WHAT IS THE SURFACE AREA OF A HEDGEHOG?

BARRY MAZUR

Notes for The Steiner Lecture— St. Johns College, September 30, 2011

Well, I don't know the answer to the question of the title and no hedgehog will be harmed, or even mentioned again, until the very end of my lecture.

It was suggested by Miss Brann that in my Steiner Lecture I address the question "What is area?" I'm delighted to do this; I'm delighted to be here, and to be among people—you, the St. Johns community—with whom it will be such a pleasure to contemplate this question.

I'll discuss the concepts of—

- *area*—how it is familiar to us, and how when we push it to the limit we get some surprises;
- *length*—since, at least at first impression it is a more primitive "prior" concept—seemingly simpler than area;
- *proportion*—crucial to the understanding of both length and area;
- *invariance*—as a way of characterizing length and area;
- *quadrature*—as a crucial "format" for expressing profound area relationships in geometry.

And I'll conclude by alluding to Archimedes' wonderful "mechanical method" where he transmutes the problem of computing *area* to computing something akin to *weight*[1] and thereby achieves the *quadrature of the parabola.* This offers us a glimpse of the power of analogy, and thought-experiment as already practiced in ancient mathematics.[2]. It also gives us the opportunity to touch, in passing, on broader issues in mathematical thought such as *analogy, heuristic, paradox, invariant* and something I'll call *characterization* (a version of *axiomatization*).

Contents

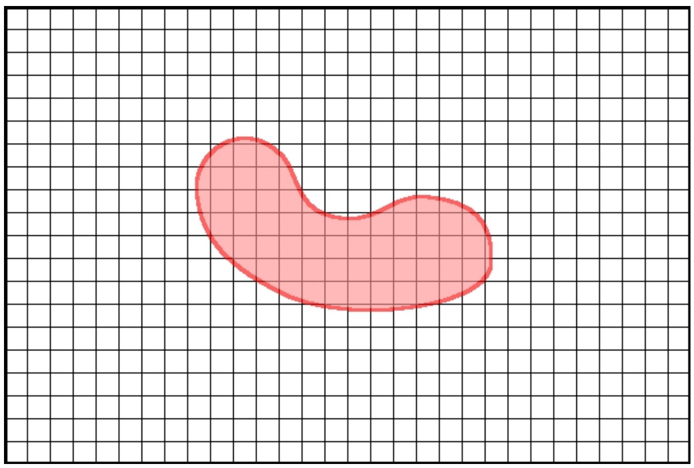
[1]Since you read Archimedes' *On the equilibrium of planes* in Freshman Lab, this may not come as a complete surprise.

[2]I am grateful to Paul VanKoughnett who drew most of the figures, and to Paul Dry for helpful and incisive comments about early drafts of these notes.

1. Area as familiar

We all know what the word *area* signifies. It often refers to a territorial cordon, as in *restricted area* or *hard-hat area* or even *area studies*. It sometimes comes as a *number* but always with a unit attached, such as square miles, square feet, square inches, acres, or if it's a bed area you're interested in, you can ask for it to be King-size or Queen-size, or a size of lesser nobility.

If you want to approximate the area of this enclosed figure on a grid with the mesh given by 1 foot by 1 foot squares,



you might count the number of 1-square-foot patches that comprise a union of squares that completely cover the figure (in the above case it is 81) and count the number of 1-square-foot patches that the figure covers completely (in the above case it is 39) and you know that the area of the figure in units of square feet is squeezed between these two numbers; that is, the area is smaller than the first number (or equal to it), and bigger than the second. In the above case we would have

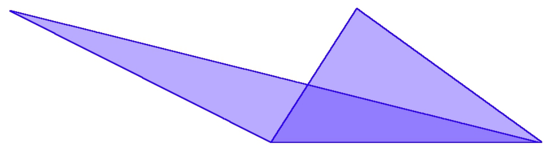
$$81 \text{ square feet} \quad \geq \quad \text{area of figure} \quad \geq \quad 39 \text{ square feet.}$$

If you want a better estimate, do the same thing with 1" by 1" squares.

If any of us were asked to calculate the square-footage of this auditorium we'd come up with some figure or other, confident that we could refine it to any degree accuracy required. And you may be painfully aware of the area of your dorm room. So, what else is there to say?

