$. We claim that this triangle is a right triangle.

Proof.

Assume that a triangle with sides of lengths a, b and c is *not* a right triangle.

Then some other triangle can be drawn with sides of length a and b, such that b meets a in a right angle, and the hypotenuse c' such that either $c' > c$ or $c' < c$.

Then by the forward theorem

$$a^2 + b^2 = (c')^2$$

But we are given that

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

so

$$(c')^2 = c^2$$

We are allowed to take the positive root only, since we are talking about lengths, which must be positive.

$$c' = c$$

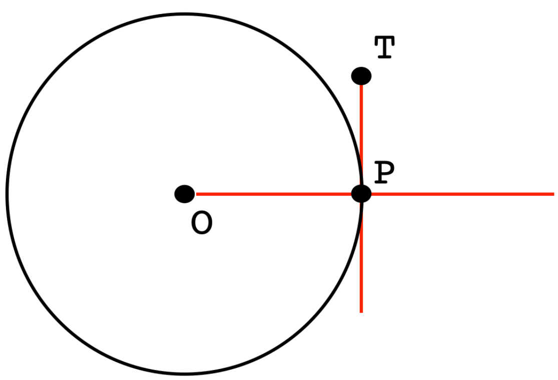
This is a contradiction, we assumed that $c' \neq c$.

Therefore $c' = c$.

$\square$

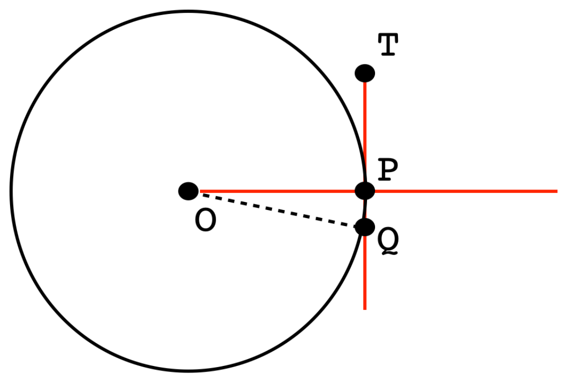
tangent to the circle

Pick a point P on a circle and draw the line segment TP through P that forms a right angle with the radius at OP. Then that line (called the tangent), touches the circle only at P.



Proof.

Pick a different point on the tangent. Let's call it Q. We suppose that Q also lies on the circle.



$\triangle OPQ$ is a right triangle and the right angle is at P. Therefore, the side opposite, OQ is the hypotenuse in a right triangle. But the hypotenuse is the longest side in a right triangle (since the greater side lies opposite greater angle. See **Euclid I.18**).

Therefore OQ is longer than OP, but OP is the radius, so OQ is longer than the radius. But Q is supposed to lie on the circle.

This is a contradiction.

Therefore, Q lies outside the circle.

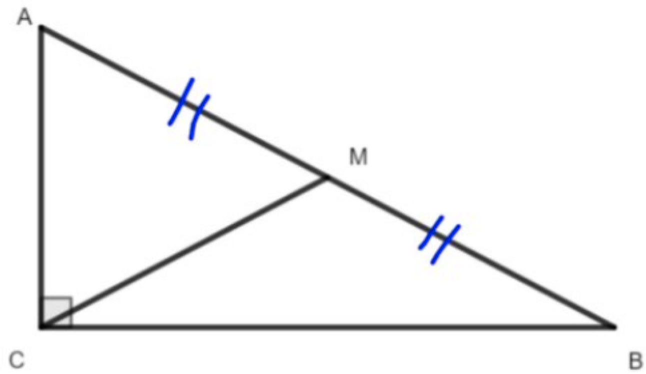
□

bisector theorem

Here is a very nice proof of the bisector theorem by contradiction.

The proof makes use of the converse of Euclid I.18: in any triangle, the greater angle is opposite the greater side.

Proof.



∘ Given $MA = MB$.

Now, make the assumption that will lead to a contradiction, that $MB < MC$.

∘ Then $\angle MCB < \angle B$ (by the converse of Euclid I.18).

∘ But $\angle MCB$ and $\angle MCA$ are complementary, as are $\angle B$ and $\angle A$. It follows that $\angle MCB < \angle B$ implies $\angle MCA > \angle A$, since both pairs sum to one right angle.

∘ So (by Euclid I.18), $MA > MC$

But we're given that $MA = MB$ and assumed $MB < MC$, so that is a contradiction.

∘ Therefore $MB \not< BC$.

Similar logic starting with $MB > MC$ also leads to a contradiction. Therefore, $MB = MC$.

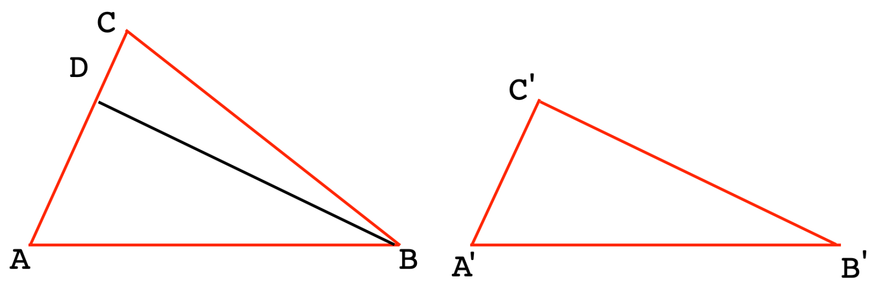
□

I saw the proof on Twitter but have lost the url. Here is a link to the author's site

https://mrhonner.com

ASA by contradiction

Euclid has a proof that ASA is sufficient for congruence (Euclid I.26). This proof is included here since it uses contradiction.



Suppose we have $\triangle ABC$ and $\triangle A'B'C'$ with equal sides $AB = A'B'$ and equal angles at vertices $A = A'$ and $B = B'$.

We claim that these two triangles are congruent.

Proof.

Suppose they are not congruent.

Then one of the other sides must differ. Suppose that only $A'C'$ is different, with $A'C' < AC$. Mark off $AD = A'C'$.

Now we have that $\triangle ABD \cong A'B'C'$ by SAS.

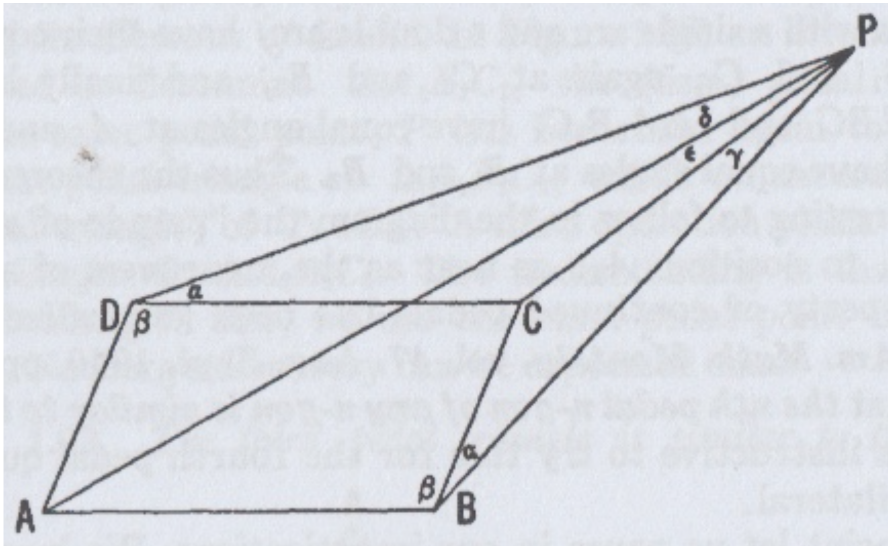
But we were given that $\angle ABC = \angle A'B'C'$ and now by congruent triangles we have that $\angle ABD = \angle A'B'C'$. So $\angle ABD = \angle ABC$

But $\angle ABD < \angle ABC$, so this is absurd.

We have a contradiction.

Therefore, ASA is sufficient to prove triangle congruence.

a hard problem



Here's a very tough problem from Coxeter. I was not able to solve it, but I'm putting it here because it uses a preliminary result that is really important.

That lemma is obtained by a proof by contradiction.

The problem statement is that $ABCD$ is a parallelogram, and P is chosen such that the angles labeled α are equal:

$$\angle PDC = \angle PBC$$

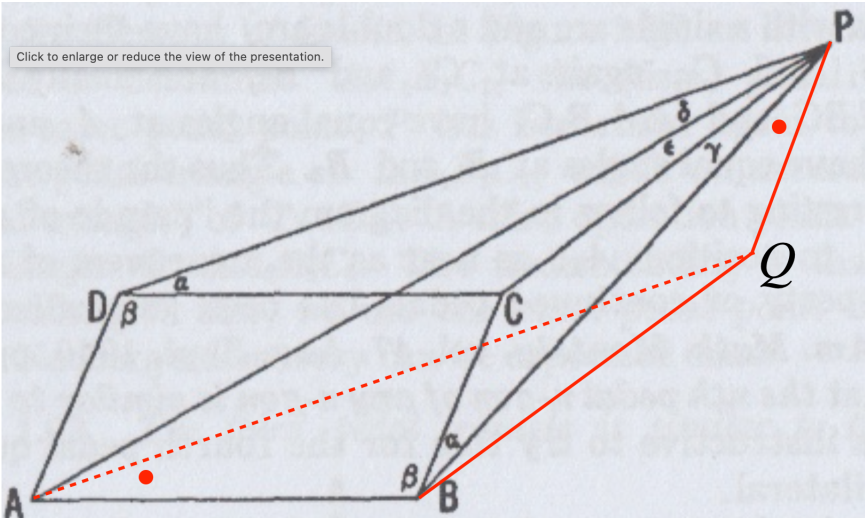
We are asked to show that

$$\angle DPA = \angle CPB$$

That is, $\gamma = \delta$.

Proof.

Like many problems, the solution starts with an inspired construction. Find Q such that $AQPD$ is a parallelogram.



Then $BQPC$ is also a parallelogram.

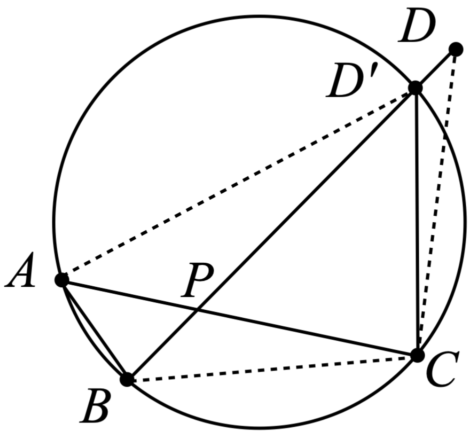
The claim is that $ABQP$ is a concyclic, a cyclic quadrilateral.

As we said, the proof involves a lemma with a proof by contradiction.

Lemma

Let $\triangle ABC$ lie on a circle.

Let point D be such that $\angle BDC = \angle BAC$.



Then D is also on the same circle.

Suppose otherwise.

Then let D' be on the circle so that $\angle BD'C$ is subtended by BC.

By the forward version of the inscribed angle theorem: $\angle BD'C = \angle BAC$

That means $\angle BDC = \angle BD'C$.

Since $\angle BCD > \angle BCD'$, we have that $\angle BD'C > \angle BDC$.

But this is absurd. So there is a contradiction.

It must be that D *is on the circle*.

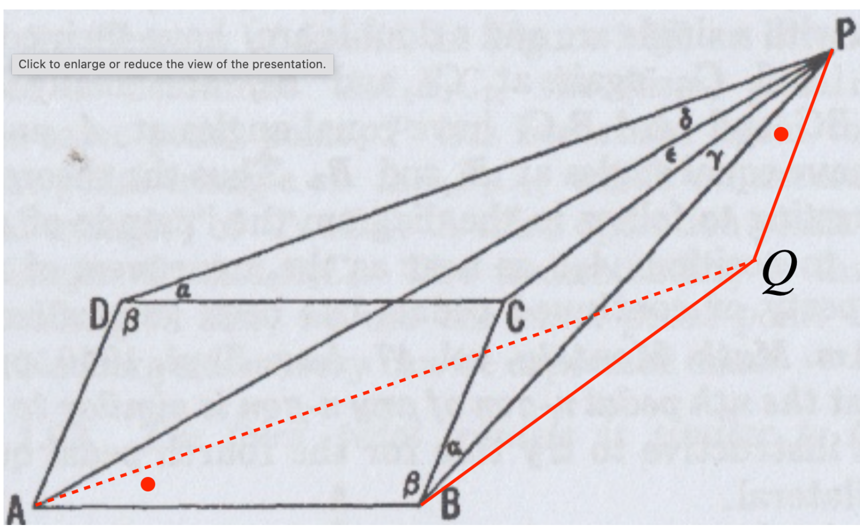
□

Also, see **here**

Proof, continued.

The arms of $\angle BAQ$ and $\angle CDP$ are both parallel, so the angles are equal.

So then they are both equal to α.



$\angle BAQ$ is subtended by chord BQ.

The diagonal of $BQPC$ forms equal $\angle BPQ$ and $\angle PBC$ where the latter is equal to α. Therefore

$$\angle BAQ = \angle BPQ = \alpha$$

So by the lemma, P is on the circle and $ABQP$ is concyclic.

By the inscribed angle theorem

$$\angle APB = \gamma + \epsilon = \angle AQB$$

because both angles are subtended by AB.

At the same time:

$$\angle AQB = \angle DPC$$

because both arms of each angle are parallel.

And $\angle CPD = \delta + \epsilon$. It follows that $\gamma = \delta$.

□

other proofs by contradiction in this book

- **Euclid I.7**
- **area of circle** (Archimedes)
- **diameter of a circle**
- **slope of tangent to parabola**
- **reflection property of ellipse**
- **isosceles angle bisectors equal**

note on proofs

If you pick up a high school geometry textbook, you will see what they call *two column* proofs, with the statement in one column and the reason in the second. I prefer *paragraph* proofs, but this is really just a matter of style.

However, another thing you will see there that I don't care for is formal statements of things that are obvious, and these will be required for *every* proof.

Rather than congruent triangles say:

- CPCTC: corresponding parts of congruent triangles are congruent.

Or rather than $AC = 2AB$ so $AB = AC/2$ say:

- division property of the equals sign"

or some such. If you must write something, say "basic arithmetic."

This seems unnecessary, unless you are taking such a course, and then it's essential.

Here are a few others:

- Reflexive property, $QR = QR$.
- Symmetric property, $QR = RQ$.
- Transitive property, if $AB = BC$ and $BC = CD$, then $AB = CD$.

Congruence of segments and angles implies all three properties.

- Angles supplementary to the same angle or to congruent angles are congruent.
- Angles complementary to the same angle or to congruent angles are congruent.

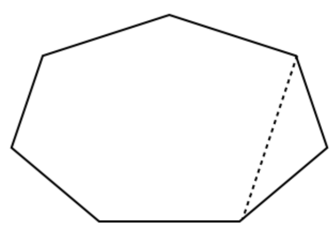
Chapter 5

Induction

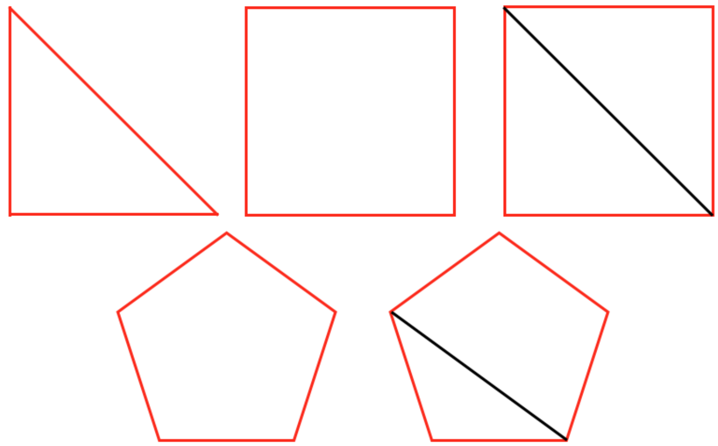
induction in geometry

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

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



The first part of the answer is to guess.



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

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

From this we guess that the formula may be:

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

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

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

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

The formula *seems* to work.

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

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

Proof.

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

$$(n-2)180° + 180°$$

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

$$((n+1)-2)180°$$

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

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

which is certainly correct.

□

squares

"Consider the number 1 represented by a dot. Round this place three other dots so that the four dots form a square ($1+3=2^2$)."

We see that the sum of the first n odd numbers is equal to n^2.

$$1 = 1^2$$
$$1+3 = 2^2$$
$$1+3+5 = 3^2$$
$$1+3+5+7 = 4^2$$

We would like to find a relationship between the last value on the left-hand side and the number that is squared on the right. Listing a few more terms we have a correspondence between

```
1  3  5  7  9 11 13 ...
1  2  3  4  5  6  7 ...
```

It appears that the first number grows like twice the second, just one less. So we guess this formula

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

and confirm it works for $n=1,2,3\ldots$.

Proof.

Now, compute

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

So if we add $2n+1$ to both sides above we have

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

and then notice that

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

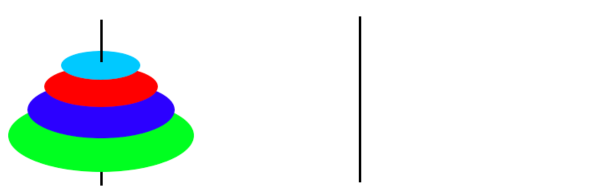
We have proved that substituting $n+1$ for n in the original formula maintains the equality. In other words, assuming the formula is correct for n, we have shown that it works for $n+1$.

This is the inductive step, and together with our tests of the base case, it proves the formula is correct for all n.

□

Towers of Hanoi

This example has a tenuous connection to geometry but it is a very clear example of induction and why it works as a method of proof.



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

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

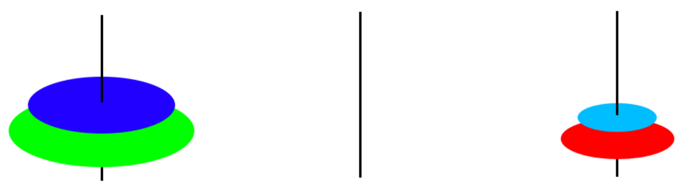
The rules are:

• Only one disk may be moved at a time.

• Each move consists of taking the upper disk from one of the pegs and sliding it onto another peg, on top of the other disks that may already be present on that peg.

• No disk may be placed on top of a smaller disk.

Here is an intermediate stage of the game:



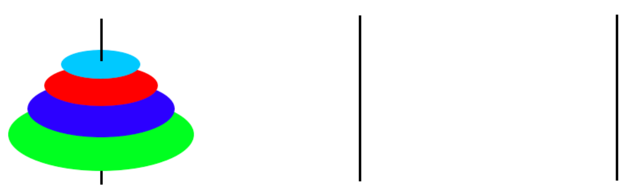
The next move is to place the blue disk on the middle peg. I think you can take it from there.

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

Proof.

By induction.

Start from the first position:



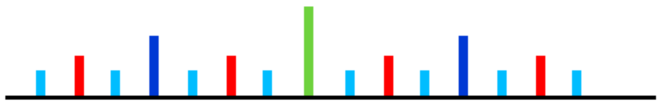
Suppose we know how to move $n-1$ disks from one peg to another. Move them to the middle peg, then move the nth disk to the right peg, then place all the $n-1$ disks on top. We have moved n disks.

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

3.

□

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



You can see the structure clearly. The first 7 steps are required to move the 3rd (blue) disk and all those above to peg no. 2. Then the green disk is placed in final position. Finally the same 7 steps are used to move the blue disk and all those above to peg no. 3, completing the puzzle.

> The puzzle was invented by the French mathematician Édouard Lucas in 1883. There is a legend about a Vietnamese temple which contains a large room with three time-worn posts in it surrounded by 64 golden disks. The monks of Hanoi, acting out the command of an ancient prophecy, have been moving these disks, in accordance with the rules of the puzzle, since that time. The puzzle is therefore also known as the Tower of Brahma puzzle. According to the legend, when the last move of the puzzle is completed, the world will end.

summary

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

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

- Graham, Knuth and Patashnik

Below is a graphic from Tom Apostol's famous calculus text. Its meaning should be obvious at this point.

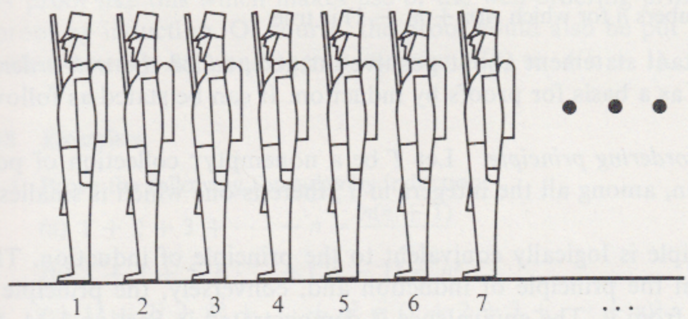


FIGURE 1.9

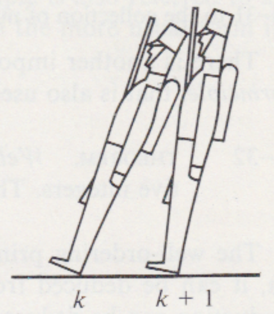


FIGURE 1.10

Binet's theorem

In the chapter on the regular **pentagon**, we had the following:

$$F_n = \frac{1}{\sqrt{5}} \cdot (\phi^n - \psi^n)$$

This is called Binet's formula, which gives the nth Fibonacci number.

We can prove this formula by induction. First, the base case. As we said $\phi - \psi = \sqrt{5}$ so $F_1 = 1$, which is certainly correct.

Since we need both F_{n-2} and F_{n-1} to compute F_n using the recurrence, we should do the next one also as part of the base case. Namely

$$\frac{1}{\sqrt{5}} \cdot (\phi^2 - \psi^2)$$
$$= \frac{1}{\sqrt{5}} \cdot \ [\ 1 + \phi - (1 + \psi)$$
$$= \frac{1}{\sqrt{5}} \cdot (\phi - \psi)$$
$$= 1$$

So then we assume that the following two equations are correct

$$F_{n-2} = \frac{1}{\sqrt{5}} \cdot (\phi^{n-2} - \psi^{n-2})$$

$$F_{n-1} = \frac{1}{\sqrt{5}} \cdot (\phi^{n-1} - \psi^{n-1})$$

and now, compute F_n:

$$F_{n-2} + F_{n-1} = \frac{1}{\sqrt{5}} \cdot (\phi^{n-2} + \phi^{n-1} - \psi^{n-2} - \psi^{n-1})$$

This will match the formula, provided that

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

Factoring, we obtain

$$\begin{aligned} \phi^n &= \phi^2(\phi^{n-2}) \\ &= (1+\phi)(\phi^{n-2}) \\ &= \phi^{n-2} + \phi^{n-1} \end{aligned}$$

And the same applies to ψ.

This completes the proof. $\square$

de Moivre

We want to mention a famous theorem by deMoivre and its proof. It is possibly too early to do this, but we mentioned this theorem in a few places and the proof is a pretty straightforward example of induction. The theorem states

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx$$

for integer n.

Probably you've seen i before as a symbol for $\sqrt{-1}$. For example, Euler has a famous formula

$$e^{ix} = \cos x + i \sin x$$

If you think of i as just being a number that when squared gives -1, that will be enough. We can't really say too much to justify it here.

There is one more rule that's important, however. If we have an equation where some terms have a cofactor of i and some do not, those without are called *real* and those with are called *imaginary*. The important (and powerful) rule is that both the real and the imaginary parts of an equation must be equal.

Here is an example. Let us multiply

$$e^{ix} \cdot e^{iy} = e^{i(x+y)}$$

But if we return to Euler's formula, the left-hand side is

$$(\cos x + i \sin x)(\cos y + i \sin y) =$$

$$= \cos x \cos y + i(\sin x \cos y + \sin y \cos x) - \sin x \sin y$$

And if that starts to look familiar, it should. The right-hand side is

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

We recall the rule that the real and imaginary parts must be equal separately. That is:

$$\cos(x+y) = \cos x \cos y - \sin x \sin y$$

$$i \sin(x+y) = i(\sin x \cos y + \sin y \cos x)$$

These are the sum of angles formulas, delivered by some alchemy.

Now, consider exponentiation. Clearly

$$(e^{ix})^n = (\cos x + i \sin x)^n$$

But by the rules of exponents we should have

$$e^{ixn} = \cos nx + i \sin nx$$

Putting the two parts together, we have de Moivre's theorem, as we said.

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx$$

An example that we have given elsewhere is the cube. If we choose $n = 3$ and consider only the real part, we can find an expression for $\cos 3x$. Expand the left-hand side:

$$\cos^3 x + 3\cos^2 x(i \sin x) + 3\cos x(i \sin x)^2 + (i \sin x)^3$$

Of course the first term is real, but the third term is also because $i^2 = -1$ so then

$$\cos 3x = \cos^3 x - 3\cos x \sin^2 x$$
$$= 4\cos^3 x - 3\cos x$$

which we obtained previously by using the sum of angles twice. Now, to prove the theorem.

Proof. (by induction, for positive integer n).

Let $n = 1$ so then

$$\cos x + i \sin x = \cos x + i \sin x$$

(Actually, $n = 0$ will work, and is legal, but no matter).

The inductive step. We assume that

$$(\cos x + i \sin x)^k = \cos kx + i \sin kx$$

So then

$$(\cos x + i \sin x)^{k+1} = (\cos kx + i \sin kx)(\cos x + i \sin x)$$

Multiplying out the right-hand side

$$= \cos kx \cos x + i \sin x \cos kx + i \sin kx \cos x - \sin kx \sin x$$

If we gather the real terms we have

$$\cos kx \cos x - \sin kx \sin x = \cos(kx + x) = \cos(k+1)x$$

while

$$i(\sin x \cos kx + \sin kx \cos x) = i \sin(kx + x) = i \sin(k+1)x$$

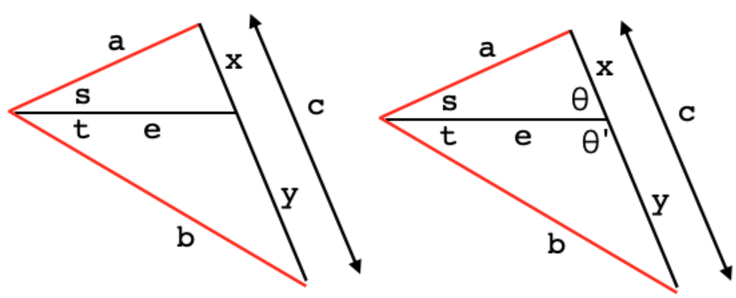
□

Chapter 6

Stewart and Apollonius

In this chapter we will continue with the theme of the **Law of cosines**. There is a general theorem and two specific cases. I think it's perhaps easier if we start with the general case.

Stewart's theorem



We have a triangle with sides a, b and c and a line segment connecting one vertex to the opposite side, dividing it into segments x and y. In the general case these two segments are not equal ($x \neq y$), and the vertex angle is not bisected ($\angle s \neq \angle t$). The two angles at the base are θ and θ'. We use the law of cosines to write:

Proof.

$$a^2 = e^2 + x^2 - 2ex\cos\theta$$

$$b^2 = e^2 + y^2 - 2ey\cos\theta'$$

Since θ and θ' are supplementary angles, $\cos\theta = -\cos\theta'$ (see **here**), so

$$b^2 = e^2 + y^2 + 2ey\cos\theta$$

We divide the first equation by x and the second by y and add which makes the term $2e\cos\theta$ disappear:

$$\frac{a^2}{x} + \frac{b^2}{y} = \frac{e^2 + x^2}{x} + \frac{e^2 + y^2}{y}$$

Get rid of the fractions:

$$\begin{aligned} a^2y + b^2x &= e^2(x + y) + x^2y + xy^2 \\ &= (e^2 + xy)(x + y) \end{aligned}$$

□

This is called Stewart's theorem.

Apollonius' theorem

The first special case is where the line segment divides the base in half, so that $x = y$. This is called Apollonius' theorem.

Proof.

We had

$$a^2y + b^2x = (e^2 + xy)(x + y)$$

Substituting $m = x = y$

$$a^2m + b^2m = (e^2 + mn)(2m)$$

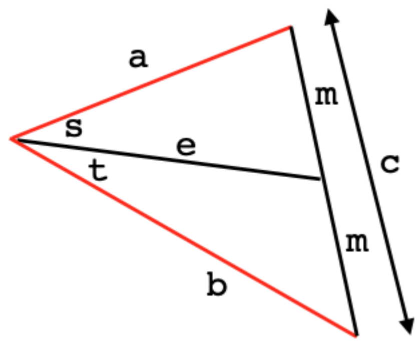
$$a^2 + b^2 = 2(e^2 + m^2)$$

$$\frac{a^2 + b^2}{2} = e^2 + m^2$$

□

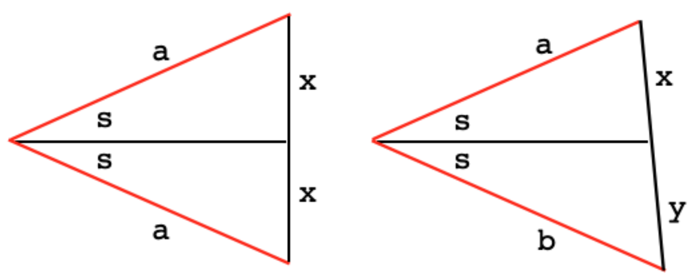
The average of the squares of the sides is equal to the median squared plus the half-base squared.

Put another way, if we imagined that the Pythagorean theorem held (it doesn't, in the general case there is not a right angle), but if it did, then this would just be the sum of $a^2 = e^2 + m^2$ and $b^2 = e^2 + m^2$.



review of angle bisector theorem

Let's go back to the generalized case of the angle bisector theorem, which we looked at (**here**).



$$\frac{x}{a} = \frac{y}{b}$$

As we turn the base c, but keep the angle bisection relationship, the ratio stays the same.

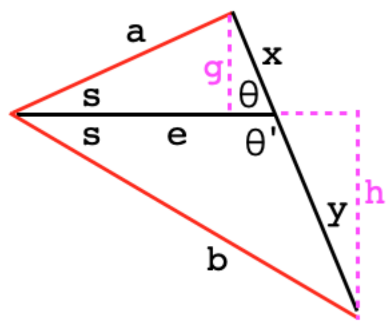
Proof.

One proof starts by recall the **area-ratio theorem**. The area of the two small triangles is in the same ratio as the lengths of the bases x and y. On the other hand, we can compute the ratio of (twice) the areas as

$$\frac{ae\cos s}{be\cos s} = \frac{a}{b} = \frac{x}{y}$$

For a different proof, compute the areas of the two smaller triangles in a different way. We will form the ratio of the areas using the same component altitudes.

In the figure below, the altitude g can be computed as $g = a\sin s$. Twice the area of the top triangle is thus $ge = ae\sin s$ (where e is the length of the bisector). Twice the area of the bottom one is b times the same factor. So the *ratio* of areas is a/b.



But the same altitudes can also be computed in terms of the sides x and y. We have that $e = x\sin\theta$ and then twice the area of the top triangle is $ge = gx\sin\theta$.

h can be computed similarly as $h = y\sin\theta' = y\sin\theta$, since θ and θ' are supplementary. So twice the area of the bottom triangle is y times the same factor. The ratio of areas is x/y.

Since the ratio must be equal no matter how it is calculated we have that

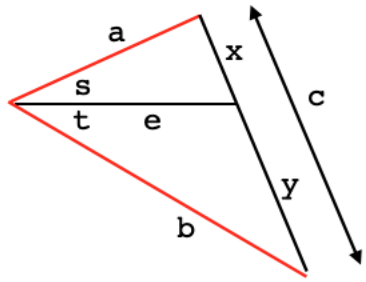
$$\frac{x}{y} = \frac{a}{b}$$

$\square$

A related proof relies on the law of sines. We have that

$$\frac{x}{\sin s} = \frac{a}{\sin\theta}$$

$$\frac{y}{\sin t} = \frac{b}{\sin\theta'}$$



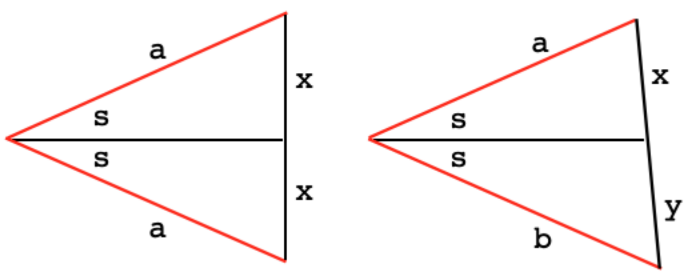
But $s = t$ so $\sin s = \sin t$ and also θ and θ' are supplementary so $\sin\theta = \sin\theta'$. We obtain

$$\frac{x}{a} = \frac{\sin s}{\sin\theta} = \frac{\sin t}{\sin\theta'} = \frac{y}{b}$$

□

bisected angle and Stewart's theorem

We use the result for the bisected angle to simplify Stewart's theorem for that case.



Proof.

For a bisected angle, the basic relationship is

$$\frac{x}{a} = \frac{y}{b}$$

which can be rewritten in two ways as

$$x = \frac{ay}{b}$$

$$y = \frac{bx}{a}$$

and we can use this to simplify Stewart's theorem for the special case.

$$a^2y + b^2x = (e^2 + xy) \cdot (x + y)$$

The substitute separately for y and x on the left-hand side:

$$a^2y + b^2x = aby + abx = ba(x + y)$$

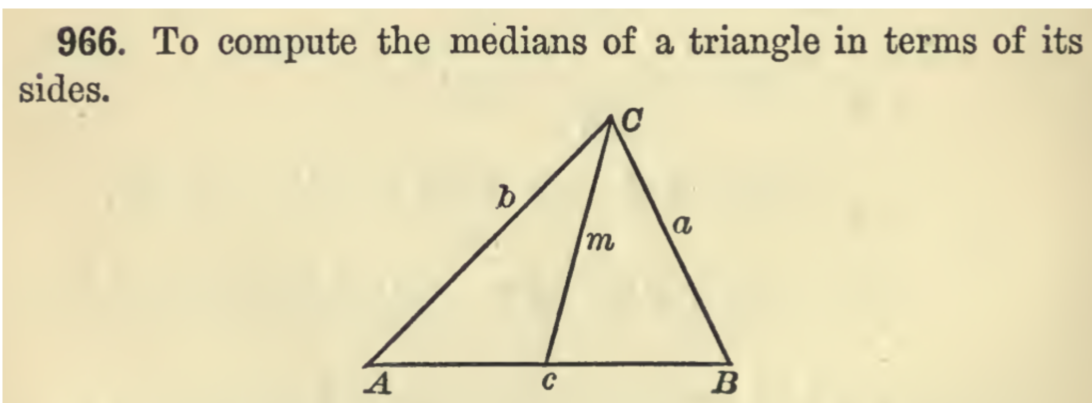
That cancels the second term on the right-hand side of Stewart's theorem, leaving:

$$ab = e^2 + xy$$

□

median in terms of side lengths

This problem is Hopkins 966.



966. To compute the medians of a triangle in terms of its sides.

We just use Apollonius' theorem.

$$a^2 + b^2 = 2(e^2 + m^2)$$

We must change symbols. We had e for the line segment dropping from the vertex, but the problem uses m for e (because it's a median). If c is the length of the base, then $c/2$ is the half-length, and we have

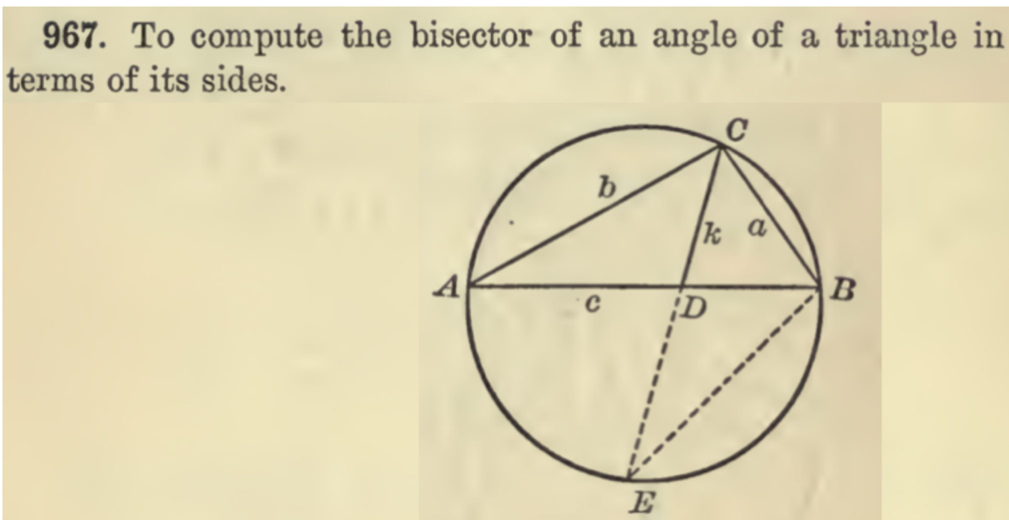
$$a^2 + b^2 = 2(m^2 + (c/2)^2)$$

Lose the fraction for a bit:

$$2(a^2 + b^2) = 4m^2 + c^2$$

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

Hopkins 967


967. To compute the bisector of an angle of a triangle in terms of its sides.

I thought, how hard can this be? If you know Stewart's theorem, it's not hard.

First of all, we notice that $\triangle ACD \sim \triangle BDE$, because $\angle A = \angle E$ (equal arcs), and there is a shared vertical angle. That helps only in the sense that it reminds us of the main result about angle bisectors that we revisited above.

Their solution starts with the statement.

$$k^2 = ab - AD \cdot DB$$

(switching from e in our notation to k). I much prefer to use x for BD and y for AD, which gives
$$k^2 = ab - xy$$

By this time we recognize this as the simplified form of Stewart's theorem for the special case of the bisected angle. Leave that for the moment.

Recall the result from the angle bisector theorem and rewrite it slightly

$$\frac{x}{y} = \frac{a}{b}$$

Add one to both sides

$$\frac{x+y}{y} = \frac{a+b}{b}$$

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

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

This manipulation was made famous by Archimedes and he used it in his work on the estimation of π.

Back to our problem:

$$y = \frac{bc}{a+b}$$

$$x = \frac{ac}{a+b}$$

Finally, we can use the original expression to write:

$$k^2 = ab - \frac{bc}{a+b} \cdot \frac{ac}{a+b}$$

which can be manipulated a bit

$$= ab\ [1 - \frac{c^2}{(a+b)^2}]$$

We have an expression for the length of the angle bisector in terms of the sides of the triangle. It is symmetric in a and b, which we should expect.

Archimedes

Back to what we were saying about Archimedes, briefly. We had

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

Inverted, this equation says that

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

We can apply this to the case where the triangle is a right triangle. In trigonometric jargon, it says that the co-tangent (inverted tangent) of the half-angle is equal to the sum of the cotangent of the whole angle plus the cosecant (inverted sine) of the whole angle.

It gives a way of computing the cotangent of the half angle knowing values for the whole angle, and then the Pythagorean theorem gives the cosecant of the half angle since

$$\sin^2\theta + \cos^2\theta = 1$$

Divide both sides by $\sin^2\theta$:

$$1 + \cot^2\theta = \csc^2\theta$$

$$\csc\theta = \sqrt{1 + \cot^2\theta}$$

The perimeter of a polygon of n sides inscribed in a circle of diameter equal to 1 is $n \sin\phi$, and for a circumscribed polygon it is $n \tan\phi$, where $\phi = \pi/n$.

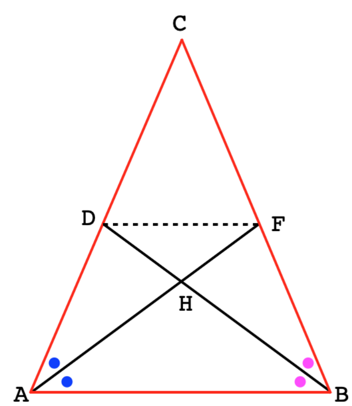
Archimedes started with a 30-60-90 triangle in a hexagon and carried out the "side doubling" four times. He obtained $223/71 < \pi < 22/7$.

Chapter 7

Steiner Lehmus

This chapter discusses a theorem about angle bisectors in an isosceles triangle. We called the forward version the **isosceles bisector** theorem.

It is easy in the forward direction, but the converse is very challenging, at least until you draw the right diagram. Then, as usual, it's not so bad.



Proof. We are given that $\triangle ABC$ is isosceles ($AC = BC$), and also that the angles at the base are both bisected. It follows that the half-angles are also equal, and thus $\triangle CDB \cong \triangle CFA$ by ASA. So the angle bisectors are equal in length: $AF = BD$. $\square$

See **here**.

That's the easy part.

Steiner-Lehmus Theorem

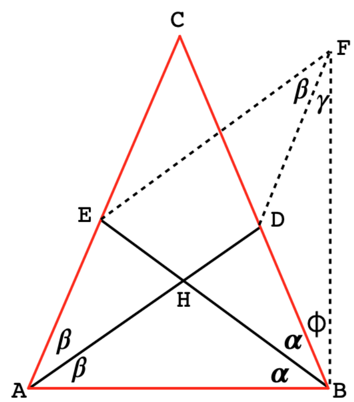
The converse theorem says that if we have angle bisectors and they are equal in length, then the triangle is isosceles.

https://en.wikipedia.org/wiki/Steiner\OT1\textendashLehmus_theorem

The problem is that, even though we can draw triangles with two sides equal, we don't know anything about the angles except for some vertical angles, which don't help.

Here is an approach which I found on the web. It's a proof by contradiction.

https://www.algebra.com/algebra/homework/word/geometry/Angle-bisectors-in-an-isosceles-triangle.lesson



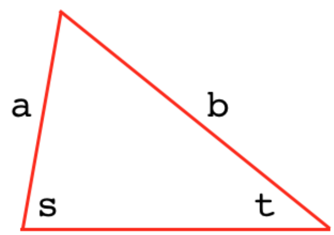
We claim that if $AD = BE$ and the angles are bisected, then $\alpha = \beta$.

Proof.

Draw $EF \parallel AD$ and $DF \parallel AEC$, meeting at F. Connect BF.

We rely on Euclid's propositions I.18 and I.19.

In any triangle if one side is larger than another, then the angle opposite the longer side is greater (I.18) and conversely, if one angle is larger than another, then the side opposite is greater (I.19).



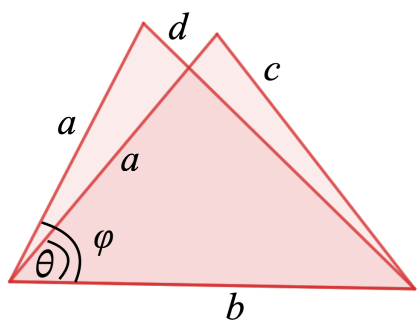
In the diagram above

$$s > t \Rightarrow b > a$$

$$b > a \Rightarrow s > t$$

We prove both these theorems **here**. We also need the following as a preliminary result.

If two triangles $\triangle ABC$ and $\triangle DEF$ have two pairs of sides equal, and the included angle is greater in one ($\phi > \theta$), then the side opposite ϕ also greater.



Proof.

Let d be opposite ϕ and c be opposite θ. Use the law of cosines:

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

$$d^2 = a^2 + b^2 - 2ab\cos\phi$$

Then if $d > c$, so $d^2 > c^2$, and

$$a^2 + b^2 - 2ab\cos\phi > a^2 + b^2 - 2ab\cos\theta$$

$$-2ab\cos\phi > -2ab\cos\theta$$

$$\cos\phi < \cos\theta$$

$$\phi > \theta$$

This chain of reasoning works just as well in reverse. So, $\phi > \theta \Rightarrow d^2 > c^2$, and then $d > c$. $\square$

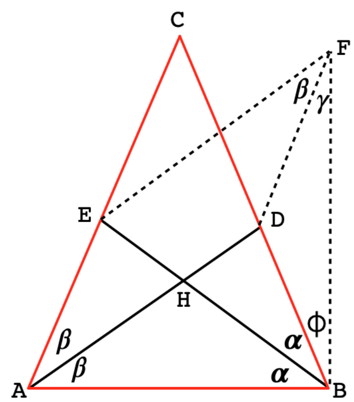
This is known as the **hinge theorem**. We showed a proof earlier. It is Euclid I.24.

$\square$

Back to our problem. We argue by contradiction. Suppose that $\alpha > \beta$.

The triangles have two sides equal and the included angle α in $\triangle ABE$ is greater than included angle β in $\triangle ADB$. Therefore, side $AE > BD$ by the lemma.

So DF, which is equal to AE, is also greater than BD.



As a result $\phi > \gamma$ by Euclid I.18. We have

$$\alpha > \beta$$

$$\phi > \gamma$$

$$\alpha + \phi > \beta + \gamma$$

This means that in triangle $\triangle BFE$, $EF > BE$, by Euclid I.19.

But since $EF = AD$ ($ADFE$ is a parallelogram) so $AD > BE$.

This is a contradiction, we were given that $AD = BE$.

Therefore, it cannot be that $\alpha > \beta$.

The reverse supposition, that $\alpha < \beta$, also leads to a contradiction by a symmetrical argument, substituting $<$ for $>$.

(Or draw the parallelogram on the other side of $\triangle ABC$ and use the same argument as previously).

Since α is neither greater than nor less than β, we conclude that $\alpha = \beta$. $\triangle ABC$ is therefore isosceles.

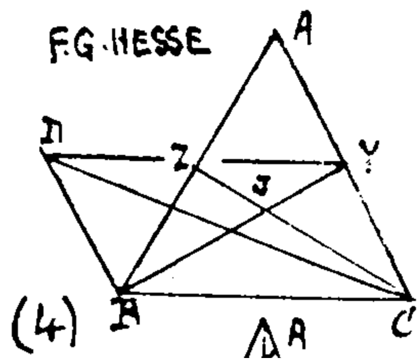
$\square$.

Hesse proof (1842)

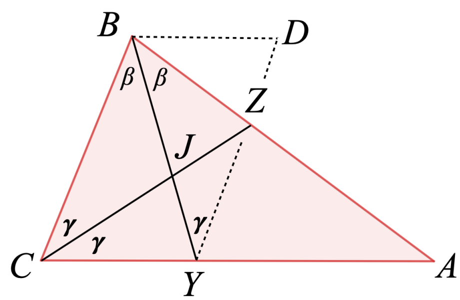
According to the internet, the Steiner-Lehmus theorem is famous for being difficult, for having many different proofs, and for some controversy over whether even one of the proofs is *direct* or not. By direct we mean, not using the technique of proof by contradiction or *reductio ad absurdum*.

I was lucky to find a (non-paywalled) review published on its centenary in 1942.

https://www.cambridge.org/core/services/aop-cambridge-core/content/view/7B625B08567935CAE06A0AC9430477C0/S0950184300000021a.pdf



The construction is to draw $YD = BC$ and $BD = BZ$.



Several sources show YZD colinear, even though this is not justified, but luckily it is not necessary to the proof. Our dotted line YD has a hole around Z. We are agnostic.

Notice that $BDYC$ *looks* like a parallelogram. We will show that it is.

The equalities of the construction are $YD = BC$ and $BD = BZ$, and $BY = CZ$ is given, so $\triangle DBY \cong \triangle BZC$ by SSS.

Thus the corresponding angles of $\triangle DBY$ and $\triangle BZC$ are equal, namely:

$$\angle DYB = \angle BCZ = \gamma$$

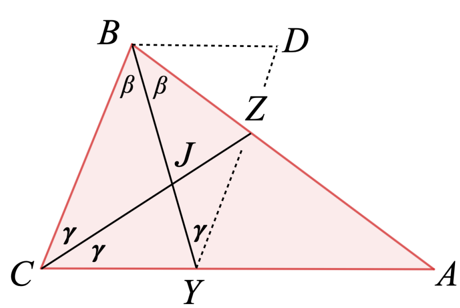
$$\angle BDY = \angle CBZ = 2\beta$$

As usual, we use Greek letters for the half-angles. Then finally

$$\angle DBY = \angle BZC = \angle A + \gamma$$

by external angles or sum of angles so

$$\angle CBD = \alpha + \alpha + \gamma + \beta = 90 + \alpha$$



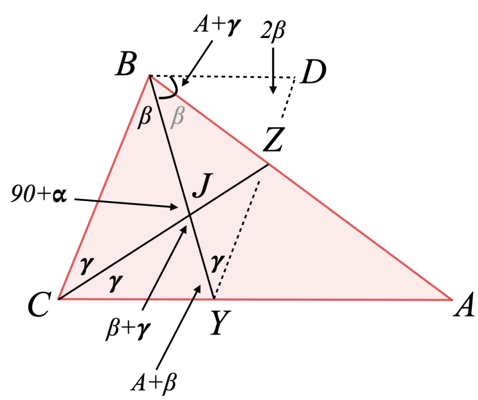
(note the source says $\angle A - \frac{1}{2}C$, but this is an error or a scanner failure).

At this point I had some trouble with the details of the proof, but all we need is to get the measure of $\angle CYD$ by some combination of external angles and sum of angles.

We find that:

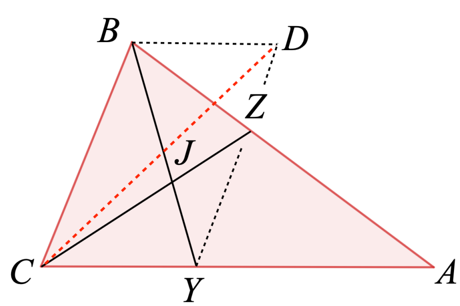
$$\angle CYD = \alpha + \alpha + \beta + \gamma = 90 + \alpha$$
$$= \angle CBD$$

Crucially, they are not only equal but obtuse.



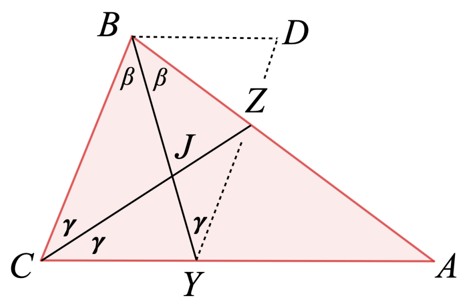
Draw CD. Comparing $\triangle CBD$ and $\triangle DYC$, we have that CD is shared, and $DY = BC$ by construction. $\angle CYD = \angle CBD$.

