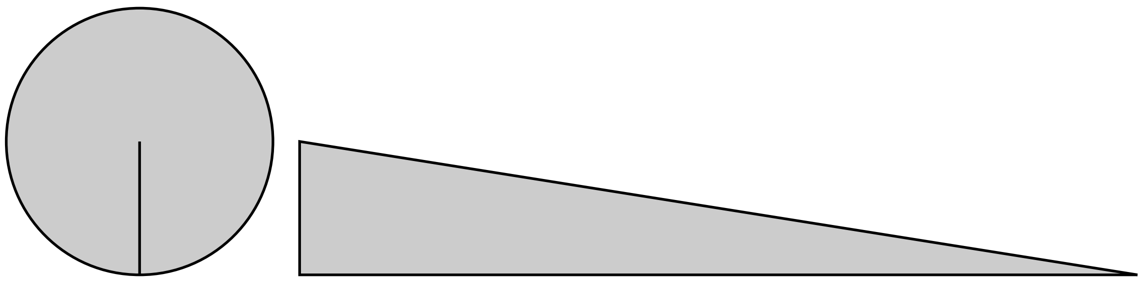
The method of exhaustion

The **method of exhaustion** is a technique that the classical Greek mathematicians used to prove results that would now be dealt with by means of limits. It amounts to an early form of integral calculus. Almost all of Book XII of Euclid's *Elements* is concerned with this technique, among other things to the area of circles, the volumes of tetrahedra, and the areas of spheres. I will look at the areas of circles, but start with Archimedes instead of Euclid.

1. Archimedes' formula for the area of a circle

We say that the area of a circle of radius r is πr^2, but as I have said the Greeks didn't have available to them the concept of a real number other than fractions, so this is not the way they would say it. Instead, almost all statements about area in Euclid, for example, is to say that one area is equal to another. For example, Euclid says that the area of two parallelograms of equal height and base is the same, rather than say that area is equal to the product of base and height. The way Archimedes formulated his Proposition about the area of a circle is that it is equal to the area of a triangle whose height is equal to it radius and whose base is equal to its circumference: $(1/2)(r \cdot 2\pi r) = \pi r^2$.



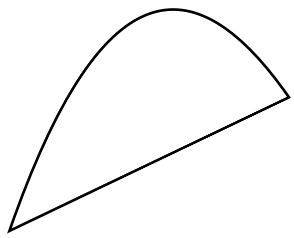
There is something subtle here—this is essentially the first reference in Greek mathematics to the length of a curve, as opposed to the length of a polygon. Now for us the length of a curve is *defined* to be a limit. In fact, it is a tricky concept—there are curves that are so wild that they have infinite length. Archimedes deals with these matters by restricting himself to a relatively simple class of curves, and by *axiomatizing* the properties of length. This is done typically by the Greeks—to formualte simple properties of soemthing that seem just about obvious, rather than attempt a direct definition. It was a wise decision, since many subtle things are involved, and a correct and rigourous theory of limits was not found until about 1820. Until then, Archimedes' techniques were the best available.

2. Convexity

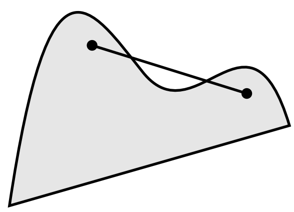
Archimedes doesn't need to know much about the length of curves, since after all a circle is a relatively simple one. One of his axioms is quite general, but the other is concerned only with a restricted class of curved paths, the **convex** ones. The best way to describe a convex path is by showing one that is convex and another that isn't. On the left a convex path, on the right one that isn't.

Convex paths bulge out, the others have dimples. Let's make this more precise.

Suppose given a curve C between two points. Make a closed path by adding in the line segment back from the end to the beginning.



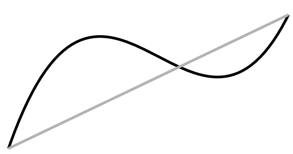
The region inside the path is called convex if it has the property that *along with any two points in it is also contained the whole line segment between them.* The first curve satisfies this condition, but the second does not.