2. How good are you at comparing areas?

I'm not very good. Here's an example. The area of the two shaded triangles in the figure below

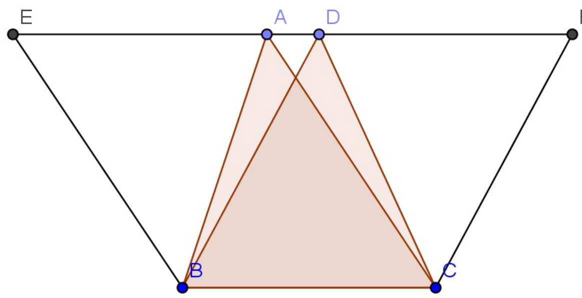


are equal. I know this thanks to Euclid's Proposition 37 of Book I of the *Elements*. You all either know that proposition now, or will after your freshman year.This is the proposition that says that triangles with the same base and height have the same area. Proposition 37 of Book I will be a recurring theme in my lecture: a marvelous piece of mathematics that demonstrates many things, including the maxim that to be profound and to be elementary are not mutually exclusive virtues.

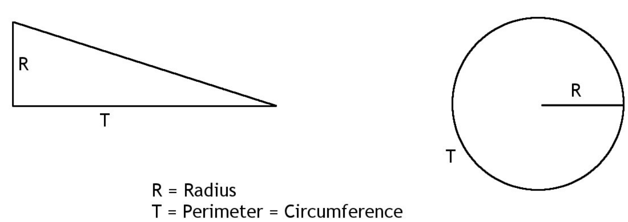
But if I simply compared those figures visually—without either explicitly remembering or somehow "internalizing" that proposition—-I would probably grossly underestimate (if that's a possible phrase) the area of the spiky triangle in comparison with the seemingly fat one. In a sense, that proposition then–embedded in my central nervous system as it is—has therefore improved (a tiny bit, not much) my ability to make off-the-cuff judgements and rough comparisons. We all have our native intuition, together with a bank of geometric experiences contributes to how effective we are in making judgements about all sorts of attributes of objects that we see. I'm guessing that our eyeball judgment of comparison of straight-line lengths is more reliable than curved lengths[3] and much more reliable than area, and volume, given the variety of possible configurations. Often, in teaching Euclid's *Elements*, one emphasizes 'logical thinking' as the great thing that students take away from it. But I also see a type of pre-logical, if I can call it that, or intuition-enriching, benefit as well; this is hard to pinpoint but comes out as a sharpening of our ways of generally thinking about, guessing about, negotiating, comparing and relating, geometric objects.

Such a gain is very different from the other valuable reward—e.g., being able to actually argue the proof of Proposition 37 of Book I by making the elegant construction

[3]Straight-limbed geometry;
In her arts ingeny
Our wits were sharp and keen.



To continue our review of how good our intuitions are, let's pass to a slightly deeper basic geometric comparison—due to Archimedes—that astonishes me now just as it must have astonished Archimedes' contemporaries. We will get into this in more depth later on[4] but consider the following striking way of *recreating* the area of any circle: the area of any circle is *equal to* the area of a right-angle triangle defined by the property that the two of its sides making the right angle have lengths equal to the radius of the circle, and to the length of the circumference of the circle, respectively.

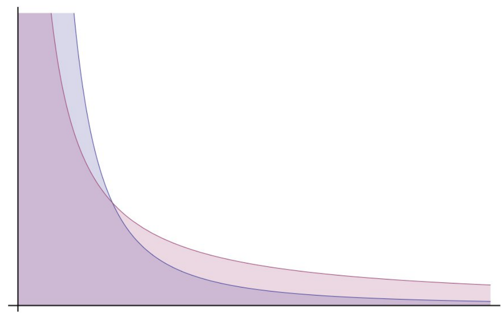


Here, as in the previous example I "see" it not visually (if this can be said) but only with the help of my memory of its proof (about which more below).

Let us push a bit further. The two examples we've just reviewed are examples of nicely enclosed—finite—figures. We are even poorer in intuition when faced with planar figures (no matter how smooth and simple-seeming boundaries they have) that "asymptote" off to infinity. It may be quite difficult, even if given long chunks of it, to extrapolate, and guess by eyeball alone, whether it is on the road to extending out to a figure with infinite area or finite area.

[4] A sketch of its proof is in the Appendix, i.e., section 20 below.

For example, consider the figure below, which was drawn as accurately as possible. I wonder whether you can guess if the area bounded by the blue curve, or the red curve has finite area. My point is that there's no reason why you should be able to do it no matter what talents of visual acuity you may possess.