We have SSA and in addition, the angle between the two known sides must be acute in both, since $\angle CYD$ and $\angle CBD$ are obtuse. It follows that $\triangle CBD \cong \triangle DYC$



This then gives

$$YC = BD = BZ$$



SSS then gives:

$$\triangle ZCB \cong \triangle YBC$$

$$\angle CBZ = \angle BCY$$

$$\angle ABC = \angle ACB$$

$$AB = AC$$

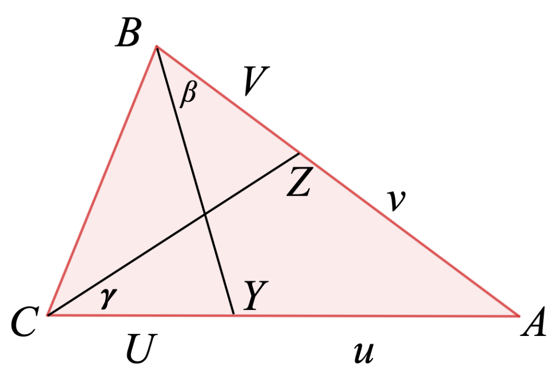
□

For more about SSA and the importance of the information about the unknown angle, see **here**.

trigonometric proof by Paul Yiu

Let the vertices and sides be labeled in the usual way. Let the half-angles be β and γ also.

Let AC (side b) be divided by BY into u and U, and let AB (side c) be similarly divided by CZ into v and V.



We will show that assuming $\beta < \gamma$ will lead to a contradiction:

$$\frac{b}{u} < \frac{c}{v} \quad \text{and also} \quad \frac{b}{u} > \frac{c}{v}$$

First consider

$$\frac{b}{u} - \frac{c}{v} = \frac{u+U}{u} - \frac{v+V}{v}$$
$$= \frac{U}{u} - \frac{V}{v}$$

By the angle bisector theorem $a/U = c/u$ so $U = au/c$ and similarly $V = av/b$. So

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

Since $c > b$ (by Euclid I.19, since $\gamma > \beta$) this expression is less than zero. Retrieving the original LHS:

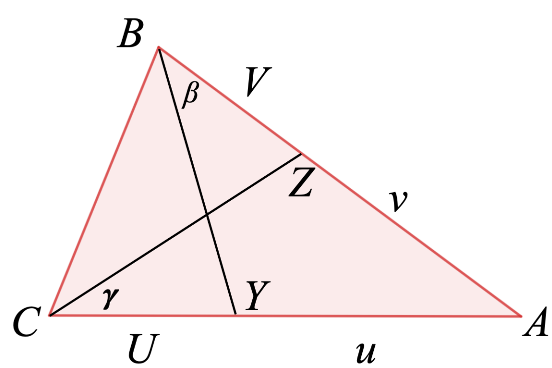
$$\frac{b}{u} - \frac{c}{v} < 0$$
$$\frac{b}{u} < \frac{c}{v}$$

The second part is

$$\frac{b}{u} \div \frac{c}{v} = \frac{b}{c}\,\frac{v}{u}$$

By the Law of Sines

$$= \frac{\sin B}{\sin C}\,\frac{v}{u}$$



By the sum of angles (really double angle)

$$= \frac{2\cos\beta\sin\beta}{2\cos\gamma\sin\gamma}\ \frac{v}{u}$$

$$= \frac{\cos\beta}{\cos\gamma}\ \frac{\sin\beta}{u}\ \frac{v}{\sin\gamma}$$

Drop the altitude h from Y to AB (side c). Then

$$\frac{h}{BY} = \sin\beta$$

$$\frac{h}{u} = \sin A \qquad \frac{h}{\sin A} = u$$

Thus $\sin\beta/u = \sin A/BY$:

$$= \frac{\cos\beta}{\cos\gamma}\ \frac{\sin A}{BY}\ \frac{CZ}{\sin A}$$

The third term has been massaged in just the same way.

And then, almost everything cancels!

$$= \frac{\cos\beta}{\cos\gamma} > 1$$

The inequality comes from the assumption that $\beta < \gamma$. Retrieve the original LHS:

$$\frac{b}{u} \div \frac{c}{v} > 1$$

$$\frac{b}{u} > \frac{c}{v}$$

We have that

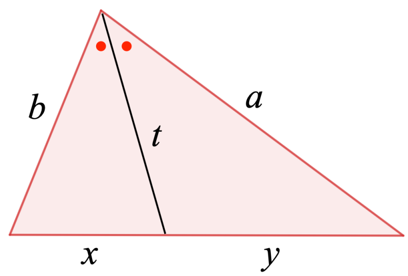
$$\frac{b}{u} < \frac{c}{v} \quad \text{and also} \quad \frac{b}{u} > \frac{c}{v}$$

This is a contradiction. It cannot be that $\beta < \gamma$. Similar logic reaches a contradiction for $\beta > \gamma$, so it must be that $\beta = \gamma$. Then it is trivial to show that $\triangle ABC$ is isosceles.

□

algebraic proof of Steiner-Lehmus

Here is another proof of the theorem. First we revisit Stewart's **theorem** in the case of the bisected angle, from the previous chapter. Substitute the symbol t (for transversal).



$$ab = t^2 + xy$$

$$t^2 = ab - xy$$

This is a general formula for the length of the transversal when it is an angle bisector. We need to eliminate the xy term. One way to write the fundamental relationship for a bisected angle is

$$\frac{a}{b} = \frac{y}{x}$$

$$\frac{a+b}{b} = \frac{x+y}{x}$$

$$x = \frac{b(x+y)}{a+b} = \frac{bc}{a+b}$$

By symmetry

$$y = \frac{a}{b}x = \frac{ac}{a+b}$$

So

$$xy = \frac{abc^2}{(a+b)^2}$$

And then

$$t^2 = ab - \frac{abc^2}{(a+b)^2}$$

The length (squared) of the transversal to side c is symmetric in a and b, as it should be.

$$= ab\ [\ 1 - \frac{c^2}{(a+b)^2}\ ]$$

$$= ab\ [\ \frac{(a+b)^2 - c^2}{(a+b)^2}\ ]$$

$$= ab\ [\ \frac{(a+b+c)(a+b-c)}{(a+b)^2}\ ]$$

We can apply this formula to the main problem, since there are two transversals which are equal. The algebra gets pretty messy. I could direct you to

https://proofwiki.org/wiki/Steiner-Lehmus_Theorem#Proof_1

As an overview, the key step is to produce (something like):

$$(b-a)(3abc + ab(a+b) + c^3 + c^2(a+b)) = 0$$

Since a, b, c are all positive, we have $(b-a)$ times something which is positive. Thus $b - a = 0$ and so $a = b$.

Let's get busy. Rewrite the t^2 expression for sides a and b:

$$t_a^2 = bc\ [\ \frac{(a+b+c)(-a+b+c)}{(b+c)^2}\ ]$$

$$t_b^2 = ac\ [\ \frac{(a+b+c)(a-b+c)}{(a+c)^2}\ ]$$

Set them equal and cancel some terms:

$$b\ [\ \frac{(-a+b+c)}{(b+c)^2}\ ] = a\ [\ \frac{(a-b+c)}{(a+c)^2}\ ]$$

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

We will obtain nine terms on each side. Start with the LHS

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

$$= (b^2 + bc - ab)(a^2 + 2ac + c^2)$$

$$= a^2b^2 + 2ab^2c + b^2c^2 + a^2bc + 2abc^2 + bc^3 - a^3b - 2a^2bc - abc^2$$

$$= a^2b^2 + 2ab^2c + b^2c^2 + abc^2 + bc^3 - a^3b - a^2bc$$

Now for the RHS (we *could* do a symbol swap with b for a and a for b, but ...):

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

$$= (a^2 + ac - ab)(b^2 + 2bc + c^2)$$

$$= a^2b^2 + 2a^2bc + a^2c^2 + ab^2c + 2abc^2 + ac^3 - ab^3 - 2ab^2c - abc^2$$

$$= a^2b^2 + 2a^2bc + a^2c^2 + abc^2 + ac^3 - ab^3 - ab^2c$$

Compare LHS with RHS for symmetry

$$= a^2b^2 + 2ab^2c + b^2c^2 + abc^2 + bc^3 - a^3b - a^2bc$$

$$= a^2b^2 + 2a^2bc + a^2c^2 + abc^2 + ac^3 - ab^3 - ab^2c$$

We are going to subtract one from the other, so first cancel two terms duplicated between the two expressions:

$$= 2ab^2c + b^2c^2 + bc^3 - a^3b - a^2bc$$

$$= 2a^2bc + a^2c^2 + ac^3 - ab^3 - ab^2c$$

Place on one side by subtracting and set equal to zero:

$$2ab^2c + b^2c^2 + bc^3 - a^3b - a^2bc - 2a^2bc - a^2c^2 - ac^3 + ab^3 + ab^2c = 0$$

Combine two pairs of like terms

$$3ab^2c + b^2c^2 + bc^3 - a^3b - 3a^2bc - a^2c^2 - ac^3 + ab^3 = 0$$

Now the really tricky part. Look for pairs that are like $bx - ax$:

$$bc^3 - ac^3 = (b - a)c^3$$

$$3ab^2c - 3a^2bc = (b - a)3abc$$

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

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

In other words the eight terms we had previously can be rewritten as

$$(b-a)\ [\ c^3 + 3abc + (b+a)c^2 + (b+a)ab\ ]\ = 0$$

$$(b-a)\ [\ c^3 + 3abc + (b+a)(c^2 + ab)\ ]\ = 0$$

And then, as we said above, we have the product of $(b-a)$ times something which is the sum of a bunch of positive terms. The only way for that to be zero is if $b-a=0$, and then $a=b$.

The two sides of $\triangle ABC$ are equal, it is an isosceles triangle.

□

Lehmus proof

One more very nice proof (reprinted in Coxeter) can be found here:

https://proofwiki.org/wiki/Steiner-Lehmus_Theorem#Proof_4

This proof is probably due to Lehmus.

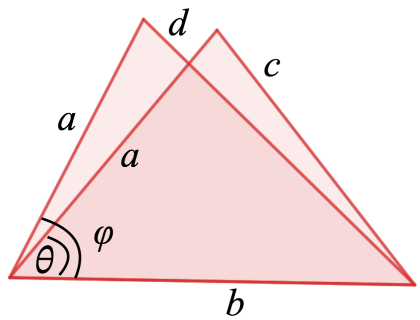
Lemma 1.

In the same circle let one chord, PR, be larger than another PQ. Then the inscribed angle subtended by PR, with its vertex in the major arc, is larger.

Coxeter says this:

> Two equal chords subtend equal angles at the center and equal angles (half as big) on the circumference. Of two unequal chords, the shorter, being farther from the center, subtends a smaller angle there and consequently a smaller acute angle at the circumference.

We could also call on the Lemma used in the first proof in this chapter (with $a=b$ equal to the radius of the circle).



□

Lemma 2.

If a triangle has two angles of different measure, the smaller angle has the longer internal bisector.

Let $\triangle ABC$ have the base angles B and C with $B < C$ and each angle bisected, by BM and CN.

We wish to prove that $BM > CN$.

Coxeter says this: take M' on BM such that $\angle M'CN = \frac{1}{2}\angle B$. Since this is equal to $\angle M'BN$, the four points N, M', B, C lie on a circle, with the same circumference subtending equal angles..

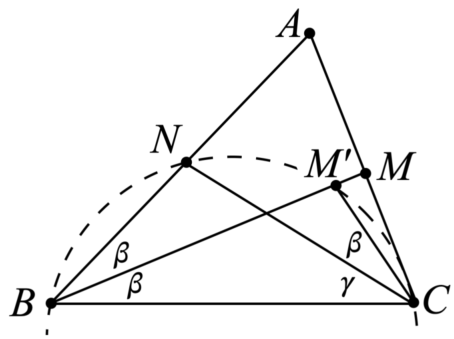
[This would be the converse to Euclid III.26. Our justification:

Draw the circle through points NBC. Draw $M'B$ to be the bisector of $\angle B$. Then by the inscribed angles theorem, $\angle NCM' = \beta$.

Since $\gamma > \beta$, CM cuts the circle to give an arc $> NM'$. Thus, M lies outside the circle so that $BM > BM'$.

This last step should be more rigorous. It's not especially elegant, but if B lies outside the circle, we can say that the area of $\triangle BMC > \triangle BM'C$. The two triangles have the same altitude. So by the area-ratio theorem the base of the first is longer, $BM > BM'$.].

For a different proof, see **here**



Then since

$$B < \frac{B+C}{2} < \frac{A+B+C}{2}$$
$$\angle CBN < \angle M'CB < 90$$

By Lemma 1, $CN < BM'$.

But $BM' < BM$. Hence $CN < BM$.

□

With the second lemma, the proof becomes simple.

Proof.

As before, we have $\triangle ABC$ with the base angles bisected and the bisectors are $BM = CN$.

Suppose $\angle B \neq \angle C$, then by Lemma 2, $BM \neq CN$. But this is a contradiction. Thus, $\angle B = \angle C$ and it follows that $\triangle ABC$ is isosceles.

□

afterward

There is some interesting discussion in Coxeter as well. According to what I can find on the web, most of the literature concerns the question of whether it is possible to provide a direct proof of the theorem. The algebraic proof, second from last, has been cited as such.

However, that proof depends on Stewart's Theorem, which as we derived it depends on the Law of Cosines, which depends in turn on the theorem of Pythagoras. And although there are several hundred proofs of Pythagoras most (all?) of them depend on the sum of angles and also on the parallel postulate, which explicitly depends on a proof by contradiction.

The question of a direct proof for Steiner-Lehmus is hard to answer conclusively. I have a write-up from John Conway claiming that it is impossible, but I don't really understand. Unfortunately, nearly all the writing in mathematics journals is paywalled and exorbitantly priced.

Chapter 8

Pizza theorem

I found this problem, called the "pizza theorem", in Acheson's *The Wonder Book of Geometry*.

Consider a circular pizza pie. Choose a point *anywhere* in the pie. Choose any orientation and draw one pair of perpendicular chords crossing at the point. Fill in with another pair of perpendicular chords on the same center, but rotated (45°).

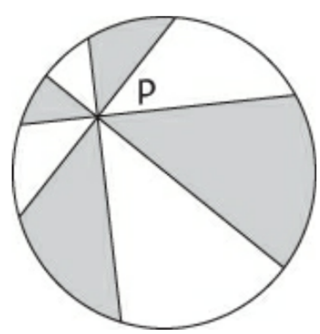


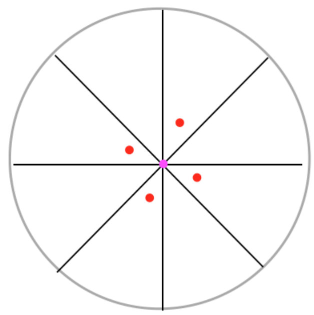
Fig. 111 The pizza theorem.

Now form the sum of the areas of alternate slices. Above, the two collections are shaded to tell them apart. The total dark area and the total light area are always equal. The pizza is evenly divided, even though there's no obvious symmetry and the slices are wonky.

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

setup

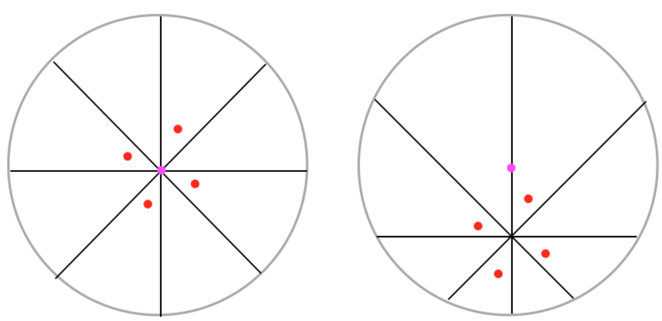
We will refer to the point where all the chords cross as the *grid center*. It is clear that if the grid center coincides with the circle center, then by the radial symmetry of the circle and the equal radial angles, all the segments will be the same. The equal area result is obvious.



For these drawings, I've chosen to mark the "shaded" slices with dots.

Now, slide the grid center away from the circle center along a diagonal of the grid, which at this point is also a diagonal of the circle. Because the circle is featureless, we can pick any diagonal, without loss of generality.

In the right panel below, the grid center has moved vertically down from the circle center, which is marked with a magenta dot.

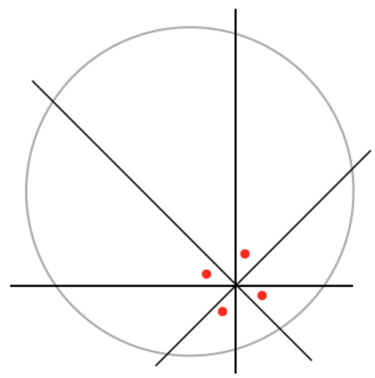


The equal area result still holds. The reason is that the figure has mirror image symmetry across the vertical line, and because of the alternation of light and dark, each pair of mirror-image shapes has one of each type.

preliminary analysis

We might try to analyze a general location for the grid center, not on a diagonal, and ask whether the areas add up properly. In fact, that is the idea of a famous proof without words, which we'll show later.

However, we will do something different and instead analyze the movement of the grid center away from the circle center. We can show that each change leaves the allocation of area between light and dark parts unchanged. Since we start with equal areas and the movements don't change area, the proof will be complete.



There are two approaches to the second movement, after a preliminary movement along a diagonal of the grid (and circle). As we said, that does not change the relative areas. One views the image radially, so having reached the desired distance from the circle center, *rotate* about the grid center to achieve the desired final orientation.

Rotation can only be performed as the second operation, since otherwise the grid center will remain on a diagonal of the circle.

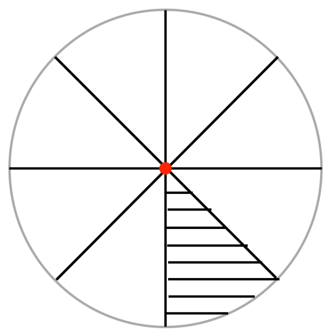
There is a second possibility, to move *horizontally* in the second phase. The grid center still runs runs parallel to a diagonal of the grid, but this is no longer on a diagonal of the circle.

The question arises whether this second method can actually achieve *any* position and orientation.

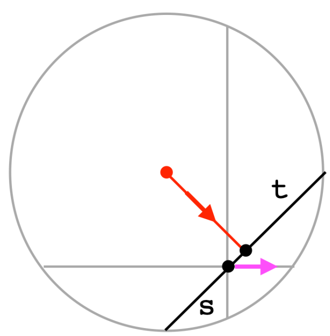
But consider the reverse movement. Start from any random position and orientation. Rotate the paper (or the circle) so that one chord is horizontal. This hasn't changed the figure and certainly doesn't change the relative areas. The picture we gave above with one chord horizontal, is still accurate.

Now, move along a line of the grid until the grid center reaches the diagonal of the circle perpendicular to the direction of movement. We have achieved equal areas.

So the crux of the analysis is that preliminary movement along a diagonal, followed by either radial or horizontal motion, can achieve an arbitrary position. We deal with horizontal motion first, since that is the part of the problem that I actually solved myself. We just note in passing that we need only move the chord center as far as the position where the chord at 45° to the vertical was originally a diagonal of the circle. Only positions in one sector of the pie matter, but that's good enough.



horizontal movement



Focus on a particular chord with arms t and s (one of the four chords is not shown). Movement of the grid center is to the right, magenta arrow.

The circle center is in red, and the second black dot marks the center of the chord. The chord center is the point where the distance to the edge of the circle is equal in both directions.

We make the following observations:

∘ The chord center is constrained to fall on the red line, its perpendicular bisector. Crucially, the bisector never changes direction, only the position of the chord center on it, i.e. the distance to the circle center.

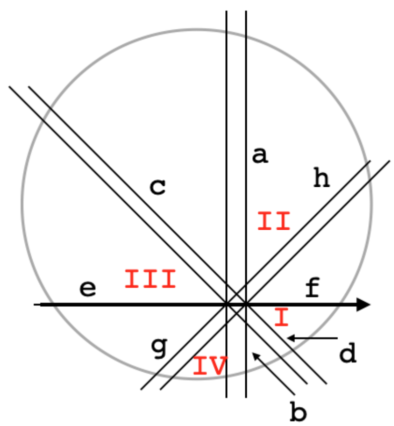
That's because in this movement the grid center doesn't rotate, the chord never changes its angle, and the final position is parallel to the beginning position.

∘ The distance from chord center to circle center changes by the amount of the horizontal movement, with an extra factor of $\sqrt{2}$ imposed by the geometry.

∘ If the region just above and left of arm t is "shaded", then that below and right of arm s is also shaded. The net change in shaded area by movement of *this* chord is t *minus* s, multiplied by a factor we will show next.

∘ The curved shapes formed by the movement at the ends of any chord are mirror images. Since one increases and the other decreases shaded area, the effects cancel. We might call these small curved shapes *wingtips*.

That last point may be appreciated from the diagram below, which shows movement to the right. Because the starting and ending positions of the chords are parallel, the wingtips are exactly the same size and of opposite shading.



Consider the total increase in shaded area due to this movement. The increase comes from rectangles formed by arms of three chords: c, b and h. The decrease in shaded area is due to rectangles on the other end of the same chords: a, d and g. The horizontal chord makes no contribution.

The net change in area for a horizontal change Δx is

$$\Delta A = \ [\ \sqrt{2}(b-a) + (c-d) + (h-g)\ ]\ \Delta x$$

The chord ab counts more because it is perpendicular to the direction of travel, so it captures more area.

Our hypothesis is that $\Delta A = 0$ so we expect to find that

$$\sqrt{2}(a-b) = (c-d) + (h-g)$$

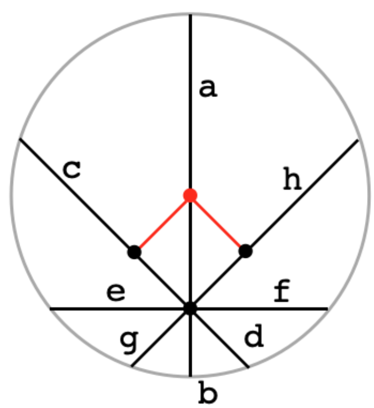
solution

The core of the argument is to connect the equation above for the change in area to something else, namely, the distances of chord centers from the grid center.

Consider chord gh, with arms g and h. Given the assignment of shaded areas above, the change in shaded area as the chord moves to the right is $(h-g)\ \Delta x$.

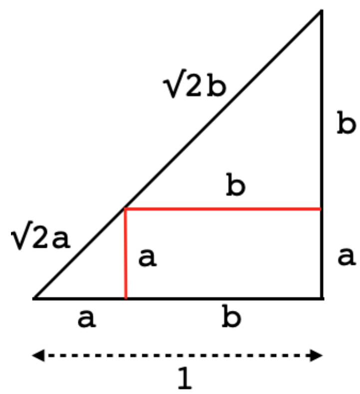
But this term $(h-g)$ is also related to the distance from the chord center of gh to the grid center. $(h-g)$ is the length of a line segment centered on the chord center, where one end is on the grid center. So $(h-g)/2$ is the actual distance from the chord center of gh to the grid center. (If this isn't clear you might review **rectangles in circles**).

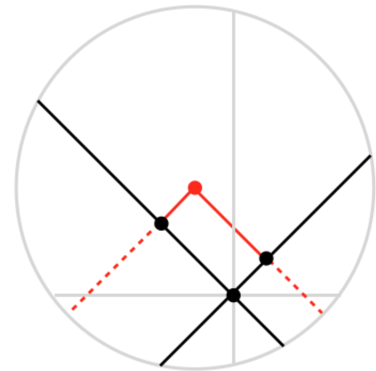
But because of all the right angles, we have rectangles, so that is also exactly the same as the distance of the center of chord cd from the center of the circle.



We can show that the net change in the latter distance for the chords cd and gh, considered together, is zero. Therefore the change in shaded area is also zero.

Proof. Consider an isosceles right triangle with sides scaled to have unit length and hypotenuse equal to $\sqrt{2}$. Pick any point along the base that is length a from the left side, and draw a rectangle (red).





The perimeter of the rectangle has length $2(a + b) = 2$, no matter which point is chosen, because by similar triangles the smaller triangles are also isosceles.

In the same way (right panel), when we slide the grid center along the horizontal grid line, the centers of the two chords at 45° are constrained to lie along the sides of an isosceles triangle (red lines).

The rectangle that is formed has a constant perimeter. The sum of distances from the center of the circle to the two chord centers is one-half that, so also constant. We can even calculate the distance as $\sqrt{2}$ times the vertical distance from the circle center to the horizontal chord.

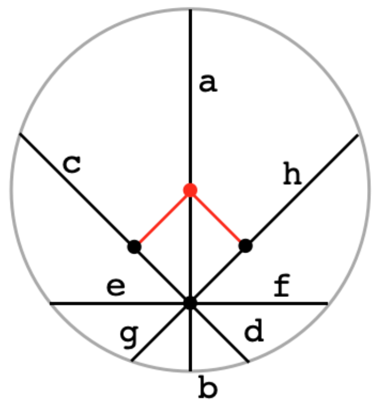
□

Call the distance from the center of a chord to the grid center δ. Take the last equation above, and divide both sides by 2. This gives an *invariant* which obviously holds when the grid center remains on a diagonal of the circle (vertical movement). In fact, by symmetry $\delta_{cd} = \delta_{gh}$ so

$$\delta_{ab} = \sqrt{2}\ \delta_{cd}$$

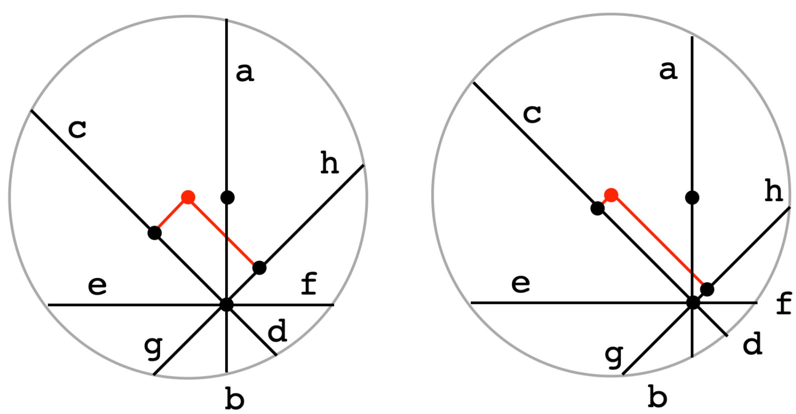
The red lines go from the center of each chord to the center of the circle (red dot). These are perpendicular bisectors of the angled chords. Because of the right angles, each one is also equal to δ for the perpendicular chord.

The vertical down from the red dot (center of the circle) to the grid center is δ_{ab}, since the center of this chord coincides with the center of the circle.



It is clear that for the central position $\delta_{ab} = \sqrt{2}\ \delta_{cd}$, as we said above. δ_{ab} is the diagonal of a square with sides of length $\delta_{cd} = \delta_{gh}$.

Next, consider two positions reached by horizontal movement:



These movements don't change δ_{ab} but they *do* change δ_{cd} and δ_{gh}. The latter two are adjacent sides in a rectangle whose four vertices are the two chord centers, plus the circle center and the grid center. It is a rectangle because the two chords are perpendicular and the red lines are perpendicular bisectors.

Both chords have successive positions lying along fixed lines connected to the circle center (red lines). The distance each moves for any given horizontal translation is the same (with opposite sign), because each center moves along the hypotenuse of an isosceles right triangle, whose base has the magnitude of the movement.

Thus, the total distance from the two chord centers to the center of the circle does not change for horizontal movement. Because of the rectangular arrangement, this distance is equal to $\delta_{cd} + \delta_{gh}$, so the invariant doesn't change either.

Recall that we need only consider the movement as far as the point where chord cd becomes a diagonal of the circle (see above). This is also the point at which δ_{cd} goes to zero.

summary

We have shown that

$$\sqrt{2} \cdot \delta_{ab} = \delta_{cd} + \delta_{gh}$$

holds regardless of motion horizontally.

In fact neither the left-hand nor the right-hand side shows a net change at all. It follows that the areas we have discussed as shaded and unshaded are also invariant under horizontal translation, even when it's not along a diagonal of the circle.

The starting position — grid center at the center of the circle — has the property of equal areas for shaded and unshaded regions, and the transformations do not change the invariant. Therefore, the ending position — grid center anywhere in the circle and with any orientation — also has the property of equal areas.

□

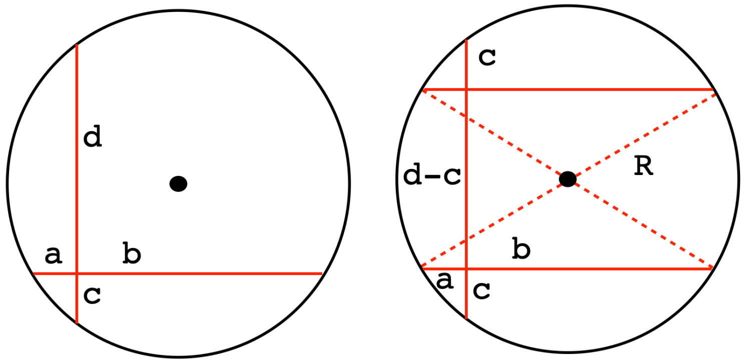
The standard proofs of this theorem analyze rotation. I found an approachable one for rotation that involves only trigonometry, here:

https://math.stackexchange.com/questions/865818/can-anyone-explain-pizza-theorem

We will need a special preliminary result.

extraordinary property

Consider two chords at right angles in a circle. Form the rectangle using a side equal and parallel to chord ab. (Only two sides of the rectangle are shown). An easy way to do this is to draw the two diagonals (dotted lines).



From a standard result about rectangles in circles, referred to above, we can conclude that the extension at the top has length c, so the side of the rectangle is $d - c$. Applying the Pythagorean theorem

$$(a + b)^2 + (c - d)^2 = (2R)^2$$

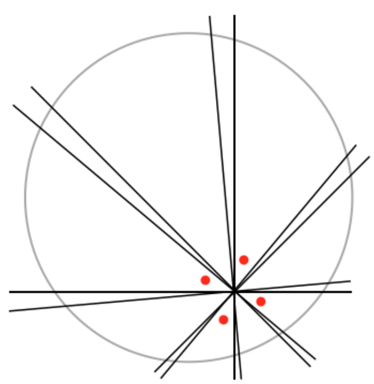
$$a^2 + 2ab + b^2 + c^2 - 2cd + d^2 = (2R)^2$$

but $ab = cd$ by the crossed chords theorem so

$$a^2 + b^2 + c^2 + d^2 = (2R)^2$$

When the components of two chords crossed at right angles in a circle are squared and summed, the result is twice the radius, squared. This is called the "extraordinary property" of the circle.

area of a sector



Rotation is connected to the special property of the circle. Consider a counter-clockwise rotation.

For a very small angle $\Delta\theta$, the area of a sector swept out by that angle is

$$A = \frac{1}{2}\, r\Delta\theta \cdot r = \Delta\theta\ \frac{r^2}{2}$$

If we focus on the sectors marked with red dots, a counter-clockwise rotation adds to the area a small wedge on the left arm, viewed from the central point and facing out. It will be seen that the relevant radii are all at right angles to one another.

These values of r are exactly those which were involved in the extraordinary property: one has a factor of $\Delta\theta/2$ times a^2, another has b^2, the next c^2 and finally d^2 (using the notation from the special property theorem).

Therefore, the total increase in area is

$$\begin{aligned}\Delta A &= \Delta\theta\ \frac{a^2 + b^2 + c^2 + d^2}{2} \\ &= \Delta\theta\ \frac{4R^2}{2} = \Delta\theta\ 2R^2\end{aligned}$$

The increase in area for four alternate segments with a small rotation depends on $\Delta\theta$ with a constant multiplier related to R for the circle (not r for the sector).

Furthermore, the shaded regions lose exactly the same amount of area at the trailing edge, along the right arms of the sectors.

The unshaded areas, although they have different radii, add and subtract exactly the same areas, and for the same reason.

A subtlety that is often neglected is that the result is only plausible and not rigorous, because we have assumed that r doesn't change *significantly* with small $\Delta\theta$.

I am not 100% happy with this geometric argument, because of the oddly shaped regions at the ends of the arms. The rate of change of r with $\Delta\theta$ is different for the different θ_i. For that reason, we'll say a little more.

One observation that might help is to notice that each angle of $\pi/4$ at the intersection is equal to the average of the two whole arcs swept out by it and its companion vertical angle. I haven't worked through the implications yet, it just seems possibly useful. There may be a simple proof in there somewhere, since it turns out that the net effect on all the θ_i together is zero.

calculus

We might use calculus for this.

https://www.math.uni-bielefeld.de/~sillke/PUZZLES/pizza-theorem

In that language we would have that (twice) the area added by a rotation is

$$\sum_i r_i^2(\phi_i) \cdot \Delta\phi = \sum_i \int_{\theta_1}^{\theta_2} r_i^2(\phi_i)\ d\phi$$

where $r_i(\phi_i)$ says that each r is a (complicated) function of the angle ϕ, for a chosen central point.

But the sum of the integrals is the integral of the sum.

$$= \int_{\theta_1}^{\theta_2} \sum_i r_i^2(\phi_i)\ d\phi$$

This is called Fubini's theorem.

https://en.wikipedia.org/wiki/Fubini%27s_theorem

And when we add them together, the sum of those factors r^2 is a constant, by the extraordinary property of the circle. So in the end we obtain $(2R)^2$ times the total of the four angular rotations. The changes in the radius simply drop out from the calculation.

I am not entirely happy with this argument either. As before, the reason is that the oddly shaped regions at the ends of the arcs are different for the different chords. In other words, the functions $r(\theta_i)^2$ must be *different* for the various i.

resolution

I finally found a solution I am happy with on the web. It shows that when you work through the algebra, almost everything really does cancel, and what's left is constant. It does use a coordinate system and polar coordinates but no real calculus.

https://math.stackexchange.com/questions/865818/can-anyone-explain-pizza-theorem

(from the answer marked as correct with a green checkmark).

We will write equations for each area that must be integrated over, each $r(\theta_i)^2$ (i.e. $[r(\theta_i)]^2$). These functions are *different* but closely related.

The integrals are polar integrals. Although they are full of sines and cosines, the θ_i are at right angles to each other. The integral will turn out to be trivial, because the great simplification referred to above, turns out to be correct.

Place the origin of coordinates at the grid center, the point where all the chords cross. Let the coordinates of the circle's center be (x, y).

Let θ be the angle that the radius of the arc r makes with the horizontal. The coordinates of any point where r meets the circle are $(r\cos\theta, r\sin\theta)$.

The basic relation of the circle is that the distance of that point on the circle from the center of the circle is R.

The squared distance is

$$R^2 = (r\cos\theta - x)^2 + (r\sin\theta - y)^2$$

$$= r^2(\cos^2\theta + \sin^2\theta) - 2rx\cos\theta - 2ry\sin\theta + x^2 + y^2$$

$$= r^2 - 2r(x\cos\theta + y\sin\theta) + x^2 + y^2$$

This is a quadratic in r. Rearrange to standard form.

We have some algebra ahead, so to make the notation simpler, let

$$P(\theta) = x\cos\theta + y\sin\theta$$

In fact, let's call it P and just remember that it is a function of θ.

$$r^2 - 2Pr + x^2 + y^2 - R^2 = 0$$

Also

$$Q = R^2 - x^2 - y^2$$

Q is a constant for any given position of the grid center. So now

The solutions are

$$r = \frac{1}{2} \cdot \; [\; 2P \pm \sqrt{4P^2 - 4Q}\;]$$

$$= P \pm \sqrt{P^2 + Q}$$

Since $R^2 > x^2 + y^2$ we have that $Q > 0$, so the second solution (i.e. minus the square root of the determinant) is < 0, and it can be ignored since we're dealing with lengths.

$$r(\theta) = P + \sqrt{P^2 + Q}$$
$$r(\theta)^2 = P^2 + 2P\sqrt{P^2 + Q} + P^2 + Q$$
$$= 2P^2 + 2P\sqrt{P^2 + Q} + Q$$

The tricky part is to figure out what these values for r^2 are for different θ. Luckily, we only have four angles to worry about, namely θ plus one each of $\{-\pi/2, 0, \pi/2, \pi\}$.

one pair of values for θ

Start with θ and $\theta + \pi$. When π is added to the angle, both the cosine and the sine switch *sign*. So

$$P(\theta + \pi) = -(x \cos\theta + y \sin\theta) = -P(\theta)$$

What's under the square root has P^2 and something that doesn't depend on θ so the only change for r is the sign of the first term.

$$r(\theta + \pi) = -P + \sqrt{P^2 + Q}$$

But that changes the sign of the mixed term in the square:

$$r(\theta + \pi)^2 = 2P^2 - 2P\sqrt{P^2 + Q} + Q$$

When we add them that pesky square root disappears:

$$r(\theta)^2 + r(\theta + \pi)^2 = 4P^2 + 2Q$$

$$= 4(x \cos\theta + y \sin\theta)^2 + 2Q$$

Pull out a factor of 2 twice

$$= 2 \ [\ 2(x^2 \cos^2\theta + 2xy \cos\theta \sin\theta + y^2 \sin^2\theta) + Q\ ]$$

two more values for θ

Adding $\pi/2$ to θ, sine and cosine switch places and also the *sign* of $\cos(\theta + \pi/2)$ becomes negative.

$$P(\theta + \pi/2) = x\cos(\theta + \pi/2) + y\sin(\theta + \pi/2)$$

$$= -x\sin\theta + y\cos\theta$$

On the other hand, subtracting $\pi/2$ gives

$$P(\theta - \pi/2) = x\cos(\theta - \pi/2) + y\sin(\theta - \pi/2)$$

$$= x\sin\theta - y\cos\theta$$

So minus the first is equal to the second, as before.

Now we need to compute $2P^2 + 2P\sqrt{P^2 + Q} + Q$.

We will not bother with the middle term. As before, the change of sign on P makes this disappear when we add them. We need only $2P^2 + Q$ for each θ.

The two terms $r(\theta \pm \pi/2)^2$ added together give

$$2(-x\sin\theta + y\cos\theta)^2 + Q + 2(x\sin\theta - y\cos\theta)^2 + Q$$

Using simpler variables to illustrate, this is (partly)

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

times another factor of 2:

$$2\ [\ 2(x^2\sin^2\theta - 2xy\sin\theta\cos\theta + y^2\cos^2\theta) + Q\ ]$$

Now, go back to what we had before for $r(\theta)^2 + r(\theta + \pi)^2$

$$2\ [\ 2(x^2\cos^2\theta + 2xy\cos\theta\sin\theta + y^2\sin^2\theta) + Q\ ]$$

There are two more cancellations coming when we add. First, the mixed term disappears. Then also there is $\sin^2 + \cos^2$ so that leaves just

$$= 2\ [\ 2(x^2 + y^2) + 2Q\ ]$$

This is the sum for all four angles:

$$\sum_{k=0}^{3} r(\theta + k\pi/2)^2 = 4x^2 + 4y^2 + 4Q$$

$$= 4x^2 + 4y^2 + 4(R^2 - x^2 - y^2)$$
$$= 4R^2$$

The sum of the integrands is independent of θ and hence invariant as θ changes.

What a beautiful simplification! That is really special.

references

I found a number of images for a "proof without words" online and I also found a reference to some related proofs, as well as a discussion on Stack Exchange:

https://math.stackexchange.com/questions/865818/can-anyone-explain-pizza-theorem

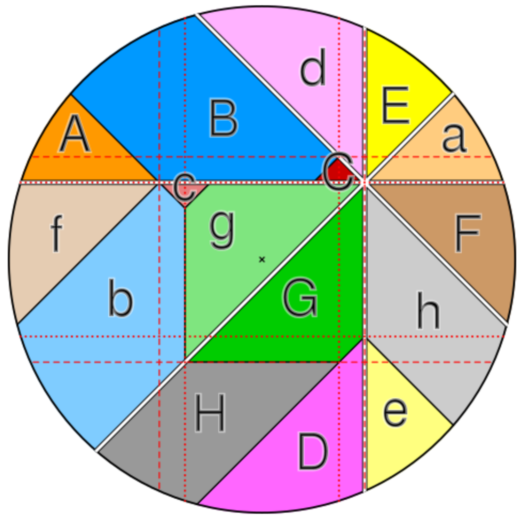
The reference says

> Sliding these arcs and chords together, we see that the chords form a right triangle with the diameter of the circle as the hypotenuse.

The various pictures of the proof without words do not explain how the shapes were arrived at. However, they do show that the complementary light and dark areas each add up to one-half of the pizza.

The original article is paywalled and the price is truly exorbitant. You have to wonder if they've ever sold even a single copy for $55.

https://www.tandfonline.com/doi/abs/10.1080/0025570X.1994.11996228


d
E
B
a
A
C
c
g
f
F
G
b
h
H
e
D

Part III

Special points

Chapter 9

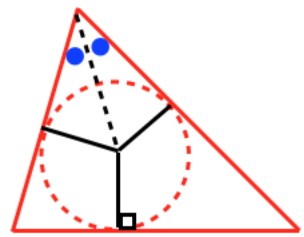
Four points

We will be talking about four special points in triangles including the orthocenter, circumcenter, incenter and centroid. The first is the point where the altitudes cross, the second is the center of the circle which includes all three vertices, and the third is the center of a circle which just fits inside the triangle. The last is formed by drawing lines from each vertex to the midpoint of the opposing side.

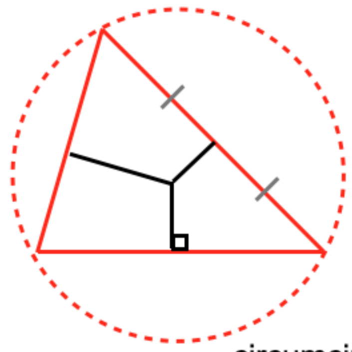
These have the common feature that three lines with certain properties cross at a single point. When this happens, they are said to be concurrent.

We have already established the general conditions under which this happens. This is the subject of Ceva's theorem.

Note that all three sides and vertices as well as the lines drawn from them in each example share the same properties. For example, when the line meeting the base forms a right angle, that is true for all three lines. Similarly, when an angle or a side is bisected, all three are bisected.

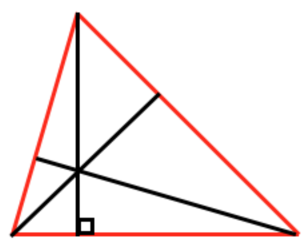


incircle/angle bisector

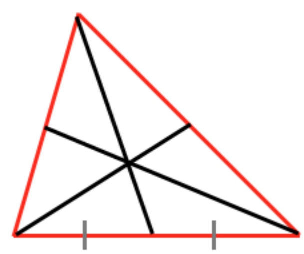


circumcircle/
perp. bisector

Two of these points relate directly to circles. The incircle is the circle that has each of the three sides as tangents, while the circumcircle is the circle that contains all three vertices.



orthocenter/altitude



centroid/median

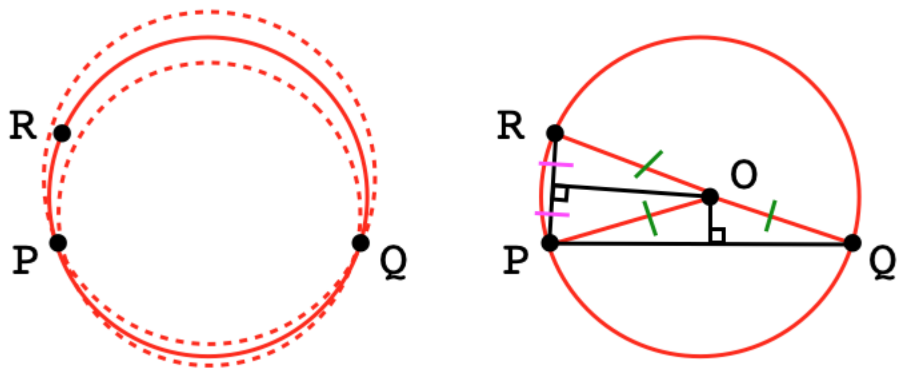
Altitudes we have seen previously as the vertical from a vertex to the opposing side. Finally, the median is the line from a vertex that bisects the opposing side, that connects to its midpoint.

A preliminary question one might ask is the following: we have drawn the three lines in each triangle meeting at a single point, but how do we know this will actually happen? How do we know the three lines are concurrent in a single point?

We will show a formal proof for each type as we come to it in this chapter. For the two involving circles, note that any three points determine a circle.

Proof.

We can make an informal argument as shown on the left.



We draw a family of circles through points P and Q. There is exactly one circle that contains all three points P, Q and R.

More formally, for the circumcircle (right panel), draw the secant PR, and then draw its perpendicular bisector. Every point on the perpendicular bisector is equidistant from R and P, by the properties of isosceles triangles. (We have congruent triangles by SAS).

Now draw the perpendicular bisector of PQ. It must cross the other one at a unique point (two non-parallel lines cross at a single point).

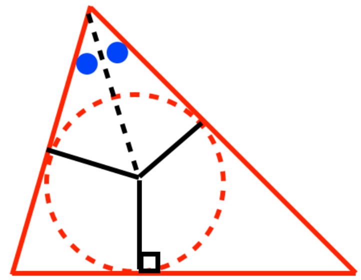
As a point on the first perpendicular bisector, O has the property that $OR = OP$. But O is also on the second perpendicular bisector, so then $OP = OQ$. Thus, all three points lie at the same distance from O.

A circle with its center at O and its radius equal to OP passes through all three points, because they satisfy the definition of points on a circle.

$\square$

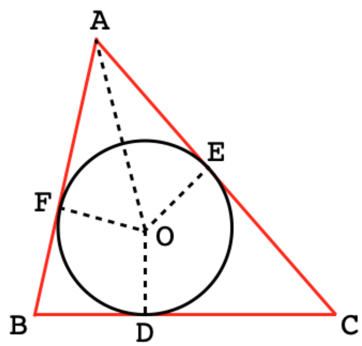
incenter: angle bisector $\rightarrow$ incenter

In the case of the incircle, we use the angle bisector. We proved previously that if a line segment is drawn from an angle bisector to the two sides of a triangle, so as to meet the sides at right angles, the distance to each side is equal.



incircle/angle bisector

The three solid black lines are all equal in length, so a circle can be drawn that includes all three. From the fact that each radius meets the side at a right angle at a single point, we can show that the sides are tangent to the circle.



Proof.

Draw the three perpendiculars from O to D, E, F, and then draw the bisector of angle A to point O.

$\triangle AOF$ is congruent to $\triangle AOE$ by AAS (right angle, angle bisector, AO is shared), so $OF = OE$.

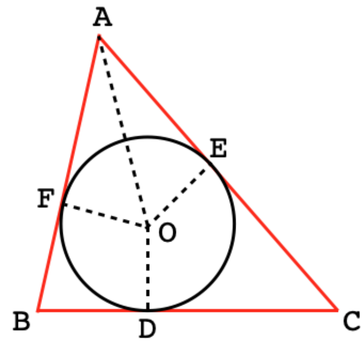
But this can be done for each vertex in $\triangle ABC$.

Thus $OE = OF = OD$.

Therefore, O is the center of a circle that includes points A, B and C (the incenter), and the line segments connecting O to the three vertices are the angle bisectors for the corresponding angles.

□

tangent of the incircle → angle bisector



Proof (Alternative).

Given $\triangle ABC$, draw a circle that just touches the sides of the triangle at D, E and F.

By definition, the tangent is perpendicular to the radius where it touches the circle. So $\triangle AOF \cong \triangle AOE$ by hypotenuse-leg in a right triangle (HL).

Thus, AO is the angle bisector of the angle at vertex A.

Similarly OB and OC bisect angles B and C, respectively.

□

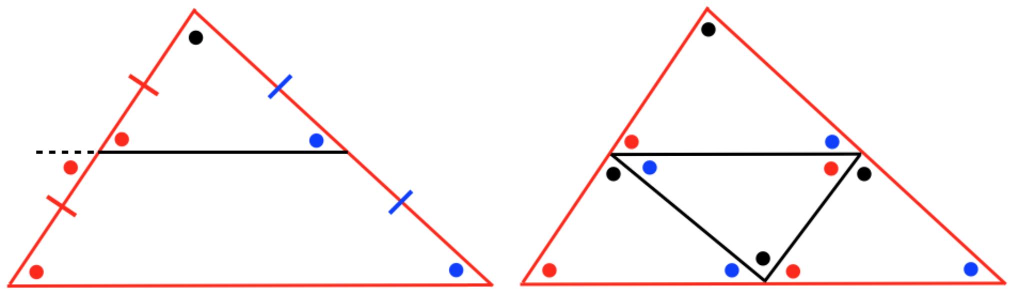
We have previously developed proofs for the orthocenter **here**, and **here**.

Finally, we develop an unusual proof that the centroid exists, and locate it on the lines to the midpoints of the sides (I found this proof in Lockhart).

The proof depends on the properties of similar triangles.

centroid

Consider the triangle in the figure below (left panel).



Draw a line segment parallel to the base and connecting to the midpoint of the left side. Then, by the alternate interior angle theorem and the vertical angle theorem, the two angles marked with red dots in the middle are equal to the red dotted angle at the base.

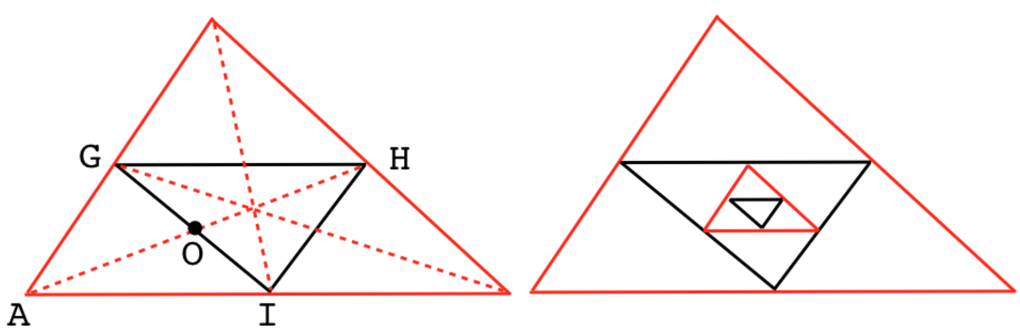
Therefore, by three angles the same, the small upper triangle is similar to the large one. The ratio of similar sides is $1:2$.