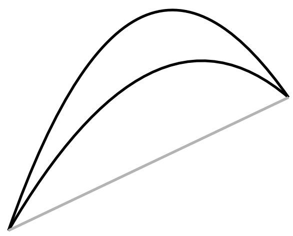
3. Archimedes' axioms

Here are Archimedes' two axioms:

- *If P and Q are two distinct points, the line segment from P to Q is shorter than any other path from P to Q.*



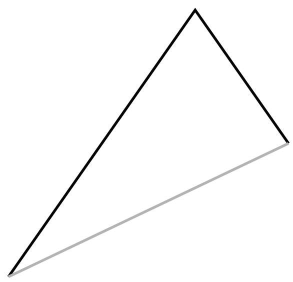
- *If we have two convex paths from P to Q, one inside the other, then the inside one is the shorter.*



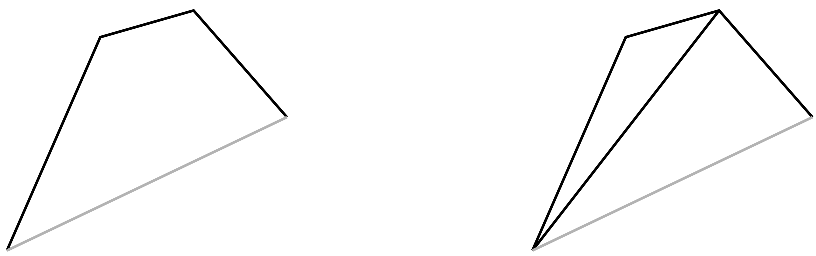
These are **axioms** for Archimedes, which means that they are assumptions, not to be proven. This is typical for the Greeks, to make an axiom something which seems to difficult to prove, but is nonetheless plausible.

But they might not be entirely plausible, and I want to motivate them to some extent by proving them for a large class of simple cases—that where the convex paths are polygons. The starting point will be a result that is actually proven in Book I of Euclid, a special case of Archimedes' axiom I.

- (Euclid I.21) *Axiom I is true if the path is made up of two sides of a triangle and the line segment is the third side.*

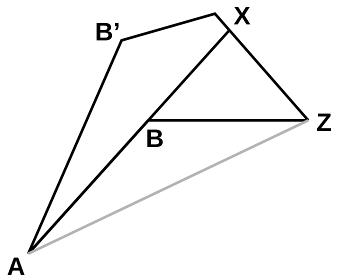


I'm not going to prove this, but assume from now on that it is true, and deduce from it the axioms of Archimedes for polygon paths. We look first at Axiom 1. Let's look at a special case, the first case more complicated case beyond Euclid's, where the convex path has three edges. On the left below is the basic diagram, and on the right is suggested how to prove Axiom 1 by applying I.21.



In the general case, the proof of Axiom 1 for a convex polygon proceeds by induction on the number of edges in the polygon, with Euclid's I.21 the first case to start with.

As for Axiom 2, the proof proceeds by induction on the number of edges in the inner polygon, and uses Axiom 1 as well. Here is the picture that illustrates the first new case:



We want to show that

$$AB + BZ < AB' \dots X + X \dots Z$$

But by Axiom 1 we have

$$AB + BZ < AB + (BX + X \dots Z) = (AB + BX) + X \dots Z = AX + X \dots Z < AB' \dots X + X \dots Z\ .$$

In the general case, you'll need to use an induction assumption as well as Axiom 1.