But... in case you are wondering, if you continued tracing the curved regions ad infinitum in the manner smoothly begun by the sketch—meaning the red curve is the graph of the function $y = 1/x$ and the blue curve is the graph of the curve $y = 1/x^2$—the (infinite) red-bounded region happens to have infinite area while the (infinite) blue-bounded region has finite area. As a side issue, many Calculus students are amazed to find that if you "construct" an infinite "trumpet" by rotating the red curve of the above figure around the x-axis, then its surface area is infinite, even though the volume it subsumes is finite. To put it into colloquial, but misleading, vibrant language: you can fill this trumpet with a finite amount of paint, but you need an infinite amount of paint to paint it.

This phenomenon illuminates much, especially for people who know Calculus, and could be the subject of a question later tonight. But I won't dwell on this; I mention it as a hint that there are things to dwell on here.

3. Pushing to the limit

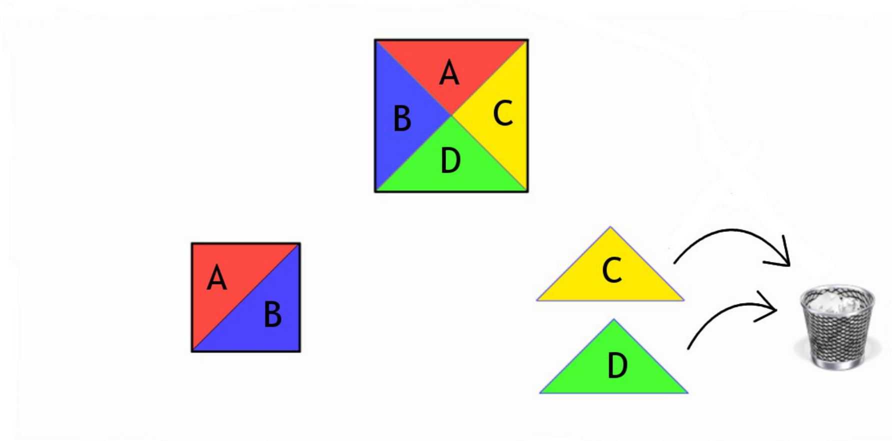
You might wonder what is to be gained by asking questions such as infinite area versus finite area, or by considering our concept *area* in various extreme contexts. Mathematics often does that sort of thing: it is a useful strategy to examine a concept when it is brought to its limit, in hopes that the tensions one has imposed on the concept will reveal important facets of it that would be hidden in tamer situations. If you push a concept to its extreme border, you may see things that would otherwise be overlooked if you always remained in the comfortable zones: for example, you learn what its precise borders are. Some of the most beautiful mathematics, and the deepest, have emerged by seeking the extremes.

The issue we have just addressed, areas of infinitely extended regions, can be broadened, for it raises the question: exactly how many subsets of the plane *deserve* to have a well-defined area? Do *all* subsets have a reasonable notion of area??

4. Area as paradoxical

To get the blood circulating, let's contemplate something that is evidently impossible: Can you

(1) take a square S in the Euclidean plane, cut it (say with a scissors) into four pieces A, B, C, D of equal area—so no two of these pieces overlap, and the four of them cover the square?



In standard notation[5]:

$$S \quad = \quad A \cup B \cup C \cup D;$$

(2) and now can you throw away two of the pieces (say C and D) and move the other two (A and B) around by Euclidean motions to get congruent shapes A', B' in the plane so that these two pieces cover the exact same square again, i.e.,

$$S \quad = \quad A' \cup B'?$$

The answer, of course, is no: you can't do this. Certainly not if the concept of *area* has the properties that we expect to have. By 'properties' I'm referring to these two self-evident axioms:

[5]The "cup" $\cup$ notation means *union*. That is if X and Y are sets, then $X \cup Y$ is the set whose members are either members of X or members of Y (or members of both).

(1) the area of a union of nonoverlapping figures is the sum of the areas of each of the figures, and

(2) the area of a figure is preserved under Euclidean motion.