But this can be done on the right side as well, and then the same for all three vertices of the original triangle (right panel).

By the triangle sum theorem and also by the alternate interior angle theorem, the angles in the interior triangle are equal to other angles as indicated. By shared sides, the four small triangles are congruent.

Now draw lines from each vertex to the midpoint of the opposing side. $GHIA$ is a parallelogram, by the angle equalities just proven.

The two diagonals of a parallelogram cross at their midpoints. Therefore O is the midpoint of the side GI and the same line that connects A to midpoint H also connects H to midpoint O.



Therefore the centroid of $\triangle GHI$ is also the centroid of the parent. This process can be repeated as many times as we please (right panel).

The triangles get smaller and eventually tend to a point. That point is on all three midpoint segments. Therefore, the centroid is a single point.

□

algebra of the centroid

We can locate the centroid by imagining that we find successive midpoints of a length from opposite ends left and right.

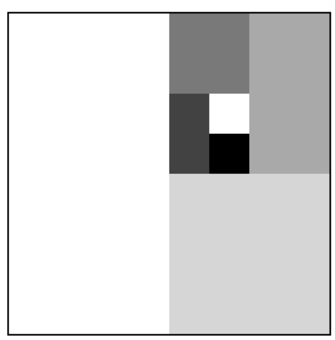
The first point is at 1/2 of the length (point O on $\triangle GHI$), the second comes back from vertex H by 1/4 so is at 0.75 (on the right edge of the small red triangle in the right panel, above). The third is at $0.5 + 1/8$ (on the left edge of the smallest black triangle).

Every second round we get closer to the centroid by advancing from the left by

$$S = \frac{1}{2} + \frac{1}{8} + \frac{1}{32} + \dots$$

Now, we can either assume this sum is finite (for now) or recognize that it is certainly smaller than

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots = 1$$



So if

$$S = \frac{1}{2} + \frac{1}{8} + \frac{1}{32} + \dots$$

then

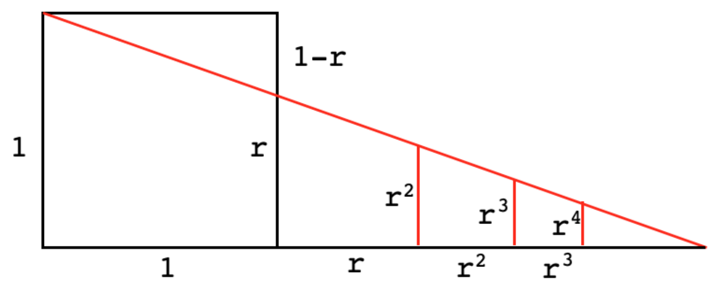
$$2S = 1 + \frac{1}{4} + \frac{1}{16} + \dots$$

and then, adding the two expressions:

$$3S = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots$$
$$= 1 + 1$$
$$S = \frac{2}{3}$$

□

Here is another proof from Nelsen's *Proof without words*. By similar triangles



$$1 + r + r^2 + \dots = \frac{1}{1-r}$$

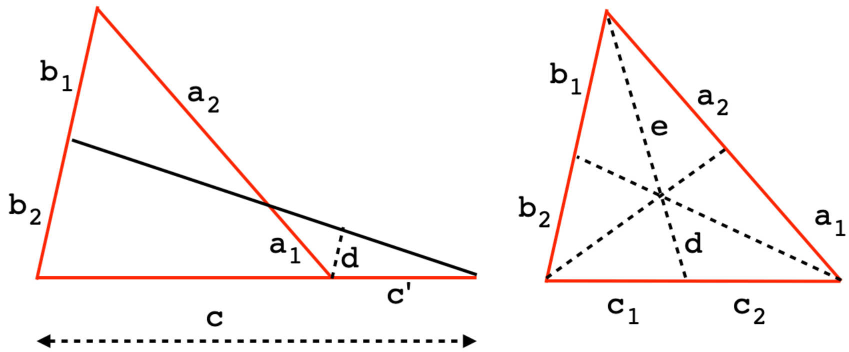
Our series is 1/2 times $1 + 1/4 + \dots$, so $r = 1/4$. The sum is 4/3 and then our series is one-half that.

centroid from Menelaus

We can apply **Menelaus' theorem** to the case of the centroid. This gives a very simple equation to identify how far along the median the point P is.

The original theorem was (left panel).

$$\frac{a_1}{a_2} \cdot \frac{b_1}{b_2} \cdot \frac{c}{c'} = 1$$



Write the same expression for the median coming down from the top vertex in the right panel and the triangle with sides b and c_1.

$$\frac{d}{e} \cdot \frac{b_1}{b_2} \cdot \frac{c}{c_2} = 1$$

$b_1/b_2 = 1$ and $c/c_2 = 2$. Therefore

$$2d = e$$

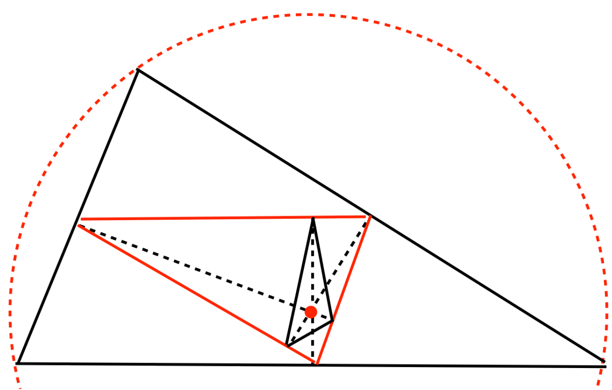
The point P lies one-third of the way up from the side to the corresponding vertex.

Chapter 10

Triangles in triangles

In the figure below, the outer black triangle lies on its circumcircle, where the circumcenter is the red point at the center.

The midpoints of the sides of the black outer triangle are joined to form a second triangle, in red.



We will show that the perpendicular bisectors of the sides of the outer triangle are the altitudes of the red triangle.

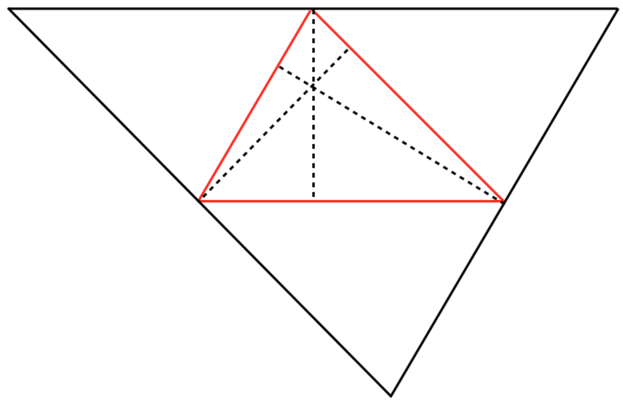
Furthermore, if the points where the altitudes meet the sides of the red triangle are joined to form a third inner triangle, those same altitudes are the angle bisectors of the inner triangle.

So for this arrangement, we have circumcenter concurrent (at the same point) with the orthocenter, which in turn is concurrent with the incenter.

Gauss and altitude

We will prove that the altitudes of a triangle are the perpendicular bisectors of a particular triangle which encloses it. The proof is due to Gauss.

Proof.



Draw the outer triangle, in black. Connect the midpoints of the sides to form an inner triangle, in red. Also draw the perpendicular bisectors of the outer triangle.

By the **midpoint theorem**, each side of the outer triangle is parallel to one side of the inner triangle and equal to twice its length. Since we have opposing sides equal and parallel, this gives three parallelograms in the figure.

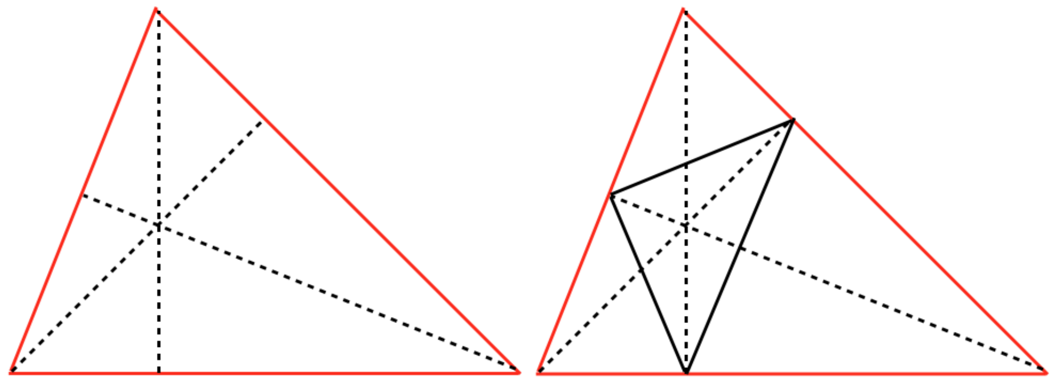
The perpendicular bisector of each black side (dotted line) is the altitude of the paired red side, because it starts from a vertex of the red triangle and meets the base at a right angle due to the parallel sides.

This shows that the circumcenter of the enclosing triangle is the orthocenter of the smaller, enclosed triangle. Since the circumcenter exists and is a single point, so is the orthocenter.

□

orthocenter

We have shown previously that the three altitudes meet at a single point, the orthocenter. The proofs include one from **Newton**, and the previous one (from **Gauss**).



Above we have drawn the altitudes (left panel) and then also connected the points where the altitudes meet the sides at right angles. We will prove that the dotted lines are the bisectors of the angles at the vertices of the small inset triangle.

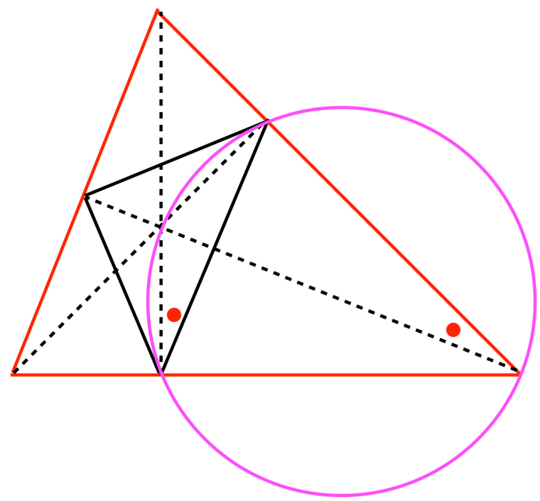
In other words, the incenter of the small triangle is the same point as the orthocenter of the bigger one.

Proof.

The key to the proof is to recognize that we can use a part of an altitude as the diameter of a circle. Draw the circle that has for its diameter the line segment connecting the orthocenter and one vertex of the large triangle.

Now consider the parts of the other two altitudes that terminate in right angles at the sides of the red triangle. I claim that these two points lie on the same circle.

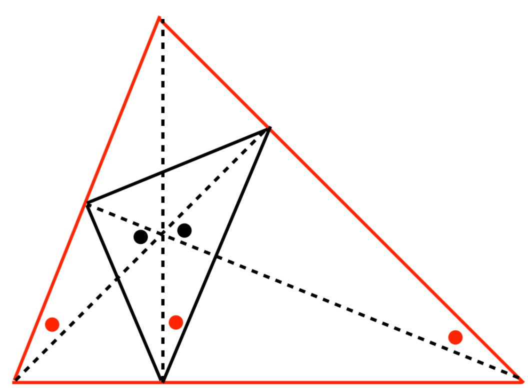
The reason is that, each one individually, taken together with the first two points, forms a right triangle. By the converse of Thales theorem, they must lie on the circle.



(The included side of the inner triangle is *not necessarily* perpendicular to the diameter, the first one looks so because the original triangle is nearly isosceles — see below).

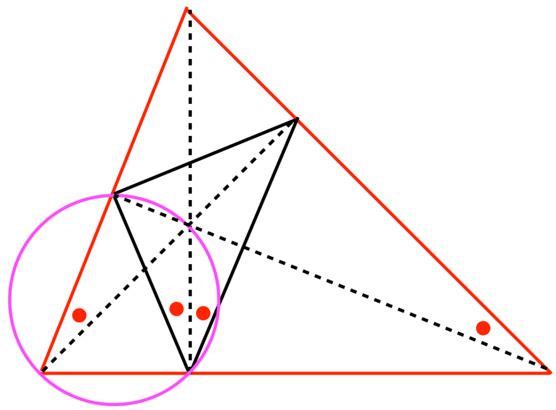
Now we can use the **theorem** about arcs that subtend an angle on the perimeter of the circle. The two angles on the circle marked with red dots correspond to the same arc of the magenta circle, so they are equal.

For the next step, we use vertical angles (marked with black dots) to show that two triangles are similar, since they also contain right angles. Therefore, we can mark a third angle as equal to the others with a red dot.



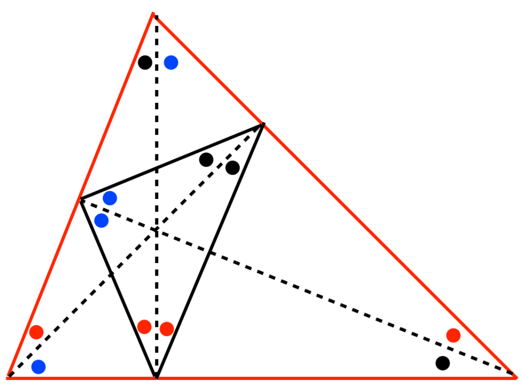
Finally we draw a circle for a different vertex. Now it is obvious that the solid black line is not necessarily perpendicular to the altitude.

Using the arc theorem, we find another equal angle, for a total of four angles marked with red dots. We see that the one vertex of the inner triangle is bisected into two equal angles marked with red dots.



But the same thing can be done for the two other vertices of the inner triangle. The pattern of the angles is the same.

With all the dots filled in:



This shows that the dotted lines are angle bisectors for the small triangle.

Thus, the orthocenter of the large triangle and the incenter of the triangle inscribed between the points where altitudes meet the base, are the same point.

□

nine-point circle

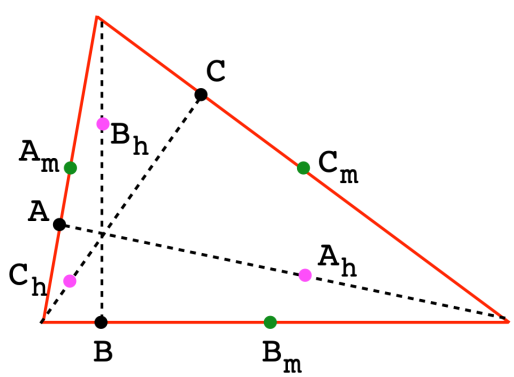
We've seen that the perpendicular bisectors of the sides and, separately, the angle bisectors of triangles, converge on (are concurrent with) points that are the centers of circles with interesting properties. These circles contain either the vertices of the triangle (circumcircle) or have the sides as tangents to the circle (incircle).

Now we investigate the altitudes and their point of convergence, the orthocenter. It turns out there is a special circle, but it does not have the orthocenter as its center.

Quite surprisingly, there are nine points on the circle. Three of these are midpoints of the sides. This means that all four categories of special points of a triangle are connected to circles of one kind or another.

The circle that goes through the midpoints of the sides also goes through the points where the altitudes of a triangle meet the sides, as well as the midpoints of that part of each altitude lying between the orthocenter and the corresponding vertex.

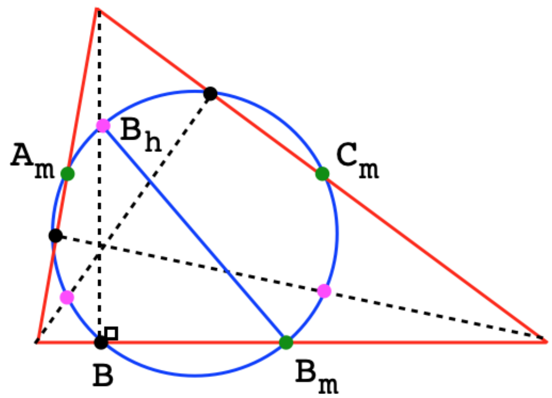
It's challenging to draw the figure. Some of the measurements may look a little off, but the logic will show that the circle indeed contains the nine points cited.



First, label the points where the altitudes meet the sides in right angles as A, B and C. Next, label the midpoints of the corresponding sides as A_m, B_m and C_m.

Place a point on each altitude such that it is the midpoint of the part of the altitude that extends from the orthocenter to the vertex. These are A_h, B_h and C_h.

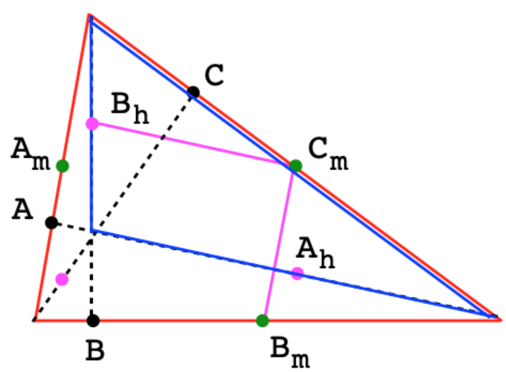
Now draw the circle that has B_mB_h as its diameter.



In the first part of the proof, we will show that this circle also goes through A_m, C_m and B. Each of the three points forms a right triangle where the hypotenuse is B_mB_h, the diameter of this circle.

Proof.

There is a triangle outlined in blue in the figure below, though its vertices aren't labeled. The top edge is the same as one side of the original triangle.



The midpoints of two of those sides are shown, they are the points B_h and C_m. By SAS similarity, it follows that B_hC_m is parallel to the altitude from A extending through A_h.

Likewise B_m and C_m are midpoints of two sides of the whole triangle, which means that B_mC_m is parallel to the third side, containing A and A_m.

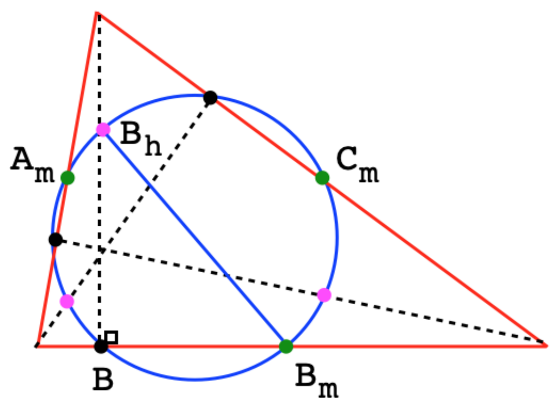
So B_hC_m is parallel to the altitude, which is then perpendicular to the side, which is in turn parallel to C_mB_m.

It follows that $B_hC_m \perp C_mB_m$ and so $\angle B_hC_mB_m$ is a right angle.

By the converse of Thales' circle theorem, C_m lies on the circle with B_mB_h as a diameter.

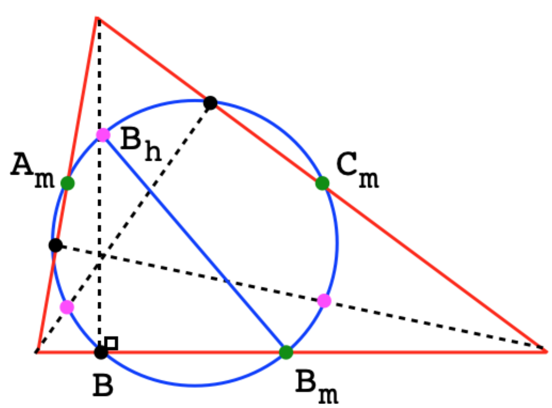
Since we started with B_hB_m as diameter, there is nothing special about the side containing C_m. Equivalent logic will show that A_m is on the same circle.

All that remains is B. But B is the right angle in $\triangle B_hBB_m$, so it is on the circle as well, for the same reason.



We have shown that the circle whose diameter is B_hB_m also goes through the three points A_m, C_m and B.

Switching perspective, we may say instead that the midpoint circle containing $A_mB_mC_m$ also contains points B and B_h.

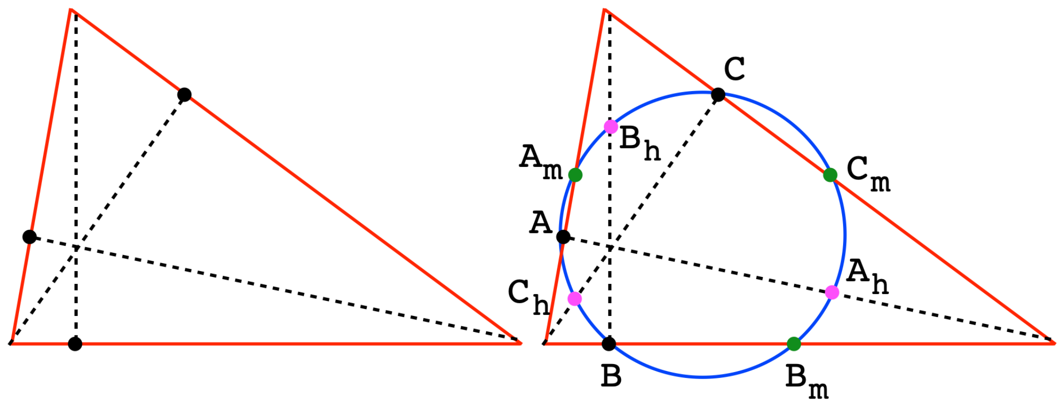


But, there is nothing special about B and B_h.

Because the midpoint circle is symmetric in containing all three midpoints, we might just as well take A_mA_h as the diameter of a circle. Then we could proceed to show that A, B_m and C_m all lie on the same circle, and after that follow with diameter C_mC_h in turn.

Instead, we just appeal to symmetry.

□



Note: It is pretty easy to show that any pair of pairs like B_h, B_m and C_h, C_m form a rectangle. That gives $B_hB_m = C_hC_m$ and so we have three equal diagonals for the circle as well as right angles at the midpoints as well as the bases of the altitudes.

Chapter 11

Euler line

It turns out that the orthocenter, centroid and circumcenter lie on a single line. This proof, due to Euler, is stunning.

For what's coming in this chapter, we need the following preliminary result: if two triangles share one angle and the two sides flanking that angle are in the same proportion, then all three sides are in the same proportion, and it follows that the two triangles are similar.

SAS for similar triangles

• If two triangles share an equal angle, and the two sets of flanking sides are both in proportion, then the two triangles are similar.

We proved this theorem **here**.

Here is a simple algebraic proof.

Proof.

Suppose the small triangle has sides a, b, c and the larger triangle has sides A, B, C with $A/a = B/b = k$. Suppose that $\angle \gamma$ flanked by a and b is also flanked by A and B.

Use the law of cosines. We have

$$c^2 = a^2 + b^2 - 2ab \cos \gamma$$

$$C^2 = A^2 + B^2 - 2AB \cos \gamma$$

Since $A = ka$ and $B = kb$

$$C^2 = (ka)^2 + (kb)^2 - 2(k^2ab)\cos\gamma$$
$$= k^2c^2$$

Hence $C/c = k$.

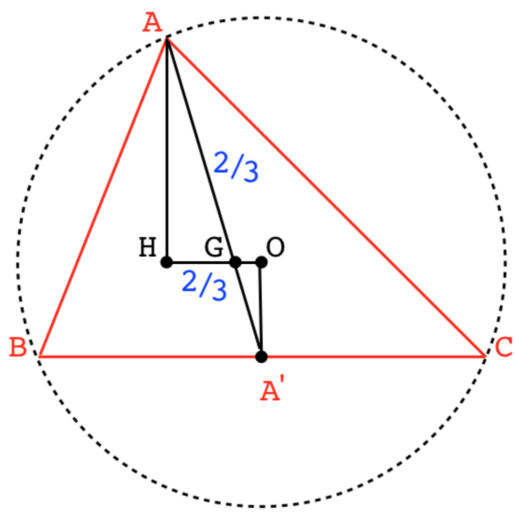
□

Euler line

We follow:

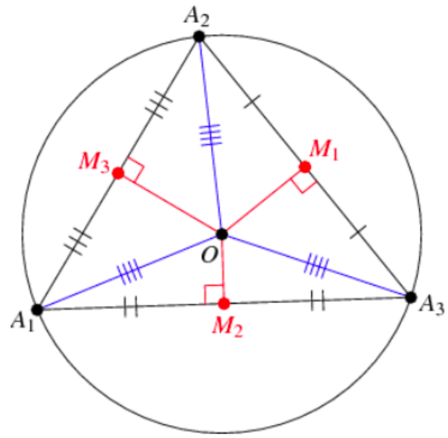
https://artofproblemsolving.com/wiki/index.php/Orthocenter

A copy of their figure:

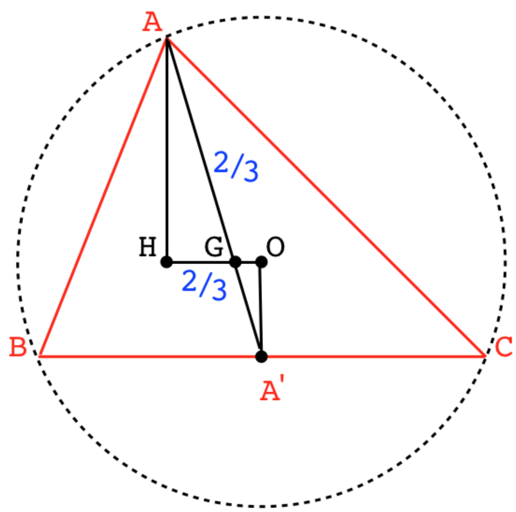


The point O is the circumcenter of the triangle: the center of the circle which contains all three vertices of the triangle.

Clearly, this circle has a center. The classic construction is to bisect each side (here BC is bisected at A'), and erect a perpendicular. The point where the three perpendiculars cross is the circumcenter, which is the center of the circle.



Assume we have done this and that point is O.



The next point, G, is the centroid. One way to find this point is to draw all three lines connecting vertices with the midpoints of the opposite side (AA'). However, if you recall, the distance from the vertex A to G is twice the distance from G to the midpoint A'. Hence we find point G using arithmetic: $AG = 2 \times A'G$.

Now, extend OG by twice its length, to H. ($2 \times OG = GH$). (It is an accident of the drawing that OH appears to be horizontal, this is not true for the general case.)

We will prove that H is the orthocenter of the original triangle.

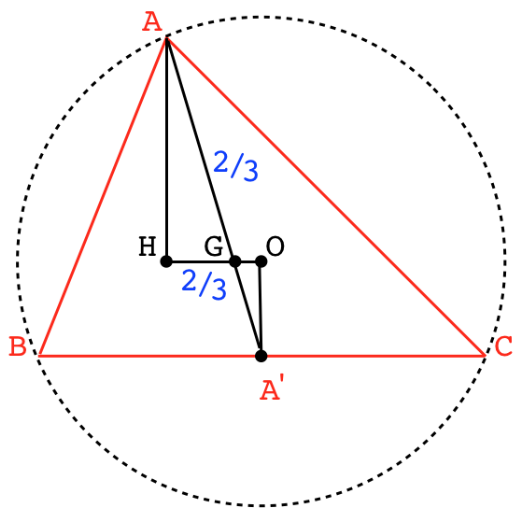
Proof.

Because AG is twice $A'G$ and GH is twice OG and the two triangles share the angle $\angle OGA'$ (equal to $\angle AGH$) between these two sides, they are similar triangles.

Therefore $\angle A'OG$ is equal to $\angle AHG$, as corresponding angles in similar triangles.

It follows that AH is parallel to $A'O$, by alternate interior angles. And since $A'O$ is perpendicular to the base, BC, the extension of AH to the base is also perpendicular to it.

Thus, AH is a part of the altitude from A to BC (the whole altitude is not shown).



We have shown elsewhere that the centroid is unique and the orthocenter is unique, and naturally, the circumcenter is unique.

The same construction could be done for either of the other two altitudes, proceeding first to O, then to G and each time ending at H.

So the argument is that the three altitudes all cross at a single point, and the line OGH crosses the three altitudes at a single point, H. It must be that H is the orthocenter, because if three lines that cross at a point, that point is unique, they can't also cross at some other point.

The orthocenter, centroid and circumcenter lie on a single line, and the distance from centroid to orthocenter is twice that from centroid to circumcenter.

□

We can also consider two other corresponding sides in the similar triangles. One extends from the vertex above down to the orthocenter, and the other extends from the midpoint of the opposing side to the circumcenter. The scale factor relating these two lengths is 2:1.

Chapter 12

Special circles

triangle area and radii for incircle and circumcircle

Acheson gives formulas that connect the area of a triangle (he uses the symbol Δ for the area), and the radius of either the incircle or the circumcircle.

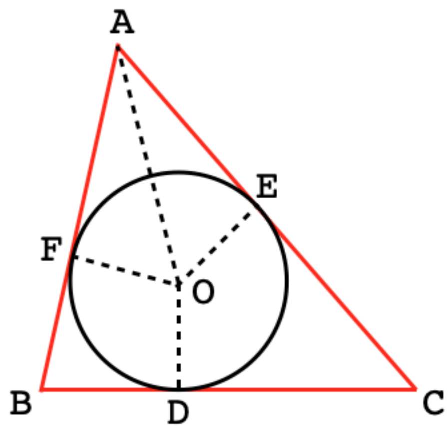
The first one is just a matter of algebra, but the second is gorgeous. It is really worth it to try to solve before you look at the answer. So, write down the answer, close the book and then try! Once again, an inspired diagram is everything.

$$r = \frac{2\Delta}{a+b+c}$$

$$R = \frac{abc}{4\Delta}$$

Let r be the radius of the incircle and a be the length of the base opposite vertex A.

Then the area of $\triangle BOC$ is equal to one-half $r \cdot a$.



So the area of the whole triangle is equal to one-half $r \cdot (a + b + c)$.

$$2\Delta = r \cdot (a + b + c)$$
$$r = \frac{2\Delta}{a + b + c}$$

Define the *semi-perimeter* s as

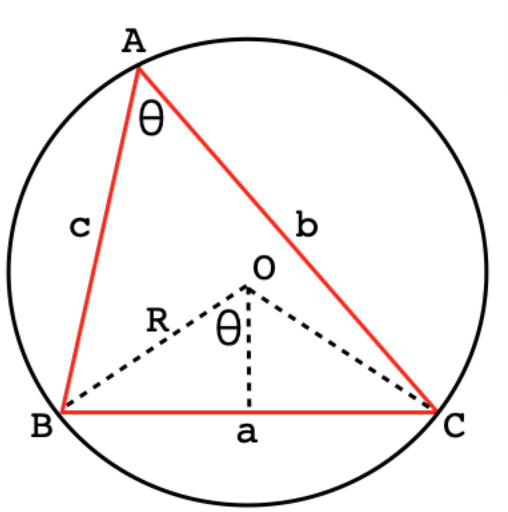
$$s = \frac{a + b + c}{2}$$

and then

$$r = \frac{\Delta}{s}$$
$$\Delta = rs$$

That's an interesting parallel, that this formula is so similar to that for the area of the circle. Here we have the radius of the incircle times the one-half the perimeter of the triangle. Of course, for a circle, we have the radius times one-half its perimeter as well.

For the second problem, the radius of the circumcircle, the key insight is in the diagram below. After that it's easy.



Proof.

The altitude to side b (not shown) has length $c \sin\theta$. So the area of the triangle is

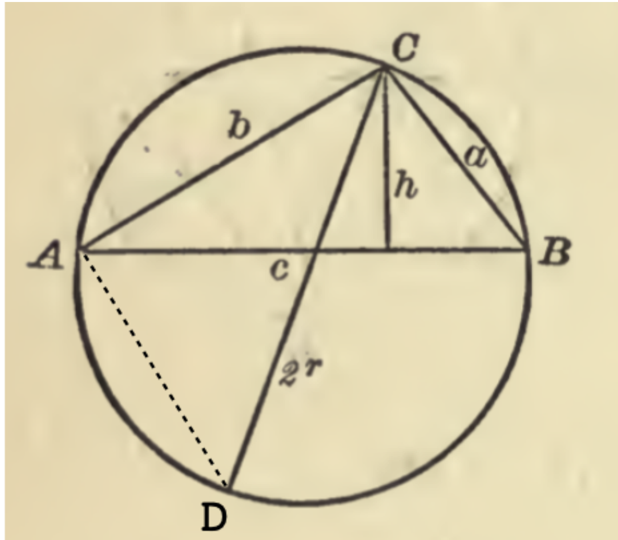
$$\Delta = \frac{1}{2} bc \sin\theta$$

But $\sin\theta = a/2$ divided by R so

$$\Delta = \frac{1}{2} bc \cdot \frac{a}{2R}$$
$$\Delta = \frac{abc}{4R}$$

□

Here is an alternate proof from Hopkins.



Proof. (Alternate).

As a preliminary matter, note that $\triangle ABC$ is any triangle, and the circle is its circumcircle, with radius r. Then the extension of the radius to D forms a right triangle $\triangle ACD$. Since $\angle B$ and $\angle D$ cut off the same arc of the circle, they are equal.

Therefore, $\triangle ACD$ is similar to the triangle formed by the altitude h and including side a. By similar triangles:

$$\frac{h}{a} = \frac{b}{2r}$$

$$h = \frac{ab}{2r}$$

Twice the area of the triangle is

$$2\Delta = ch$$

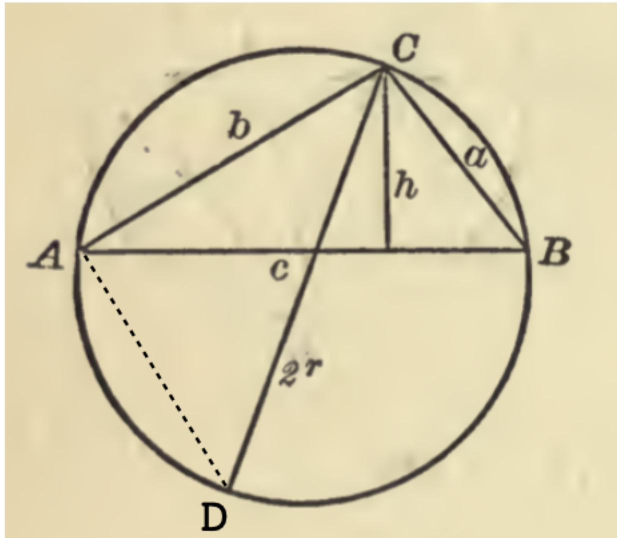
$$\Delta = \frac{abc}{4r}$$

□

Hopkins also notes that this result can be expressed purely in terms of the side lengths by using **Heron's formula** (which we introduced above and will say more about soon):

$$A^2 = s \cdot (s-a) \cdot (s-b) \cdot (s-c)$$

(where $s = (a+b+c)/2$).



So

$$r = \frac{abc}{4\sqrt{s \cdot (s-a) \cdot (s-b) \cdot (s-c)}}$$

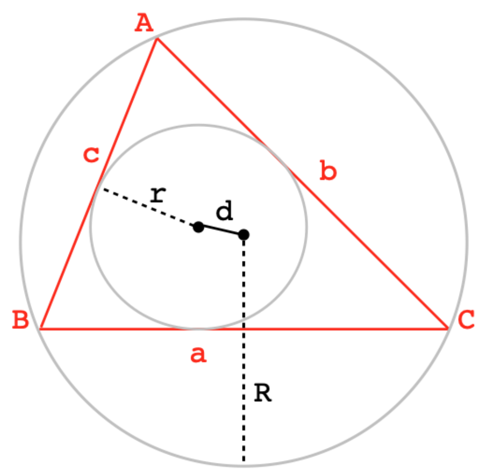
You should be able to show that

$$r = \frac{abc}{\sqrt{(a+b+c)(a+b-c)(a+c-b)(b+c-a)}}$$

Chapter 13

Euler's theorem

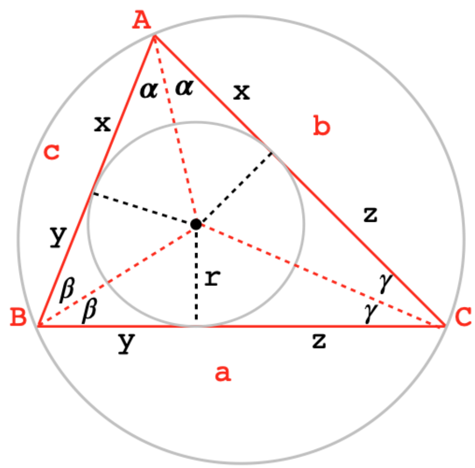
Euler has a huge number of important results. Euler's theorem in *geometry* relates to the circumcircle and incircle of a triangle. As before, the circumcircle is the circle that contains the three vertices, while the incircle is that circle which just fits inside the triangle. The respective radii will be R and r.



We have already seen that these circles exist and are unique. The circumcircle is constructed using the perpendicular bisectors of the sides. The incircle is drawn tangent to all three sides of the triangle; it has its center (the incenter) the same distance r away from any side.

We also showed previously that the angle bisectors of angles $\angle A$, $\angle B$ and $\angle C$ go through the incenter. If we divide the side lengths into two parts on either side of the tangent point to the incircle, then it is easy to show that the area of the triangle, label it K, is equal to $r(x+y+z)$. There are six smaller triangles, three pairs with equal area, of the form $rs/2$, where s is one of x, y, z.

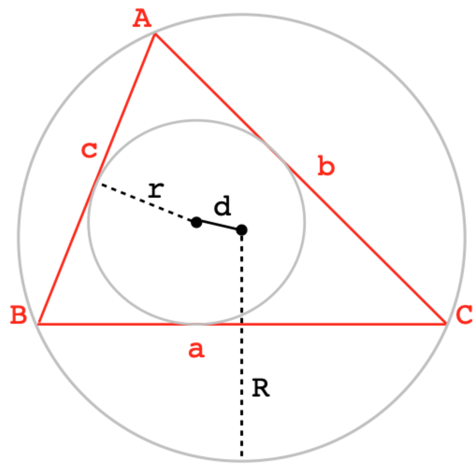
$$K = r(x+y+z)$$



We had another result about the area of the triangle earlier, relating K to the side lengths and to the radius of the circumcircle R. Using a standard formula for area we have that $2K = ab \sin C$. But $\sin C = c/2R$. So $2K = abc/2R$.

$$4KR = abc$$

Euler's result concerns the relationship between R and r. Going back to the first figure



We will (eventually) show that

$$d^2 = R(R - 2r) = R^2 - 2rR$$

But we will start by proving the inequality $R \geq 2r$, which is easily checked against the formula. Since d^2 is not negative, we have that $R(R - 2r) \geq 0$, and since R is positive, it must be that $R - 2r \geq 0$, which implies $R \geq 2r$,

derivation

We follow Nelsen (2008)

https://www.maa.org/sites/default/files/Nelsen2-0859469.pdf

We will first derive the formula:

$$xyz = r^2(x + y + z)$$

To begin with, Nelsen builds a rectangle with scaled triangles, just as for the sum of angles theorem, as well as one proof of the Pythagorean theorem and Ptolemy's theorem.

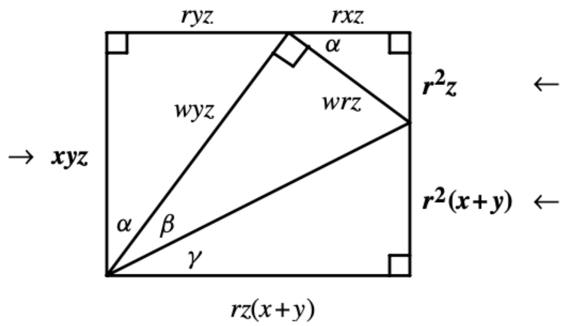
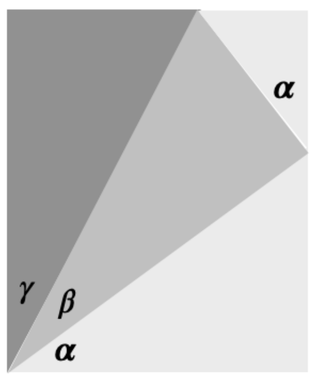


Figure 3 $xyz = r^2(x + y + z)$

I am going to modify the proof a bit. For starters, we'll run the triangles from α at the bottom, then β and γ. This simply amounts to a reflection and rotation of the version he has, and makes the orientation consistent with the way we've done it before.

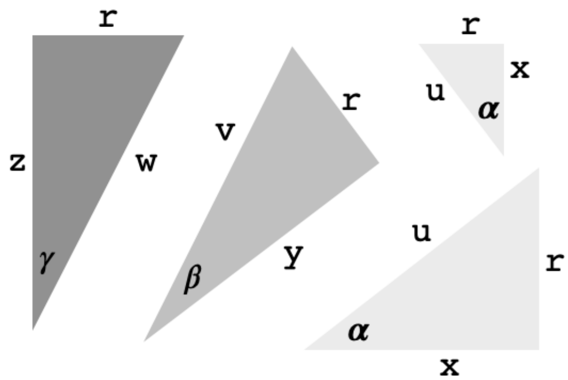
More important, we are going to define w differently, since I want to have symbols for the hypotenuse of all three triangles.



Let's construct the box containing four triangles. We know the three half-angles, α, β and γ, add up to one right angle.

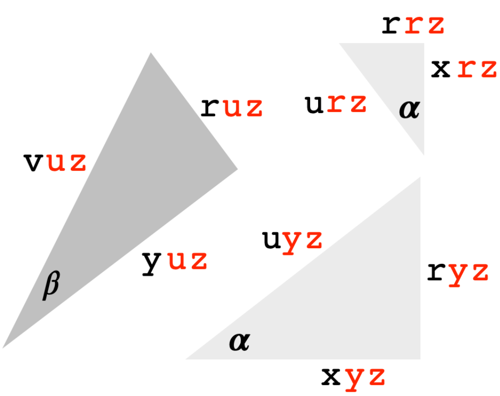
We also know that the tangents of these angles are r/x, r/y and r/z, respectively. It is simple to show that the fourth triangle has angle α and we will distinguish the two triangles with angle α as $\triangle_{\alpha 1}$ and $\triangle_{\alpha 2}$.

You can see that u, v and w are labels for each hypotenuse.



As always, the trick here is to choose the right scale factors. The identity that we're looking for comes from a comparison of the bottom of the rectangle to the top.

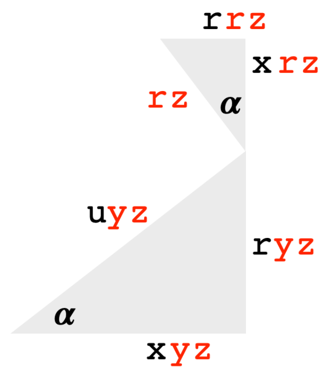
The bottom of the rectangle consists of side x of $\triangle_{\alpha 1}$. We want that to be xyz (compare with the formula at the beginning of this section), hence the appropriate factor is yz.



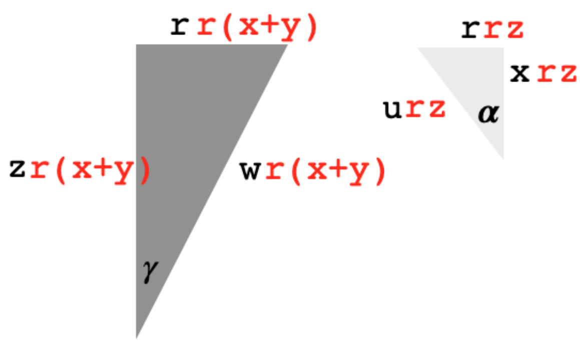
Next, the hypotenuse of $\triangle_{\alpha 1}$ is also the base of $\triangle_{\beta}$, so they must be equal. We see then, that the scale factor for the second triangle must be uz, so that the base of $\triangle_{\beta}$ does match the hypotenuse of $\triangle_{\alpha 1}$.

That, in turn gives the scaled hypotenuse for the small $\triangle_{\alpha 2}$ as ruz (or urz), which means that its scale factor must be rz.

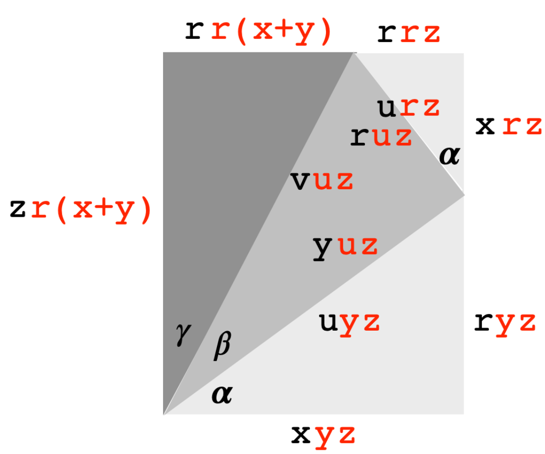
If you look at what we have for the two lengths on the right side of the rectangle, from the first $\triangle_{\alpha 1}$ we get ryz and from the second we get xrz. The sum is $rz(x+y)$.



The final triangle ($\triangle_\gamma$) must be scaled to match. Looking at its base (left side of the rectangle), we want $rz(x+y)$ so we deduce the necessary scale factor is $r(x+y)$.



Assembling the parts:



Comparing the top and bottom sides (which must be equal, since this *is* a rectangle), we obtain

$$xyz = r^2(x + y + z)$$

checking our rectangle

It you look at the last (admittedly messy) figure, you may notice that there is one length we did not explicitly write or try to match. That is the hypotenuse of $\triangle_\gamma$.

The original triangle had hypotenuse w and the scale factor is $r(x + y)$. This needs to match what is written on the opposite side of the line segment, namely

$$wr(x + y) = vuz$$

Somehow this must also be an equality.

We don't explicitly need it for our proof (the other scaling takes care of everything), but it's a loose end that I'd like to tie up.

We notice that $r/u = \sin\alpha$, and even more, if we multiply both sides by r we can get $r/w = \sin\gamma$, so let's try doing that and then rearranging a bit

$$\begin{aligned}\frac{r}{u} \cdot \frac{r}{v} \cdot (x + y) &= \frac{r}{w} \cdot z \\ &= z \sin\gamma\end{aligned}$$

γ is complementary to $\alpha + \beta$ so

$$\sin\gamma = \cos\alpha + \beta$$

which by sum of angles is

$$\sin\gamma = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

Let's rewrite this as

$$\frac{r}{w} = \frac{x}{u} \cdot \frac{y}{v} - \frac{r}{u} \cdot \frac{r}{v}$$

That's a great simplification since we can cancel u and v on the bottom of the original left-hand side, giving

$$r^2(x + y) = z(xy - r^2)$$

But this is just our formula back again.

$$r^2(x+y+z) = xyz$$

$\square$

We scaled the triangles by matching equal sides on different triangles, guided by our desire to have the distance on the left and right sides match.

We obtain the desired equation by noting that since it's a rectangle, the top and bottom sides must also be equal. Alternatively, when we consider the hypotenuse shared by $\triangle\beta$ and $\triangle\gamma$, the same equation comes out. It is all consistent.

combining results

So far we have three preliminary results or lemmas

$$K = r(x+y+z)$$

$$4KR = abc$$

$$xyz = r^2(x+y+z)$$

Combine the first and third

$$xyz = Kr$$

The challenge now is to connect the side lengths abc with the components xyz.

The geometric mean comes to our rescue. Recall that the arithmetic mean ($A.M.$) is always greater than or equal to the geometric mean ($G.M.$).

$$A.M. \geq G.M.$$

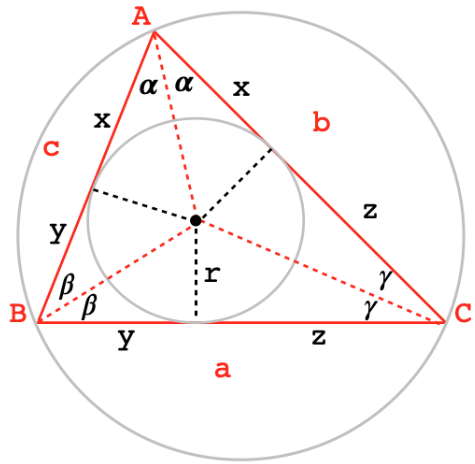
For example

$$\frac{x+y}{2} \geq \sqrt{xy}$$

Putting all three pairwise combinations together

$$\frac{(x+y)(y+z)(z+x)}{8} \geq \sqrt{xy} \cdot \sqrt{yz} \cdot \sqrt{zx}$$

The right-hand side is just xyz. Referring to the figure, we see that the numerator on the left-hand side is simply cab or abc.



It follows that

$$abc \geq 8 \cdot xyz$$

So from the second lemma ($4KR = abc$):

$$KR \geq 2 \cdot xyz$$

From the combined first and third ($xyz = Kr$):

$$KR \geq 2Kr$$

which yields

$$R \geq 2r$$

□

Show that $R = 2r$ for the equilateral triangle. We leave this as an exercise.

Hint: Draw the incircle, then draw the right triangle with hypotenuse extending from the incenter to any vertex, and one side as the adjoining radius r.

The triangle is a 30-60-90 triangle so the ratio of the hypotenuse to r is $2 : 1$. If you go back to our previous work, we showed that the same same line segment (the hypotenuse of this triangle) is equal to R.

equality form of Euler's theorem

We claimed that

$$d^2 = R(R - 2r) = R^2 - 2rR$$

Which can be written in various ways

$$R^2 - d^2 = (R+d)(R-d) = 2rR$$

and even more elegantly as

$$\frac{1}{R-d} + \frac{1}{R+d} = \frac{1}{r}$$

since

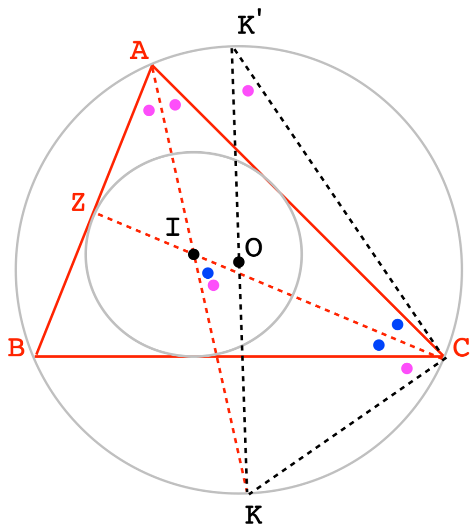
$$R + d + R - d = 2R = \frac{R^2 - d^2}{r}$$

Our proof follows

https://www.cut-the-knot.org/triangle/EulerIO.shtml

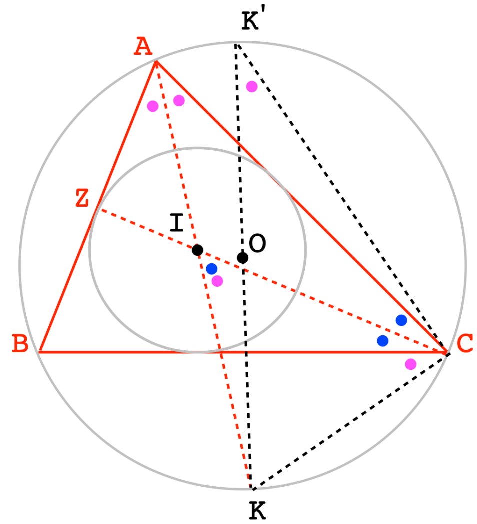
We draw $\triangle ABC$ with its incircle and incenter I. O is the circumcenter (center of the circumcircle). d is the distance between I and O.