Exercise 1. *Finish the proofs of Axioms 1 and 2 for polygons. You'll need to label more vertices carefully.*

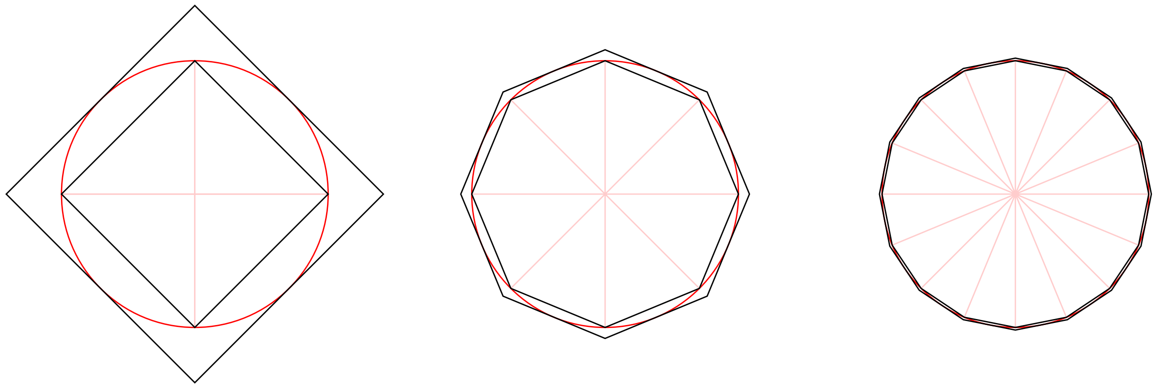
The consequences of these axioms that Archimedes uses are these:

- *If a polygon is inscribed in a circle, its length is less than the circumference of the circle.*
- *If a polygon is circumscribed outside a circle, its length is greater than the circumference of the circle.*

These follow immediately from Axioms 1 and 2, respectively, by summing over the arcs of the circle whose endpoints are where the polygon meets the circle.

4. The proof of Archimedes' formula

The basic idea is to approximate the area of a circle from above and below by circumscribing and inscribing regular polygons of a larger and number of sides.



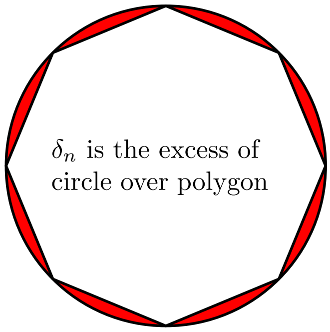
Each of the polygons is a union of triangles, and hence for them it is easy to see that the area is equal to one half the product of the radius and the circumference, or equal to the area oif a triangle whose height is equal to the radius and whose base is the circumference.

The rest of the proof amounts in our terms to showing that the limit of the areas of the polygons is equal to that of the circle. How does Archimedes phrase this? In a phrase, by the **method of exhaustion**.

Let C be the area of the circle, A that of the triangle. for each n, let I_n be the area of the regular polygon of n sides inscribed in the circle, O_n that of the polygon circumscribing the circle. The method of exhaustion is a proof by contradiction. It proves first that we cannot have $C > A$, and then that we cannot have $C < A$. The only other possibility is that $C = A$.

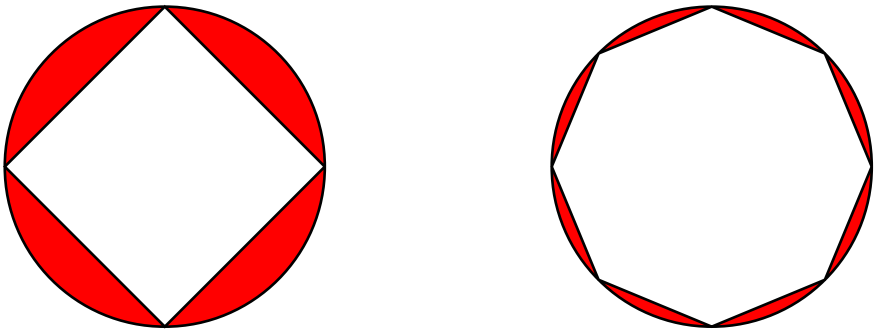
(1) *We cannot have* $C > A$. If $C > A$, let $d = C - A$, which is a positive magnitude.

We know that $I_n < A$ for all n. Let $\delta_n = C - I_n > 0$, the amount by which the area of the circle exceeds that of the triangle.