For, if you could do this, then our two formulas for S displayed above will be giving contradictory answers to the question: *what is the area of S?*

With this in mind, consider the following strange fact about *spherical* rather than Euclidean geometry. Let $\mathcal{S}$ be the surface of a ball (that is, what mathematicians call *the two-dimensional sphere*). There is a way of separating $\mathcal{S}$ into four sets $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$, no two of which overlap, such that each of these sets are—in an evident sense—*congruent* to any of the others. (This means that, for example, there is a way of rotating the sphere that brings $\mathcal{A}$ precisely to the position that $\mathcal{B}$ occupied (before that rotation)—and similarly, there are ways of rotating the sphere to bring $\mathcal{A}$ to $\mathcal{B}$ and to $\mathcal{C}$ and to $\mathcal{D}$. NEVERTHELESS, you can throw two of them away (say, $\mathcal{C}$ and $\mathcal{D}$) and find a way of rotating the sphere so that $\mathcal{A}$ is brought to a set $\mathcal{A}'$ and a (different) way of rotating the sphere so that $\mathcal{B}$ is sent to a set $\mathcal{B}'$ and these maneuvers have the strange property that $\mathcal{A}'$ and $\mathcal{B}'$ together cover the sphere; i.e.,

$$\mathcal{S} = \mathcal{A}' \cup \mathcal{B}'.$$

This is called the *Banach-Tarski Paradox.* Despite first appearances this is not actually a paradox although there is indeed a subtlety lurking in the way in which I worded things[6]. It is merely a para-dox, that is: *contrary to expected opinion.*

I brought this seeming-paradox up not to confuse you but rather to point out, at the very outset, that

- even though *area* is a concept we tend to feel perfectly at home with, to get closer to its essence is to appreciate more keenly its complexity, and so
- in our discussion about area this hour we had better start from the very beginning, by noting that
- despite its reputation for having what are called 'proper foundations', mathematics doesn't seem to have a 'beginning.'

[6]For people who are familiar with group theory a fairly complete description of what is going on here can be found in http://en.wikipedia.org/wiki/Banach%E2%80%93Tarski_paradox.

5. Length in everyday life

Nevertheless, let's begin with something seemingly a bit simpler than area: plain old length of straight line segments. I say "seemingly" because as often happens in mathematics, the simpler-seeming concept (in this case, length) contains, in a more visible form, lots of the essential aspects of its complicated companions (e.g., area and volume).

We all know what is meant when someone says "*a ten foot pole.*" This is a relative statement, comparing the pole to some foot-long ruler, and claiming that we can lay ten copies of our measuring device onto the pole, covering it completely with no overspill or overlap. Usually, of course, the speaker of this phrase has some thing on his mind other than this length-measuring thought-experiment.

In small contrast, when we are told that the circumference of this cup is eight inches long and want to verify this directly, we must set aside our rigid, calibrated, ruler and use something like a tape-measure, wrapping it around the rim of the cup[7]. Of course, there is also a well-known indirect way of verifying this measurement which starts by using our rigid ruler to calculate the diameter of the cup—but this indirection already involves a certain amount of mathematical experience with π.

In even greater contrast, when we are told that the star cluster NGC 1929 within the Large Magellanic Cloud (a satellite galaxy of our own Milky Way) is $179,000$ light-years away from us, other measuring devices are required, and—given relativistic issues—what *distance* means is already a subtle business.

[7]A delightful book that discusses modes of calculation of this sort, and of more theoretical sorts, see *How Round is Your Circle? Where Engineering and Mathematics Meet,* by John Bryant and Chris Sangwin, Princeton University Press, 2011.

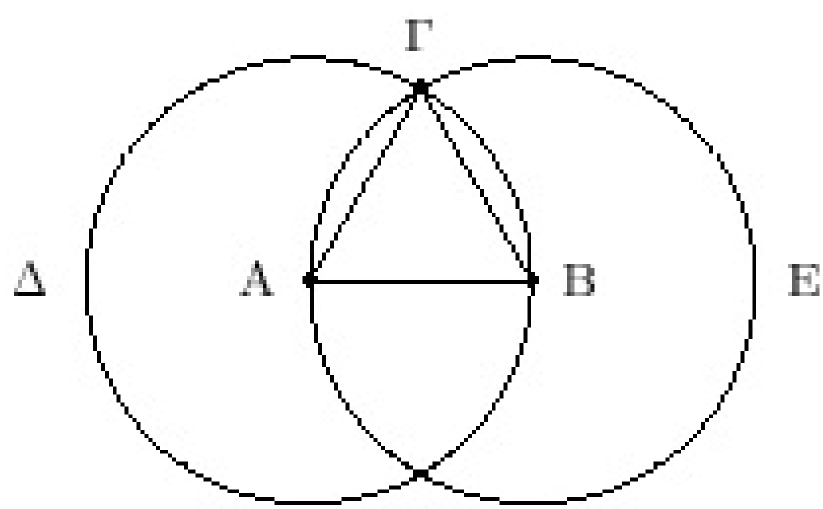
6. Equality [of Length] in Euclid

The concept *length* occurs—in a somewhat cryptic form—early in the *Elements*. It appears ($\mu\hat{\eta}\kappa os$) in the definition of *line* (Definition 2):

A line is breadthless length.