Proof.

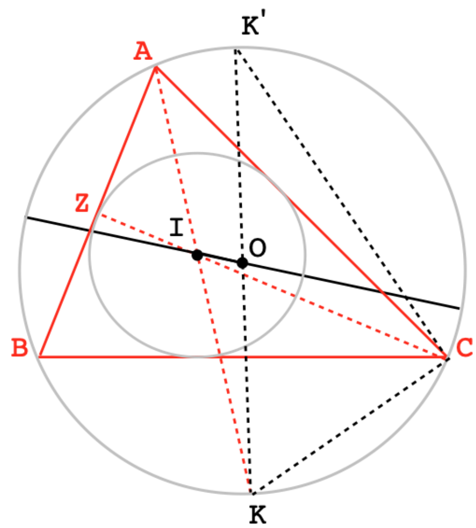


We begin by considering some angles in the figure above. As we showed elsewhere, the incenter is the point where the angle bisectors for the vertex angles of the original $\triangle ABC$ meet. The half-angles at A are labeled with magenta dots, while those at C are labeled with blue.

Two other angles are labeled with magenta by the inscribed angle theorem, one at K' and one next to vertex C. Then, we notice something interesting, that $\angle CIK$ is the external angle for $\triangle AIC$, so it gets a blue plus a magenta dot. Thus $\triangle ICK$ is isosceles, with $IK = KC$.



Draw the diameter of the circle that passes through I and O. I divides this diameter into two parts of lengths $R + d$ and $R - d$ (since the sum must be $2R$).



By our theorem on intersecting chords, the products of the parts are equal, namely

$$(R + d)(R - d) = AI \cdot IK$$

Since the left-hand side is $R^2 - d^2$, what remains to be proved is that the right-hand side is equal to $2Rr$.

We substitute for IK in the previous expression $AI \cdot IK = AI \cdot KC$.

Now, $\triangle AIZ$ and $\triangle KK'C$ are both right triangles containing the half-angle at A. Hence they are similar, with ratios:

$$\frac{AI}{IZ} = \frac{KK'}{KC}$$

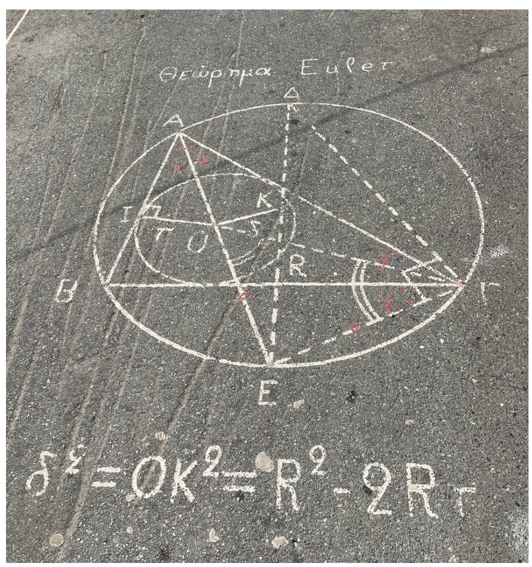
Cross-multiplying

$$AI \cdot KC = KK' \cdot IZ = 2R \cdot r$$

$\square$.

$$R^2 - d^2 = (R + d)(R - d) = 2rR$$

Drawn on a small street in Athens.



That is indeed (part of) our proof.

Part IV

Vectors without coordinates

Chapter 14

Simple vectors

length and direction

In this chapter and the next we use vectors to solve several problems in plane geometry. Some of these are difficult to solve by other methods.

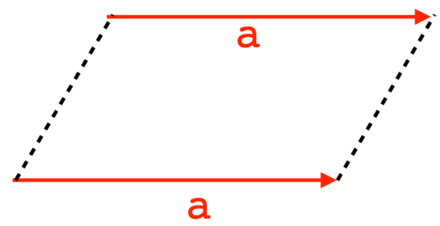
Your basic vector is a mathematical construct that represents both a length and a direction in the plane (2D), or in space (3D). It can be helpful to think of them as arrows extending from a tail up to the head.

A vector might represent an object's position, or its velocity, or the direction and strength of the electric field at some location.

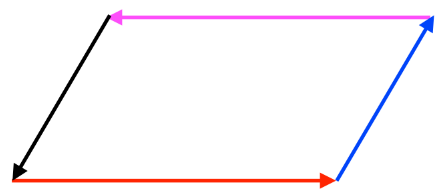
Normally, this idea is only introduced after defining a coordinate system, such as the orthogonal x- and y-axes in the plane. Using a coordinate system, the length and direction can be computed as the difference in coordinates between head and tail. If the difference in the x-coordinate is $\Delta x = x' - x$ and the difference in the y-coordinate is $\Delta y = y' - y$, then $\mathbf{v} = \langle \Delta x, \Delta y \rangle$.

But it is possible to get very interesting results even without a coordinate system, as we will see. Hence the motivation for including this chapter in a book on plane geometry.

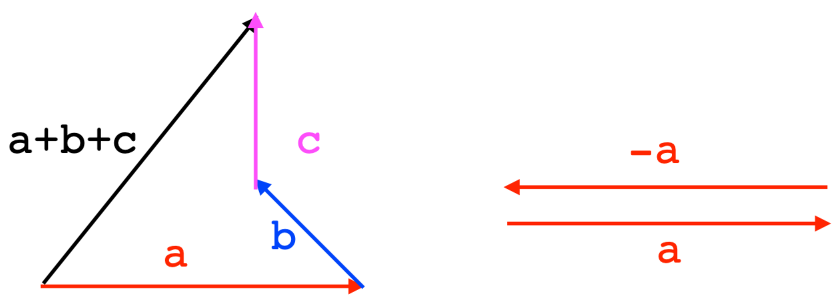
In physical applications it is probably more usual to think of the vector as fixed in space, whereas for mathematical ones, it may be moveable. For example, suppose a vector in the plane is duplicated and the copy moved anywhere else (without changing the orientation). When the ends are connected to form a quadrilateral, the result is always a parallelogram.



We will look at polygons, and use vectors to represent the lengths of sides, and also their relative orientations. For this purpose, we'll put vectors together head to tail to form the figure.



∘ Vector *addition* is defined as placing the tail of the second vector coincident with the head of the first.

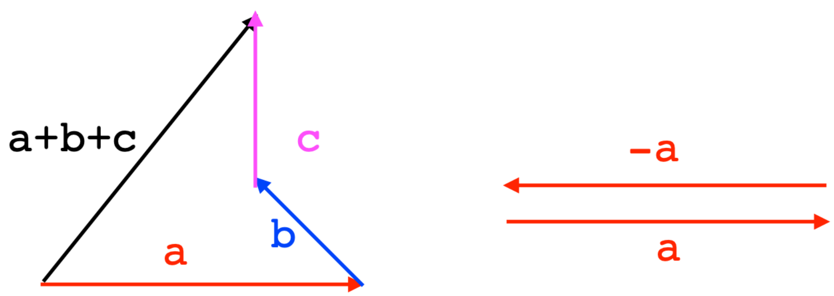


Addition of vectors is like a path that you might walk along, where the final vector connects the point where your journey started and the one where you end up. A bird might fly straight along $\mathbf{a} + \mathbf{b} + \mathbf{c}$ — "as the crow flies" — while you meander through some medieval town, puzzling at your map.

∘ The sum of the vectors comprising a closed polygon is zero.

Walking all the way around a polygon gets you back to where you started.

Applying this new idea to the parallelogram from before, we see that opposing sides are the negatives of each other and add to zero.



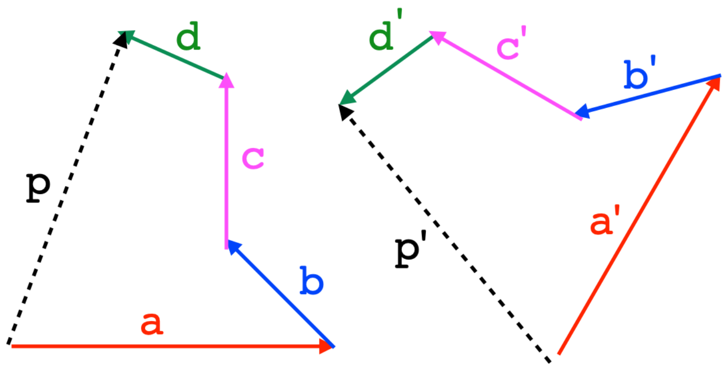
In the figure above, if we defined $\mathbf{d}$ to join the head of $\mathbf{c}$ with the tail of $\mathbf{a}$, closing the quadrilateral, then

$$\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} = 0$$

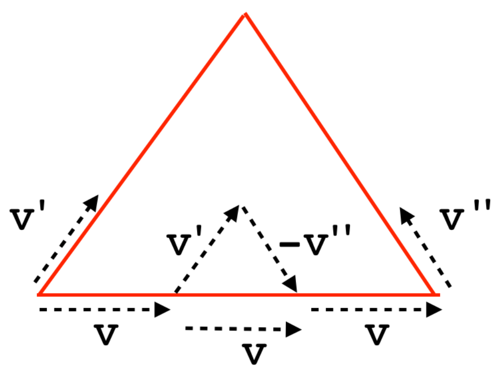
○ The negative of a vector is defined to be $-\mathbf{a} + \mathbf{a} = \mathbf{0}$.

If you walk backward along the same path you'll end up where you started. And if we want to know whether $\mathbf{u} = \mathbf{v}$, we can ask whether $-\mathbf{u} + \mathbf{v} = 0$.

○ A transformation such as rotation can be carried out by rotating each component vector of a figure, one after the other. If a vector $\mathbf{p} = \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}$ is transformed by the rotation denoted by $'$ then $\mathbf{p}' = \mathbf{a}' + \mathbf{b}' + \mathbf{c}' + \mathbf{d}'$.



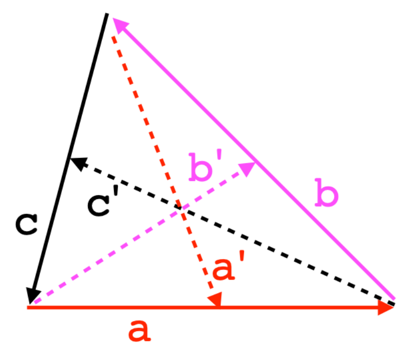
In the figure below we have an equilateral triangle. $\mathbf{v}$ is defined as 1/3 of the side length, and its orientation the same as the bottom side, from left to right. The operation of rotation through an angle is represented by a prime. Here, prime is a 60 degree rotation counter-clockwise. In another problem it will be defined to be a rotation of 90 degrees.



Using this figure for the definition, you should be able to see that $\mathbf{v}'''$ (*triple*-prime) is equal to $-\mathbf{v}$.

vector sums

In an arbitrary triangle, draw the three vectors that connect the vertices with the sides opposite. These vectors are called medians in geometry, and they cross at a unique point called the centroid. For now, we just label them with the same letter as the side they meet, adding a prime.



To prove: the sum of the three primed vectors is zero.

As we said above, the sum of the paths for a closed polygon is zero:

$$\mathbf{a} + \mathbf{b} + \mathbf{c} = 0$$

Scaling by a constant k doesn't change the result:

$$k\mathbf{a} + k\mathbf{b} + k\mathbf{c} = 0$$

Following closed paths, we can write the following equations:

$$\mathbf{a}' + \mathbf{a}/2 + \mathbf{b} = 0$$

$$\mathbf{b}' + \mathbf{b}/2 + \mathbf{c} = 0$$

$$\mathbf{c}' + \mathbf{c}/2 + \mathbf{a} = 0$$

Adding these equations, we find that taking the third term from each gives zero. Similarly, taking the second term from each also gives zero, replacing $k = 1/2$. So finally

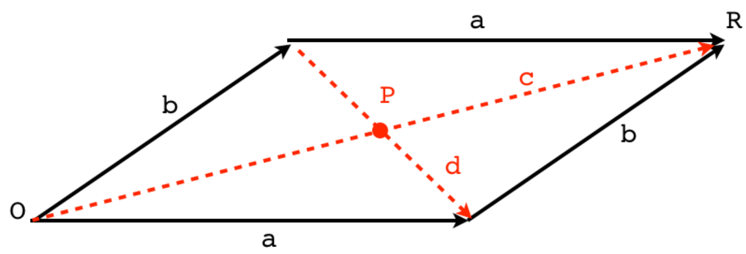
$$\mathbf{a}' + \mathbf{b}' + \mathbf{c}' = 0$$

□

parallelogram

To prove: the two diagonals of a parallelogram cross at their midpoints.

Suppose $\mathbf{c}$ and $\mathbf{d}$ are diagonals as shown below.



Clearly, $\mathbf{a} + \mathbf{b} = \mathbf{c}$.

Finding $\mathbf{d}$ takes some thought. $\mathbf{d}$ is that vector, which when added to $\mathbf{b}$, gives $\mathbf{a}$.

$$\mathbf{b} + \mathbf{d} = \mathbf{a}$$

So

$$\mathbf{d} = \mathbf{a} - \mathbf{b}$$

Follow $\mathbf{b} + \mathbf{d}/2$ to P. That is

$$\mathbf{b} + \frac{\mathbf{a} - \mathbf{b}}{2} = \frac{\mathbf{a} + \mathbf{b}}{2}$$

But that is one-half of $\mathbf{c}$.

Similarly

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

□

Varignon

Varignon's theorem concerns a general quadrilateral, with four vertices whose positions can be anywhere.

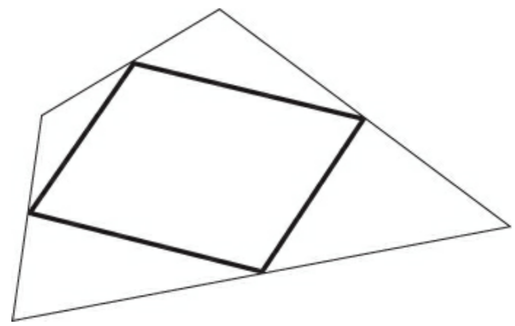


Fig. 50 Varignon's theorem.

The theorem states that when the *midpoints* of the sides of the quadrilateral are connected, the result is a parallelogram.

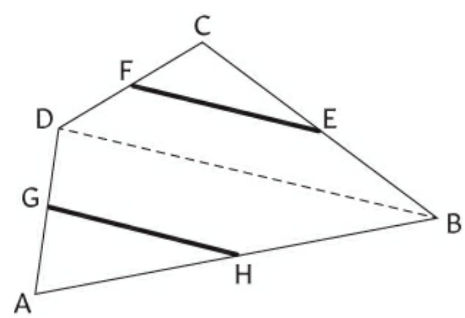


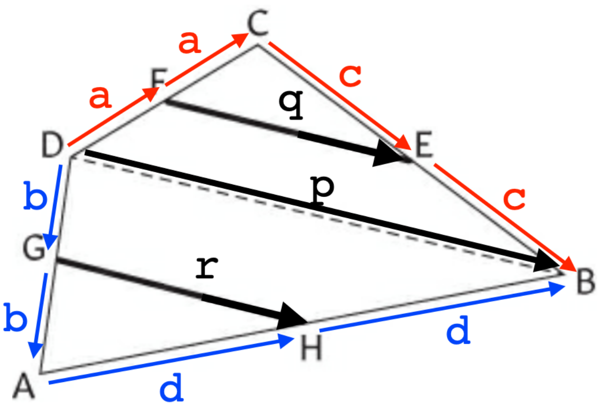
Fig. 51 Proof of Varignon's theorem.

$EFGH$ is a parallelogram. With the midpoint theorem from similar triangles, the proof is trivial.

$$GH \parallel DB, \qquad FE \parallel DB \quad \Rightarrow \quad GH \parallel FE$$

and $GH = FE = DB/2$.

Let's suppose we don't have the midpoint theorem and apply vectors to this problem.



We use the fact that E, F, G, and H are midpoints to draw the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$, as well as $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$.

For the lower $\triangle ABD$, we get from point D to point B by one of three paths: $\mathbf{p}$, $\mathbf{b}+\mathbf{r}+\mathbf{d}$ or $2\mathbf{b}+2\mathbf{d}$. Since these all go from point D to point B, they must be *equal*.

We have two additional paths that are also equal to the starting three: $\mathbf{a}+\mathbf{q}+\mathbf{c}$ or $2\mathbf{a}+2\mathbf{c}$.

Equate

$$2\mathbf{b}+2\mathbf{d} = 2\mathbf{a}+2\mathbf{c}$$

$$\mathbf{b}+\mathbf{r}+\mathbf{d} = \mathbf{a}+\mathbf{q}+\mathbf{c}$$

Subtract one-half of the first equation from the second to obtain

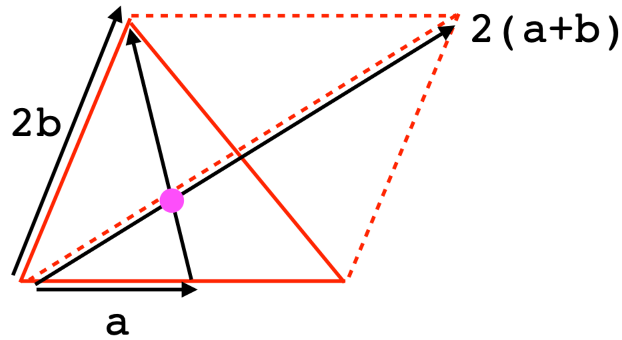
$$\mathbf{r} = \mathbf{q}$$

The opposing sides of the new figure, FE and GH, are equal *vectors*. That means FE and GH point in the same direction and are the same length. It follows that $EFGH$ is a parallelogram.

$\square$

Ceva

Ceva's theorem says that if we draw a line from each vertex of a triangle to the midpoint of the opposite side, the lines cross at a single point called the centroid, and that point is 1/3 of the distance from the side and 2/3 of the distance from the vertex.



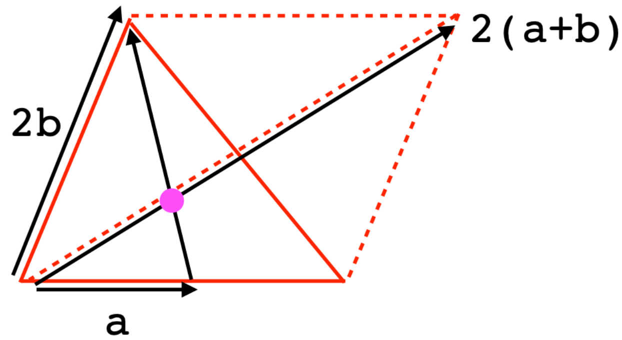
In this diagram the point in magenta is the centroid of the red triangle; as we said, it is 1/3 of the way up from the base.

We can derive this result using vectors as follows. The origin is at the bottom left. The base of the triangle is $2\mathbf{a}$, so $\mathbf{a}$ extends to the midpoint of the side. The second side is $2\mathbf{b}$.

The vector going up from the base, starting from the "head" of $\mathbf{a}$, and passing through the centroid, has the vector formula $2\mathbf{b} - \mathbf{a}$. To find this, I simply ask, what is the vector that when added to $\mathbf{a}$, gives $2\mathbf{b}$?

When a rotated version of the triangle is added (dotted lines), the diagonal passes through the midpoint of the third side, by a standard property of parallelograms. In vector notation, that diagonal is $2(\mathbf{a}+\mathbf{b})$, or $2\mathbf{a}+2\mathbf{b}$, so the diagonal to the midpoint is just $\mathbf{a} + \mathbf{b}$.

So then we say: there is no reason to think that one side of the triangle is any closer to the centroid than the other two. We expect then that the ratio of the distance from the side to the centroid compared to the length of the whole median, is the same for all three medians. Let us call that ratio r.



We compare two paths to the centroid, they must be equal:

$$\mathbf{a} + r(2\mathbf{b} - \mathbf{a}) = (1 - r)(\mathbf{a} + \mathbf{b})$$

$$\mathbf{a} + 2r\mathbf{b} - r\mathbf{a} = \mathbf{a} + \mathbf{b} - r\mathbf{a} - r\mathbf{b}$$

$$2r\mathbf{b} = \mathbf{b} - r\mathbf{b}$$

$$2r = 1 - r, \qquad r = \frac{1}{3}$$

The same result can be obtained using as one of the paths **2b** and back down to the centroid.

$$\mathbf{a} + r(2\mathbf{b} - \mathbf{a}) = \mathbf{b} + r(2\mathbf{a} - \mathbf{b})$$

$$\mathbf{a} + 2r\mathbf{b} - r\mathbf{a} = \mathbf{b} + 2r\mathbf{a} - r\mathbf{b}$$

$$\mathbf{a} + 3r\mathbf{b} = \mathbf{b} + 3r\mathbf{a}$$

$$3r = 1, \qquad r = \frac{1}{3}$$

□

Those are simple, pretty results. We look at two spectacular applications in the next chapter.

Chapter 15

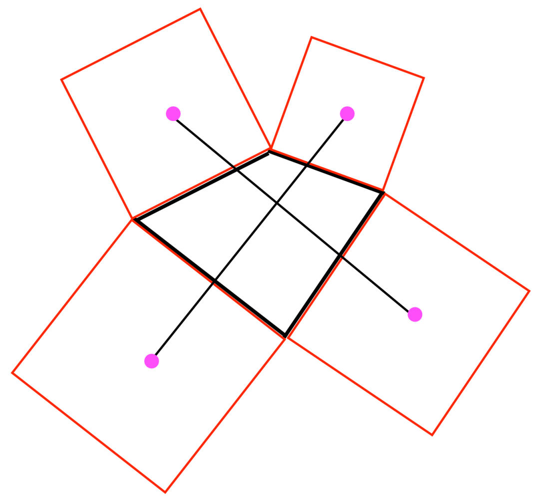
Napoleon's theorem

We continue looking at vectors, without coordinates, and apply them to two difficult problems.

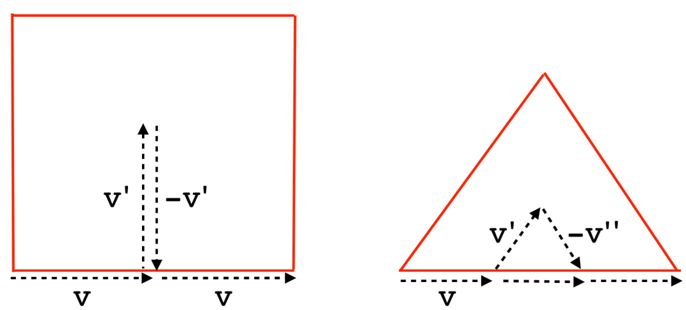
sides of a quadrilateral

We choose any four points in the plane to form a generalized quadrilateral, so the vertices could be anything, and the vectors corresponding to the sides could be anything as well.

Draw a square on each side of the quadrilateral and connect the centers of opposing squares, as shown in the diagram below. The resulting vectors will be perpendicular and equal in length, no matter which points are chosen initially.

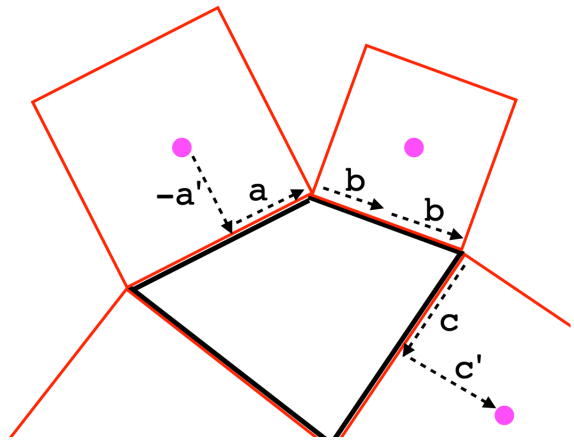


We are going to reason about these paths using vectors but not coordinates. For this first problem, we are concerned with squares. The path from vertex to center and back to the next vertex is shown in the left panel, below.



v′ is defined to be **v** transformed by a quarter-turn counter-clockwise. −**v**′ is **v**′, but in the opposite direction.

Since the sides can be any length, rather than **v**, we will use vectors **a** through **d** for the four different sides.



So the first path, from center to center is

$$\mathbf{p} = -\mathbf{a}' + \mathbf{a} + \mathbf{b} + \mathbf{b} + \mathbf{c} + \mathbf{c}'$$

And the second is

$$\mathbf{q} = -\mathbf{b}' + \mathbf{b} + \mathbf{c} + \mathbf{c} + \mathbf{d} + \mathbf{d}'$$

Notice that the first path goes roughly left-to-right, while the second goes top-to-bottom.

In order to check whether the two paths are at right angles, we should rotate the first path by one-quarter turn counter-clockwise. If the theorem is correct, $\mathbf{p}'$ should run in the exact opposite direction as $\mathbf{q}$, with the same length. They should add to give zero.

First path, rotated:

$$\mathbf{p}' = -\mathbf{a}'' + \mathbf{a}' + \mathbf{b}' + \mathbf{b}' + \mathbf{c}' + \mathbf{c}''$$

Each component has gained a prime ($'$). That's all it takes to rotate a vector path, just rotate each component.

Now, $-\mathbf{a}'' = \mathbf{a}$ and $\mathbf{c}'' = -\mathbf{c}$ (rotating 180 is the same as minus). Also, it makes no difference in which order you do the operations, whether first minus and then prime, or vice-versa.

Added together

$$\mathbf{p}' + \mathbf{q} = \mathbf{a} + \mathbf{a}' + \mathbf{b}' + \mathbf{b}' + \mathbf{c}' - \mathbf{c} - \mathbf{b}' + \mathbf{b} + \mathbf{c} + \mathbf{c} + \mathbf{d} + \mathbf{d}'$$

The question is, do these add up to zero? If so, we can conclude that the original two vector paths were orthogonal and the same length.

We can knock out the full path of the quadrilateral, as we end up where we started so that sum should be zero.

We subtract $\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} = 0$, leaving:

$$\mathbf{p}' + \mathbf{q} = \mathbf{a}' + \mathbf{b}' + \mathbf{b}' + \mathbf{c}' + -\mathbf{c} + -\mathbf{b}' + \mathbf{c} + \mathbf{d}'$$

Similarly $\mathbf{a}' + \mathbf{b}' + \mathbf{c}' + \mathbf{d}' = 0$. Subtract again:

$$\mathbf{p}' + \mathbf{q} = \mathbf{b}' - \mathbf{c} - \mathbf{b}' + \mathbf{c}$$

What's left is zero, so the whole thing is zero.

We conclude that the vectors which connect opposing centers are orthogonal and the same length.

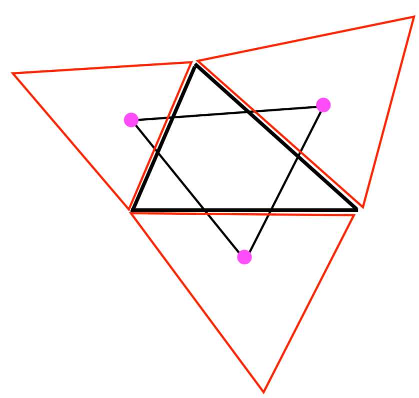
□

That's pretty amazing. The theorem is called van Aubel's theorem (1878). I learned about this proof and the theorem here:

https://www.youtube.com/c/MathyJaphy/videos

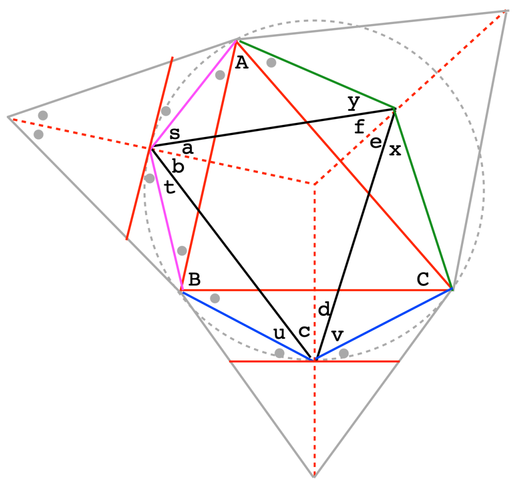
Among other things, the animation is terrific.

sides of a triangle

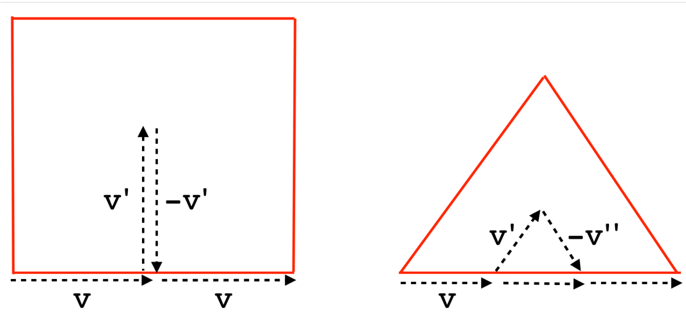


This problem is similar to the last one. On the three sides of any triangle, we construct equilateral triangles. Find the "center" of each triangle, and connect them. The result is an equilateral triangle.

I saw this problem in another book and worked on it for quite a while (days) without getting anywhere.



The vector approach provides a simple answer. The right panel, below, shows the idea. First, we need to find the center of an equilateral triangle. Luckily, all the centers — centroid, circumcenter, orthocenter and incenter — are the same point for an equilateral triangle.



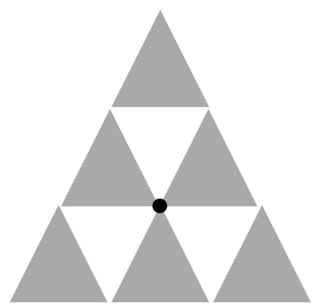
That point lies on a central line of the triangle and is one-third of the way up from the base.

To get there, we will use rotations that are 60 degrees, and define those as the transformation called prime. As shown, $\mathbf{v}'$ is a 60 degree rotation counter-clockwise from any starting vector $\mathbf{v}$.

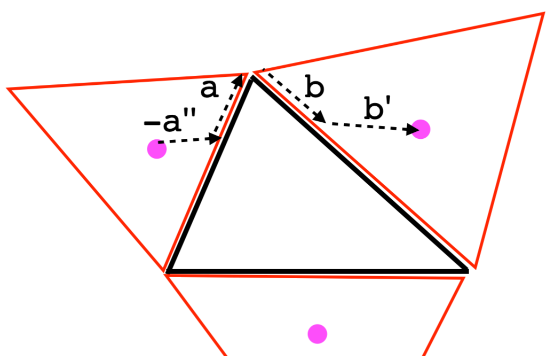
A clever idea is to go to the centroid by addition of one-third of a side plus the same vector, primed. The path along one edge is as shown above.

For any triangle, the centroid is the average of the three vertices. (This gives a different simple proof for r in the centroid problem, using coordinates).

From one vertex, take one-third of a step toward each of the other two vertices. Here is a visual proof without words.



Using this approach, we construct two paths as shown. One is from the center of the triangle with sides constructed from $\mathbf{a}$ to the one with sides $\mathbf{b}$.



$$\mathbf{p} = -\mathbf{a}'' + \mathbf{a} + \mathbf{b} + \mathbf{b}'$$

The other is (by symmetry)

$$\mathbf{q} = -\mathbf{b}'' + \mathbf{b} + \mathbf{c} + \mathbf{c}'$$

The big question is then: how do we decide when two vectors form the sides of an equilateral triangle? One set of criteria is that the two sides should be the same length, and flank an angle of 60 degrees.

If you look at the graphic, you'll see that $\mathbf{p}$ points toward the vertex of interest, while $\mathbf{q}$ points away from it (vectors are head to tail). Therefore, if we turn the first vector $\mathbf{p}$ counter-clockwise, it should point in the exact opposite direction from $\mathbf{q}$, and then the sum $\mathbf{p}' + \mathbf{q}$ will be $\mathbf{0}$. Let's see.

Starting with

$$\mathbf{p} = -\mathbf{a}'' + \mathbf{a} + \mathbf{b} + \mathbf{b}'$$

After the rotation, each one has acquired a prime. The first component, $-\mathbf{a}''$, turns into $-\mathbf{a}''' = \mathbf{a}$.

$$\mathbf{p}' = \mathbf{a} + \mathbf{a}' + \mathbf{b}' + \mathbf{b}''$$

Add together:

$$\mathbf{p}' + \mathbf{q} = \mathbf{a} + \mathbf{a}' + \mathbf{b}' + \mathbf{b}'' - \mathbf{b}'' + \mathbf{b} + \mathbf{c} + \mathbf{c}'$$

Now, $\mathbf{a} + \mathbf{b} + \mathbf{c} = 0$ and similarly for the primed versions, so those terms all drop out, leaving $\mathbf{b}'' - \mathbf{b}''$, which is of course, zero!

We have proven that the vectors corresponding to two sides of our construct are the same length, and have an angle of 60 degrees between them. Thus, they form an equilateral triangle.

□

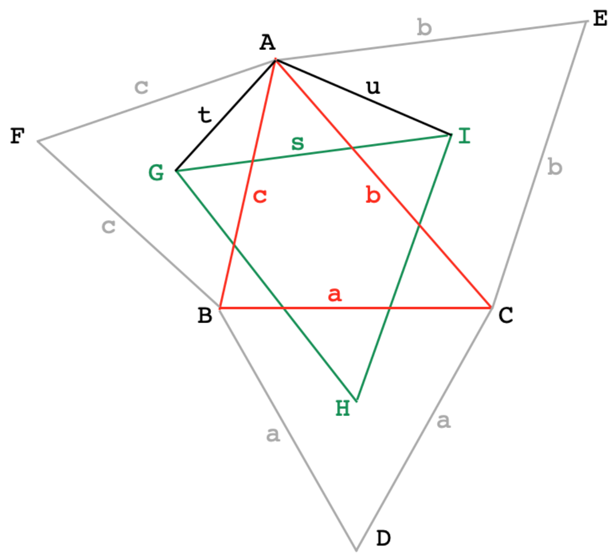
This theorem is called Napoleon's theorem. It seems likely that the association with Napoleon is apocryphal.

https://en.wikipedia.org/wiki/Napoleon%27s_theorem

algebraic proof

To simplify the notation, let the angle at vertex A be A, the side opposite be a with length a, and so on.

The medians of the equilateral triangles erected on the sides will be H, I and G, while s, t and u are sides of $\triangle AGI$ with vertices being two of the medians and vertex A.



We use the law of cosines, several times. First, recall that the line from the centroid of an equilateral triangle to any vertex bisects the angle at the vertex. So side u bisects $\angle CAE$ and side t bisects $\angle BAF$, and the angle between t and u is $A + 60$.

Then, by the law of cosines

$$s^2 = t^2 + u^2 - 2tu\cos(A + 60)$$

A second key point is that we can obtain sides t and u in terms of the original sides c and b.

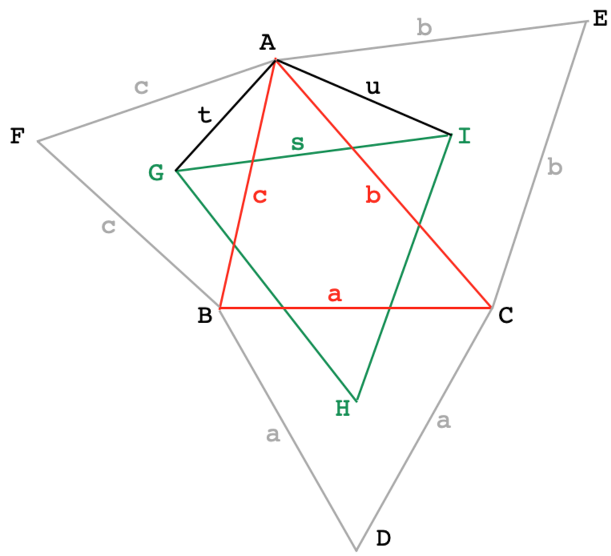
The median of an equilateral triangle forms a triangle with sides in the ratio 1-2-$\sqrt{3}$. The triangle with hypotenuse u is similar to this triangle, so the ratio of the long side to the hypotenuse is:

$$\frac{b/2}{u} = \frac{\sqrt{3}}{2}, \qquad b = \sqrt{3}u, \qquad b^2 = 3u^2$$

Also, $c^2 = 3t^2$.

Going back to the first equation and multiplying by 3 we obtain:

$$3s^2 = c^2 + b^2 - 2bc\cos(A + 60)$$



This could be done for any of the other two sides, by symmetry. For example, we can relate HI to sides a and b:

$$3HI^2 = a^2 + b^2 - 2ab\cos(C + 60)$$

We can connect these two expressions through the line segment BE.

Looking toward vertex A and employing the law of cosines again, we have:

$$BE^2 = b^2 + c^2 - 2bc\cos(A + 60)$$

Looking instead toward vertex C we have

$$BE^2 = a^2 + b^2 - 2ab\cos(C + 60)$$

These two expressions are equal, and therefore the things we found equal to them previously are also equal. Namely:

$$3s^2 = 3HI^2$$

$$s = HI$$

The same thing could be done for either of the other pairs of sides.

$$s = GI = GH = HI$$

$\triangle GHI$ is equilateral.

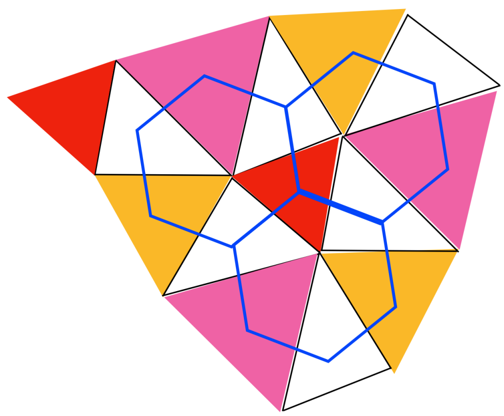
□

We just note in passing that since $AD^2 = a^2 + b^2 - 2ab\cos(C + 60)$, and BE^2 is equal to the same expression, we have that $AD = BE$ and by extension, both are equal to CF.

Also, connect each vertex of the original triangle with that of the equilateral triangle opposite: forming AH, BI and CG. These lines cross at *Fermat's point*.

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

Apparently, there is a tiling pattern in the figure. This graphic isn't perfect (my poor skills with the program), but I think you get the idea.



There is a hexagonal pattern that contains the centroids of the equilateral triangles, and a three-fold rotational symmetry around each of them.

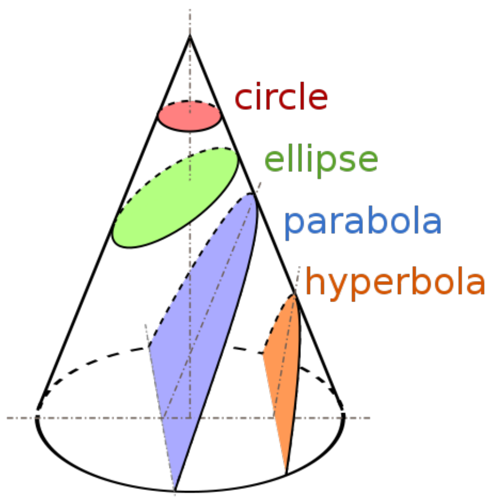
Part V

Curves

Chapter 16

Parabola

The parabola is one of a larger class of geometric figures called the conic sections, which include the circle, ellipse, parabola and hyperbola.



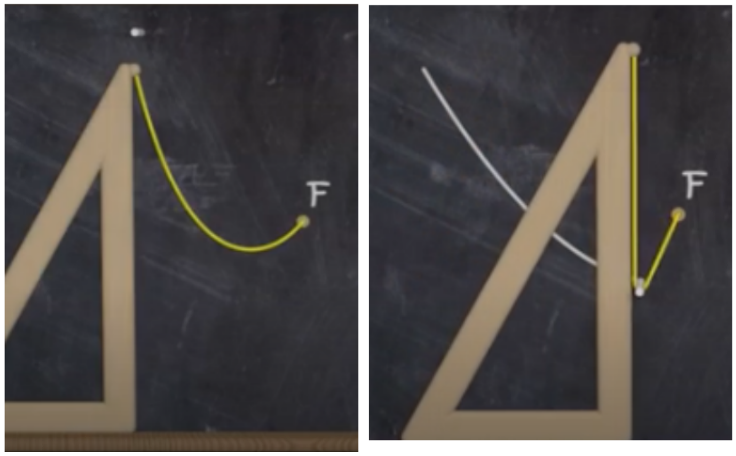
These can be viewed as the intersections of a plane and a (hollow) cone. The parabola is the particular intersection where the plane is parallel to the edge of the cone.

construction

There are methods to draw each of these shapes using a string and pencil. The most familiar is probably the one for the ellipse (**here**). I found a demonstration on the web

https://imaginary.org/film/mathlapse-constructions-by-pin-and-string-conics

Here are two screenshots for the parabola:



On the left we see the a string suspended between the top of a set square and a point called the focus (F). In the right panel, the pencil holds the string taut against the vertical side of the square, and as the square is moved horizontally to the right, the pencil traces out a parabola.

This arrangement captures the geometrical constraint for a parabola. Imagine moving the square to the left. As the distance to the focus increases, the length of the string from the top of the square decreases by the same amount, so that the distance of the pencil from the initial minimum increases exactly the same amount as the distance to the focus increases.

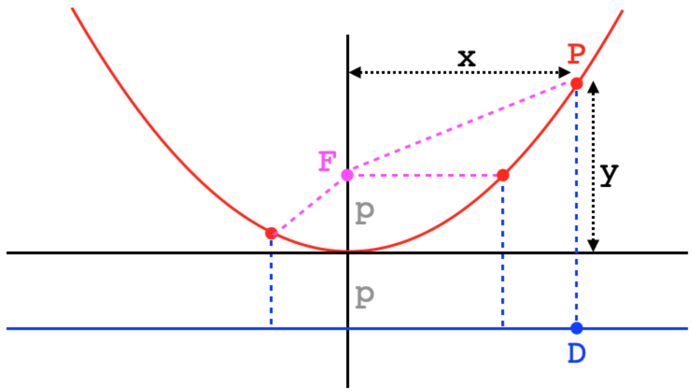
directrix and focus

It is pretty complicated to look at parabolas in the way that the Greeks did. We will do most of our work later on using methods from analytic geometry, which allows us to draw the curve based on the equation $y = ax^2$, where a is a constant. But for the moment, we are trying to say as much as we can without introducing those ideas.

The geometric definition of a parabola is this. Draw two lines that are perpendicular (colored black in the figure below). Let us call these two lines the horizontal and vertical axes, and the point where they cross the origin.

On the vertical axis choose point F a distance p above the origin. This point is colored magenta in the figure.

Then draw another horizontal line which intersects the vertical axis the same distance p below the origin. This line is called the directrix (colored blue).



The parabola consists of all those points whose distance to the focus is equal to the vertical distance to the directrix.

We look at a general point P.

Now, $PF = PD$ for all points on the parabola. To simplify the notation, let the distance from the vertical axis out to PD be called x, and the distance up from the horizontal axis to point P be y.

Then $PD = y + p$ while PF is the hypotenuse of a right triangle with sides x and $y - p$. Since the lengths are equal, the squared lengths are also equal. Thus

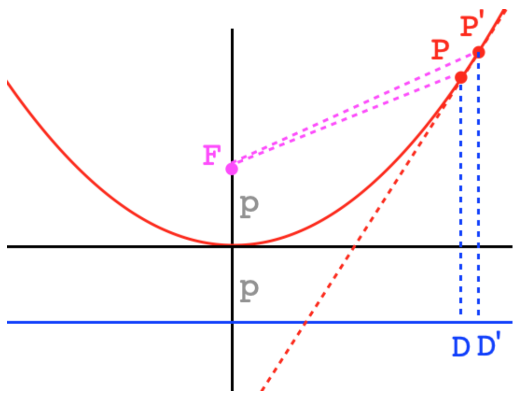
$$(y+p)^2 = x^2 + (y-p)^2$$
$$2yp = x^2 - 2yp$$
$$y = \frac{1}{4p}x^2$$

The equation corresponding to a parabola is a quadratic, it contains x^2. If we define $a = 1/4p$ then $y = ax^2$.

If $p = 1/4$, then $y = x^2$. For a larger p, the constant factor $1/4p$ gets smaller, meaning that y will be smaller for each corresponding x, so the parabolas' arms become more shallow.

slope of the tangent

Now, we want to introduce the idea of the tangent to the curve and think about its slope. Look at the figure below, where the tangent has been drawn to the curve at point P. Obviously, the slope changes as x changes, becoming steeper as we move up the curve.



Consider a point P' very close to P. In moving from point P to P', we are "essentially" moving along the tangent line. By definition, the tangent only touches at a single point, but the tangent at P goes very near to P'.

The invariant of the parabola (the thing that does not change) is that the distance to any point from F must be equal to the vertical distance to the same point from the directrix.

In moving from P to P', it seems clear that to maintain the invariant, we must, as it were, take a step which combines two directions: it is partly in the same direction as F to P, continuing beyond P, and also partly beyond P in the same direction as from D to P. If we average these two directions, the equality of distance will be maintained. This idea is due to Roberval.

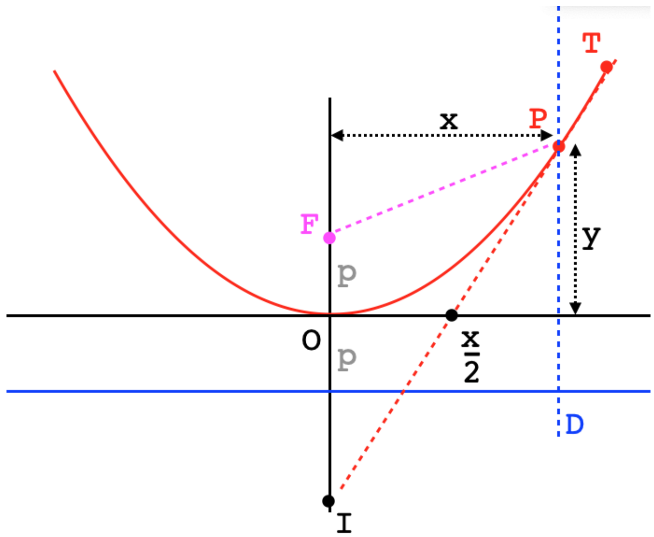
Averaging is easy because $FP = DP$. Let us move a small fraction k of the length FP with $x-$ and $y-$components $\langle kx, k \cdot (y - p)\rangle$, together with a movement of equal distance in the direction of DF of $\langle 0, k \cdot (y + p)\rangle$. Together, these add up to $\langle kx, 2ky\rangle$.

The slope of the tangent is simply the ratio of these $y-$ and $x-$ components.

$$\Delta y/\Delta x = 2y/x = 2x^2/4px = x/2p$$

For a parabola defined as $y = ax^2$, the slope is $2ax$. As you will see later on, this result is literally the first result in calculus.

Since the horizontal distance between I and P is x, the vertical distance between I and P is $2y$. Therefore, the intersection of the tangent line with the vertical axis (at I) will occur a distance y below the horizontal axis. $IO = y$.



We have sides of equal length y plus vertical angles in a right triangle, so there are two congruent triangles by ASA. Therefore the tangent cuts the horizontal axis a distance $x/2$ from the origin.

Since $IO = y$, $IF = y + p$, but this is equal to PD, and since $IF \parallel PD$, it follows that $IFPD$ is a parallelogram. And since $FP = PD$, it is a regular parallelogram, with four sides equal.

In a regular parallelogram, the diagonals cross at right angles. Thus, IP is the perpendicular bisector of FD and conversely, FD is the perpendicular bisector of IP.

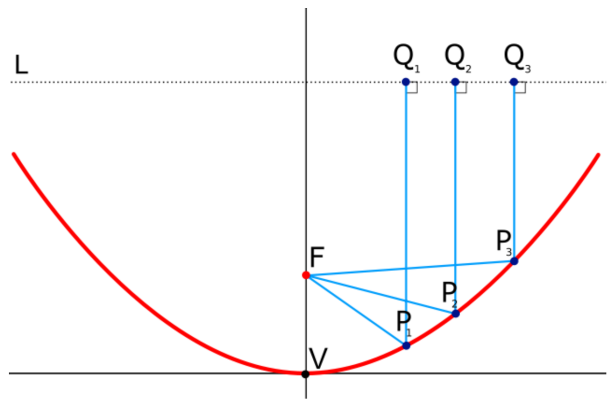
For any given value of x (distance of P from the $y-$axis), there is a corresponding point D on the directrix, and the perpendicular bisector of FD contains all points equidistant from both F and D. Therefore, the point P is at the intersection of the perpendicular bisector of FD and the vertical from D.

headlight property

Since IP is a diagonal of the parallelogram, it bisects the angle FPD. Therefore, the angle that FP makes with the tangent is equal to the angle that the tangent makes with the vertical above point P, by vertical angles. (This can be obtained in other ways, for example from the fact that $\triangle IFP$ is isosceles, or by alternate interior angles).