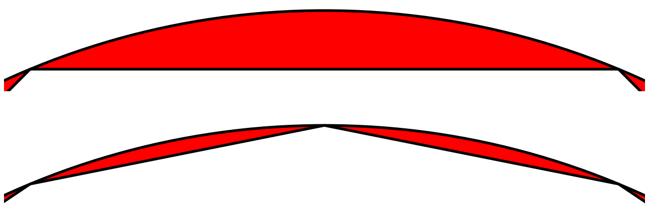


The crucial part of the argument is now this: we have $\delta_{2n} < \delta_n$. I'll prove this in a moment. Let's suppose it to be true for the moment, and see where that gets us. First of all, we know by 'Archimedes' axiom' that if we keep halving any given magnitude, we get arbitrarily small magnitudes eventually. In particular, for large enough n we'll have $\delta_n < d$. But now $I_n < A < C$, which means that $d = C - A < C - I_n = \delta_n$, a contradiction.

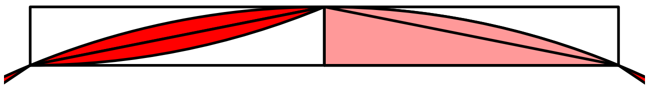
It remains to verify that $\delta_{2n} < \delta_n/2$, or in other words why the excess of the circle over a regular polygon s more than halved in doubling the number of sides.



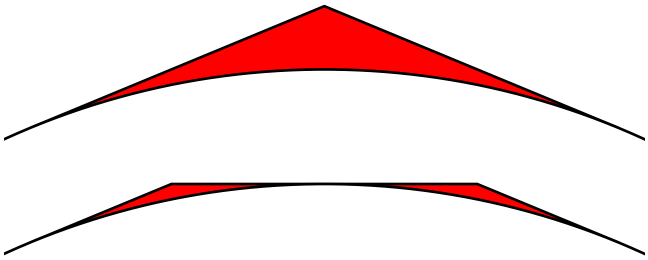
This amounts to looking at what happens when just one arc of the circle is subdivided. The excess is visibly cut by a factor of more than two, but how can this be proved?



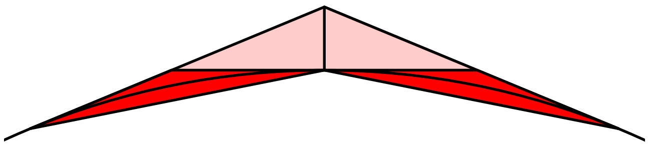
The trick appears already in Euclid's proof of XII.2. We add a rectangle to the figure, bisect it, and then show the excesses like this:



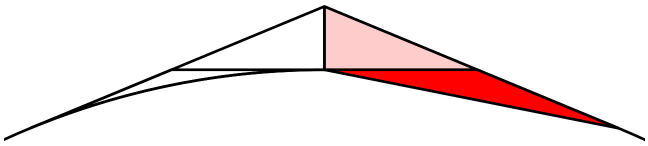
(2) *We cannot have $C < A$.* If $C < A$, let $d = A - C$, which is a positive magnitude. From here on the argument is almost the same, except that it works with circumscribed polygons. The only difference is the final argument, which compares the excesses in these figures:



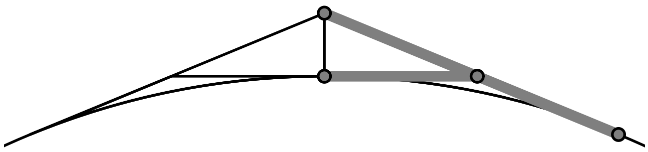
Again we want to show the one is less than half the other. It suffices to show that the area of the upper triangle is greater than the sum of the areas of the two lower triangles in this figure:



Or, in other words, that the area of the upper shaded triangle below is greater than that of the lower one:



But the bases of these two triangles may be taken to be on the same oblique line. The follwoing picture suggests why the base of the lower triangle is less than that of the upper one, since tangents to a circle from any given point have equal length.



5. References

Archimedes, *Measurement of the circle.*

Archimedes, *Sphere and cylinder I.*

Euclid, *Book I, Proposition 21.*