The concept reappears as the term $\delta\iota\alpha\sigma\tau\eta'\mu\alpha$ (translated often as *distance* but meaning, more specifically, *interval* or *gap*) as something of a surprise. Euclid slips it into the discussion in Book I in the definition of **circle** (Def. 15) which is described as—and I'll put it in modern vocabulary—a figure bounded by a curve the points of which are *equidistant* from a given point.

Thanks to this definition and the ability we have—given to us by Postulate 3—of *drawing a circle with any center and any radius*—we can begin to construct many line segments that—in Euclid's terms—are "equal," (meaning: *are of equal length*). Even better, we are supplied with tools for establishing *equality.* Euclid wastes no time making use of these tools: the very first proposition (Proposition 1 of Book I) goes straight to the task of constructing an equilateral triangle on any line segment, i.e., a triangle where all three sides are "equal." And we're off and running (at least as far as understanding *equality* of length).

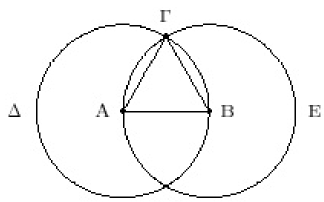


Euclid's Proposition 1

Ἐπὶ τῆς δοθείσης εὐθείας πεπερασμένης τρίγωνον ἰσόπλευρον συστήσασθαι.

Ἔστω ἡ δοθεῖσα εὐθεῖα πεπερασμένη ἡ ΑΒ.

Δεῖ δὴ ἐπὶ τῆς ΑΒ εὐθείας τρίγωνον ἰσόπλευρον συστήσασθαι.



Κέντρῳ μὲν τῷ Α διαστήματι δὲ τῷ ΑΒ κύκλος γεγράφθω ὁ ΒΓΔ, καὶ πάλιν κέντρῳ μὲν τῷ Β διαστήματι δὲ τῷ ΒΑ κύκλος γεγράφθω ὁ ΑΓΕ, καὶ ἀπὸ τοῦ Γ σημείου, καθ᾽ ὃ τέμνουσιν ἀλλήλους οἱ κύκλοι, ἐπὶ τὰ Α, Β σημεῖα ἐπεζεύχθωσαν εὐθεῖαι αἱ ΓΑ, ΓΒ.

Καὶ ἐπεὶ τὸ Α σημεῖον κέντρον ἐστὶ τοῦ ΓΔΒ κύκλου, ἴση ἐστὶν ἡ ΑΓ τῇ ΑΒ: πάλιν, ἐπεὶ τὸ Β σημεῖον κέντρον ἐστὶ τοῦ ΓΑΕ κύκλου, ἴση ἐστὶν ἡ ΒΓ τῇ ΒΑ. ἐδείχθη δὲ καὶ ἡ ΓΑ τῇ ΑΒ ἴση: ἑκατέρα ἄρα τῶν ΓΑ, ΓΒ τῇ ΑΒ ἐστιν ἴση. τὰ δὲ τῷ αὐτῷ ἴσα καὶ ἀλλήλοις ἐστὶν ἴσα: καὶ ἡ ΓΑ ἄρα τῇ ΓΒ ἐστιν ἴση: αἱ τρεῖς ἄρα αἱ ΓΑ, ΑΒ, ΒΓ ἴσαι ἀλλήλαις εἰσίν.

Ἰσόπλευρον ἄρα ἐστὶ τὸ ΑΒΓ τρίγωνον, καὶ συνέσταται ἐπὶ τῆς δοθείσης εὐθείας πεπερασμένης τῆς ΑΒ.

[Ἐπὶ τῆς δοθείσης ἄρα εὐθείας πεπερασμένης τρίγωνον ἰσόπλευρον συνέσταται]: ὅπερ ἔδει ποιῆσαι.

There it is in full glory ending with a triumphant *hoper edei poiaisai*—i.e., "as was to be constructed".