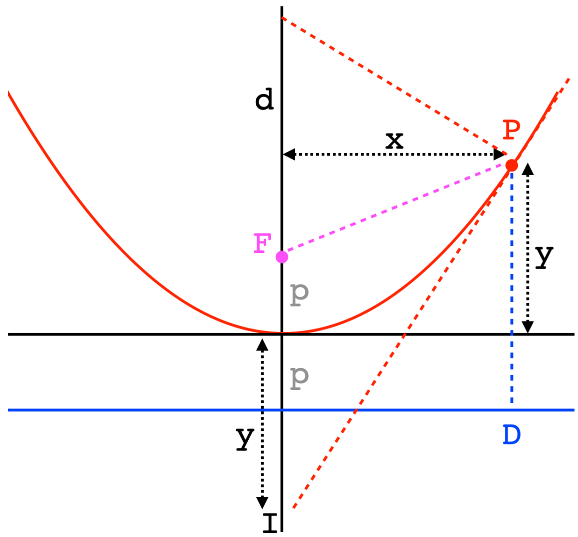
What this means is that if a light ray enters the parabola and comes down vertically in the plane we've drawn, no matter where P is, it will be reflected to F. This is due to the law which says that the angle of incidence is equal to the angle of reflection. Hence the name of F, it is called the focus.

Here is an illustration from wikipedia:



normal and its projection

We showed above that $OI = y$ and the vertical distance between I and P is $2y$.



Now draw the line perpendicular to the tangent at P, which is called the normal. What is the vertical distance d above P where the normal intercepts the vertical axis?

The normal, the tangent line and the vertical form a right triangle that is similar to all the other right triangles in the figure (except those involving FP). This includes the small one with sides x and d, and the large one with sides x and $2y$.

By similar triangles we have that

$$\frac{d}{x} = \frac{x}{2y}$$

$$d = \frac{x^2}{2y} = \frac{x^2 \cdot 4p}{2 \cdot x^2}$$

$$d = 2p$$

So the normal hits the vertical axis a distance of $2p$ above the point, regardless of where the point is.

slope of the tangent: alternative derivation

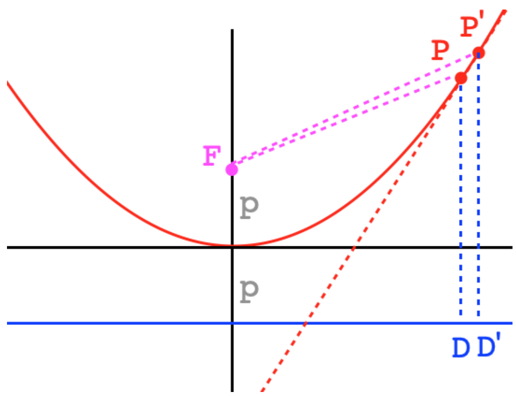
By the definition of the parabola, any point P on the parabola lies on a vertical up from the directrix at D such that $PF = PD$. That equality means that the same point P must lie on the vertical bisector of FD, by the properties of isosceles triangles.

For any given value of D, there can be only one such point P, because there is only one point where two lines cross, provided they do cross and are not the same line. But this follows from the placement of F, for all points except the vertex.

We will prove that the perpendicular bisector of FD is also the tangent at P.

Proof.

Suppose the perpendicular bisector of FD were not the tangent. Then, perhaps it would not cut the parabola at all. But this would mean that no point on the parabola would satisfy the invariant. And we know that PI is not parallel to PD, so they must cross.



So, if it is not the tangent, then the perpendicular bisector of FD must be a secant and cut through two points on the parabola. Let them be P and P'. In that case, P' has two contradictory properties.

On one hand, P' is on the bisector of FD. It is therefore equidistant from F and D.

This means that

$$FP' = P'D$$

On the other hand, P' is on the parabola. The invariant says that

$$FP' = P'D'$$

with P'D' perpendicular to the directrix. But PD is also perpendicular.

As a result

$$P'D = P'D'$$

But this would mean that $P'D$ is the hypotenuse in a right triangle with base $P'D'$, so $P'D$ must be larger.

This is a contradiction.

Therefore the perpendicular bisector *is* the tangent.

$\square$

Since the tangent lies at a right angle to FD, the product of the tangent's slope and the slope of FD is equal to -1 (see **here**).

But the slope of FD is just $-2p/x$. Therefore the slope of the tangent is

$$m = \frac{x}{2p}$$

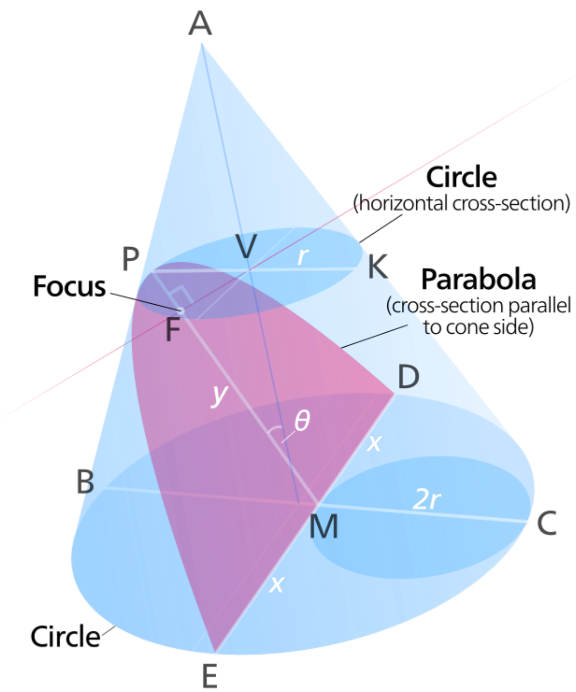
where the equation of the parabola is

$$y = \frac{1}{4p}x^2$$

There are various other results, for example the relationship between pairs of tangents, and Archimedes' *quadrature* of the parabola, by which he found the exact area bounded by the curve of the parabola and any secant. But we will wait for those until we get a real coordinate system established.

conic view

Here is a picture from wikipedia which shows some of the details for a parabola as the cross section of a cone.



It is possible to show that the distance labeled y in the diagram is equal to some constant a times x^2. This is true for every parabola.

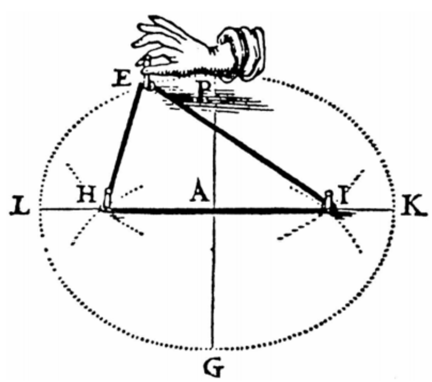
Chapter 17

Ellipse

In the chapter on orbits, we assumed that the planetary orbits are circular. They are actually ellipses.

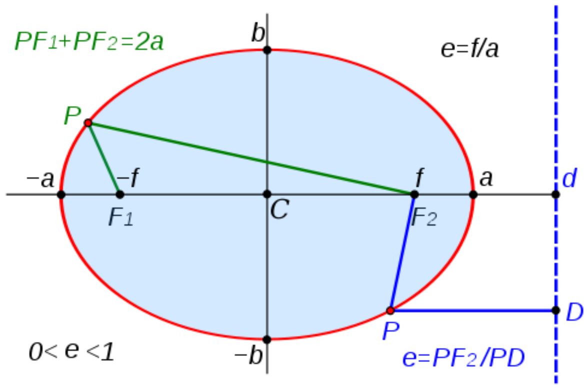
We will defer most discussion of the ellipse until we can use equations found using analytical geometry. However there are a few things we can say now.

construction



Learning how to draw an ellipse using two pins and a circular piece of string holding a pencil is an early adventure in mathematics. The ellipse is the set of all points whose combined distance to the two pins (foci) is the same.

The drawing above in Acheson is reproduced from a 17th century book.



The pin positions with respect to the origin or center are called the foci, lying at the points shown in the second figure as $(\pm f, 0)$.

We will use the notation c: the focus in the first quadrant is at the point $(c, 0)$.

The lengths of the axes (called semi-major and semi-minor) are usually labeled a and b.

Consider the situation when the pencil is at the point $P = (0, a)$. The distance to the left focus is $c + a$, so the length L of the string is twice that

$$L = 2(c + a)$$

The combined distance from each point on the ellipse to the two foci is the length of the string minus the distance between the two foci

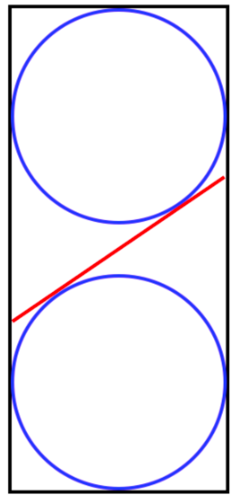
$$L - 2c = 2(c + a) - 2c = 2a$$

Starbird

Here is a neat approach to the ellipse that I saw in one of Michael Starbird's lectures.

Imagine a glass cylinder, shown here in cross-section and colored black. The cylinder has been sliced through at an angle by a plane, and we suppose a flat piece of glass in the shape of an ellipse is glued between the two halves.

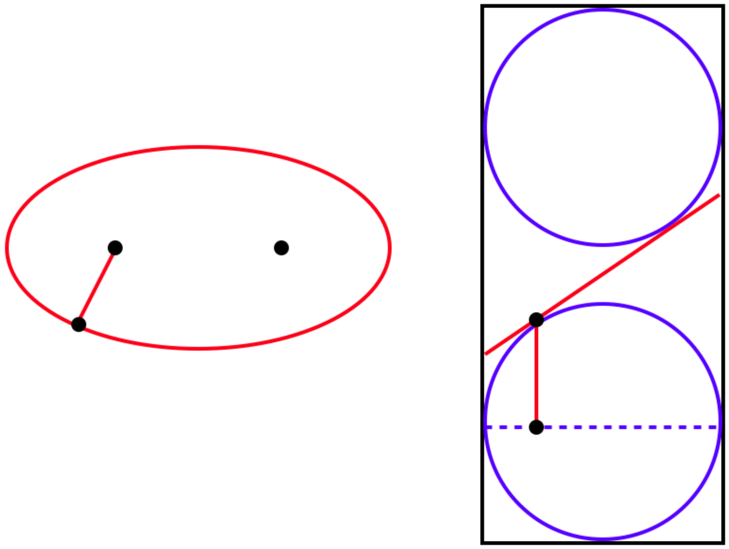
The elongated region in red (formed at the plane of the cut) is the ellipse, and the cylinder is oriented so that at each horizontal position going across the page, the two points on the ellipse are at the same vertical position. We see the plane of the cut edge-on.



Two spheres that fit snugly inside the cylinder lie above and below the ellipse, just touching it. The planar surface of the ellipse is tangent to the spheres, touching each one at a single point.

We claim that the points where the spheres touch the ellipse are the foci of the ellipse.

By the nature of the construction, the two spheres just fit inside the cylinder. That means the intersection where the spheres touch the cylinder is a circle, the lower one is shown with a dotted blue line in the next figure.

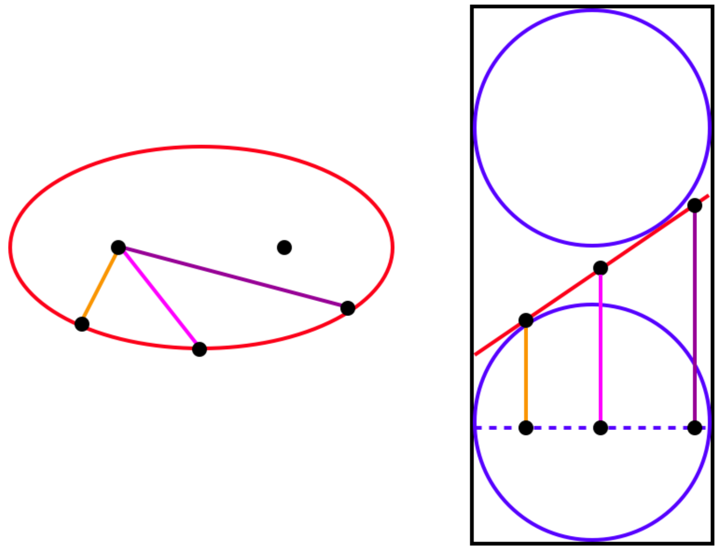


Now consider any point on the ellipse. On the left, we see one point on the ellipse together with two interior points we claim are foci, with a line drawn from our point to one of the foci.

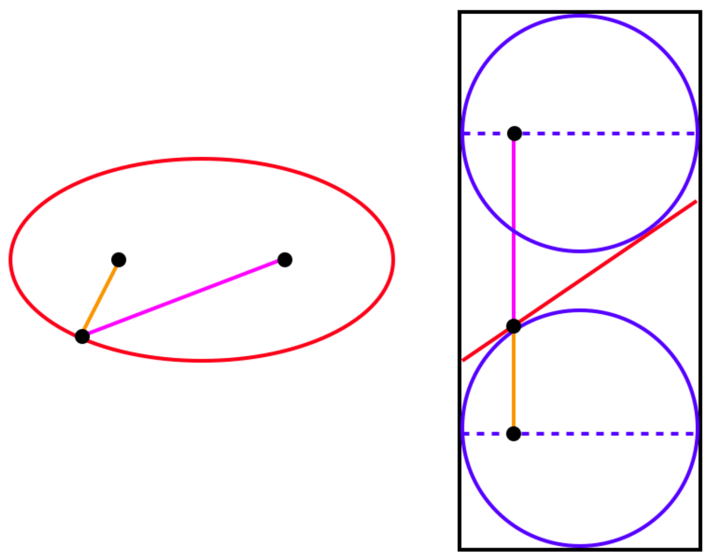
We said that this point is the point where the ellipse touches the lower sphere. We conclude that the line we've drawn from the edge of the ellipse to the focus is a tangent to the sphere.

A second tangent of interest is the perpendicular dropped vertically down the surface of the cylinder, shown in the right panel. Since they are both tangents, this line is the same length as the line to the focus.

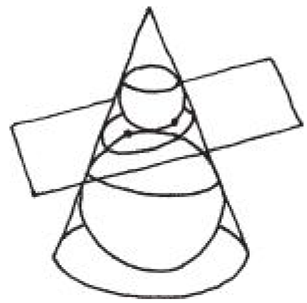
But the construction, and this equality, holds for any point on the ellipse, as shown in the next figure.



Finally, this is true for both spheres (below). The sum of the perpendicular tangents for any point is a constant.



Thus, the points where the spheres touch the ellipse are its foci, because the sum of the distances to any point on the ellipse, which is equal to the sum of the vertical tangents, is a constant.



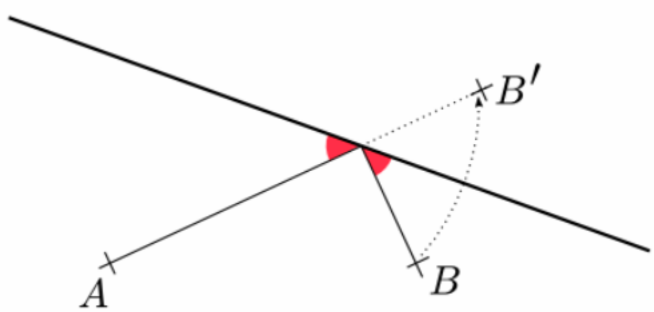
According to Lockhart, the same argument can be used to prove that the cross sections of a cone are ellipses (which seems strange at first since we've been demonstrating that the cross-sections of cylinders are also ellipses).

reflected rays

In any ellipse, the segments from the foci to any point on the ellipse make equal angles with the tangent. This means that light rays emitted from one focus and striking anywhere on the ellipse will pass through the other focus upon reflection. It is the principle behind "whispering galleries."

Here is a simple geometric proof.

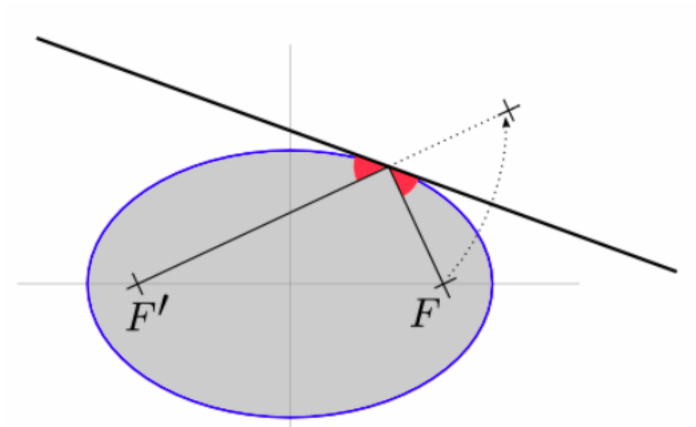
We consider the problem of the "shortest path."



The problem is to go from A to the line and then back to B by the shortest path. The clever solution is to place B' on the other side of the line at the same distance away. By definition (see Euclid) the shortest path A to B' is a straight line.

We can use vertical angles (or supplementary angles twice) and then similar triangles to prove that the two angles colored red are equal.

Now consider an enhanced diagram of the same situation:



We draw the tangent to the ellipse. By definition, the tangent has only a single point on the curve. This point lies at a distance $2a$ from the combined foci. All other points on the line are farther away from the two foci than the point of intersection. (You would have to make the string bigger to draw the ellipse that goes through any of those points).

Therefore, the path shown is the shortest path from F' to the tangent and then to F. But we know that for the shortest path the angles colored red are equal.

http://math.stackexchange.com/questions/1063977/how-to-geometrically-prove-the-focal-property-of-ellipse

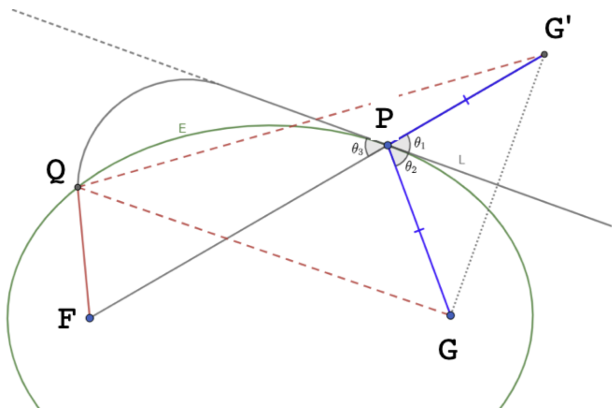
alternate proof reflected rays

Here is a slightly more elaborate argument from the same StackExchange question. It is a proof by contradiction.

Consider the curve of ellipse E.

To simplify the notation we will label the foci F (F_1 in the original drawing) and G (F_2). Consider an arbitrary point P on E.

We will prove that FP and PG make equal angles with the tangent to E at P.



Extend FP to G' (chosen so that $PG = PG'$).

Draw the perpendicular bisector L of GG'. Since $\triangle PGG'$ is isosceles, $\theta_1 = \theta_2$, the perpendicular bisector of the base also bisects the angle at P. By vertical angles, $\theta_3 = \theta_1$ so $\theta_3 = \theta_2$.

We claim that L is the tangent line to E at P.

Proof.

Suppose L is not the tangent.

Then, L also meets E at another point Q.

As a point on L, Q is also equidistant from G and G' with $QG = QG'$.

By the definition of the ellipse $FQ + QG = FP + PG$.

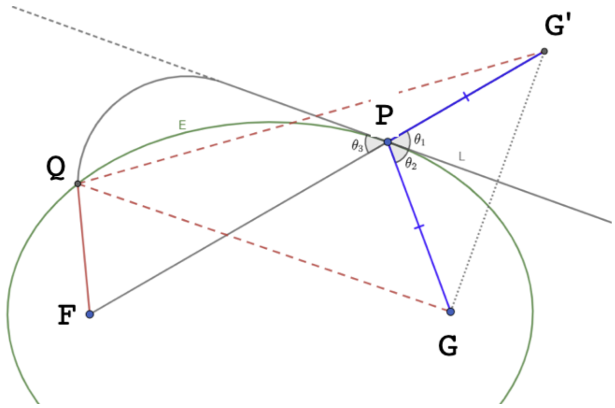
So then

$$FQ + QG' = FQ + QG = FP + PG = FP + PG' = FG'$$

Consider just the first and last terms:

$$FQ + QG' = FG'$$

This violates the **triangle inequality** for $\triangle FQG'$ and is a contradiction.



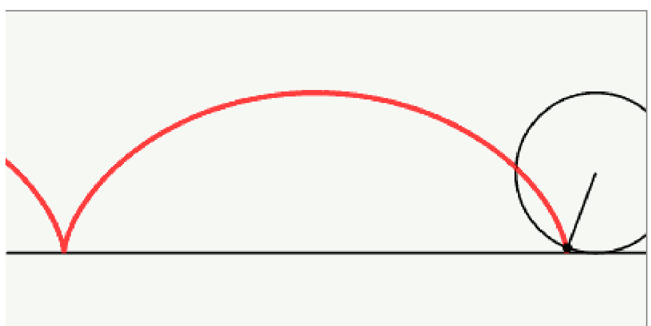
Therefore L is the tangent line to E at P and $\theta_3 = \theta_2$.

$\square$

Chapter 18

Cycloid intro

Imagine a bicycle tire marked at a particular point on the rim, say with fluorescent paint or a small light. Time starts at $t = 0$ with that point P in contact with the x axis at $(0, 0)$. Then the bike rolls to our right. As the tire rotates the fixed point P on the rim traces a curve



Roberval

The cycloid curve was not known to the ancients, as far as our sources have anything to say. The first mention is around 1500, and by the time of Galileo a century later, it was something he thought about quite a bit and encouraged others to study.

Galileo is said to have used Archimedes' method of cutting out shapes and weighing them to show that the area was approximately 3 times that of the generating circle. But he was apparently not prepared for it to be *exactly* equal, so was skeptical.

The cycloid became very popular in the decades before Newton. Roberval came up with a clever application of Cavlieri's principle to find the area under the curve.

The great idea is to draw a second curve, called the *companion curve*, by sliding values derived from a half-circle to add them to the x-values on the cycloid curve, as shown. We can imagine sliding them from the left, as above, but it works also to slide them from the right.

The areas marked by horizontal lines are equal, by Cavalieri's principle. For example $EF = PQ$. But that total extra area is just $\pi a^2/2$, one-half of the circle.

https://maa.org/sites/default/files/pdf/cmj_ftp/CMJ/January%202010/3%20Articles/3%20Martin/08-170.pdf

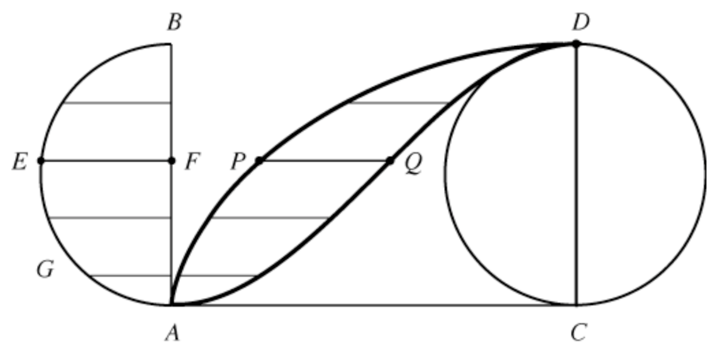


Figure 4. The companion curve to the cycloid.

Notice that the companion curve AQD divides the rectangular area $ACDB$ in half, by symmetry. Each piece at a distance h from the top is matched by a piece of equal length at a distance h from the bottom, by the vertical symmetry of the half circle.

The rectangle has width πa and height $2a$, so its area is $2\pi a^2$, and one-half that is πa^2.

The area under the cycloid curve is then three-quarters of the total, which is $(3/2)\pi a^2$, and the area under one complete lobe is twice that or $3\pi a^2$.

slope

Descartes gave the slope of the tangent line to the cycloid by the following construction:

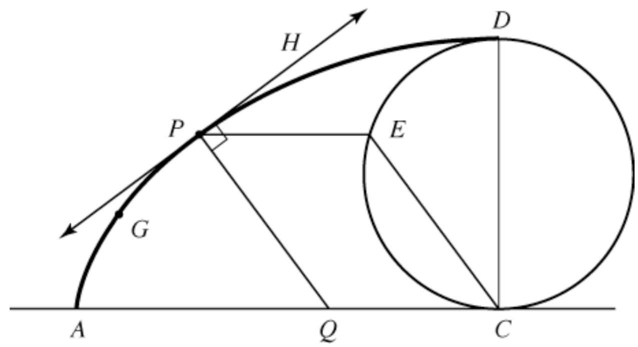


Figure 6. Descartes' tangent construction.

Given P on the cycloid, draw PE horizontally to find E on the generating circle, then draw EC and also PQ parallel to EC. The tangent is perpendicular to those two line segments.

I like this construction a lot because it reminds me of what Roberval did with the parabola. For any t, the distance moved horizontally is equal to the distance along the rim of the generating circle.

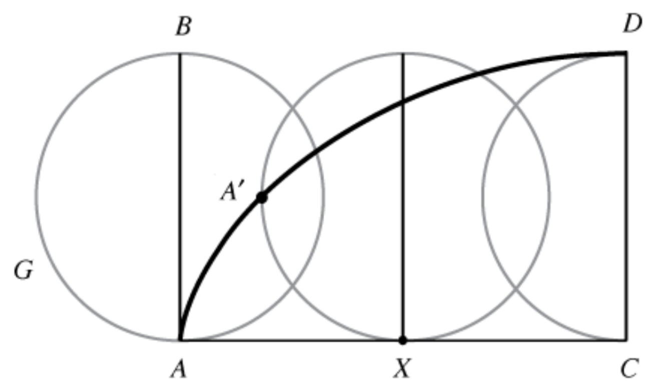
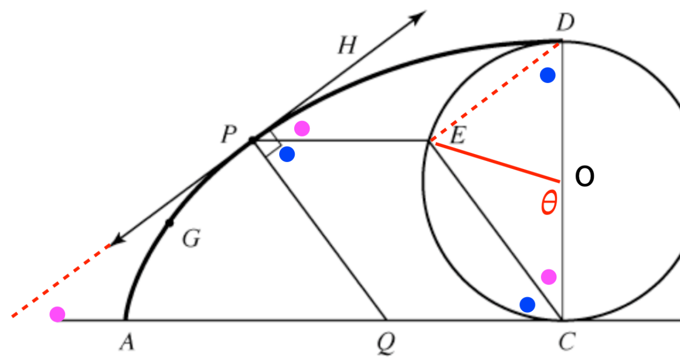


Figure 3. Roberval's definition of the cycloid.

Indeed, in Roberval's view of the curve (see Martin) ““Let the diameter AB of the circle AGB move along the tangent AC, always remaining parallel to its original position until it takes the position C D, and let AC be equal to the semicircle AGB. At the same time, let the point A move on the semicircle AGB, in such a way that the speed of AB along AC may be equal to the speed of A along the semicircle AGB."

Therefore, since no slippage occurs, movement along the tangent should be in the direction of the sum of the two vectors. The tangent is the average of the tangent to the circle at E and the horizontal.

Since the latter is zero, the result is one-half of the former.



We are given that PE is horizontal, so it is parallel to QC. Also, $PQ \parallel EC$. Therefore, $PECQ$ is a parallelogram. $\angle DEC$ is right. Angles marked with blue and magenta dots are complementary.

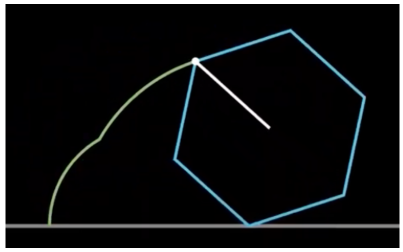
$\angle\theta$ at the circle's origin is supplementary to two magenta dots, so $\theta/2$ is complementary to one. That is, $\theta/2$ is complementary to the angle of the tangent, so the tangent's slope is the cotangent of $\theta/2$.

arc length

According to John Martin, it was Descartes who had the idea of approximating the cycloid curve as the (bumpy) curve generated by a polygon.

https://maa.org/sites/default/files/pdf/cmj_ftp/CMJ/January%202010/3%20Articles/3%20Martin/08-170.pdf

Here is a frame from an animation

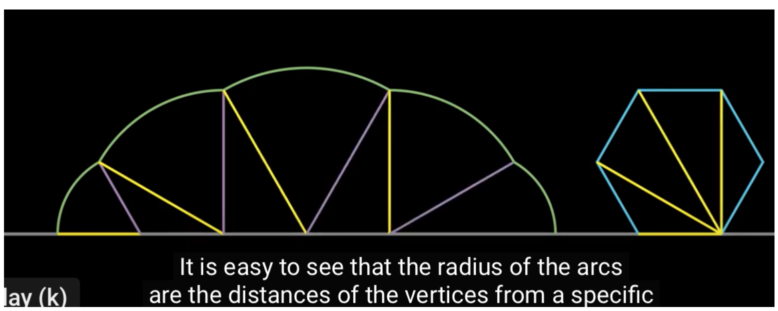


The Youtube video gives a derivation for the length of the curve, first solved by Christopher Wren, although the proof given is not the same as his.

https://www.youtube.com/watch?v=3yK2tJZR3Js

For a polygon of n sides, we will have n-1 circular arcs. The arc generated by rotating a polygon around one of its vertices subtends the same angle for each vertex. *Proof.* The angle of rotation is just the exterior angle of the polygon. $\square$

What does change is the radius of each arc element. The first one is the side length, but the second one is longer. For example, with the square it is the diagonal of the square. And in general, the length of the arm that sweeps out the curve is the length of a chord that connects $k = 1, 2, 3 \dots$ faces.



Notice that the angle subtended by any face (or multiple adjacent faces) of a polygon is the same regardless of the position of the face with respect to the vertex. *Proof.* The central angle for any two adjacent vertices is the same. So the peripheral angles are all equal, by the inscribed angle theorem. $\square$

But these arms are also chords! And that is how we will get their lengths. A single side is the chord for a central angle $2\pi/n$. Two sides for twice that. k sides for $2k\pi/n$.

Recall that if t is the central angle subtended by any chord, then the chord length is twice the sine of the half-angle, times the radius of the circle.

$$c = 2a \sin \frac{t}{2}$$

So if we divide into n sides and we are on the kth arc with the chord of k sides as the arm

$$c_k = 2a \sin \frac{k\pi}{n}$$

The chord length is also the radius of the circular sector. We get the arc length as the radius times the angle. The kth side has arc length

$$2a \sin \frac{k\pi}{n} \cdot \frac{2\pi}{n}$$

Sum them all

$$\sum_{1}^{n-1} 2a \sin \frac{k\pi}{n} \cdot \frac{2\pi}{n}$$

We can add an extra term at the bottom and the top (because $\sin 0 = \sin \pi = 0$:

$$\sum_{0}^{n} 2a \sin \frac{k\pi}{n} \cdot \frac{2\pi}{n}$$

$$\frac{4\pi}{n} \sum_{k=0}^{n} \sin \frac{k\pi}{n}$$

Now we need to find out what happens to $\sin \frac{k\pi}{n}$ as $n \to \infty$.

Lagrange

We use Lagrange's trigonometric identity. (We follow the video even though I found another way to do it, and a puzzle. — see the next section).

Lagrange's Identity

$$\sum_{k=0}^{n} \sin(k\theta) = \frac{\sin\left(\frac{n\theta}{2}\right) \sin\left(\frac{n\theta}{2} + \frac{\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)}$$

$$\sum_{k=0}^{n} \sin k\theta = \frac{\sin \frac{n\theta}{2} \cdot \sin(\frac{n\theta}{2} + \frac{\theta}{2})}{\sin \frac{\theta}{2}}$$

The argument to the sine function $k\theta$ is $\frac{\pi}{n}$. That is, $\theta = \frac{\pi}{n}$.

The first term in the numerator is

$$\sin \frac{n\theta}{2} = \sin \frac{n\pi}{2n} = \sin \frac{\pi}{2} = 1$$

And the second is

$$\sin(\frac{n\theta}{2} + \frac{\theta}{2}) = \sin(\frac{\pi}{2} + \frac{\pi}{2n})$$

Recall that $n \to \infty$. Hence this expression is also equal to 1.

For the denominator, we use the small angle approximation, since as n becomes large, $\frac{\theta}{2} = \frac{\pi}{n}$ becomes very small.

So $\frac{\theta}{2} \approx \sin \frac{\theta}{2}$ and then

$$\sum_{0}^{n} \sin k\pi/n = \frac{1}{\frac{\pi}{2n}} = \frac{2n}{\pi}$$

Multiply by what was in front to obtain the answer $8a$ (for a circle or set of polygons with n sides and radius a).

puzzle with the proof

However, there is a puzzle with the proof as presented. According to wikipedia

https://en.wikipedia.org/wiki/List_of_trigonometric_identities#Lagrange's_trigonometric_identities

The well-known formula for the sine is

$$\sum_{k=0}^{n} \sin k\theta = \frac{\cos \frac{1}{2}\theta - \cos((n + \frac{1}{2})\theta)}{2 \sin \frac{1}{2}\theta}$$

Derivation:

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

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

Subtract the first from the second:

$$\cos(A - B) - \cos(A + B) = 2 \sin A \sin B$$

$$2 \sin B \sum \sin A = \sum \cos(A - B) - \cos(A + B)$$

If $B = \frac{1}{2}$ and A consists of integral θ as in $A = k\theta$, then the series will consist of terms like $\cos(k - \frac{1}{2})\theta$ followed by $-\cos((k + 1 + \frac{1}{2})\theta)$. These cancel each other, forming a *telescoping series*.

The only terms that survive come from the first and last:

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

(Actually for this derivation it is more convenient to start the sum from $k = 1$).

We have $\theta = \frac{\pi}{n}$. As $n \to \infty$, the first term is $\cos 0 = 1$ and the second term is $-\cos(\pi + 0) = 1$.

The sum is 2 which cancels the 2 on the bottom. The rest of the denominator is

$$\sin \frac{1}{2}\theta \approx \frac{\pi}{2n}$$

which inverts and multiplies to give exactly the same answer as before.

One can also find online proofs using Euler's identity.

https://math.stackexchange.com/questions/1349466/calculating-sum-k-0n-sink-theta

which gives the first identity, where the terms are all sines, the right hand side has a multiplication, and there is no factor of 2.

I have written out a full version of both of these derivations separately.

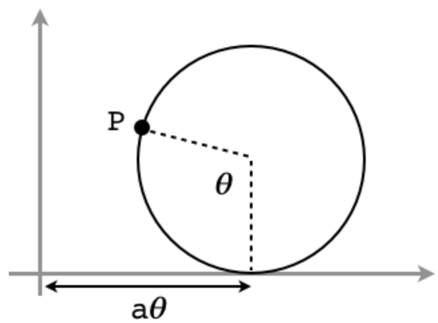
https://github.com/telliott99/short_takes/blob/master/Lagrange%20trig.pdf

I also found a proof that the two formulas are equivalent, which is included.

standard parametric approach

We want to find equations that give the point P as a function of time. We will parametrize the curve, yielding parametric equations $x(t)$, $y(t)$.

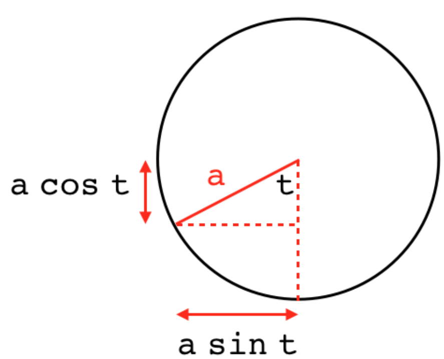
The diagram below shows the angle through which the wheel has turned as θ, but we will use t for θ here.



The x displacement of the vertical straight down from the center of the tire is just at, where a is the radius of the wheel, it is equal to the arc on the circumference of the wheel from the point which is currently in contact with the ground, going around up to P.

It is reasonably easy to derive the desired parametric equations, using vectors, especially once you know the answer. For x, we have the vector that goes from $(0, 0)$ to the contact point with the ground. As indicated in the figure, that is at.

We need to subtract the distance $a\ sin\ t$ from that. Basically the rationale is that the motion is a standard parametric circle which has been rotated by 90 degrees clockwise and then inverted. The rotation changes cosine to sine, and the inversion brings the subtraction.



It's easier to see for $t < \pi/2$, but it is true always. Check some other values of t like π or $3\pi/2$ to confirm. This is the usual circular motion.

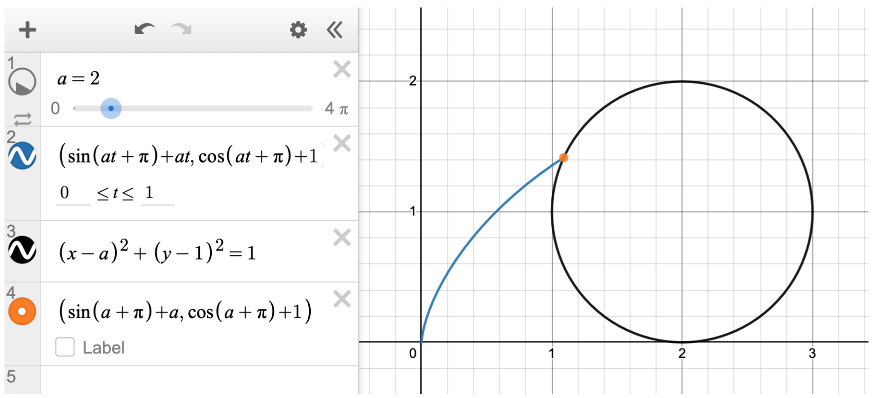
For y, we have a constant factor of a above the x axis, then the additional displacement is $-a\ cos\ t$. So for $t = 0$ we have the additional displacement is -a (we were on the ground), for $t = \pi/2$ it is zero, and for $t = \pi$ it is plus a for a total of $2a$.

The parametric equations are then

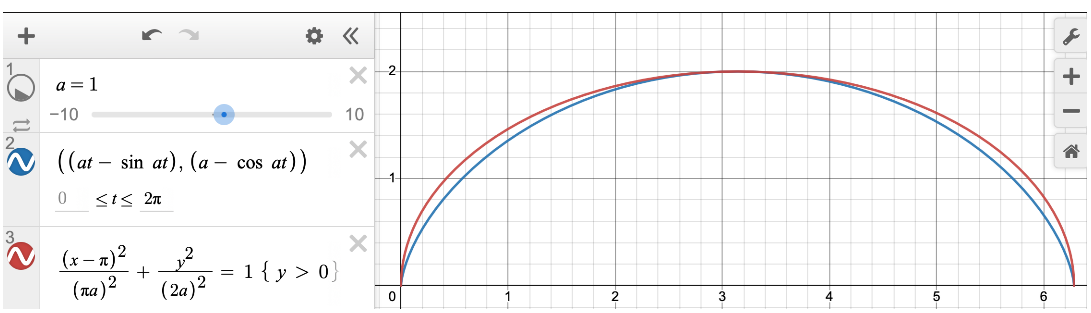
$$x(t) = at - a\sin t$$

$$y(t) = a - a\cos t$$

Here is an animation from Desmos, if you want to set it up and play with it.



It is almost, but not quite, a semi-ellipse.



If you look closely at the animation, we had

$$\sin(at + \pi) + at = at - \sin at$$

$$\cos(at + \pi) + 1 = 1 - \cos at$$

where we have used the addition formulas

$$\sin A + B = \sin A \cos B + \sin B \cos A$$

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

So we see that the Desmos formulas are actually the same as what we've written above (and used in the second screenshot, regarding the semi-ellipse.

Once we know some calculus, we can get various expressions like the slope of the curve. For example, the derivatives are:

$$x'(t) = a - a \cos t$$

$$y'(t) = a \sin t$$

We can get the slope of the tangent line, by simple division:

$$y' = \frac{dy}{dx} = \frac{a \sin t}{a - a \cos t}$$

$$= \frac{\sin t}{1 - \cos t}$$

Simmons uses half-angle formula to make something simpler, like so:

$$\sin t = 2 \sin \frac{t}{2} \cos \frac{t}{2}$$

$$\cos t = \cos^2 \frac{t}{2} - \sin^2 \frac{t}{2}$$

$$= 1 - 2 \sin^2 \frac{t}{2}$$

So the ratio is

$$y' = \frac{2 \sin \frac{t}{2} \cos \frac{t}{2}}{2 \sin^2 \frac{t}{2}} = \cot \frac{t}{2}$$

We saw this before, when we looked at the geometry of the tangent.

aside about Archimedes

It struck me that $\cot t/2$ is one of the terms in the relationship Archimedes used in approximating π. There we had

$$\cot 2\theta + \csc 2\theta = \cot \theta$$

$$\cot t + \csc t = \cot t/2$$

So somehow, if what we have above is correct

$$\frac{\sin t}{1 - \cos t} = \cot t/2 = \cot t + \csc t$$

Factor the right-hand side

$$= \frac{1}{\sin t}(\cos t + 1)$$

Then

$$\sin^2 t = (1 - \cos t)(\cos t + 1)$$
$$= 1 - \cos^2 t$$

and one more step gives our favorite identity.

Let's see if we can figure out a parametric equation for the companion curve. $x(t)$ for the cycloid was $x(t) = at - a\sin t$. The companion curve gets an additional length in the x-direction of $a\sin t$. So $x(t) = at$. That was easy!

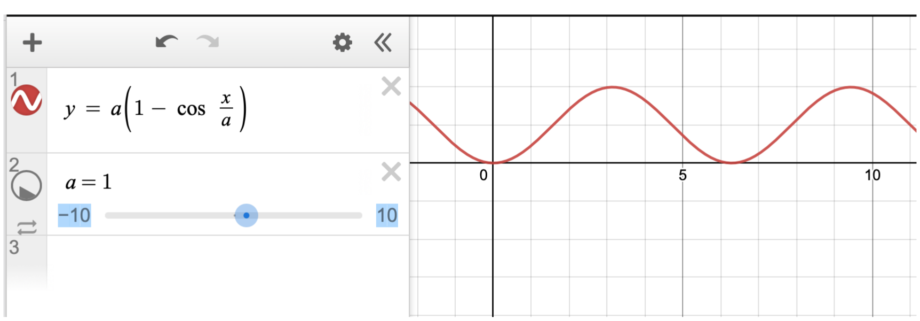
If we want $y = f(x)$ we must do

$$y = a - a\cos t$$

and then substitute $t = x/a$ so

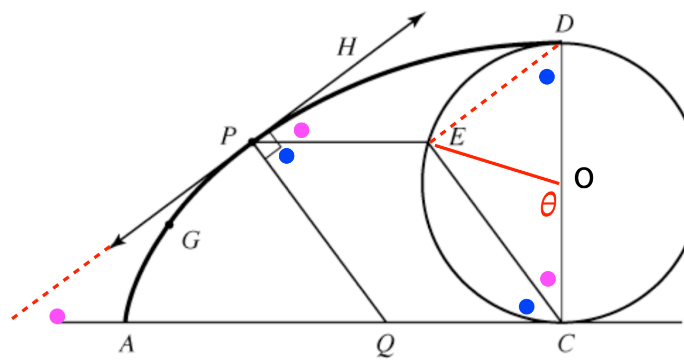
$$y = a(1 - \cos x/a)$$

which goes like the cosine of a scaled version of x, shifted up



Since the latter is zero, the result is one-half of the former.

Referring to the figure below, and repeating what we said before, $\angle\theta$ at the circle's origin is supplementary to two magenta dots, so $\theta/2$ is complementary to one. That is, $\theta/2$ is complementary to the angle of the tangent, so the tangent's slope is the cotangent of $\theta/2$.



Let us also take the origin of the circle temporarily as the origin of coordinates $(0,0)$. Then let point E be

$$x' = -a \sin t$$

$$y' = a - a \cos t$$

Since $C = (0, -a)$ we have that the slope of $CE \parallel PQ$ is

$$\frac{\Delta y}{\Delta x} = \frac{(a - a\cos t) - a}{(-a \sin t) - 0} = \frac{\cos t - 1}{\sin t}$$

The slope of the tangent is the negative inverse

$$\frac{\sin t}{1 - \cos t}$$

which matches. We can also show that ED has the same slope as the tangent, since it makes a right angle with CE.

Part VI

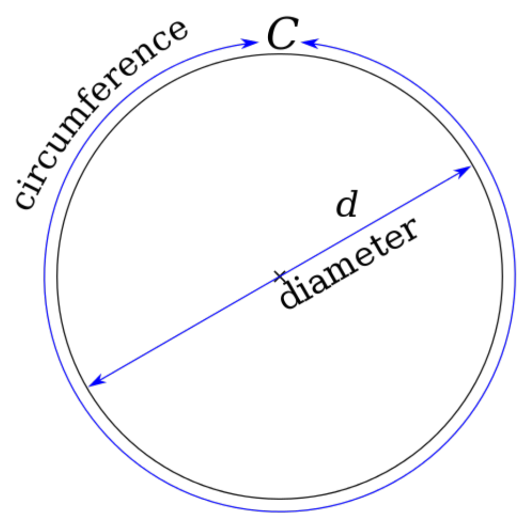
Pi

Chapter 19

Area of a circle

circumference

A fundamental result about circles is that the ratio of the circumference of a circle to its diameter is independent of the size of the circle. All circles have the same shape.



The proportionality constant was named in the early 1700s and popularized by Euler a few decades later:

$$\pi = \frac{C}{d}$$

Since the radius is one-half the diameter, $2r = d$ and

$$2\pi r = C$$

This is usually stated as a self-evident fact, that π is a *constant*, but it is actually a theorem to be proved. We defer that one for the moment.

area of a circle

Figures in the plane have area: triangles, squares and rectangles, and straight-sided polygons, so-called rectilinear figures. But also, circles and ellipses and parabolas.

Then there are solid figures, like cubes and pyramids, and cones and spheres, that have volume.

For a rectangular figure, it is easy to see why the definition of area as length times width makes sense. For a cube, the volume is the length times width times height.

One of the miracles of calculus is that it can give us areas and volumes of curved figures. But some of those results were available from Greek geometry, even before calculus. We'll see a bit of that here. As we start reasoning about circles, we recognize that the area of a circle is going to be something *squared*, because it occupies space in the plane.

According to wikipedia

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

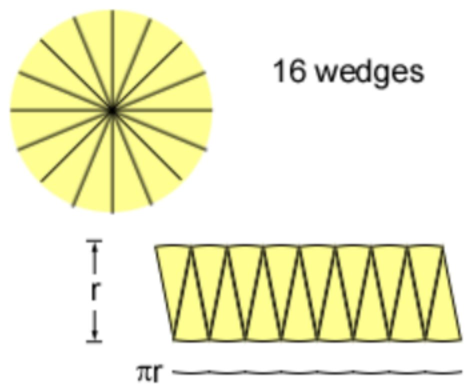
Eudoxus of Cnidus, born in the 5th century (408 BCE), proved that the area of a circle, like that of regular polygons, is proportional to both horizontal and vertical dimensions, and thus is proportional to the radius squared.

The area of a circle is equal to π times the radius squared.

$$A = \pi r^2$$

pizza proof

Imagine dividing a circle into wedges, like you might do with a pizza. Here, the pie has been divided into 16 parts.



Since the pieces are triangular, it is easy to stack them next to each other with the bases and tips alternating, as shown. Of course the bases are not straight, but have the same curvature as the edge of the circle.

The length of the short side is the radius, r, although it is angled. The original perimeter or circumference is divided into the top and the bottom of the figure, so the length of the long side is approximately one-half the circumference and thus, with length times width, we obtain

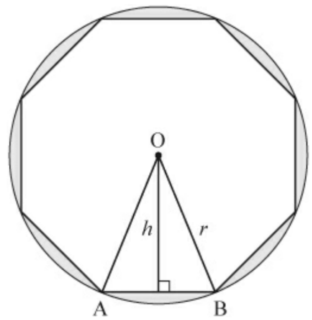
$$A = r \cdot \frac{1}{2} \cdot 2\pi r = \pi r^2$$

The trick is to imagine what happens if we subdivide the circle into many slices. The more slices, the more vertical the side, and the straighter the edges. If there are infinitely many slices, the edges will be perfectly straight and this calculation becomes exact.

The pizza proof is very much like one attributed to Leonardo da Vinci, among others.

apothem

What follows is a more formal statement of the pizza proof. Consider a regular polygon inscribed in a circle. Each side of the polygon is the base of a triangle with its vertex at the center of the circle.

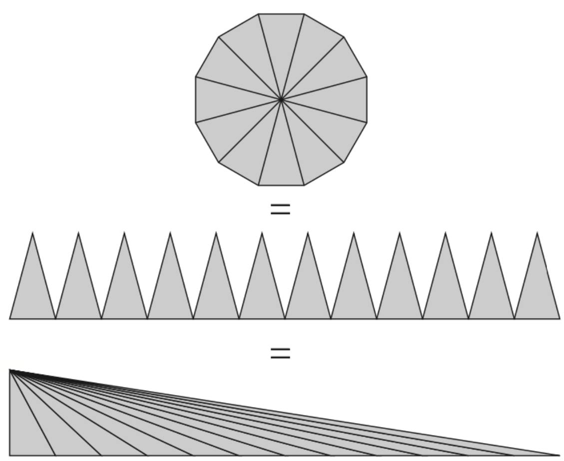


The height of these triangles, all the same, is called its apothem. We calculate both the height and the base for specific examples elsewhere. Our point here is that if the polygon has n sides, then the total area is

$$A = n \cdot \frac{1}{2} hb$$

and this is *approximately* the area of the circle.

A fun way of visualizing this is to stack the triangles next to each other and then slide the top vertices all the way to the left (this doesn't change the area of any of the triangles).



Now imagine that we let the number of sides get very large. As n increases, $nb \to 2\pi r$, the circumference of the circle.

$$A = 2\pi r \cdot \frac{1}{2}h = \pi r h$$

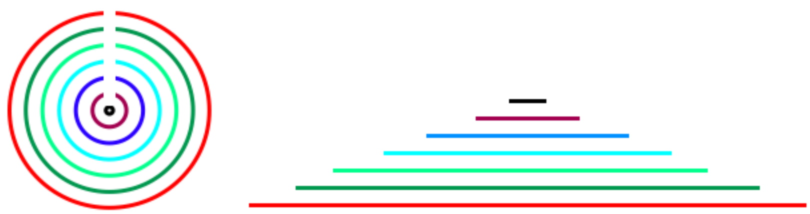
What happens to h? As n increases, $h \to r$ so then

$$A = \pi r^2$$

This only happens in the limit, when $n \to \infty$.

concentric rings

Another idea is to remove concentric strips from the edge and stack them.



This is actually the same calculation as we did previously, just based on a different idea. Here, we imagine the concentric strips infinitely thin.