When later mathematics takes on the issue of length, things proceed quite differently from Euclid. Modern mathematics throws a spotlight on *transformations*, in a way that ancient mathematics did not. Nowadays, as we introduce a new concept, or new type of structure, often—at the same time—we make explicit the types of *transformations*, or mappings between exemplars of this structure, that we are willing to consider (that we are willing to *allow*). These *allowed transformations* are the ones that respect the inner coherence of the structure we are studying. In Euclidean geometry, the *allowed transformations* are the mappings of the Euclidean plane onto itself that preserve the notion of congruence. The Euclidean allowed transformations consist of rotations about points in the plane, translations, and also those transformations that can be viewed as the composition of a "flip," (i.e., a symmetry about some straight line) with a translation, or rotation. In one of the modern formats, the concept of Euclidean *length* and the collection of *the allowed transformations* of Euclidean geometry are yoked concepts, working in tandem:

- The allowed transformations are precisely those transformations that preserve length of all line-segments, while

- two line-segments have equal length if and only if there are allowed transformations bringing any one of them onto the other.

In effect, these notions ({length between points in Euclidean geometry} and {the transformations that preserve Euclidean geometry}) are yoked—chicken & egg style—in that each can be used to begin the discussion and *characterize*—i.e., explicitly determine—the other. Think of it this way: we could invoke each of these concepts to provide the vocabulary for a system of axioms for a geometry, and the other concept would then be one of the many features of that geometry. You can have length as your basic concept and stipulate the transformations that preserve your geometry to be those that preserve length, or you may start with the stipulation of transformations of your geometry and derive length as one of its invariants—in effect, deriving the entire geometry from its group of symmetries. (The second viewpoint represents a celebrated shift of emphasis, known as the *Erlangen Program*.) But there is also an important difference of mood between 'axiomatization' which sets up a theory starting from one direction or the other and presenting things in a balanced way, where each concept 'characterizes' the other.

7. Proportions

Length is, at bottom, a *relative* concept: i.e., "length compared to what?" is a bona fide question. What are the units? Inches? Feet? Miles?[8]. That is, when we deal with length, we are dealing—unavoidably—with a *proportion*[9].This puts us in the mood of Euclid's Book V, a work that deals exclusively with proportions of magnitudes.

Say we are interested in the length of our ten-foot pole P. We compare it to our one-foot ruler F, we might emphasize the proportional aspect of length by recording the answer symbolically this way:

$$(*) \qquad P : F \quad \text{“}=\text{”} \quad 10 : 1$$

[8]Two cowboys:

A: "My ranch is so big I can ride Old Paint from morning to night and still not cover it."

B: "I know exactly how you feel. My horse is like that too!"

[9]This issue is taken up by Kant from a slant perspective (that's typical for Kant). In Book I Sections 25, 26 of *The Critique of Judgment* in discussing what he calls the *mathematical sublime* he points out that in comprehending in our imagination a specific magnitude (say, this pole is ten feet long) one is engaging in two acts, of different natures: there is the mathematical one of counting a number of feet (and comprehending that act of counting) and then there is the essentially aesthetic one of comprehending—or internalizing in some way or other—what a *foot* is. From Kant's perspective, then, considering a proportion, per se, is an act that extracts the purely mathematical aspect of "comprehension of a magnitude" from the underlying, and otherwise unavoidable, aesthetic aspect: comprehending the unit. Of course, it is the latter that interests him.

I've put quotation-marks around the equality sign to emphasize that it is indeed a serious abbreviation of thought, turning what began as an analogy (P is to F *as* 10 is to 1) to an equality (the relationship that P has to F *is* the relationship that 10 has to 1) turning an *as* into a straight *is* (which is a curious transition). The older notation for our *quotation-marks equality sign* is a double-colon,

$$(**) \qquad P : F \ :: \ 10 : 1,$$

capturing equally well—I believe—the "*as*" aspect of the relationship. That a proportion of lengths *is interpretable as* a proportion of numbers may well be self-evident, but that it is an *interpretation* is worth bearing in mind.