We obtain a triangle of height r and base $2\pi r$ so its area is

$$\frac{1}{2}\ 2\pi r \cdot r = \pi r^2$$

Archimedes proof

A formal proof that this triangle has the same area as the circle was given by Archimedes and is found in his *Measurement of a Circle*, proposition 1. However, many sources, including

http://www.math.tamu.edu/~dallen/masters/Greek/eudoxus.pdf

attribute the proof to Eudoxus, who was perhaps the second most famous mathematician of antiquity, and a colleague of Plato in Athens.

We will sketch the idea and discuss it more formally **here**.

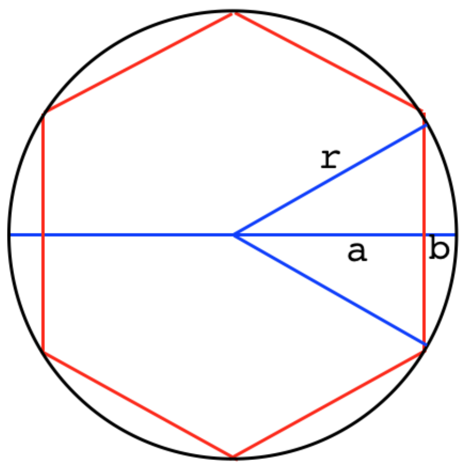
The method used is called proof by contradiction, *reductio ad absurdum*, which Hardy called one of the mathematicians finest weapons

One begins with an assumption. A slightly strained analogy might be, turning into a very narrow street and, having missed the sign, assuming it is one-way the direction you want to travel. Later, when you meet a semi trailer head-on you know there is a problem somewhere with your assumption.

inscribed polygon

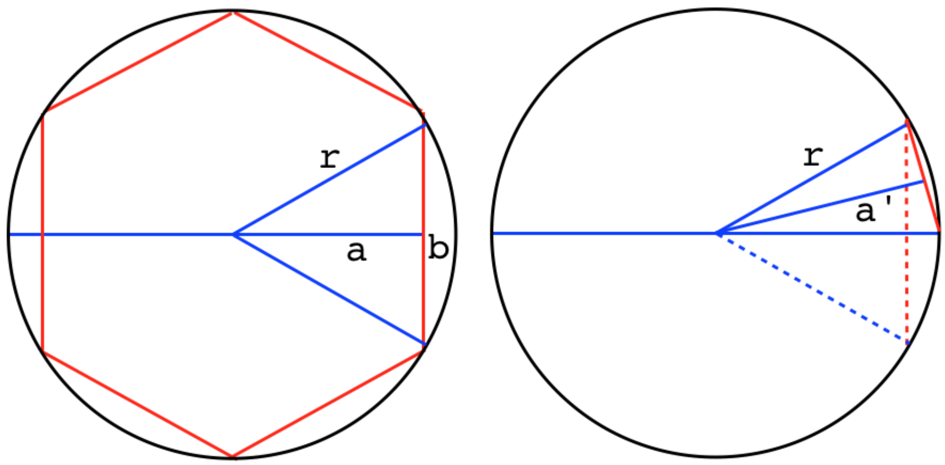
Draw a circle. Call the actual, correct, yet still unknown value for the area of the circle A.

The idea is to draw a regular polygon (all sides equal length) inside the circle.



Here we have drawn a hexagon (6-sides). If the lines from the center to the vertices have length r, then they divide the hexagon into six triangles, each with height a and base length b, where $a < r$ and b is less than the arc length. Clearly the area of the triangle is less than the area of that sector of the circle.

The base of the triangle is closest to the center (and farthest from the edge) at the center of the base, where the line segment from the center is called the apothem (labeled a). Using trigonometry, it isn't hard to calculate the area of the polygon, but we don't need to. Just call it P.



Archimedes now says, let us double the number of sides (right panel). What happens then?

The new 12-sided polygon and its component triangles will have a larger apothem a', and the total of all the bases all the way around will be larger and closer to the true circumference of the circle. There is obviously less white space between the triangle's base and the perimeter of the circle.

Thus the new P will be larger, and closer to A.

This doubling trick is easy to do. We don't even have to carry it out, just imagine doing so. We can make the difference between the area of the polygon P and the circle's true area A as as small as you please.

If your boss decides it's not close enough, just double the number of sides *again*.

That was all setup, here is the punchline.

proof

Archimedes says, let us suppose that the true area of the circle A is *not* actually equal to T (which is exactly πr^2) but is larger. Just suppose. In symbols, we are assuming that

$$T < A$$

We've already seen that $P < A$

We know we can make P as close to A as we please.

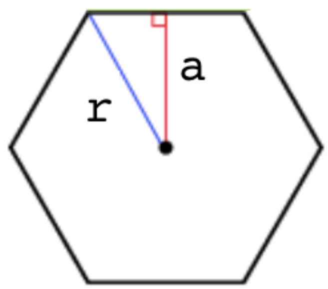
And therefore (the key point) we can make P *closer* to A than T is. The meaning of $T < A$ is that there must be some daylight between T and A. The side-doubling operation can get us into that window.

So now, by doubling, we have obtained a new P that is larger than T. We have established that

$$T < P < A$$

If $T < A$ there is really no other choice.

But look at the figure below. No matter how many sides our many-sided polygon has, $a < r$ and and the base must be less than the circumference of the circle so clearly $P < T$ for any polygon.



This is a contradiction. We have shown by two arguments which are both logically correct that $P < T$ and also that $P > T$. There is something wrong.

The resolution is that assumption we made above, that $A > T$, cannot be right. Therein lies our problem. A is not greater than T. It is either less than or equal to T.

But now try it the other way around. Circumscribe the circle with a hexagon that goes around the outside and run the argument again, and you will find that it cannot be true that $A < T$ either.

But if A is neither less than nor greater than T there is only one possibility, equality:

$$A = T = \pi r^2$$

$\square$

Plutarch, talking about Archimedes:

> It is not possible to find in all geometry more difficult and intricate questions, or more simple and lucid explanations. Some ascribe this to his natural genius; while others think that incredible effort and toil produced these, to all appearances, easy and unlabored results. No amount of investigation of yours would succeed in attaining the proof, and yet, once seen, you immediately believe you would have discovered it; by so smooth and so rapid a path he leads you to the conclusion required.

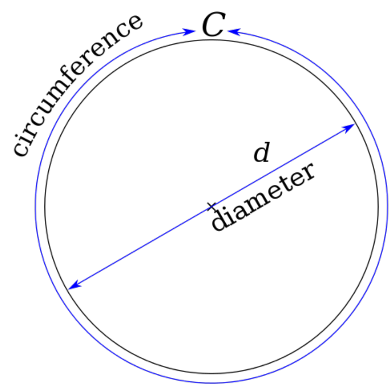
There is one additional point. Archimedes actually provides a way of calculating the improved area of each successive polygon (or its perimeter, it is really the same problem) obtained by side-doubling.

Each cycle gives a smaller and smaller improvement, which means that there is a limiting value of this process.

That value is π, when talking about the perimeter for a unit diameter, or equivalently when talking about the area, for a unit radius. We will see how this works later, suffice it to say that the side-doubling trick gives us a way to calculate the value of π to any accuracy we have the patience to compute.

Chapter 20

Pi is a constant



We began the book with a bold claim: the ratio of the circumference of a circle to its diameter is a constant, independent of the length of the diameter:

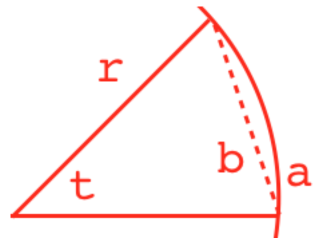
$$\pi = \frac{C}{d} = \frac{C}{2r}$$

We did not prove this theorem at the time but will do so now.

We need the idea of limits, which was introduced previously, and a property of similar triangles. The theorem is: if two triangles are similar, then their sides are proportional to each other.

Consider an arc a of a circle and two radii.

The triangle corresponding to that arc has base b. We can restate Archimedes argument about inscribed polygons by saying that, in the limit, as the inscribed polygon gets very close to being the same as the circle, $b \to a$.

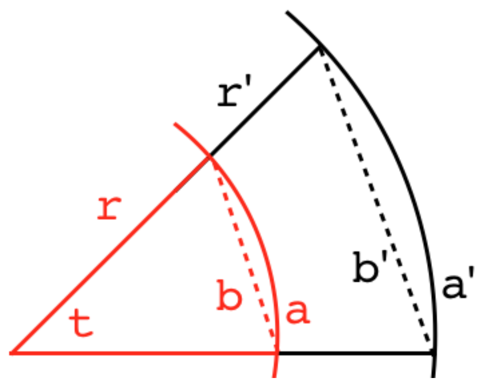


So if there are n pieces ($t = 2\pi/n$), the ratio of the circumference to the arc is just n and we have

$$n = \frac{na}{a} = \frac{C}{a} = \frac{C}{b}$$

The last step is "in the limit."

Now draw a larger arc. In the same way $b' \to a'$.



and

$$n = \frac{C'}{b'}$$

since t and n haven't changed:

$$n = \frac{C'}{b'} = \frac{C}{b}$$

But by similar triangles the ratios are equal:

$$\frac{r}{b} = \frac{r'}{b'}$$

so

$$\frac{C'}{C} = \frac{b'}{b} = \frac{r'}{r}$$
$$\frac{C'}{r'} = \frac{C}{r}$$

Suppose for a moment that $C = 2\pi r$ and $C' = 2\pi' r'$ and we don't know how π compares to π':

$$\frac{2\pi' r'}{r'} = \frac{2\pi r}{r}$$

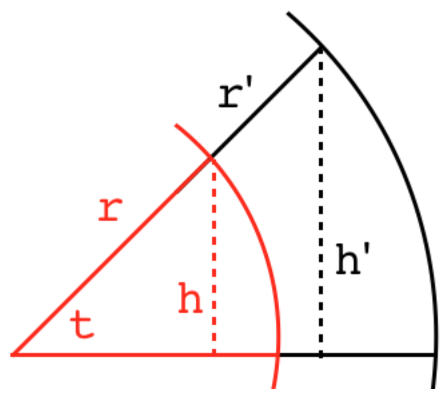
We have that $\pi = \pi'$.

□

Second proof:

Here is a simple variant which assumes something we will prove in the section on sine and cosine. If this is confusing, it can easily be skipped.

Drop the altitude h in each of the two similar triangles. The ratio h/r is equal to $\sin t$, but the arc length is equal to t, measured in radians.



In the limit that $n \to \infty$, the ratio between t and $\sin t = h/r$ is equal to our "special limit":

$$\lim_{n\to\infty} \frac{t}{\sin t} = 1$$

If the ratio to the sine is equal to 1, so is the ratio to its inverse and thus the ratio s/r is constant, which is what we wanted to prove.

□

Chapter 21

Estimating Pi

Here is a short Python program that carries out the estimation process which Archimedes invented to find the value of π. It computes the cotangent (the inverse tangent) of the half-angle by adding the cotangent and the cosecant (the inverse sine) of the parent angle.

```
from math import sin, tan, radians, sqrt

def f(csc,cot):
    Cot = csc + cot
    return sqrt(1 + Cot**2),Cot

phi = radians(60)
x = 1/sin(phi)
y = 1/tan(phi)
n = 3

for i in range(10):
    x,y = f(x,y)
    n *= 2
    print("%4d: %3.8f" % (n,round(1/x*n,8)))
}
```

This algorithm uses an inscribed polygon in a circle of diameter 1. When multiplied by the number of sides of the polygon, we get an estimate for π. I started with 60° so it would print the result for a hexagon after round one.

This estimate bounds π from below. Here is the output:

```
> p3 script.py
   6: 3.00000000
  12: 3.10582854
  24: 3.13262861
  48: 3.13935020
  96: 3.14103195
 192: 3.14145247
 384: 3.14155761
 768: 3.14158389
1536: 3.14159046
3072: 3.14159211
>
```

A similar approach with the tangent bounds π from above. Archimedes stopped at 96, after four rounds of doubling a hexagon.

Of course, Archimedes used fractions. He gave this part of his result as $\pi > 223/71$. That's about 3.140845. The upper bound was $\pi < 22/7$, which is about 3.142857.

Ptolemy used $377/120 = 3.141666...$ as an estimate for π.

In the 5th century AD, Zu Chongzhi used $355/113 = 3.1415929...$ His computation is strictly based on Pythagoras and is rather unwieldy compared to Archimedes. However, he carried out the calculation to nine rounds, and because of that he realized that 355/113 is a better estimate. It is amazingly accurate.

The answer as to how Ptolemy missed it seems to be that he didn't know the true value of π well enough to believe that 355/113 was any better than 377/120. The other possibility is that he was led to his value by some logic, and didn't check nearby values systematically.

I wrote about that here:

https://telliott99.blogspot.com/2021/04/pi-again.html

Part VII

Later geometry

Chapter 22

The Almagest

The Almagest is a treatise written by Ptolemy in the second century AD which laid out a geocentric (earth-centered) picture of the universe. Among other things, it relies on his calculation of a table of lengths of chords formed by angles between 0 and 180 degrees in very small increments.

Chords are related to the sine function, as we've seen. The ratio of the chord length to the diameter is equal to the sine of the half-angle.

So far, the only two angles for which we've calculated the sine and cosine are those which are either one-half or one-third of a right angle. We get another one (two-thirds of a right angle) by the properties of complementary angles.

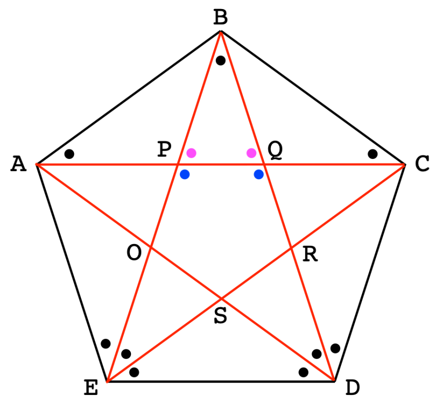
Here, we start by extending our calculation to angles that are one-fifth or two-fifths of a right angle. These come naturally from the study of the pentagon.

regular pentagon

The regular pentagon has five sides of equal length. We looked at its properties **here**.

The first step was to show that each vertex angle is divided into three equal parts by the internal chords of the pentagon.

Let the measure of that angle be t. Each one has a measure of $t = 1/5 \cdot \pi$ or $36°$. (From now on we will suppress the degree symbol in angle measurements, so the angle is just 36). Five of them add up to one triangle or 180. Only some of the equal angles are marked in this diagram.



Another major result is that the figure contains five parallelograms (actually, each is a rhombus, with four equal sides). Each of the five internal chords is the base of one of five congruent triangles containing two sides and a vertex, so they are all equal in length.

Since entire chords are equal and the sides of the parallelograms are all equal, each of the small parts remaining, such as BP, are equal to their counterparts.

The small, central pentagon is a regular pentagon, so all of its vertex angles are equal, and their measure is $3t$. Therefore the supplementary angles marked in magenta are equal and their measure is $2t$. Thus, $\triangle BPQ$ is isosceles and is similar to $\triangle BED$ by SAS similarity.

Let the side length of the small, central pentagon (e.g. PQ) be equal to 1 and the length BP be equal to x, so the ratio of the long side to the base in $\triangle BPQ$ is x.

BE is equal to $2x + 1$ and $DE = AQ$ is equal to $x + 1$ so by equal ratios in similar triangles

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

$$x^2 + x = 2x + 1$$

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

We will re-label x as ϕ. The solution to this quadratic equation is $\phi = (1 + \sqrt{5})/2$. The second solution is negative, so we will not worry about it here, where we are concerned with ratios of lengths.

In any isosceles triangle with vertex angle 36, the ratio of either of the equal sides to the base is ϕ.

algebra with ϕ

We first verify that ϕ solves the equation

$$\phi^2 = \frac{1}{4}(6 + 2\sqrt{5}) = \frac{1}{2}(3 + \sqrt{5})$$

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

So far we have

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

and

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

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

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

If we continue, we'll get the Fibonacci sequence. Divide by ϕ and rearrange:

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

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

And finally

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

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

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

Some expressions involving ϕ I have only been able to figure out by working backward from the answer. We'll see examples in a bit.

a little trigonometry

We have that the ratio of BE to ED is equal to ϕ, so the inverse ratio is $1/\phi$.

One-half of that is the sine of one-half of t, $36/2 = 18$.

$$\sin 18 = \frac{1}{2\phi}$$

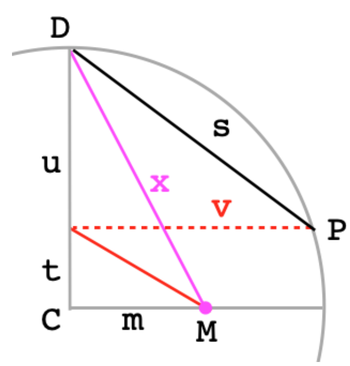
If the pentagon is inscribed in a circle then the length of the side s is the chord formed by an inscribed angle of 36 or $s = 2r \sin 36$.

The double angle formulas allow us to get to $\sin 36$ and so find s in terms of r, but we also have some constructions that work. We look at one of those now.

construction 1

Wikipedia gives two methods to construct a regular pentagon. I have redrawn their first figure.

Inscribe the pentagon in a unit circle ($CD = 1$). Draw perpendicular radii (or diameters) and divide the right horizontal radius in half at M, so length $m = 1/2$.



Draw DM. Bisect $\angle CMD$.

Extend the bisector to the vertical diagonal CD. Finally, draw the horizontal to intersect the circle at P. We label the lengths with single letters to make the algebra more intuitive.

We claim that DP or s is one side of a regular inscribed pentagon. We will first verify that the construction gives the correct side length.

calculation

Let the length of DM (magenta line) be x. By the Pythagorean theorem

$$x^2 = m^2 + 1^2 = (\frac{1}{2})^2 + 1^2 = \frac{5}{4}$$

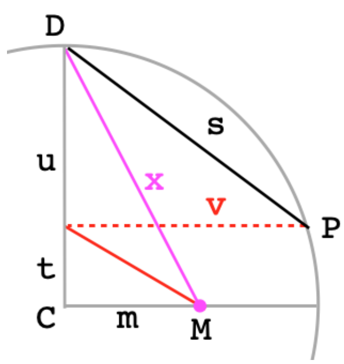
$$x = \frac{\sqrt{5}}{2}$$

Next we invoke the angle bisector theorem:

$$\frac{t}{m} = \frac{u+t}{x+m}$$

$$\frac{t}{1/2} = \frac{1}{\sqrt{5}/2 + 1/2}$$

$$t = \frac{1}{1+\sqrt{5}} = \frac{1}{2\phi}$$



We note in passing that t has the value of the sine of 18. Next:

$$v^2 = 1^2 - t^2$$

We haven't drawn the hypotenuse for the above triangle, but its base is the dotted red line. Then

$$u^2 = (1-t)^2 = 1 - 2t + t^2$$

and

$$s^2 = u^2 + v^2$$

$$= 2 - 2t = 2 - \frac{1}{\phi}$$

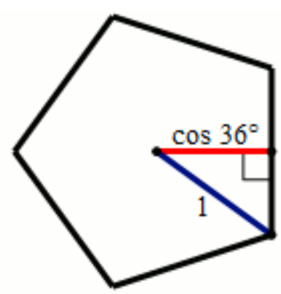
Since $1/\phi = \phi - 1$, the right-hand side is $2 - (\phi - 1) = 3 - \phi$ and the result is finally:

$$s = \sqrt{3 - \phi}$$

Now consider the $\angle CDP$, that is one-half of the vertex angle of a pentagon, namely, one-half of 108 or 54. The cosine of that angle is side s divided by the diameter, or $s/2$.

The same value is also the sine of the complementary angle, 36. I get that sine 36 is equal to $\sqrt{3 - \phi}/2$. We will use trigonometry to derive the same result.

Another way to look at this is to actually draw the whole pentagon inscribed into a unit circle, and then draw its apothem, the line that splits the side in half. Since there are five sides, the central angle is 72 and one-half that is 36. So the red line is the cosine of 36 and the sine of 36 is one-half the side length.



Observation. Note that $\cos 36° = \frac{1+\sqrt{5}}{4}$.

The figure claims that

$$\cos 36 = \frac{1 + \sqrt{5}}{4}$$

It's a simple challenge to derive it directly:

$$\cos^2 36 = 1 - (\frac{\sqrt{3 - \phi}}{2})^2$$

$$= 1 - \frac{3 - \phi}{4}$$

$$= \frac{1 + \phi}{4} = \frac{\phi^2}{4}$$

$$\cos 36 = \frac{\phi}{2}$$

We can also check by squaring both results and adding:

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

but $\phi^2 - \phi = 1$ so the result is just 1, naturally.

Writing what we have so far all in one place:

- $\sin 18 = 1/2\phi = (\phi - 1)/2$
- $\cos 18 = \sqrt{2+\phi}/2$ (see below)
- $\sin 36 = \sqrt{3-\phi}/2$
- $\cos 36 = \phi/2$

Recall the half-angle formula for cosine:

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

so

$$\cos^2 18 = \frac{1 + \phi/2}{2}$$

$$\cos 18 = \sqrt{\frac{2+\phi}{4}} = \frac{\sqrt{2+\phi}}{2}$$

We can also check this using our favorite identity, as follows:

$$(\frac{1}{2\phi})^2 + \frac{1 + \phi/2}{2}$$

$$= \frac{1}{4}(\frac{1}{\phi^2} + 2 + \phi)$$

Go back to the section on algebra:

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

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

Substituting into what's in the parentheses above immediately yields the correct answer, 1.

We can also confirm that

$$\sin 36 = 2 \sin 18 \cos 18$$
$$= 2 \cdot \frac{1}{2\phi} \cdot \frac{\sqrt{2+\phi}}{2} = \frac{\sqrt{2+\phi}}{2\phi}$$

This isn't close to what we had, but algebra with ϕ can get pretty weird. Let's just go backwards from the answer.

$$\frac{\sqrt{2+\phi}}{2\phi} = \frac{\sqrt{3-\phi}}{2}$$

squaring

$$2 + \phi = (3-\phi)\phi^2$$
$$= 3\phi^2 - \phi^3$$
$$= 3(\phi+1) - (2\phi+1)$$
$$= 2 + \phi$$

which checks.

Let's do one more:

$$\sin 54 = \sin 36 + 18 = \sin 36 \cos 18 + \sin 18 \cos 36$$
$$= \frac{\sqrt{3-\phi}}{2} \cdot \frac{\sqrt{2+\phi}}{2} + \frac{1}{2\phi} \cdot \frac{\phi}{2}$$
$$= \frac{1}{4}\ (\sqrt{(3-\phi)(2+\phi)} + 1)$$
$$= \frac{1}{4}\ (\sqrt{6+\phi-\phi^2} + 1)$$

It looks like a bit of a mess. But we know the answer must be simple because $\sin 54 = \cos 36 = \phi/2$. In other words, the expression in parentheses must be equal to 2ϕ.

$$2\phi \stackrel{?}{=} \sqrt{6+\phi-\phi^2} + 1$$

But $\phi^2 = 1 + \phi$ so the right-hand side is

$$\sqrt{6+\phi-1-\phi} + 1 = \sqrt{5} + 1$$

which is, indeed, 2ϕ!

trigonometry

So now we actually have all the pieces and might follow Ptolemy's method.

I would rather digress to show a bit of new trigonometry. It will connect to another place in the book where we talk about de Moivre's theorem.

And there is a fancy trick that's specific to 18 degrees. Later we'll adjust to 36.

First, the new general result: the formula for the cosine of three times the angle.

cos 3A

We use the standard angle sum formula in a new version:

$$\cos 3A = \cos 2A \cos A - \sin 2A \sin A$$

so then a second application of the formula gives

$$\cos 3A = (\cos^2 A - \sin^2 A) \cos A - (2 \sin A \cos A) \sin A$$

$$= (2 \cos^2 A - 1) \cos A - 2 \cos A(1 - \cos^2 A)$$

Just count up the terms. We have

$$\cos 3A = 4 \cos^3 A - 3 \cos A$$

□

This result can also be obtained by **de Moivre's theorem**:

$$\cos nx + i \sin nx = (\cos x + i \sin x)^n$$

We have

$$\cos 3x + i \sin 3x = (\cos x + i \sin x)^3$$

$$= \cos^3 x + 3i \cos^2 x \sin x + 3i^2 \cos x \sin^2 x + i^3 \sin^3 x$$

We only need the real part, which is

$$\cos 3x = \cos^3 x - 3 \cos x \sin^2 x$$

$$= \cos^3 x - 3 \cos x(1 - \cos^2 x)$$

which simplifies to the same result.

sine of 18 degrees

Let $A = 18$. Then

$$5A = 90$$

$$2A = 90 - 3A$$

$$\sin 2A = \cos 3A$$

Plug in the previous result

$$\sin 2A = 4\cos^3 A - 3\cos A$$

$$2\sin A\cos A - 4\cos^3 A + 3\cos A = 0$$

Each term contains one copy of $\cos A$, and since that is non-zero, we simply multiply by $1/\cos A$ on both sides, giving

$$2\sin A - 4\cos^2 A + 3 = 0$$

$$2\sin A - 4(1 - \sin^2 A) + 3 = 0$$

$$4\sin^2 A + 2\sin A - 1 = 0$$

Now we have a quadratic in $\sin A$. The roots are

$$\sin A = \frac{-2 \pm \sqrt{4 + 16}}{8}$$

$$= \frac{-2 \pm 2\sqrt{5}}{8}$$

We take the positive root because $\sin A > 0$ in the first quadrant.

$$\sin 18 = \frac{-1 + \sqrt{5}}{4}$$

$$= \frac{1}{2} \cdot \frac{\sqrt{5} - 1}{2}$$

This is almost ϕ.

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

so

$$\sin 18 = \frac{\phi - 1}{2}$$

and since we showed earlier that

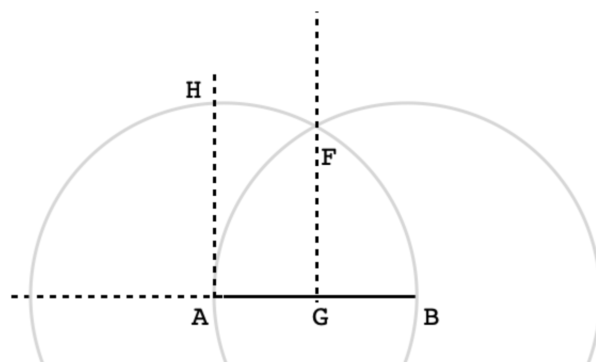
$$\frac{1}{\phi} = \phi - 1$$

the previous result can be re-written as

$$\sin 18 = \frac{1}{2\phi}$$

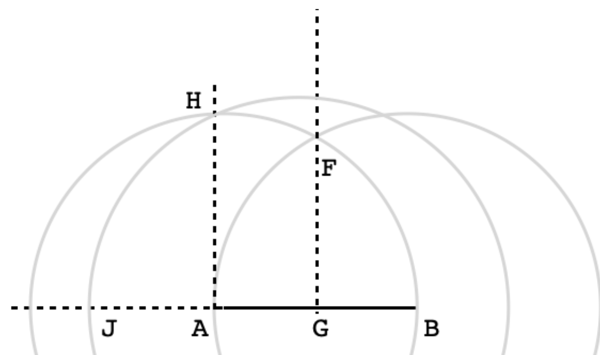
construction 2, starting from a side length

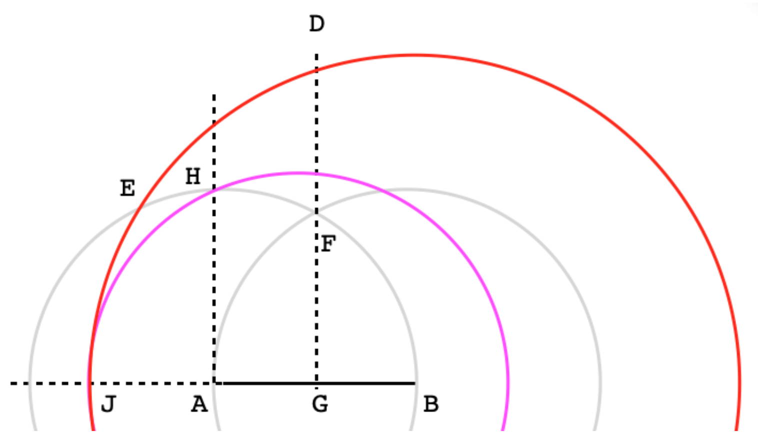
Wikipedia gives another method for constructing a regular pentagon, this time starting from a given side length, AB. The first part is straightforward.



Draw two congruent circles with a radius of length AB centered at A and B. Construct the perpendicular bisector FG (not shown). Construct the perpendicular from A to meet the circle centered on A, at H.

We will construct two more circles. The first one is centered at G with radius GH, intercepting the extension of AB at J.





The second one is centered at B with radius BJ. The intersection with the previous circle (at E) and with the perpendicular at D are vertices of the pentagon.

The last vertex, at C, can be constructed by laying off arcs of the distance DE from D and B, or by repeating the construction with a large circle centered on A.

why this works

It is claimed that

$$\frac{BJ}{AB} = \frac{AB}{AJ} = \phi$$

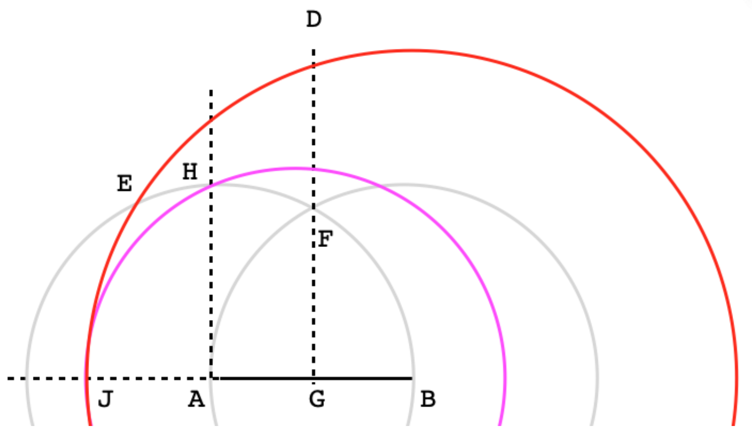
Let's see.

$$AB = 1, \qquad AG = \frac{1}{2}$$

$$AH = AB = 1$$

$$GJ = GH = \sqrt{1^2 + (\frac{1}{2})^2} = \frac{\sqrt{5}}{2}$$

$$BJ = GJ + \frac{1}{2} = \phi$$

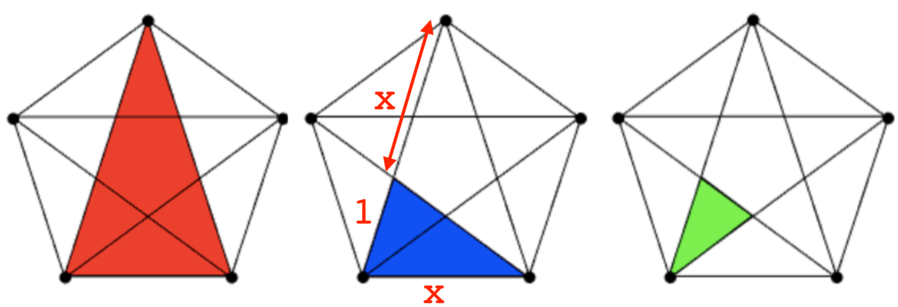


and

$$AJ = BJ - 1 = \phi - 1$$

$$\frac{AB}{AJ} = \frac{1}{\phi - 1} = \frac{\phi}{\phi^2 - \phi} = \phi$$

So we can confirm that the ratios are correct. Now we just need to connect the lengths in the diagram to sides of triangles in our view of the pentagon with internal chords.



The top vertex is easy.

$$AD = BD = BJ = \phi$$

while $AB = 1$ so the ratio is ϕ, which matches the red triangle.

For vertex E, we must show that $BE = \phi$. But $BE = BJ = \phi$.

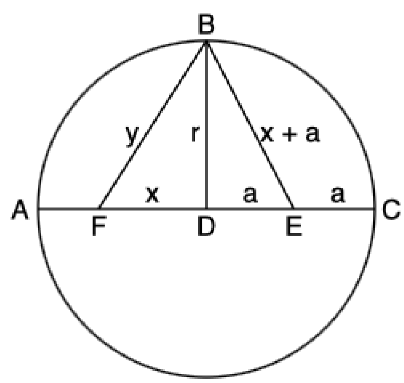
We verify that D is placed correctly, as follows. D lies on the perpendicular bisector of AB and is also a distance ϕ away from both A and B, forming an isosceles triangle with long sides ϕ and short side 1.

We could try to verify that D is placed correctly by proving that the base $DE = 1$ to complete the triangle. But this seems difficult, and it's easy to show that $\triangle ABE$ has two sides of length 1 and one, BE, of length ϕ. Since this triangle is congruent to the short fat ones we find in a pentagon of side length 1, we're done.

$\square$

construction 3

The third construction I have is said to be contained in Ptolemy's book *The Almagest* (link below).



Once again, we divide the radius in half: $a = r/2$. The hypotenuse BE is marked out so that $EF = BE$. Then the second hypotenuse, BF is claimed to be the side of a pentagon inscribed in a unit circle and BE is the side of the decagon.

If we let $r = 1$ and $a = 1/2$ then

$$x + a = \sqrt{(\frac{1}{2})^2 + 1)} = \frac{\sqrt{5}}{2}$$

so

$$x = \frac{\sqrt{5}}{2} - \frac{1}{2} = \phi - 1$$

and then BF or

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

but $\phi^2 - \phi = 1$ so we have

$$y = \sqrt{3 - \phi}$$

$BF = y$ is supposed to be the side of a pentagon and that matches what we had before. $BE = x$ is supposed to be the side of a decagon. We check that by recalling that the side of the pentagon is twice the sine of 36, which matches.

So the side of the decagon is twice the sine of 18 which is simply $1/\phi$. We must show that this is equal to $\phi - 1$. But we did this already, back near the beginning.

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

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

And that's it.

more about Ptolemy

This completes the first part of Ptolemy's *The Almagest*, described in the link below.

https://hypertextbook.com/eworld/chords/#table1

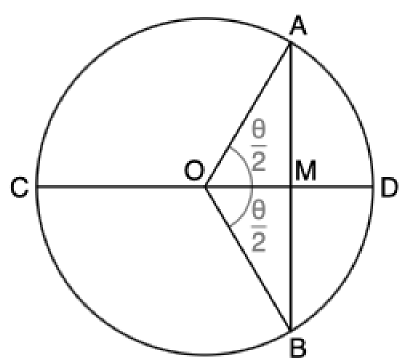


Figure 1

This figure is just to remind us that the discussion in the link describes values for the chord AB of a central angle θ, or inscribed (peripheral) angle $\theta/2$. In modern language, we would say that the chord length $AB = 2r \sin \theta/2$ or

$$\frac{AB}{CD} = \sin \theta/2$$

When Ptolemy associates a value to a particular angle θ we should find the sine of one-half θ for comparison.

Also, Ptolemy deals in ratios and his units are degrees, minutes and seconds. In particular, the diameter CD is also divided into degrees, namely, 120°.

Table 1: Chords of the special angles

angle	crd
36°	37°4'55"
60°	60°
72°	70°32'3"
90°	84°51'10"
108°	97°4'56"
120°	103°55'23"
144°	114°7'37"
180°	120°

On the first line, the angle is given as 36° and from what we said, we need to look for the sine of 18°, which we have as

$$\frac{\phi - 1}{2} = \frac{\sqrt{5} - 1}{4} = 0.3090$$

This result must be converted to degrees for comparison. I don't want to do arithmetic in degrees, etc., so I multiply $120 \times 3600 = 432000$. The diameter is divided into that many parts.

What we have is 0.309017 times that or 133495.34. I divide by 3600 and the whole part is 37, the modulus is 295. Dividing by 60 I then get 4 with 55 and a bit left over. That's a match.

The second part of *The Almagest* involves what we know as the sum of angles formulas. We actually used Ptolemy's theorem and his derivation of those relationships as our preferred proofs of them (**here**).

He uses the sum of arcs (i.e. angles), difference of arcs and half-arc formulas to fill out a table for every angle up to 180° in increments of 3/4°.

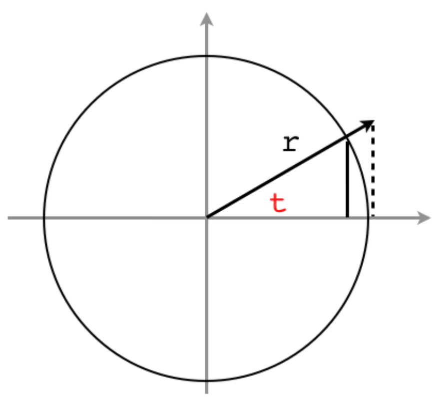
So he went from $36 \to 18$ and $30 \to 15$ and then $18 - 15 = 3$ and half 3 is 1.5 and half again is 3/4. That's a lot of work.

approximation

The last part of the book involves an approximation for the sine of very small angles. This allows not just small values to be calculated, but also fine-grained interpolation for the table as a whole.

The basic idea is that as θ gets small, the sine becomes approximately equal to the arc traced by the angle. For a hemisphere the arc is π and the chord is the diameter, which I am calling 2, for a unit circle. For one-quarter circle the arc is $\pi/2$ and the chord is $\sqrt{2}$. So the first ratio is π divided by 2 and the second is π divided by $2\sqrt{2}$. Clearly, the denominator is increasing.

To put it another way



For $0 < t < \pi/2$, we have $\sin t < t < \tan t$, but as $t \to 0$ they all become equal.

If you already know about power series, you know (or can look up) that

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

so

$$\sin x \approx x$$

when x is sufficiently small.

The justification used by Ptolemy involves a relationship called Aristarchus' inequality, which says that for $\alpha < \beta$:

$$\frac{\sin\alpha}{\sin\beta} > \frac{\alpha}{\beta}, \qquad \frac{\sin\alpha}{\alpha} > \frac{\sin\beta}{\beta}$$

If α is smaller than β, $\sin\alpha$ gets proportionately closer to α than $\sin\beta$ is to β.

$$\sin\alpha \cdot \frac{\beta}{\alpha} > \sin\beta$$

An aid to memory: when the fraction multiplying the sine or chord is greater than one, then we need the greater than symbol on that side of the inequality.

Suppose that $\angle\alpha$ is one-half $\angle\beta$. Then the chord corresponding to $\angle\alpha$ or the sine of α is *more than* one-half the corresponding value for $\angle\beta$.

lemma and proof

I want to give a proof of Aristarchus' inequality. Rather than use the one he or Ptolemy would use, we'll do something more modern.

It's based on a proof in wikipedia, and this gives a chance to introduce a trigonometric identity we've not seen until now.

https://en.wikipedia.org/wiki/Aristarchus%27s_inequality

Let's start with a lemma. Recall the formulas for the sine of the sum and difference of two angles:

$$\sin x + y = \sin x \cos y + \sin y \cos x$$

$$\sin x - y = \sin x \cos y - \sin y \cos x$$

Adding them makes the second term disappear and gives

$$\sin(x + y) + \sin(x - y) = 2 \sin x \cos y$$

Now, the neat idea. We do something close to what we want and then go back and fix it. For starters, let

$$A = x + y, \qquad B = x - y$$

Then

$$A + B = 2x, \qquad A - B = 2y$$

so

$$\sin A + \sin B = 2 \sin \frac{A+B}{2} \cos \frac{A-B}{2}$$

We actually want something slightly different. We want the sine term to have $B - A$. So go back and instead let

$$-A = x - y, \qquad B = x + y$$

Then

$$\begin{gathered} B - A = B + (-A) \\ = x + y + (x - y) = 2x \\ x = (B - A)/2 \end{gathered}$$

and

$$\begin{gathered} B + A = B - (-A) \\ = x + y - (x - y) = 2y \\ y = (B + A)/2 \end{gathered}$$

and

$$\sin(x + y) + \sin(x - y) = 2 \sin x \cos y$$

becomes

$$\sin B + \sin -A = 2 \sin \frac{B-A}{2} \cos \frac{B+A}{2}$$

Finally, since $\sin -\theta = -\sin\theta$

$$\sin B - \sin A = 2 \sin \frac{B-A}{2} \cos \frac{B+A}{2}$$

We proceed to the main part of the proof.

Proof.

Consider two angles and let $A < B$, A is smaller than B. We are to prove that

$$\frac{\sin B}{B} < \frac{\sin A}{A}$$

The meaning of this is that, as we said above, if A is smaller than B then proportionally, $\sin A$ is closer to A than $\sin B$ is to B.

$$\sin A > \frac{A}{B} \cdot \sin B$$

The first step is an algebraic trick that we've seen several times. First, rearrange what we are to prove like this:

$$\frac{\sin B}{\sin A} < \frac{B}{A}$$

Subtracting 1 from both sides doesn't change the inequality

$$\frac{\sin B}{\sin A} - \frac{\sin A}{\sin A} < \frac{B}{A} - \frac{A}{A}$$

$$\frac{\sin B - \sin A}{\sin A} < \frac{B - A}{A}$$

$$\frac{\sin B - \sin A}{B - A} < \frac{\sin A}{A}$$

This statement is equivalent to what we need to prove. If we can show that this is true, we'll be done.

We proceed by finding a value in between the two expressions.

It turns out that $\cos A$ will work.

$$\frac{\sin B - \sin A}{B - A} < \cos A < \frac{\sin A}{A}$$

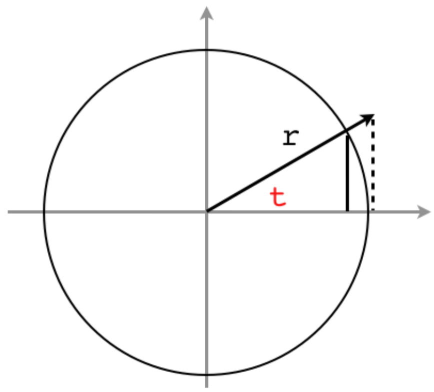
The second part is easy

$$\cos A < \frac{\sin A}{A}$$

This is just

$$A < \frac{\sin A}{\cos A} = \tan A$$

which can be shown by considering areas. Compare the triangle with the dotted line with the area of the sector in the figure below.



The complete result is $\sin A < A < \tan A$. (Recall that the area of the whole unit circle is π, while the entire angle is 2π, so the area of a sector swept out be angle θ is $A = \theta/2$.)

Now we look more closely at the left-hand side.

$$\frac{\sin B - \sin A}{B - A} < \cos A$$

First, from our lemma, the numerator is

$$\sin B - \sin A = 2 \sin \frac{B-A}{2} \cos \frac{B+A}{2}$$

We will make a simple substitution that make this term larger. Then we'll show that the resulting inequality really is correct.

Since we've made the left-hand side *bigger* and the relationship still holds, we know that the original expression is also valid.

$B > A$ so the difference $B - A$ is positive but smaller than B.

For a small positive angle, the sine of the angle is smaller than the angle itself so we can replace $\sin(B - A)/2$ by $(B - A)/2$. This makes the whole left-hand side larger.

$$2 \cdot \frac{B-A}{2} \cos \frac{A+B}{2}$$

We notice that we can cancel the original denominator (which was $B - A$), as well as that factor of 2. As a result, the whole inequality is reduced to

$$\cos \frac{B+A}{2} < \cos A$$

But the average of two angles $(B + A)/2$ is larger than the smaller angle A:

$$A < B$$

$$2A < B + A$$
$$A < \frac{B + A}{2}$$

And, as the angle gets larger, the cosine gets smaller. Hence, $\cos(B+A)/2$ is smaller than the cosine of A.

In other words, the inequality is correct. Since our manipulation made the numerator bigger and the inequality definitely holds, the original expression is valid.

$$\frac{\sin B}{B} < \frac{\sin A}{A}$$

□

examples

The example from the book uses known values for the *chord* of an angle of $3/4°$ as well as 1-1/2°. As we said, the chord corresponds to the sine of the half-angle.

The values are given in degrees, minutes and seconds of arc. Let s and t be those values:

$$s = 0° \ 47' \ 08'' = crd \ 3/4°$$
$$t = 1° \ 34' \ 15'' = crd \ 1.5°$$

It is immediately apparent that $2s > t$ but not by much, only one second of arc. It should not be surprising that we end up with a linear approximation.

The diameter is divided into 432000 parts (120 degrees, 60 minutes, 60 seconds). A calculator (and some fiddling) gives the following value for the sine of $3/4°$ in parts of 432000: 5654.705.

A similar calculation for the sine of $3/8°$ gives: 2827.413. The whole part is exactly one-half, while the fractional part is slightly larger than half for the smaller angle.

Doing the modular arithmetic, I get 47′ 07” for the first and 1° 34′ 14” for the second. Our source says they ”didn't believe in rounding up, ever”, as we do today for fractions larger than one half. I suppose the error may come from the previous calculations.

In any event, let us estimate the value for the chord c of 1° as Ptolemy would. The angle $\theta = 1^\circ$ is $4/3 \cdot s$ so the calculation goes:

$$c < \frac{4}{3} \cdot 0^\circ\ 47'\ 08'' = 1^\circ\ 2'\ 50''$$

On the other hand, $\theta = 1^\circ$ is $2/3 \cdot t$ so

$$c > \frac{2}{3} \cdot 1^\circ\ 34'\ 15'' = 1^\circ\ 2'\ 50''$$

(Remember, the greater than symbol goes with the fraction larger than one).

Clearly, we have a good approximation for c. The results seem to be exactly the same, though they cannot really be. It just appears so because of round-off error. If we do the same calculation with our modern precise decimals we get

$$c < \frac{4}{3} \cdot 2827.413 = 3769.884$$

$$c > \frac{2}{3} \cdot 5654.705 = 3769.803$$

The true value is something like $1^\circ\ 2'\ 49''$ plus 800 and some thousandths, which I cannot help but note is smaller than the reported value by (truncated to 49") by one second.

According to the source:

> The remainder of the Almagest consists of astronomical calculations: the position of the sun, moon, and planets at various times relative to the fixed stars. The Table of Chords played an important role in their compilation.

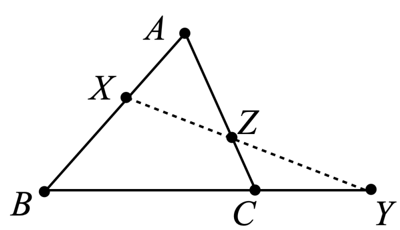
Chapter 23

Pappus

more about Menelaus

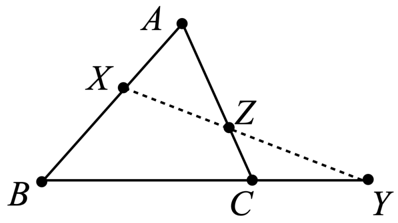
We gave two proofs of **Menelaus's Theorem** previously.

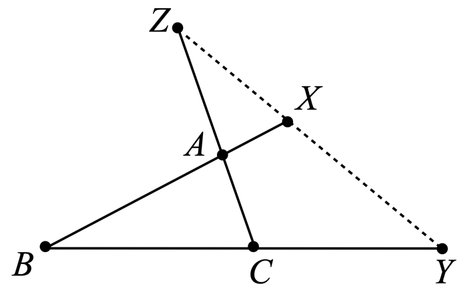
There is a somewhat more sophisticated version of the theorem which views the path around a triangle as consisting of *directed* (that is, *signed*) line segments.



Let the triangle be $\triangle ABC$, then the transversal is XZY, which meets the extension of BC at Y. Since Y does not lie between B and C, the line segments BY and YC point in opposite directions. The ratio $BY : YC$ thereby acquires a minus sign.

Every transversal has this property. Here (on the right) is a transversal that does not go through the triangle at all.





These are the only two possibilities if we exclude that the transversal goes through a vertex.

The path around the triangle has three parts:

A to X to $B \Rightarrow AX : XB$

B to Y to $C \Rightarrow BY : YC$

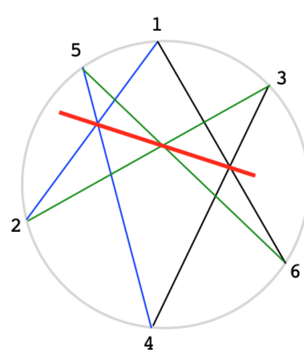
C to Z to $A \Rightarrow CZ : ZA$

Each ratio has a minus sign so the total product also has a minus sign. Menelaus's Theorem says:

$$\frac{AX}{XB} \cdot \frac{BY}{YC} \cdot \frac{CZ}{ZA} = -1$$

Pappus' theorem

As a way to explore the usefulness of Menelaus's theorem, let us look at a remarkable theorem of Pappus.

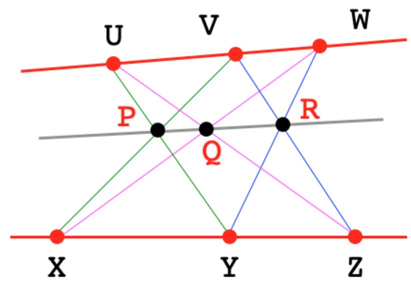


One version says that if we inscribe a hexagon in a circle and then connect some of the vertices in a certain way, then certain points will be co-linear. The hexagon we are talking about has its vertices connected in the order 123456, it is a *degenerate* hexagon.

The points of interest are the intersections of opposing sides. The point where 12 cuts 45 is on the red line, the next is 23 and 56, and the third is 34 and 61.

It seems like the red line does connect three co-linear points. We will not prove this version of the theorem.

This hexagon can be deformed into two lines. In fact we can still think of it as a hexagon, although the order of the vertices is strange: $UYWXVZ$.



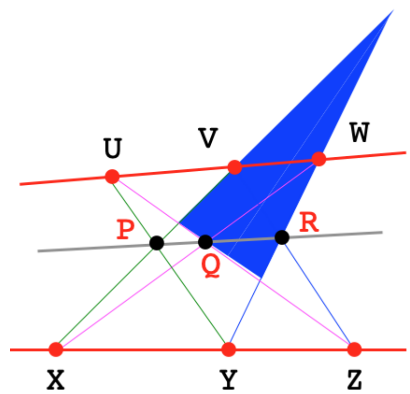
We claim that if UVW and XYZ are collinear, then the three points PQR are also collinear. These points are formed by the intersections of corresponding lines. We count the intersection of UY and VX at P, for example.

Since we started with a hexagon inscribed in a circle and ended with two sets of collinear points UVW and XYZ, collinearity of the starting position is not a requirement.

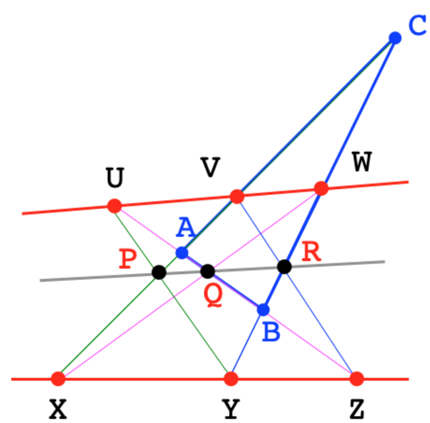
Nevertheless, we'll stick with what the source says.

Again, the idea is that $UYWXVZ$ is a (degenerate) hexagon. The points of interest are at the intersections of opposite sides: UY cuts XV at P, YW cuts VZ at R, and WX cuts ZU at Q.

We will use some of the unlabeled intersections in our proof. We focus on one particular triangle.



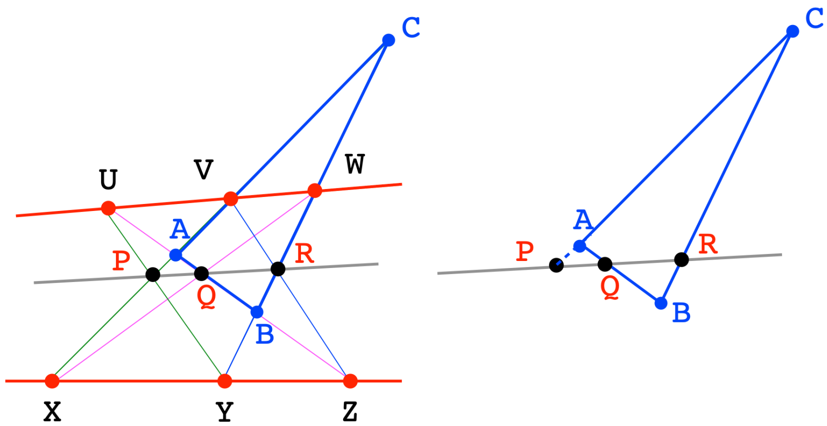
I have labeled the vertices ABC.



You will notice that the points PQR apparently lie on a transversal of $\triangle ABC$. In particular, if we can prove that the product of Menelaus's ratios is equal to -1 for PQR in $\triangle ABC$, we will have established co-linearity of the points PQR, in other words, that it really is a transversal.

We will show that

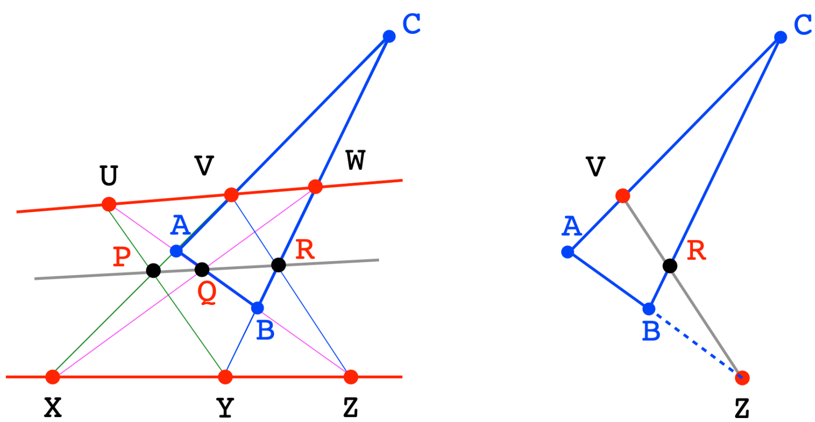
$$\frac{AQ}{QB} \cdot \frac{BR}{RC} \cdot \frac{CP}{PA} = -1$$



We will achieve this by applying the forward theorem to five other transversals.

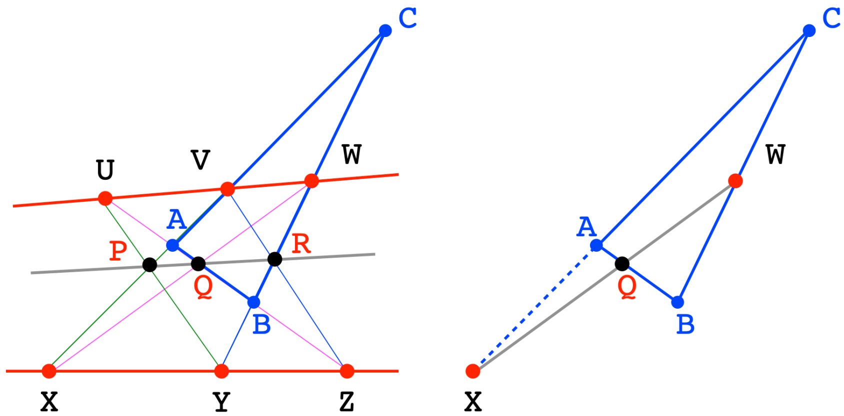
first three

The first one is VRZ. Just read it off the diagram below:



$$\frac{BR}{RC} \cdot \frac{CV}{VA} \cdot \frac{AZ}{ZB} = -1$$

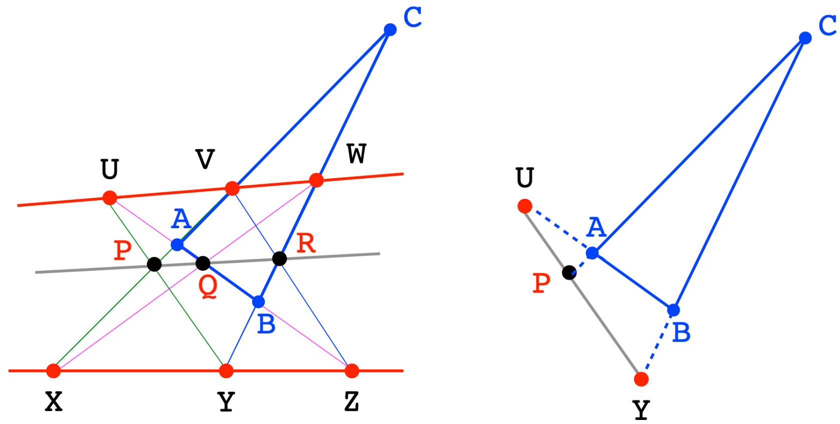
The second is XQW:



$$\frac{AQ}{QB} \cdot \frac{BW}{WC} \cdot \frac{CX}{XA} = -1$$

And the third is UPY:

$$\frac{CP}{PA} \cdot \frac{AU}{UB} \cdot \frac{BY}{YC} = -1$$



It is harder in this case, since all three points of the transversal lie outside the triangle. The key is to take the points in order: vertex-intermediate-vertex.

For each ratio, start and end at a vertex of $\triangle ABC$. The ratios go like vertex $\rightarrow$ point divided by point $\rightarrow$ vertex.

Thus

$$\frac{BR}{RC} \cdot \frac{CV}{VA} \cdot \frac{AZ}{ZB} = -1$$
$$\frac{AQ}{QB} \cdot \frac{BW}{WC} \cdot \frac{CX}{XA} = -1$$
$$\frac{CP}{PA} \cdot \frac{AU}{UB} \cdot \frac{BY}{YC} = -1$$

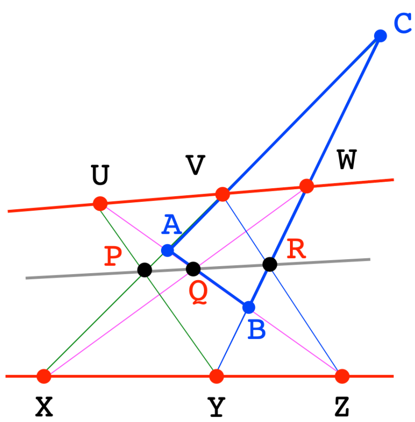
Recall that we need

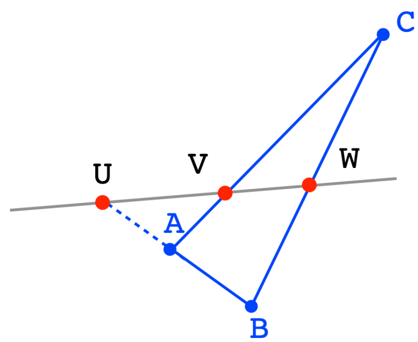
$$\frac{AQ}{QB} \cdot \frac{BR}{RC} \cdot \frac{CP}{PA} = -1$$

finishing up

In each of the products above the first term is what we seek. The others need to disappear. The fourth transversal is UVW.

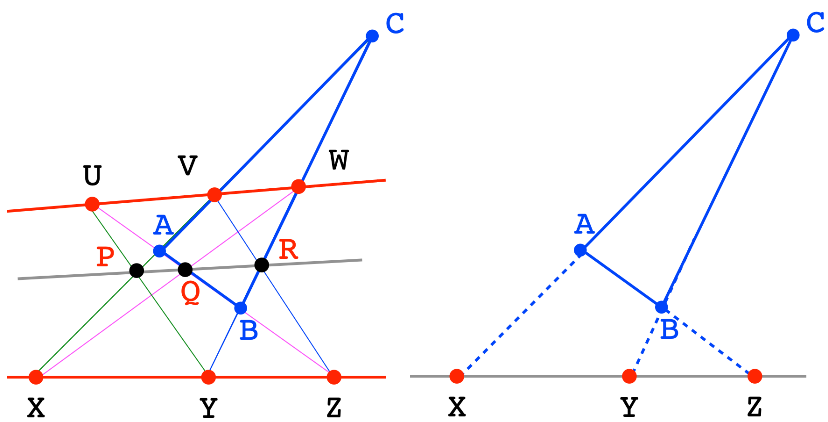
$$\frac{BW}{WC} \cdot \frac{CV}{VA} \cdot \frac{AU}{UB} = -1$$





And finally, the last one is:

$$\frac{AZ}{ZB} \cdot \frac{BY}{YC} \cdot \frac{CX}{XA} = -1$$



If we multiply together all five of these expressions, all fifteen terms, on the right we obtain just -1. Instead, we will construct a numerator with the first three of them and a denominator with the last two. Again we obtain -1 on the right-hand side.

$$\frac{\frac{BR}{RC} \cdot \frac{CV}{VA} \cdot \frac{AZ}{ZB} \cdot \frac{AQ}{QB} \cdot \frac{BW}{WC} \cdot \frac{CX}{XA} \cdot \frac{CP}{PA} \cdot \frac{AU}{UB} \cdot \frac{BY}{YC}}{\frac{BW}{WC} \cdot \frac{CV}{VA} \cdot \frac{AU}{UB} \cdot \frac{AZ}{ZB} \cdot \frac{BY}{YC} \cdot \frac{CX}{XA}}$$

Then it's time for some cancelation.

$$\frac{\frac{BR}{RC} \cdot \cancel{\frac{CV}{VA}} \cdot \cancel{\frac{AZ}{ZB}} \cdot \frac{AQ}{QB} \cdot \cancel{\frac{BW}{WC}} \cdot \cancel{\frac{CX}{XA}} \cdot \frac{CP}{PA} \cdot \cancel{\frac{AU}{UB}} \cdot \cancel{\frac{BY}{YC}}}{\cancel{\frac{BW}{WC}} \cdot \cancel{\frac{CV}{VA}} \cdot \cancel{\frac{AU}{UB}} \cdot \cancel{\frac{AZ}{ZB}} \cdot \cancel{\frac{BY}{YC}} \cdot \cancel{\frac{CX}{XA}}}$$

We are left with simply:

$$\frac{BR}{RC} \cdot \frac{AQ}{QB} \cdot \frac{CP}{PA} = -1$$

which is what we wanted.

By the converse of Menelaus's theorem, we have that PQR are collinear.

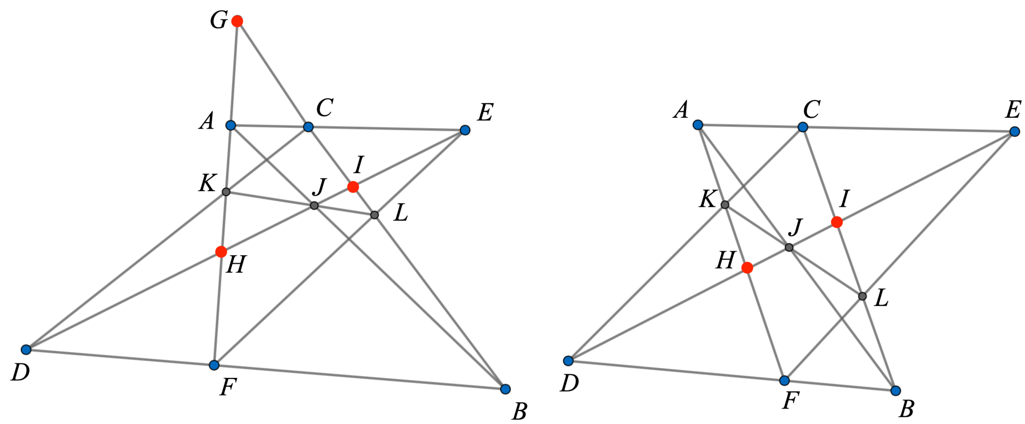
□

This proof is a constructive proof. It shows that for this diagram, the relationship holds and the points PQR are collinear.

the parallel case

What would happen if $AC \parallel BC$, or put another way, $VX \parallel WY$? Suppose also $UY \parallel VZ$? For some arrangements, we would not be able to draw the triangle we need.

Construct the diagram in an online tool like Geogebra and then play with it by sliding points around. One sees that it is easy to alter a diagram where we could not draw the triangle, to one in which we can, and the lines and intersections retain their relationships when this is done.



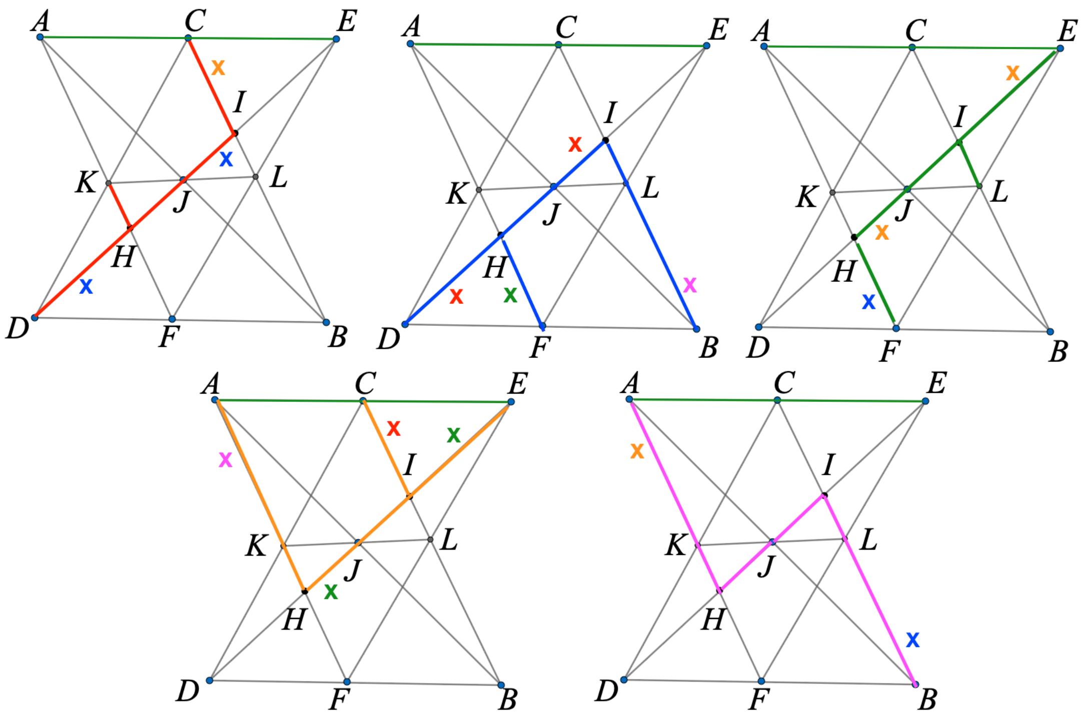
The diagram above was redrawn from

http://cut-the-knot.org/pythagoras/Pappus.shtml

which has a wonderful discussion of this and many other theorems. Here the "hexagon" is $ABCDEF$, the points to be proved collinear are KJL, and the triangle whose transversals are analyzed is $\triangle GHI$.

One way Bogolmony proposes to deal with the diagram on the right, the case of $AF \parallel BC$ and $CD \parallel EF$, is to use the parallel lines to find similar triangles.

There are a lot of them. There are five that we will use, you can see them in this diagram. For each pair of similar triangles, we show only the sides whose ratios we need. Also shown is the origin of the side that will cancel the given side.



The first two are:

$$\triangle DHK \sim \triangle DIC \Rightarrow \frac{DH}{DI} = \frac{HK}{IC}$$

$$\triangle DHF \sim \triangle DIB \Rightarrow \frac{DH}{DI} = \frac{HF}{IB}$$

Thus

$$\frac{HK}{IC} = \frac{HF}{IB}$$

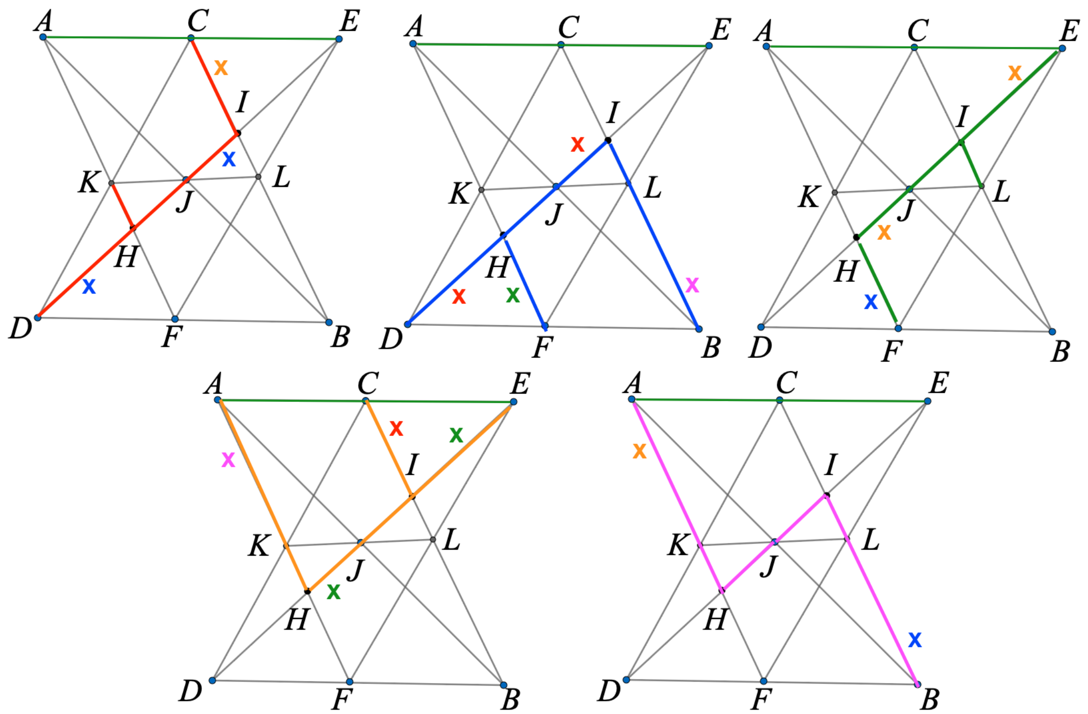
The next two are:

$$\triangle EHF \sim \triangle EIL \Rightarrow \frac{HE}{IE} = \frac{HF}{IL}$$

$$\triangle EHA \sim \triangle EIC \Rightarrow \frac{HE}{IE} = \frac{HA}{IC} = 1$$

So

$$\frac{HF}{IL} = \frac{HA}{IC}$$



Combining these results:

$$HF \cdot IC = HA \cdot IL = HK \cdot IB$$

$$\frac{HK}{IL} = \frac{HA}{IB}$$

Finally:

$$\triangle HAJ \sim \triangle IBJ \Rightarrow \frac{HJ}{IJ} = \frac{HA}{IB} = 1$$

Hence

$$\frac{HJ}{IJ} = \frac{HK}{IL}$$

Since they have two sides in proportion and also $\angle KHJ = \angle LIJ$, it follows that

$$\triangle HKJ \sim \triangle ILJ$$

But this can only be if KJL are collinear. If we didn't have collinearity we would also not have the equal vertical angles $\angle HJK = \angle IJL$.

□

We won't prove that last part, but it seems clear how to approach a proof by contradiction.

This isn't completely dispositive, since we cannot prove that this handles *all* cases. But it's sufficient for now.

Projective geometry solves this problem formally by showing that any arrangement can be transformed to one in which this proof holds, without changing the concurrences or collinearity properties.

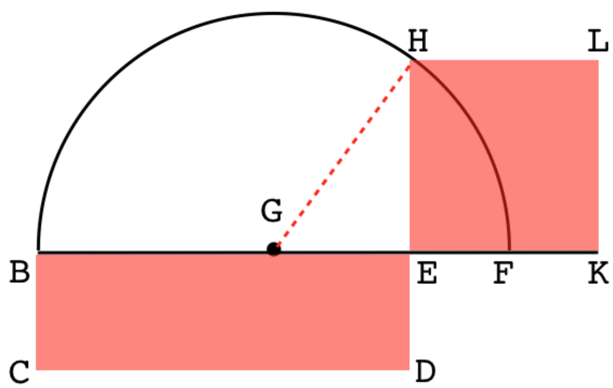
Ultimately, projective geometry is about believing that two parallel lines intersect at a *point at infinity*. One is not supposed to ask whether such an answer makes sense, but only whether it is consistent. Many words have been written about whether this reveals a limitation of Euclidean geometry, or is just some kind of parallel universe.

Projective geometry is well beyond what we can do now, so let's just breathe a sigh of relief that it all came out as we were promised.

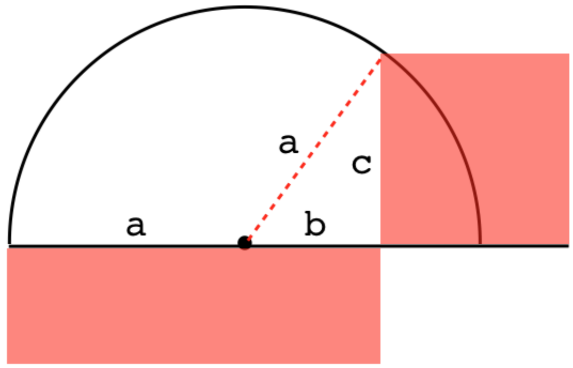
Chapter 24

Hippocrates

Hippocrates of Chios (470-410 BC) was a major figure in Greek geometry. (Not to be confused with the physician of the same name, from Kos). Hippocrates focused on quadrature, the process of constructing (with straight-edge and compass) a square with area equal to a given geometric figure, particularly curved figures, called lunes. Here is one of the first of these—construction of the square equivalent to a given rectangle.



The construction says to: (i) extend BE horizontally, (ii) mark off the same distance as DE to construct EF, (iii) find the midpoint G of BF, (iv) draw the half-circle of radius BG, (v) extend DE up to meet the circle at H, construct the square of side the same length as EH.



As suggested by the dotted line in the figure, the proof invokes the Pythagorean theorem. The long side of the rectangle is $a + b$, while its short side is $a - b$, so the area is

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

but Pythagoras says that is equal to c^2.

$\square$

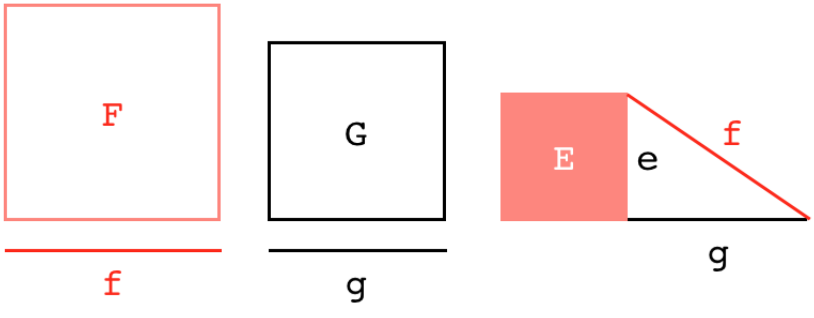
This is a slight restatement of our proof about the geometric mean.

The side of the square, c is the geometric mean of the sides of the rectangle.

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

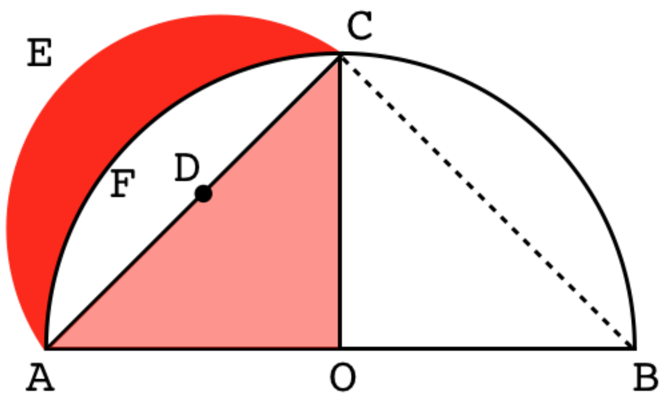
other constructions

Hippocrates "squared" rectangles, triangles and polygons. A lot of his constructions depended on Pythagoras as suggested by this figure:



where two squares resulting from manipulation of part of a polygon need to be subtracted to obtain the final result.

Hippocrates moved on to curves, trying to find squares with area equal to that under or between two curves. That turns out to be a class of problems where few have solutions (in fact, only five, according to Dunham). Famously, it is impossible to square the circle. However, here is one that is possible, it is an example of (the) quadrature of the lune.



We will prove that the two shaded regions are equal in area.

Consider the smaller semicircle with base ADC, which is also the hypotenuse of the right triangle. Let radius AD be equal to r. Let the large semicircle have radius AO equal to R. Pythagoras tells us that

$$R^2 + R^2 = (2r)^2$$

$$R^2 = 2r^2$$

Let the area of the triangle be T. (Its value is $R^2/2$ but that's not needed).

The segment of the larger semicircle (white) is the area of the quadrant minus the area of the triangle

$$\pi \frac{R^2}{4} - T$$

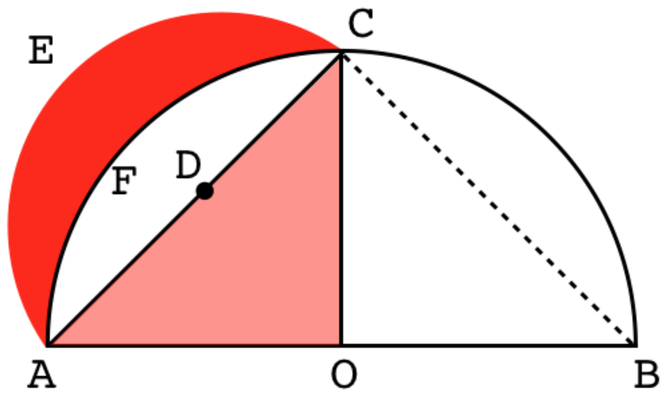
The area of the red lune is the area of the small semicircle minus the white area

$$\pi \frac{r^2}{2} - \ [\ \pi \frac{R^2}{4} - T\ ]$$

$$= \pi \frac{r^2}{2} - \pi \frac{2r^2}{4} + T$$

$$= T$$

Which is just the area of the triangle.

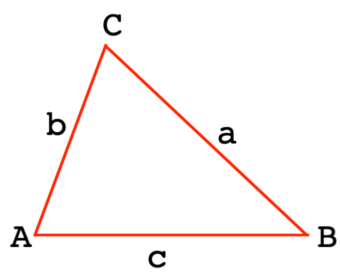


Chapter 25

Heron and Brahmagupta

Heron (or Hero) of Alexandria lived in the first century AD. He was primarily an engineer, but is also remembered for Heron's Formula, which can be used to compute the area of a triangle from the lengths of its sides. It is a simple formula that does not explicitly include the altitude h or the components of side c.

Heron's formula was later found to be a special case of a similar formula for quadrilaterals, discovered by Brahmagupta, which we'll study later. Here's a standard triangle with vertex A opposite side a and so on



Let s be one-half the perimeter, called the semi-perimeter:

$$s = \frac{1}{2}(a + b + c)$$

$$2s = a + b + c$$

then Heron says that the area, call it Δ, is

$$\Delta = \sqrt{s \cdot (s-a) \cdot (s-b) \cdot (s-c)}$$

$$\Delta^2 = s \cdot (s-a) \cdot (s-b) \cdot (s-c)$$

simple derivation

We use a bit of trigonometry. The triangle shown above has area Δ in terms of angle B:

$$\Delta = \frac{1}{2} ac \sin B$$

$$16\Delta^2 = (2ac)^2 \sin^2 B$$

$$= (2ac)^2 \; [\; 1 - \cos^2 B \;]$$

From the law of cosines, substitute for $\cos B$:

$$16\Delta^2 = (2ac)^2 \; [\; 1 - (\frac{a^2+c^2-b^2}{2ac})^2 \;]$$

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

Some algebra with differences of squares follows:

$$16\Delta^2 = (2ac + a^2 + c^2 - b^2)(2ac - a^2 - c^2 + b^2)$$

$$= \; [\; (a+c)^2 - b^2 \;] \; [\; (b^2 - (a-c)^2 \;]$$

$$= (a+c+b)(a+c-b)(b+a-c)(b-a+c)$$

Let $2s = a+b+c$. Then

$$16\Delta^2 = 2s \cdot 2(s-a) \cdot 2(s-b) \cdot 2(s-c)$$

$$\Delta^2 = s(s-a)(s-b)(s-c)$$

□

We now explore a justification for why each term is present in the equation. To begin, note that the equation is symmetrical in a, b and c. This is expected, since there is no reason to distinguish among the sides.

Levi

Mark Levi has a short proof of Heron's formula, linked on this page:

https://www.marklevimath.com/sinews

The url I have for this quote no longer points to the correct document, but I still like it:

> The area-squared is obviously a symmetric and homogeneous polynomial of degree 4 in a, b, c, divisible by (a + b - c)(a + c - b)(b + c - a), since degenerate triangles have zero area.
>
> Hence the area-squared divided by (a + b - c)(a + c - b)(b + c - a) is a symmetric and homogeneous polynomial of degree 1 in a, b, c, and so is (a + b + c) times some constant that must be 1 by considering, say, the 90, 45, 45 triangle.

Let's just play with the formula. Take what is under the square root above:

$$s \cdot (s-a) \cdot (s-b) \cdot (s-c)$$

Multiply each term by 2

$$\begin{gathered} 2s \cdot (2s-2a) \cdot (2s-2b) \cdot (2s-2c) \\ = (a+b+c)(b+c-a)(a+c-b)(a+b-c) \end{gathered}$$

According to the formula above, $16A^2$, and hence the area itself, will be zero when

◦ $a+b+c=0$

that is, when the sum of all three sides is equal to zero. Since lengths are always positive, this means that $a=b=c=0$, or

◦ one of the other terms is zero, e.g. $a+b-c=0$.

that is, when one side length is equal to the sum of the other two.

These are all "degenerate" triangles, where the shape has collapsed either to a point (the first case) or to a line segment.

The factor of 16 may be deduced from an example, e.g., an equilateral triangle with unit sides, altitude equal to $\sqrt{3}/2$ and area of $\sqrt{3}/4$.

Suppose we do not know the factor, so let it be k (rather than 16):

$$k \cdot (\frac{\sqrt{3}}{4})^2 = (a+b+c)(b+c-a)(a+c-b)(a+b-c)$$

$$= 3 \cdot 1 \cdot 1 \cdot 1 = 3$$

Clearly, $k = 4^2 = 16$.

Or, for an isosceles right triangle with sides 1, 1, $\sqrt{2}$:

$$k \cdot (\frac{1}{2})^2 = (2+\sqrt{2})(\sqrt{2})(\sqrt{2})(2-\sqrt{2})$$

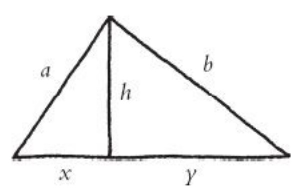
$$= (4-2)(2) = 4$$

The proof starts with the deduction that the area squared is "a polynomial of degree 4 in a, b, c", and he works through why that is so. It makes sense, since area is itself the product of two lengths, each of which must be proportional somehow to the lengths of the sides.

Lockhart

Here is a long-winded algebraic proof, from Lockhart, but there is a point! It does not depend explicitly on trigonometry. Instead, it hides what is in effect, a derivation of the law of cosines.

Proof.



Side c is split into x and y. We can write three equations:

$$x^2 + h^2 = a^2$$

$$y^2 + h^2 = b^2$$

$$x + y = c$$

Our objective is an equation that contains only a, b and c. From the first two:

$$a^2 - b^2 = x^2 - y^2$$

and from the third:

$$y^2 = c^2 - 2xc + x^2$$

so

$$\begin{aligned} a^2 - b^2 &= x^2 - c^2 + 2xc - x^2 \\ &= 2xc - c^2 \end{aligned}$$

then

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

Finally a slight rearrangement:

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

This says that to find the point where c is divided into x and y, we move from the center $c/2$ a distance of $(a^2 - b^2)/2c$.

The corresponding equation for y is

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

which is easily checked by adding together the final two equations, obtaining $x+y = c$.

For the area, we will need h somehow. It is easier to use h^2.

$$\begin{aligned} h^2 &= a^2 - x^2 \\ &= a^2 - \frac{(c^2 + a^2 - b^2)^2}{(2c)^2} \end{aligned}$$

The area squared is

$$\begin{aligned} \Delta^2 &= \frac{1}{4}c^2h^2 \\ &= \frac{1}{4}c^2a^2 - \frac{1}{4}c^2\frac{(c^2 + a^2 - b^2)^2}{(2c)^2} \end{aligned}$$

Lockhart:

the algebraic form of this measurement is aesthetically unacceptable. First of all, it is not symmetrical; second, it's hideous. I simply refuse to believe that something as natural as the area of a triangle should depend on the sides in such an absurd way. It must be possible to rewrite this ridiculous expression...

Here's a start:

$$16\Delta^2 = (2ac)^2 - (c^2 + a^2 - b^2)^2$$

This is much better. It is still problematic, in that a and c do not appear symmetric with b.

However, we immediately notice that it is a difference of squares. First

$$16\Delta^2 = \ [\ 2ac + (c^2 + a^2 - b^2)\ ]\ [\ 2ac - (c^2 + a^2 - b^2)\ ]$$

And that has within it two squares, namely $(a+c)^2$ in the first term on the right-hand side, and $(a-c)^2$ in the second.

$$16\Delta^2 = \ [\ (a+c)^2 - b^2)\ ]\ [\ b^2 - (a-c)^2\ ]$$

A second difference of squares. Thus

$$16\Delta^2 = (a+c+b)(a+c-b)(b+a-c)(b-a+c)$$

At this point, we introduce the semi-perimeter $2s = a + b + c$ and then obtain after several steps

$$\Delta = \sqrt{s \cdot (s-a)(s-b)(s-c)}$$

□

And that is symmetric in each of the three sides, as we hope and expect.

check

As a simple example, if we have a right triangle with sides 3,4,5, then the area is one-half of 3 times 4 = 6. The semi-perimeter is s

$$s = \frac{(3+4+5)}{2} = \frac{12}{2} = 6$$

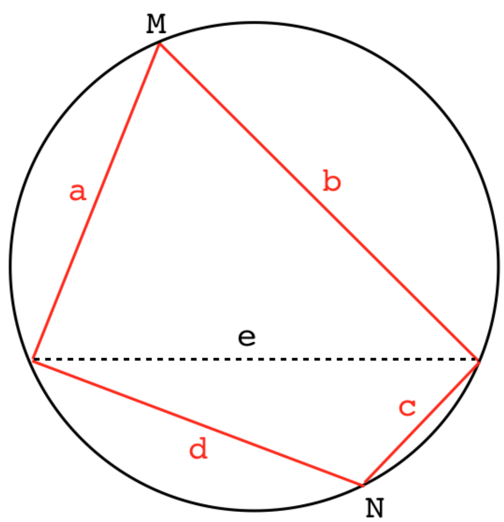
We have

$$\Delta = \sqrt{6(6-5)(6-4)(6-3)} = \sqrt{6(1)(2)(3)} = 6$$

Brahmagupta

Brahmagupta was an Indian mathematician who lived in the 7th century AD in a region of India called Bhinmal, which is in Rajastan. He completed the square to obtain the quadratic equation, and did many other amazing things in trigonometry and arithmetic, as well as this example from geometry.

We consider a quadrilateral inscribed into a circle. This is a special case, where the fourth point fits into the same circle determined by any three of the points.



We will prove that the area of this quadrilateral is given by Brahmagupta's formula.

$$A = \sqrt{(s-a)\cdot(s-b)\cdot(s-c)\cdot(s-d)}$$
$$A^2 = (s-a)\cdot(s-b)\cdot(s-c)\cdot(s-d)$$

Heron's formula is thus a special case where $d = 0$.

$$A = \sqrt{s\cdot(s-a)\cdot(s-b)\cdot(s-c)}$$

preliminary

We need two preliminary results. If M and N are supplementary angles, then

$$\sin M = \sin N, \qquad \cos M = -\cos N$$

Supplementary angles have mirror image symmetry across the y-axis. This becomes obvious if you plot them.

Then, draw the line connecting the two opposing vertices which are not M and N. Using the law of cosines we can write two equal expressions for e^2, namely:

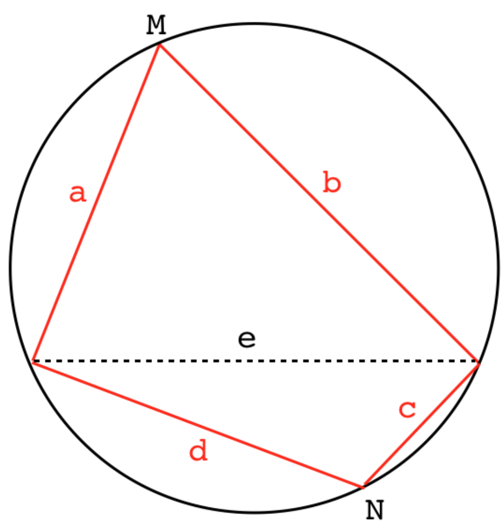
$$e^2 = a^2 + b^2 - 2ab\cos M$$

$$e^2 = c^2 + d^2 - 2cd\cos N = c^2 + d^2 + 2cd\cos M$$

Equating the two and grouping terms:

$$a^2 + b^2 - c^2 - d^2 = 2(ab + cd)\cos M$$

Look at the diagram again.



The triangle above the dotted line has area $(1/2)\ ab\sin M$ and similarly for the one below so the total area is

$$A_1 = \frac{1}{2}ab\sin M$$

$$A_2 = \frac{1}{2} cd \sin N = \frac{1}{2} cd \sin M$$

Adding, the total area is:

$$A = \frac{1}{2}(ab + cd) \sin M$$

$$4A = 2(ab + cd) \sin M$$

algebra

Square the two main equations so far:

$$(a^2 + b^2 - c^2 - d^2)^2 = \; [\; 2(ab + cd) \;]^2 \cos^2 M$$

$$16A^2 = \; [\; 2(ab + cd) \;]^2 \sin^2 M$$

and add

$$16A^2 + (a^2 + b^2 - c^2 - d^2)^2 = \; [\; 2(ab + cd) \;]^2$$

Rearrange

$$16A^2 = \; [\; 2(ab + cd) \;]^2 - (a^2 + b^2 - c^2 - d^2)^2$$

As before, we proceed to factor two differences of squares.

First:

$$16A^2 = \; [\; 2(ab + cd) + (a^2 + b^2 - c^2 - d^2) \;] \; [\; 2(ab + cd) - (a^2 + b^2 - c^2 - d^2) \;]$$

$$= \; [\; (a + b)^2 - (c - d)^2 \;] \;\; [\; (c + d)^2 - (a - b)^2 \;]$$

Second

$$= (a + b + (c - d))(a + b - (c - d)) \;\; (c + d + (a - b))(c + d - (a - b)) \;]$$

$$= (a + b + c - d)(a + b - c + d)(c + d + a - b)(c + d - a + b)$$

If the semi-perimeter is s then

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

So we have

$$16A^2 = (2s - 2d)(2s - 2c)(2s - 2b)(2s - 2a)$$

$$A^2 = (s-d)(s-c)(s-b)(s-a)$$

So lastly

$$A^2 = (s-a)(s-b)(s-c)(s-d)$$

$$A = \sqrt{(s-a)(s-b)(s-c)(s-d)}$$

In comparing the two proofs, it's clear that the latter proof draws on the ideas of (i) using the semi-perimeter and (ii) difference of squares, which are in Heron's proof. The main new ideas are trigonometric: the law of cosines and the cancellation of $\sin^2 x + \cos^2 x$.

We might see this by rewriting the proof of Heron's formula in the style of the Brahmagupta proof. (Note: Brahmagupta does not give proofs or say how he obtained his results).

But we don't have to! That's essentially what the proof from Twitter does, above. First, the law of cosines. Let α be the angle opposite side a:

$$a^2 = b^2 + c^2 - 2bc\cos\alpha$$

$$(b^2 + c^2 - a^2)^2 = (2bc)^2\cos^2\alpha$$

And the area is

$$A = (1/2)bc\sin\alpha$$

$$4A = 2bc\sin\alpha$$

$$16A^2 = (2bc)^2\sin^2\alpha$$

Adding

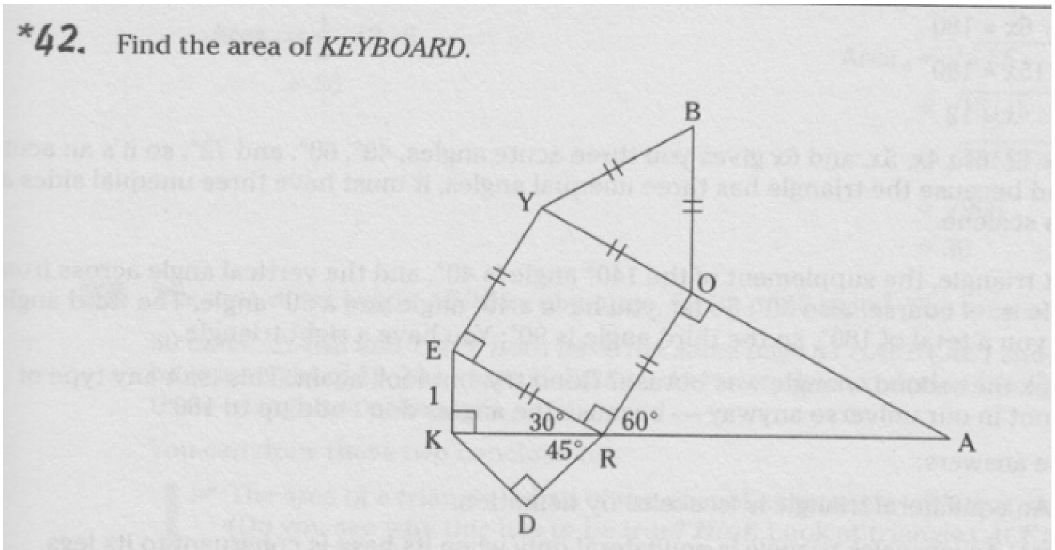
$$16A^2 + (b^2 + c^2 - a^2)^2 = (2bc)^2$$

$$16A^2 = (2bc)^2 - (b^2 + c^2 - a^2)^2$$

The rest is exactly as before, it's just a matter of two differences of squares.

example

Here is a problem where we can use Heron's formula:



The smaller 30-60-90 right triangle has a side labeled 1.

Since $\sin 30° = 1/2$, the side of the square has length 2 so the square has area 4.

Using Heron's formula, the equilateral triangle has area

$$A_{eq} = \sqrt{3\ (1)^3} = \sqrt{3}$$

We get the base of the largest right triangle from the tangent of 60°.

$$\tan \pi/3 = \frac{\sin \pi/3}{\cos \pi/3} = \frac{1/2}{\sqrt{3}/2} = \frac{1}{\sqrt{3}}$$

so

$$\frac{2}{\text{base}} = \frac{1}{\sqrt{3}}$$

The base is $2\sqrt{3}$ and the area is then

$$A_{\ bigT} = \frac{1}{2}\ 2\sqrt{3} \cdot 2 = 2\sqrt{3}$$

Next is the small $\triangle EKR$. Its base is

$$\text{base} = 2\ \cos \pi/6 = 2\ \frac{\sqrt{3}}{2} = \sqrt{3}$$

$$A = \frac{1}{2}\ \sqrt{3} \cdot 1 = \frac{\sqrt{3}}{2}$$

Finally, the last triangle is isosceles. We know its diagonal is $\sqrt{3}$ Let the side be x, then

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

$$x = \frac{\sqrt{3}}{\sqrt{2}}$$

The area is

$$A = \frac{1}{2}x^2 = \frac{1}{2} \cdot \frac{3}{2} = \frac{3}{4}$$

The total is

$$4 + \sqrt{3} + 2\sqrt{3} + \frac{\sqrt{3}}{2} + \frac{3}{4}$$

which equals something.

Part VIII

Celestial applications

Chapter 26

Equal area, equal time

Tycho Brahe was a Danish nobleman who made extraordinarily detailed measurements of planetary motion (by eye).

Johannes Kepler was a German mathematician who went to work for Brahe and later inherited all of Brahe's observations. Kepler is most famous for his three laws of planetary motion:

◦ K1: The planetary orbits are ellipses with the sun at one of the foci.

◦ K2: A line segment joining a planet and the Sun sweeps out equal areas during equal intervals of time.

◦ K3: The square of a planet's orbital period is proportional to the cube of the length of the semi-major axis of its orbit.

Newton explained K1 as a consequence of the $1/r^2$ dependence of the gravitational force. I have written about this in my calculus book.

We talked about a derivation of K3 using the assumption of spherical orbits in a previous chapter. (Some orbits are nearly spherical).

In this chapter we'll discuss how Newton made a geometric derivation of K2: equal areas in equal times. It depends on some simple geometry and the idea that the force of gravity is directed at the sun.

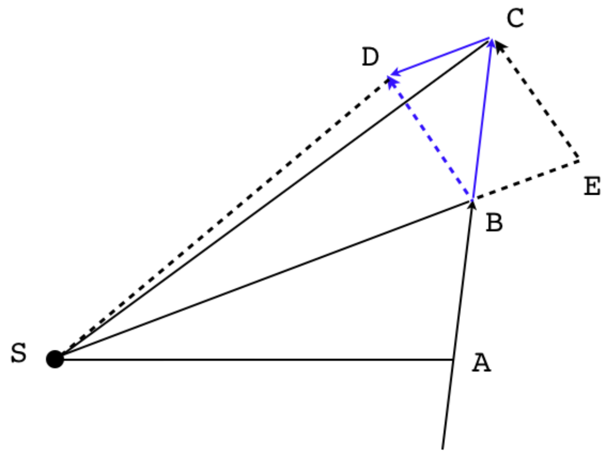


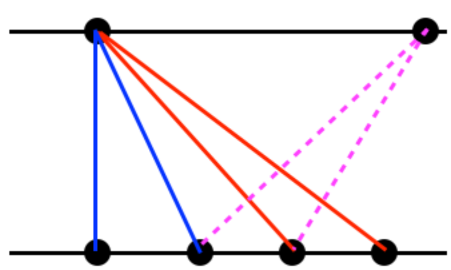
Diagram the sun S and a planet at A.

Imagine that the force toward the sun is applied discretely. That is, during a small interval Δt, the planet travels from A to B at constant velocity and if undisturbed, would travel to C in the next unit of time.

In the absence of a force, the velocity would be constant and the length of AB the same as that of BC, and then since AB is on the same line as BC, the area of $\triangle ABS$ is equal to the area of $\triangle BCS$.

Proof.

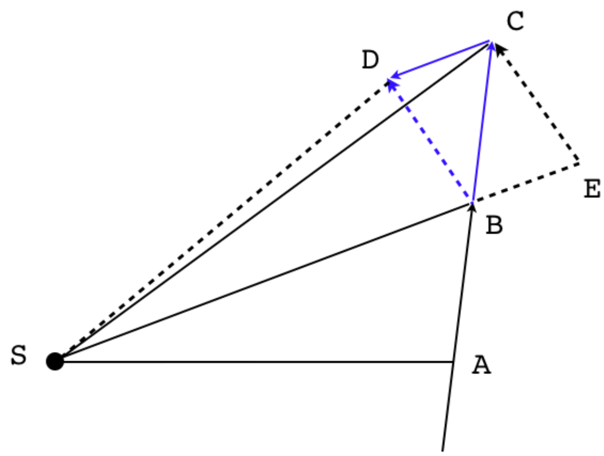
Draw the vertical line from S to the line containing ABC. The area of either triangle is one-half the length of that altitude times the distance, either AB or BC. The principle is illustrated in the next figure.



Given two parallel lines separated by a distance h, pick two points on one line separated by a distance d and *any* point on the other line. The triangles drawn using those points will all have equal area, namely $(1/2)dh$.

$\square$

Now, suppose the force is applied at B *toward the sun* along EBS.



As a result, the trajectory BC is modified by the change in velocity resulting from application of the force toward the sun. The part of the path resulting from the change in velocity is the velocity resulting from the application of the force of gravity, times Δt.

Call that length CD and add it to BC to give the actual trajectory, BD.

CD is parallel to SBE. Therefore, every point on CD has an altitude with respect to SBE of the same length. So any point on CD can be used to draw a triangle with the same base SB and the result will have the equal area no matter which point is chosen.

In particular, the area of $\triangle BDS$ is equal to the area of $\triangle BCS$, which was found earlier to be equal to the area of $\triangle ABS$. Since the two triangles from the actual motion have the same area, the area is constant.

Chapter 27

Aristarchus

Al-Biruni

I found another method for measuring the size of the earth in Acheson's geometry book.

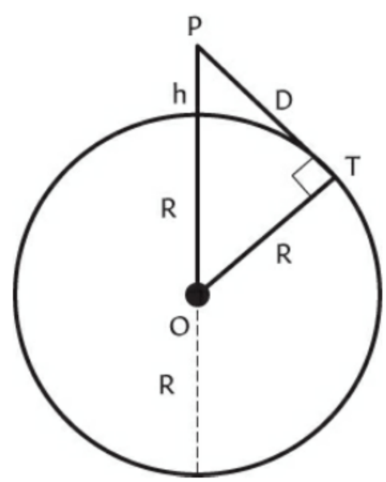


Fig. 68 Measuring the Earth.

In the figure, the circle is the earth, of radius R, h is the height of a convenient mountain, and D is the distance to the horizon, which is tangent to the earth's radius.

Recall from the tangent-secant theorem

$$D^2 = h(2R + h)$$

We neglect h^2 compared to the other term so

$$D^2 \approx 2Rh$$

About 1019 C.E., Al-Biruni obtained a value for R equivalent to 3939 miles.

Note: I have to look into why Acheson says this. The standard treatment of the Al-Biruni's method uses the distance to the horizon plus the angle between a ray to the horizon and the horizontal or vertical axis at the point on the mountain.

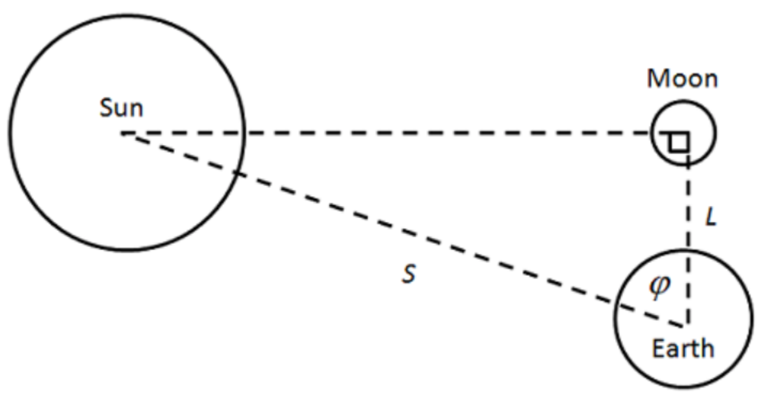
Aristarchus

Aristarchus of Samos (310-230 BCE) wrote a famous book in which he calculated the relative sizes of the sun and the moon and their distances from earth.

One straightforward observation is that the apparent size of the sun and moon in the sky is about the same. This can be seen during a solar eclipse, or observed at any other time by holding a disk up at a fixed distance from the eye, (while taking care to block most of the sun's rays). The value is approximately one-half degree.

Since the distance to the sun is much greater than that to the moon (see below), we can infer that the sun is much larger than the moon.

The central idea of Aristarchus is that, at half moon, the geometry of the three orbs is like this:



In other words, when the phase is half moon and that moon is exactly overhead, the sun has not yet set, but is a bit above the horizon.

If S is the distance to the sun and L is that to the moon, he estimated that

$$18 < \frac{S}{L} < 20$$

with the same ratio for their sizes. Unfortunately, this is not a particularly good estimate. The true value is about 390. Aristarchus obtained a value of 20 for the Earth-Moon distance in Earth radii. The correct value is about 60. Much better estimates were obtained later, by Hipparchus and Ptolemy.

However, Aristarchus made up for this by being the first person to propose a heliocentric theory of the solar system: that the earth and planets rotate around the sun.

https://en.wikipedia.org/wiki/On_the_Sizes_and_Distances_(Aristarchus)

quick estimate

Here is an estimate for the earth-moon distance based on a lunar eclipse.

One measures the time it takes for a complete, total eclipse. From the first shadow of the earth on the moon to the last, that time is about 3 hr. The moon has moved approximately 1 earth diameter in its orbit in that time.

However, we must correct for the fact that the first and last shadows occur on opposite edges of the moon. It was noted that the shape of the eclipse suggests the earth's diameter (at that distance) is about 2.5 moon diameters. So the moon has actually moved $(2.5 + 1.0)/2.5 = 1.4$ earth diameters in the given time. The relevant time becomes 2.14 hr.

Any correction for the true size of the earth's diameter is minimal because the earth-moon system is so far from the source of illumination.

The other piece of information we need is the time for a full revolution, one lunar cycle. This part is tricky. Naively, you'd look for the moon to be in the same place against the fixed stars (27 days, c. 8 hr). This is off because the earth has moved in the meantime — there is a parallax error. As a rough correction, mutliply by 360/330 degrees. The result in hours is 715.

The circumference of the orbit is then

$$715/2.143 = 333$$

earth diameters.

This gives a radius of 53 earth diameters, which is not too far from 60.

Chapter 28

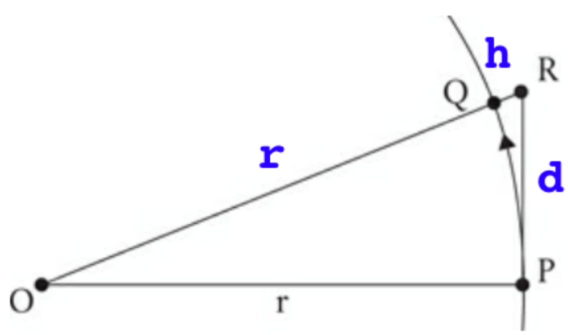
Circular orbits

Pythagoras and Newton

Here, we explore one use of the Pythagorean theorem to provide a taste of orbital mechanics, which is a particular focus of calculus. Newton made early calculations similar to these, which increased confidence about his famous inverse-square law and inspired the mathematics that led to the explanation of elliptical orbits.

Although the orbits of the planets around the sun are ellipses, they are very nearly circular and we will make that approximation for what follows here.

We use the Pythagorean Theorem to make another approximation. Using r for the (fixed) radius of the orbit for the moment, because the construction has capital letters for the points, including the symbol R:



$$r^2 + d^2 = (r+h)^2 = r^2 + 2rh + h^2$$
$$d^2 = 2rh + h^2$$

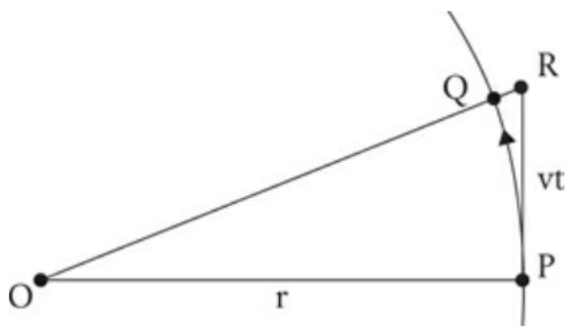
If $h << r$ then we can ignore the very small quantity h^2 and obtain

$$d^2 = 2rh$$
$$r = \frac{d^2}{2h}, \qquad h = \frac{d^2}{2r}$$

If the planet were not accelerated, then it would move from P to R, a distance d, and this is equal to the velocity $\times$ time:

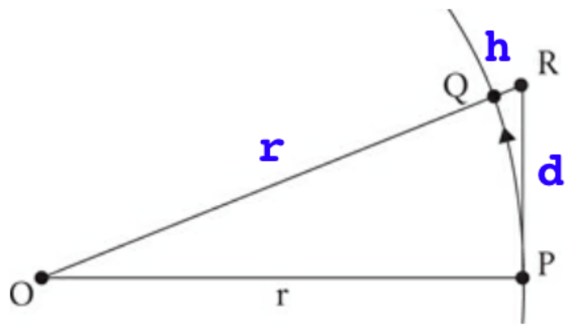
$$d = vt$$

At this point, we use an idea from calculus. *For a small enough segment of the orbit*, this distance PR is the same as the arc length PQ.



So we substitute for $d^2 = (vt)^2$ into the equation from above

$$h = \frac{d^2}{2r} \approx \frac{(vt)^2}{2r}$$



Also, for a small enough part of the orbit (again), h and d are perpendicular to each other as well.

At this point we use the additional assumption that the force is directed toward the sun. We might say that the distance *fallen* by the planet in this short time is h.

By the standard equation of motion, under gravitational acceleration g is related to h and the time t by this equation:

$$h = \frac{1}{2}gt^2$$

We combine the two different expressions for h

$$h = \frac{1}{2}gt^2 \approx \frac{(vt)^2}{2r}$$
$$g \approx \frac{v^2}{r}$$

Note: we have not covered this yet. If this idea (dependence on t^2) is completely new to you, you may want to come back to this part after going through a basic introduction to calculus.

The equation $a = v^2/r$ also comes even more easily with a little bit of calculus and the use of vectors.

Kepler's Third Law

The famous mathematician Johannes Kepler (of whom much more later also), working with observational data from Tycho Brahe, had the following values for the radius

R of the (assumed circular) orbit and the period T (time for completion of one orbit), for five planets.

Orbital data for the six planets known in Kepler's time

	$\overline{r}$ (units of $\overline{r}$ Earth)	T (years)
Mercury	0.387	0.241
Venus	0.723	0.615
Earth	1.000	1.000
Mars	1.524	1.881
Jupiter	5.203	11.862

On the basis of this data, Kepler published his **third law** (in 1619, about 10 years after the first two). K3 states that

$$T^2 = kR^3$$

The square of the period is proportional to the cube of the radius of the orbit. The data in the table has been scaled so that $k = 1$.

For a circular orbit, the orbital speed, the magnitude of the velocity $v = |\mathbf{v}|$, is constant.

The period times the speed is equal to the circumference.

$$vT = C = 2\pi R$$
$$T = \frac{2\pi R}{v}$$

K3 above says that

$$R^3 = T^2$$
$$= \frac{(2\pi)^2 R^2}{v^2}$$

Hence

$$v^2 \approx \frac{1}{R}$$

We showed above that the acceleration for a circular orbit is

$$a = \frac{v^2}{R} = v^2 \cdot \frac{1}{R}$$

so we conclude that that

$$g = a \approx \frac{1}{R} \cdot \frac{1}{R} = \frac{1}{R^2}$$

if the acceleration of gravity g is directed toward the sun, with a magnitude that is inversely proportional to the square of the distance, then we can explain Kepler's third law by running this chain of reasoning in reverse.

comparing the moon to an apple

Earlier we worked out that the acceleration is

$$a = \frac{v^2}{R}$$

Let's figure out the acceleration of the moon. We make a decision to work in English units for this one.

The moon averages about 237 thousand miles from earth (221.5 - 252.7 thousand miles). The earth's circumference is about 24.9 thousand miles so its radius is about 3.96 thousand miles. Thus, the ratio of the moon's distance to the center of the earth, compared to my distance to the center of the earth, is about 60 : 1 (ranging between 56-64).

What is the moon's velocity? The distance it travels in one complete orbit (in feet) is:

$$2\pi \cdot 2.4 \times 10^5 \cdot 5280$$

The time that takes in seconds is about

$$28 \cdot 24 \cdot 3600$$

$$v = \frac{2\pi \cdot 2.4 \times 10^5 \cdot 5280}{28 \cdot 24 \cdot 3600}$$

The acceleration is v^2/R so we square everything except the radius.

$$a = \frac{(2\pi)^2 \cdot 2.4 \times 10^5 \cdot 5280}{(28 \cdot 24 \cdot 3600)^2} = 0.0085$$

That's in feet per second.

We compare this value to the acceleration measured at the surface of the earth, which is 32.2 in the same units. The ratio is 3788, which is just over $(61.5)^2$.

Newton:

> I began to think of gravity extending to the orb of the Moon . . . and computed the force requisite to keep the Moon in her Orb with the force of gravity at the surface of the earth . . . & found them answer pretty nearly. All this was in the two plague years of 1665-1666. For in those days I was in the prime of my age for invention & minded mathematicks and Philosophy more than at any time since.

Part IX

Archimedes

Chapter 29

Pyramid and cone

- Albert Einstein

> Any fool can know. The point is to understand.

The early version of this book (and its parent calculus book) started with the results developed by Archimedes, in this chapter and the next, as well as some ideas about the area of circles and the value of π. I've moved that matrerial here, so it won't be too much of a shock to a beginning student, but it really contains some of the most marvelous ideas in geometry, and was my motivation for writing the book.

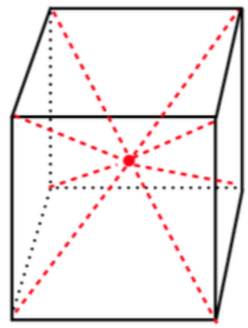
In the next chapter, we develop the most famous of Archimedes geometrical contributions, a theorem on the volume of the sphere.

Before we get there we talked briefly about circles (a topic to which he also contributed) and now look at the volume of cones and pyramids. These are topics in geometry that come before the volume of the sphere.

We need a formula for the volume of a cone in order to find the volume of the sphere. To find the cone's volume, let's start with something simpler, a pyramid with a square base.

Below is a cube with all eight edges having length s. Each of the six faces is a square with sides of length s and area s^2.

Label the central point inside the solid as P. Draw lines connecting each of the 8 external vertices to P, something like this.

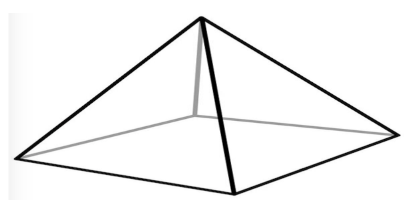


Now we imagine slicing on planes that connect adjacent pairs of lines.

You can't do this in real life by slicing up a single cube or rectangular solid, because the cuts to form one surface would ruin some of the other pieces. The cuts must enter the solid at a corner and then pivot on a line ending at the exact center.

Perhaps you could do it with a *light saber* since the beam comes to a point.

The result is 6 identical pieces (right square pyramids) looking something like this



The procedure described generates pyramids with height $s/2$ on a base of side length s. I admit they are a little squat, but just hang on.

Since we started with a cube, and work from a central point, the six resulting solids are identical.

You can either have six pieces come out exactly the same, as we've done, or start with an elongated base, so then two of the pieces will come out with equal base and height, but you can't do both at the same time by this construction.

Let the six identical pyramid volumes each be V. Their sum is equal to the volume of the cube that we started with.

$$6V = s^3$$
$$V = \frac{1}{6}s^3$$

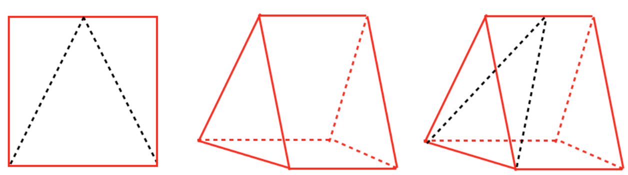
The height $h = s/2$ so

$$V = \frac{1}{3}\ hs^2$$

This is the volume for each pyramid with base area s^2 and height $h = s/2$.

By changing the area of the base or the height, you can alter the shapes obtained. But all other constructions give at least two different types of pyramids, which means we can't simply divide by 6 as we did. This is, unfortunately, a fatal defect.

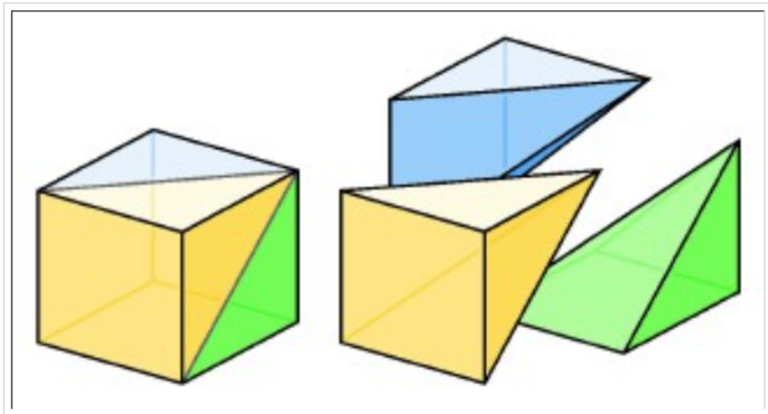
Here is another way to think about the result. Start with a cube and view it from the side. We make two cuts starting at the middle of the top and cutting to the bottom left and right corners. Together, those two cuts remove exactly half of the volume. The result is shown in the middle panel.



Now, turn the remaining solid 90 degrees, and make exactly the same two cuts. One of these is shown in the right panel. Together these two cuts remove not half but one-third of the volume. The reason is that each cut takes more from the top, where the slice is narrower. If you worked it out you would find that this combined cut removes only one-third, leaving 2/3 of 1/2, or 1/3 of the original cube.

better way

Here is a an even better way to slice a cube.

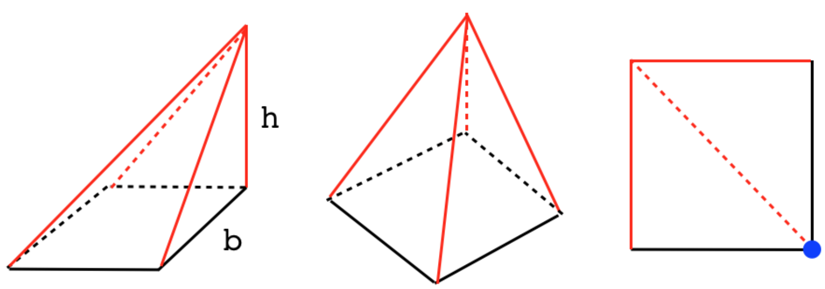


Three congruent pyramids meet along a diagonal of a cube.

When I first saw this, I thought it was a trick. We have produced 3 identical pyramids (they are called oblique because the apex is not in the center).

http://www.math.brown.edu/~banchoff/Beyond3d/chapter2/section02.html

Each pyramid has a square base, two vertical sides that were halves of the original sides, now cut along the diagonal. The fourth and fifth sides slant down from the apex to the opposing sides. Here are three views of one pyramid.



On the right is a simple schematic way to draw pyramids, with perspective from the top. The blue dot indicates the placement of the apex.

a cheese pyramid

I found a fun method to do the demonstration easily and safely. I was going to cut some wood on the table saw, but this is much better.

Get a thick piece of cheese and cut out a cube as large as you can make it and with everything squared as close as you can.

Then cut straight down on a diagonal all the way through the cube, resulting in two identical pieces.

If you now take each of the pieces and orient them with the new angled surface resulting from the cut facing up, you can then make another diagonal cut straight down for each (that cut has just been made, in the figure).

You can see that the second cut results in a large piece that has a square base (currently oriented to the left in the photo above), and a smaller piece with a triangular base, which is facing down and stays that way in the final arrangement.

The two small pieces are mirror images that can be glued together into a single shape identical to each of the large pieces.

By this means we deconstruct the cube into three identical pyramids. Good luck! It is OK to eat the demonstration afterwards.

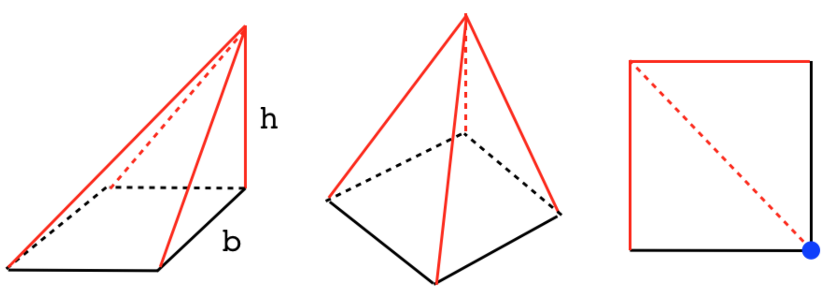
perspectives

Our dissection proves that a pyramid with a square base and the apex placed directly above one of the vertices of the base has

$$V = \frac{1}{3}\, Bh$$

where B is the area of the base and h is the height.

Here is that sketch again. The vertex lies directly above the edge labeled h.

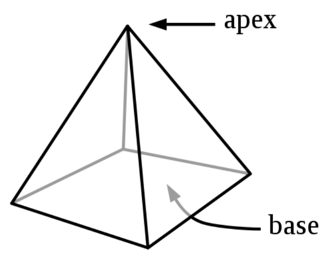


Solid lines outline the base, and the color of the line indicates whether that side meets the plane of the base in a right angle (black) or on a slant (red). Because of the right angle, the solid black line hides the diagonal along the side up to the vertex.

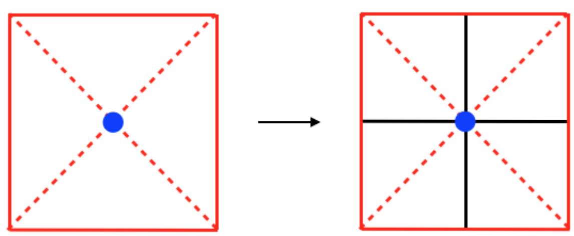
The other kind of diagonal (not overlying a side of the base) is shown as a dotted line.

moving the apex

It is natural to look at other shapes for the base, and other placements for the apex. This is a right square pyramid.



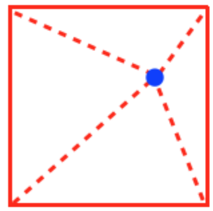
The apex lies above the center of the base. Here is the perspective view (below, left panel).



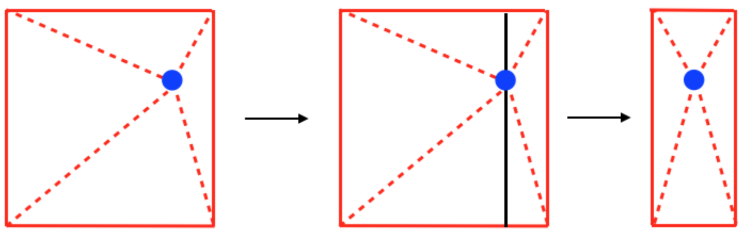
The point is that the right pyramid can be cut into four identical cheese pyramids (right panel). Therefore the volume of each is one-fourth, and so is the base.

We have proved the formula works for a right pyramid.

The next question is how to prove that we can put the apex anywhere and still have the same volume.

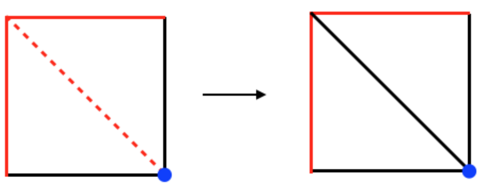


At the moment, I know how to turn that into a regular rectangle, but not into a square:



triangular pyramid

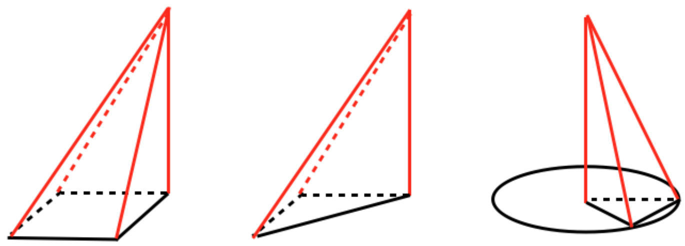
It is also easy to show that a triangular base with the apex directly above a vertex obeys the volume rule.



We have two identical triangular pyramids, and half the base for each one.

If we had rectangles, by using them we could produce a triangular base of any shape. If necessary, we can put two together to form an isosceles base.

In the figure below, on the left, is a view of that same pyramid cut from cheese, drawn with the height extended to make it easier to see.



In the middle is one of two identical halves constructed by bisecting the square pyramid. Its base has become triangular. Clearly it is still the case that $V = Bh/3$, since halving the base gave two identical shapes each with half the volume.

On the right side, we have a triangular pyramid inscribed in a cone. The volume of the cone will be the summed volume of the inscribed pyramids, whose bases, as Archimedes showed, become the area of a circle.

Both the triangular pyramid cut from cheese and the shape we need on the right have two vertical sides, at right angles to the plane of the base. The difference between them is that one vertex on the base is a right angle, whereas for the one we need, the angles of the base of the pyramid on the right depend on the size of the slice.

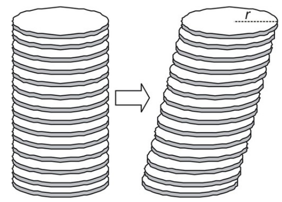
We are close, but not there yet.

Euclid proves the volume formula in book XII of the *Elements*. However, the proofs are much too long for us, and ultimately, this problem has a trivial solution in our first bit of calculus.

cone

A pyramid is not a cone. But we believe, based on these ideas, that the volume is independent of the shape of the base. It just depends on the area.

There is another hand-waving argument from the age before calculus. It is called Cavalieri's principle, also the *method of indivisibles*, or the stack of quarters argument.



https://en.wikipedia.org/wiki/Cavalieri's_principle

Knowing basic calculus allows us to see easily where the factor of one-third comes from in the formula for the volume of a pyramid or a cone. It comes from integrating $x^2\ dx$ and obtaining $x^3/3$.

algebraic derivation of the constant 1/3 for a cone

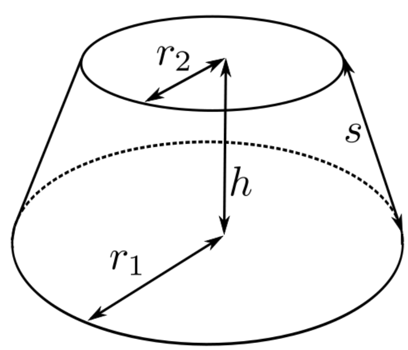
I found an algebraic derivation on the web at

https://web.maths.unsw.edu.au/~mikeh/webpapers/paper47.pdf

Let us assume for this proof that the volume of a cone is proportional to both the area of the base and the height: $V = cAh$; our objective is to find the constant of proportionality.

It takes a bit of algebra to see, but gives the value for the proportionality constant as $1/3$.

Consider a conical frustum, a cone with the top lopped off.



Suppose the area of the base is A and the height of the frustum is h.

Calculate the volume of the frustum as the difference between that of a larger cone with base A and height $h + e$ (e for extra), and that of a small cone (the part cut off to form the frustum) with base area a and height e.

$$V = cA(e+h) - cae$$

Now, the area of the base of a cone is π times the radius squared, and the radius is proportional to the height (depending on how sharply the side slants). Hence

$$a = \pi r_2^2 = ke^2$$

The area of the small base is proportional to the height squared, and so is the large one, with the same proportionality constant:

$$A = k(e+h)^2$$

thus

$$\frac{a}{A} = \frac{ke^2}{k(e+h)^2}$$

$$\frac{\sqrt{a}}{\sqrt{A}} = \frac{e}{e+h}$$

Let us manipulate this expression to find e in terms of h. It just requires a bit of facility with square roots:

$$\frac{\sqrt{A}}{\sqrt{a}} = \frac{e+h}{e} = 1 + \frac{h}{e}$$

$$\frac{h}{e} = \frac{\sqrt{A}}{\sqrt{a}} - 1$$

$$= \frac{\sqrt{A} - \sqrt{a}}{\sqrt{a}}$$

$$e = \frac{\sqrt{a}}{\sqrt{A} - \sqrt{a}} \cdot h$$

And then

$$e + h = \left[\frac{\sqrt{a}}{\sqrt{A} - \sqrt{a}} + 1 \right] h = \frac{\sqrt{A}}{\sqrt{A} - \sqrt{a}} \cdot h$$

Substituting into what we had above for the volume:

$$V = cA(e+h) - cae$$

$$= cA \ [\frac{\sqrt{A}}{\sqrt{A} - \sqrt{a}} \cdot h \ ] - ca \ [\frac{\sqrt{a}}{\sqrt{A} - \sqrt{a}} \cdot h \ ]$$

$$= c \ [\ \frac{A\sqrt{A} - a\sqrt{a}}{\sqrt{A} - \sqrt{a}} \ ]h$$

This really looks like a mess.

But suppose we let $m = \sqrt{A}$ and $n = \sqrt{a}$ so

$$\sqrt{A} - \sqrt{a} = m - n$$

then the numerator above is really just $m^3 - n^3$.

We can factor that, we get

$$m^3 - n^3 = (m - n)(m^2 + mn + n^2)$$

which you can confirm by multiplying back out. So the first term $(m - n)$ cancels the denominator. We now have:

$$V = c(m^2 + mn + n^2)h$$

$$V = c(A + \sqrt{A}\sqrt{a} + a)h$$

Here's the point: consider what happens as a gets larger and closer to A.

We say: let $a \to A$.

The expression in parentheses becomes $3A$. Hence:

$$V = c(3A)h$$

But if $a = A$, the frustum has become a cylinder, whose volume we know. It is equal to Ah.

$$V = c(3A)h = Ah$$

Therefore $c = 1/3$.

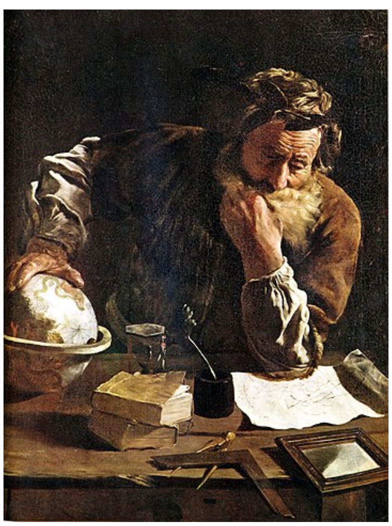
□

Chapter 30

Archimedes and the sphere

biography

Archimedes is often ranked as the greatest of the Greek mathematicians. He stands with Newton, Euler and Gauss, the best of the moderns.



Here are two images, the first is a painting by Fetti from 1620 that is in the wikipedia article on Archimedes:

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

The other is the image on the famous Fields medal, which is sometimes described as the "Nobel prize" for mathematics.

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

Archimedes lived and died (c.287-212 BC) in the beautiful city of Syracuse, found on the southeastern coast of modern Sicily. He is famous for many inventions, derivations and discoveries, but was evidently proud of the formula for the volume of the sphere.

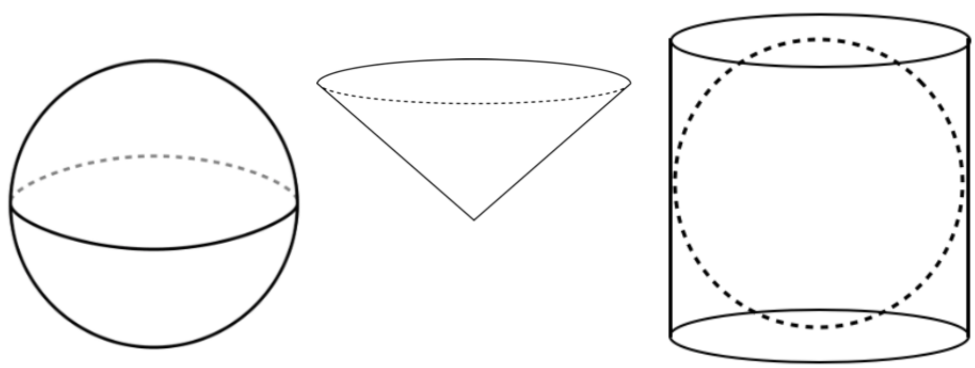
The very simple result is that the volume is two-thirds that of a cylinder that just encloses the sphere.

Because of his discovery, there was a sculpture of the sphere and cylinder carved on his gravestone, located near the Agrigentine gate of Syracuse. The grave was rediscovered by the Roman orator Cicero, covered by brush after 137 years of neglect. It is now lost again. (It is so lost that I didn't bother to try to find it when I was there. That may have been a mistake.)

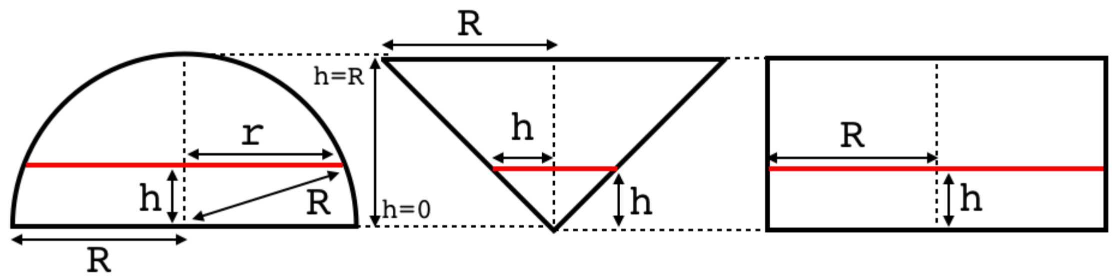
slices of solids

The following is Archimedes simple but subtle argument.

We compare a half-sphere and an inverted cone to a cylinder.



Below is a diagram showing a **vertical** cross-section through the center of each solid so we can visualize the geometry. The radius R is the same for all three. In addition, the cone and cylinder have overall height equal to R.



Now, imagine making a **horizontal** slice through each solid at an arbitrary but constant height h, shown by the red lines. I hope you can visualize each of these red slices, which are perpendicular to the page.

Each slice is a circle. Any cross-section of a sphere is a circle.

For the cylinder and cone, cross-sections perpendicular to the central axis are circles as well.

The question we ask is: what is the area for each slice?

To answer that, we need to determine the radius for each red circle.

Moving right-to-left, the radius of the cylinder is just R. For the cone, the radius at each height h is equal to h, since $R = H$. And for the sphere, we use the Pythagorean theorem to find that

$$r^2 + h^2 = R^2$$

$$r^2 = R^2 - h^2$$

For more on this theorem see **here**.

The first insight of the proof is to recognize that the radius squared for the sphere's slice (r^2), plus the radius squared for the cone (h^2) is equal to R^2, the radius squared for the cylinder.

Since the area of each circle is proportional to to the radius squared (namely $A = \pi r^2$ and so on) and

$$\pi r^2 + \pi h^2 = \pi R^2$$

so the areas add too. Our just famously and remarkable simple result: *sphere plus cone equals cylinder.*

invariance

The second crucial insight of the proof is to recognize that this property is invariant, it does not depend on which height we choose to make the slice. The three slices obtained at any height h add up like this. So if we imagine making a bunch of slices for each solid and adding them all up to find the volume, the volumes will add too.

This idea is now called Cavalieri's principle, though it was called the "method of indivisibles" before that.

The volume of the cylinder is simply πR^3. The volume of the cone is known to be one-third the area of the base times the height, or $1/3\ \pi R^3$.

Subtract to find that the area of the half-sphere is $2/3\ \pi R^3$, and therefore the volume of the whole sphere is

$$V_{\text{sphere}} = \frac{4}{3}\pi R^3$$

There is a bit of a trick here to hide the idea introduced in calculus, which makes this thinking rigorous. The sphere and cone have variable widths, which means that the radius will be different on the top of a slice compared to the bottom. Therefore, the slices have to be made very thin. In calculus they become infinitely thin, but we add up infinitely many of them.

Moreover, Archimedes was subject to certain limitations (discussed by Bressoud), which lead him to formulate the argument in terms of moments (masses of the solids and their centers). I have left that complication out of this discussion.

knowledge before proof

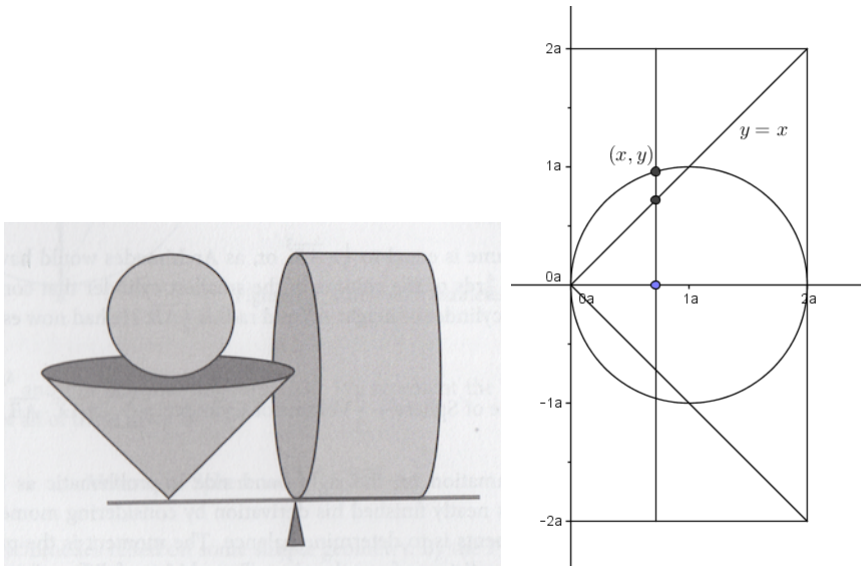
Archimedes said that he discovered the correct result by balancing the three objects on a fulcrum.

According to Archimedes (in the Method, translation by Heath)

> For certain things which first became clear to me by a mechanical method had afterward to be demonstrated by geometry...it is of course easier, when we have previously acquired by the method some knowledge of questions, to supply the proof than it is to find the proof without any previous knowledge. This is a reason why, in the case of the theorems the proof of which Eudoxus was the first to discover, namely, that the cone is a third part of the cylinder, and the pyramid a third part of the prism,

having the same base and equal height, we should give no small share of the credit to Democritus, who was the first to assert this truth...though he did not prove it.

From his description, what Archimedes actually balanced is a set-up like that shown below:



https://proofwiki.org/wiki/File:SphereVolume.png

We have a:

- sphere with radius R
- cone with radius $2R$ and height $2R$

Their combined volumes:

$$V = \frac{4}{3}\,\pi R^3 + \frac{1}{3}\,\pi(2R)^2 \cdot 2R$$
$$= \frac{12}{3}\,\pi R^3 = 4\pi R^3$$

balanced against

- a cylinder with radius $2R$ and height $2R$ with

$$V = \pi(2R)^2 \cdot 2R = 8\pi R^3$$

However, the moment of the cylinder is at one-half the distance from the fulcrum as that of the sphere-cone combination, so it balances.

Chapter 31

Platonic solids

This last part doesn't exactly fit in the book, because it's solid geometry and we don't have much of that, but it's one of the most beautiful proofs in Greek geometry and I would like to include it. So here goes.

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

> In three-dimensional space, a Platonic solid is a regular, convex polyhedron. It is constructed by congruent (identical in shape and size) regular (all angles equal and all sides equal) polygonal faces with the same number of faces meeting at each vertex. Five solids meet these criteria.

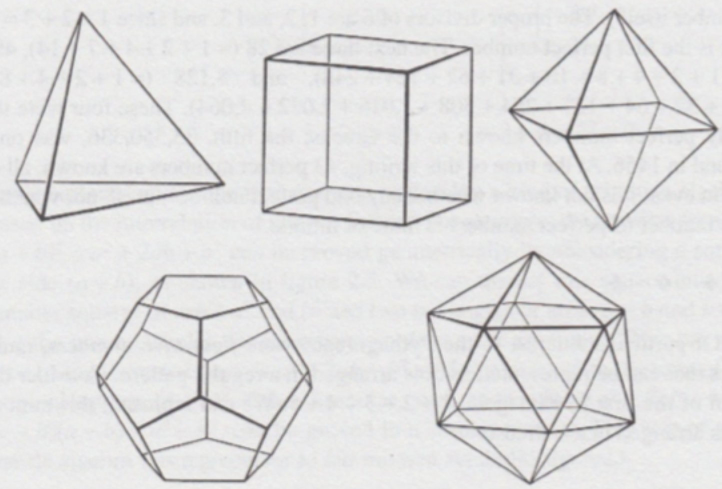


Figure 2.3. The five regular or Platonic solids

These are: (i) tetrahedron, (ii) cube, (iii) octagon, (iv) dodecagon, and (v) icosahedron.

There is a wonderful, simple proof that there are only five of them. Any solid requires at least three sides meeting at each vertex, otherwise the joint between two sides can just flap, like a hinge. Furthermore, the total of all the vertex angles added up must be less than 360 degrees, since otherwise the figure would be planar, not 3-dimensional.

∘ Three equilateral triangles total $60 \times 3 = 180$, four total $60 \times 4 = 240$ and five total $60 \times 5 = 300$. Six would be a hexagon lying in the plane.

∘ Three squares total $90 \times 3 = 270$, while four give a square array in the plane.

∘ Finally, three pentagons give $108 \times 3 = 324$. And that's it. Three hexagons would give $120 \times 3 = 360$, which gives an array in the plane.

Proving that all the angles and side lengths come out correctly, so that the possible solids actually can be constructed is another matter, however. Euclid devotes book XIII of *The Elements* to this:

https://mathcs.clarku.edu/~djoyce/elements/bookXIII/bookXIII.html#props

Part X

Addendum

Chapter 32

Pythagorean triples

The simplest right triangle with integer sides is a $3, 4, 5$ right triangle:

$$3^2 + 4^2 = 5^2$$

but of course any multiple k will work

$$(3k)^2 + (4k)^2 = (5k)^2$$

However, that's not so interesting.

The triples which are not multiples of another triple are called *primitive*. There is a small table of triples in this discussion of Euclid X:29 by Joyce:

https://mathcs.clarku.edu/~djoyce/elements/bookX/propX29.html

	1	3	5	7	9	11	13
3	3 : 4 : 5						
5	5 : 12 : 13	15 : 8 : 17					
7	7 : 24 : 25	21 : 20 : 29	35 : 12 : 37				
9	9 : 40 : 41	27 : 36 : 45	45 : 28 : 53	63 : 16 : 65			
11	11 : 60 : 61	33 : 56 : 65	55 : 48 : 73	77 : 36 : 85	99 : 20 : 101		
13	13 : 84 : 85	39 : 80 : 89	65 : 72 : 97	91 : 60 : 109	117 : 44 : 125	143 : 24 : 145	
15	15 : 112 : 113	45 : 108 : 117	75 : 100 : 125	105 : 88 : 137	135 : 72 : 153	165 : 52 : 173	195 : 28 : 197

We can see that the entries in each column are similar. For example, in the first column

$$(3, 4, 5) \quad (5, 12, 13) \quad (7, 24, 25) \quad (9, 40, 41)$$

The values a_n differ by a constant: In column 1, Δ is 2:

$$a_{n+1} = a_n + 2$$

In the second and third columns, the Δ for a_n is 6 and 10, respectively.

For b and c the Δ at each step is the same for b and c, but the scale ratchets upward. For the first column, $c = b + 1$ and the step rule is

$$\Delta = 4 + 4k$$

You can think of the first entry for b as being generated by $k = 0$.

The first difference is 8, then 12, then 16 and so on.

It is conventional to write a as odd. We observe that, in the table, b is even and c is odd.

the basic rule

Euclid had a formula for triples:

$$b = 2mn, \qquad a = n^2 - m^2, \qquad c = n^2 + m^2$$

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

Using the formula, we look at some examples:

```
m  n   b   a   c
1  2   4   3   5
2  3  12   5  13
3  4  24   7  25
4  5  40   9  41
5  6  60  11  61

1  3   6   8  10
1  4   8  15  17
1  5  10  24  26
1  6  12  35  37

2  4  16  12  20
```

2	5	20	21	29
2	6	24	32	40
3	5	30	16	34
3	6	36	27	45
4	6	48	20	52

Notice that when the sum of m and n is even, the triple is not a primitive one. Another one that we don't want is constructed from $m = 3, n = 6$, where the factor of 3 is obvious. It seems that m and n should be co-prime, they should have no common factors (other than 1).

factors

The elementary rule about squares is this: if n is even then so is n^2, while if n is odd then so is n^2.

To see this, write $n = 2k$ for $k \in 1, 2, 3 \ldots$ as the definition of an even number. Then $n^2 = 4k^2$, which is even.

On the other hand, if n is odd, write $n = 2k + 1$ with $k \in 0, 1, 2 \ldots$, so $n^2 = 4k^2 + 4k + 1$, which is odd. Since there are only these two cases, we can conclude that the converse is also true: an even square comes from an even number and an odd square from an odd number.

As a result, we find that for the triples we care about, a and b are not both even, because then a^2 and b^2 would be even, and so would c^2. So then c would be even, and the triple would not be primitive.

even and odd

Let us go back to

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

We said that a and b cannot both be even, because then c would be even. Or rather, they can, but in that case we are not interested.

The other possible cases are, either a and b both odd, or one is even and one odd. In the first case we have that c is even because odd plus odd is even. So for that case, a and b both odd and c even:

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

The left-hand side is not evenly divisible by 4, but the right-hand side is. This is impossible. Hence one of a and b is even and one odd. As they say, *without loss of generality*, let a be odd, as we saw in the table above.

more about factoring

Rearrange the equation:

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

Since b is even, we can write $b = 2t$

$$4t^2 = (c+a)(c-a)$$

Now we come to an argument about common factors. There are some basic facts we should obtain first. Let

$$p + q = r$$

Now, suppose that p and q share a common factor, f. So then

$$f \cdot i + f \cdot j = f \cdot (i+j) = r$$

By the fundamental theorem of arithmetic, if f is a factor of the left-hand side, it is also a factor of r.

In a similar way, suppose that p and r share a common factor, f. Then

$$r - p = f \cdot k - f \cdot i = f \cdot (k-i) = q$$

and again, all three must have the common factor. But we have agreed that these cases do not interest us.

The same argument applies to squares, since if there is a common factor, it will be there as f^2 when it is present.

We conclude that a, b and c must be all relatively prime. No two of them can share a common factor.

Let us now go back to

$$4t^2 = (c+a)(c-a)$$
$$t^2 = \frac{(c+a)}{2} \cdot \frac{(c-a)}{2}$$

Recall that a and c are both odd, so their sum and difference are both even. Therefore the two factors on the right-hand side are integers, while t^2 is a perfect square, namely, that of t.

The sum and difference for these terms are:

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

On the supposition that $(c+a)/2$ and $(c-a)/2$ did have a common factor, then they would share that common factor with both c and a. Since we know that a and c (at least the particular ones we're interested in) do not have a common factor, neither do these two terms.

So the two terms have no common factor and yet multiply together to give a perfect square.

So then

$$t^2 = xy$$

where x and y are co-prime.

Suppose that x and y are not perfect squares. Then

$$t^2 = (x_1 \cdot x_2 \dots)(y_1 \cdot y_2 \cdots)$$

But this is impossible. Each factor of t^2 must be present an even number of times on the right-hand side, since t^2 is a perfect square. Since no factors are shared between x and y, it must at least be that

$$t^2 = x_1^2 \cdot y_1^2$$

but possibly

$$t^2 = x_1^2 \cdot x_2^2 \dots y_1^2 \cdot y_2^2$$

Therefore both factors are themselves perfect squares.

That is, there exist integers m and n such that

$$m^2 = \frac{(c-a)}{2}$$

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

with $n > m$.

Adding

$$m^2 + n^2 = c$$

Subtracting

$$n^2 - m^2 = a$$

Go back again to

$$4t^2 = (c+a)(c-a)$$
$$= m^2n^2$$
$$2t = mn = b$$

We have not limited m and n in any way except to say that they are not equal so one is larger than the other and arbitrarily suppose $n > m$. Every primitive triple must have an integer m and n with these properties:

$$c = m^2 + n^2, \qquad a = n^2 - m^2, \qquad b^2 = 2mn$$

So finally not only do m and n exist with these properties for any triple, but any integer m and n will satisfy the Pythagorean condition, since:

$$a^2 + b^2 = (n^2 - m^2)^2 + (2mn)^2$$
$$= n^4 - 2n^2m^2 + m^4 + 4n^2m^2$$
$$= n^4 + 2n^2m^2 + m^4$$
$$= (n^2 + m^2)^2 = c^2$$

Any integer m, n, with $n > m$ will work.

For 3-4-5, $n = 2$, $m = 1$.

This is a proof that this formula gives all Pythagorean triples.

$\square$

another derivation

Start with our favorite:

$$\sin^2 x + \cos^2 x = 1$$

$$\tan^2 x + 1 = \frac{1}{\cos^2 x}$$

$$\cos^2 x = \frac{1}{1 + \tan^2 x}$$

The double-angle formula for sine:

$$\sin 2s = 2 \sin s \cos s$$

$$= 2 \frac{\sin s}{\cos s} \cos^2 s$$

$$= 2 \tan s \ \frac{1}{1 + \tan^2 s}$$

Let $a = \tan s$, then

$$\sin 2s = \frac{2a}{1 + a^2}$$

cosine

$$\cos 2s = \cos^2 s - \sin^2 s$$

$$= \ [\ \frac{\cos^2 s}{\cos^2 s} - \frac{\sin^2 s}{\cos^2 s} \ ] \ \cos^2 s$$

$$= \ [\ \frac{1 - \tan^2 s}{1 + \tan^2 s} \ ]$$

so

$$\cos 2s = \frac{1 - a^2}{1 + a^2}$$

In general, a can be anything.

But if a is a rational number, then we can obtain the corresponding sides of a right triangle with rational lengths as well.

The sides are: $2a, 1-a^2$ with the hypotenuse:

$$\sqrt{4a^2+(1-2a^2+a^4)}$$

$$\sqrt{1+2a^2+a^4)}$$

$$=1+a^2$$

Suppose $a=\frac{2}{3}$. Then, we have side lengths: $\frac{4}{3}=\frac{12}{9}, \frac{5}{9}$, and $\frac{13}{9}$, which can be converted to integers: $12, 5, 13$.

In general, if $a=\tan s=p/q$ then the sides are

$$\frac{2p}{q}, \qquad 1-\frac{p^2}{q^2}, \qquad 1+\frac{p^2}{q^2}$$

which as integers will be

$$2pq, \qquad q^2-p^2, \qquad q^2+p^2$$

examples

Make a series with $m+1=n$:

$$b=2mn$$

$$a=n^2-m^2=2m+1$$

$$c=n^2+m^2$$

So then $m=1,2,3\ldots$:

$$(3,4,5) \quad (5,12,13) \quad (7,24,25)$$

Chapter 33

References

- Acheson. *The Wonder Book of Geometry.*
- Byer, Lazebnik and Smeltzer. *Methods for Euclidean Geometry.*
- Coxeter and Greitzer. *Geometry Revisited.*
- Densmore. *Euclid's Elements* (annotated Heath translation).
- Dunham. *Journey Through Genius.*
- Dunham. *The Mathematical Universe.*
- Gelfand and Saul. *Trigonometry.*
- Hopkins. *Inductive Plane Geometry.*
- Jacobs. *Geometry: Seeing, Doing, Understanding.*
- Kiselev. *Geometry, Book I. Planimetry.*
- Lehocsky and Rusczyk. *the Art of Problem Solving, Volume 1: the Basics.*
- Lockhart. *Measurement.*
- Nelsen. *Proofs Without Words.*
- Posamentier. *The Joy of Geometry.*
- Simmons. *Precalculus Mathematics in a Nutshell.*
- Surowski. *Advanced High School Math*

https://www.math.ksu.edu/~dbski/index.htm