The legacy of the Pythagoreans offers us yet another interpretation for the versatile notion of a proportion of lengths:

> *As the length is to the length,*
> *So the heard tone is to the heard tone.*

After this discussion it is safe to remove the quotation marks in formula (*) displayed above, and write

$$\frac{P}{F} = \frac{10}{1}.$$

We thereby see arithmetic in geometry (i.e., by going from left to right in the above equation). That is, we have an *arithmetic* (of proportions of straight line segments) that mirrors "ordinary" arithmetic (of ratios of numbers). For example, you can add proportions of lengths of line segments:

$$A \cdot - - - - - - - \cdot B - - - - - - - - - - - - - - - - \cdot C$$

$$D \cdot - - - - - - - \cdot E$$

$$\frac{AB}{DE} + \frac{BC}{DE} = \frac{AC}{DE}$$

and multiply proportions of lengths of line segments:

$$A \cdot - - - - - - - \cdot B$$

$$C \cdot - - - - - - - - - - - - - - - - \cdot D$$

$$E\cdot - \cdot F$$

$$\frac{AB}{CD} \quad \times \quad \frac{CD}{EF} \quad = \quad \frac{AB}{EF}$$

and we have a natural interpretation of *inequalities* between these proportions. These behave formally "just like fractions" as the notation indicates, and we have a veritable algebra of geometrical proportions.

8. Common Measures, and uncommon measures

All this makes perfect, and natural, sense and conforms to the most elementary basic ideas we have about arithmetic as long as we treat proportions of lengths that "admit a common measure."

That is, imagine that you are given two intervals,

$$A\cdot - - - - - - - \cdot B$$

$$C\cdot - - - - - - - - - - - - - - \cdot D$$

and you know that there is a certain unit measure, say given by another interval EF,

$$E\cdot - - \cdot F$$

such that AB and CD are measured by (whole) number multiples of EF. For example, say AB is seventeen EFs long and CD is four hundred ninety one EFs long; so we may write:

$$\tfrac{AB}{EF} = \tfrac{17}{1} \text{ and } \tfrac{CD}{EF} = \tfrac{491}{1}.$$

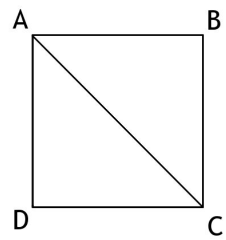
We then say that EF is a **common measure** for the line segments AB and CD.

And, in this particular case, we then comfortably write

$$\frac{AB}{CD} = \frac{17}{491}.$$

But the fun, as I think you all know, is already there at the very outset of geometry for one of the most fundamental of geometric proportions, *diagonal* to *side* of a square:

$$\frac{\text{the diagonal } AC}{\text{the side } AB},$$



was shown to have *no common measure*[10] and nevertheless the proportion $\frac{AC}{AB}$ (alias: $\frac{\sqrt{2}}{1}$) was *still* regarded as a genuine object of study, with the consequence that forced us—by the analogy between proportions of lengths and proportions of numbers—to extend our very idea of what it means to be a number. It is worth thinking about what it means for *geometry* to guide us in our evolving concept of *number*.

For, all this is a beginning of one of the great analogies,

$$\textbf{arithmetic} \leftrightarrow \textbf{geometry},$$

where each profoundly influences the other. This type of thinking goes against a view held by Aristotle (a view often referred to as *purity*), namely:

> ... we cannot in demonstrating, pass from one genus to another. We cannot, for instance, prove geometrical truths by arithmetic.
>
> Aristotle, *Posterior Analytics,* 75a29-75b12

But much of the deepest mathematics goes counter to that view!

---

[10] at least if we insist that both intervals be measured by a *whole number* of multiples of the chosen "common measure."

9. The ubiquity of "analogy" in mathematical thought

In the previous section we have been working through the idea that straight line segments stand *in relation to each other* "just as" numerical quantities stand in relation to each other; i.e., we are faced with—as we've mentioned—one of the primordial analogies between geometry and arithmetic.

That this "just as" relation *is* an analogy and not a direct equality takes convincing. A curious phenomenon occurs with many of the mathematical analogies once they get embedded in our thought. If A is seen to be analogous to something else, B, there is the impetus to think of A and B as, somehow, special cases of, or aspects of, a single more encompassing C; and somehow to rethink the analogy as *equality.* This switch is a form, but not the only form, of *abstraction* that is indigenous to mathematical sensibility. Versatile switching of viewpoints is one of the reasons for the power of a mathematical frame of thought. This replacement of a pair of analogous contexts for a single encompassing context occurs so often that people with experience in mathematics have this type of thought engrained in them, as second nature[11].

As I mentioned, replacing the two parts of an analogy by a common generalized concept is powerful and occurs often in mathematics, but in other contexts of thought it might seem a strange thing to do. One rarely does this kind of generalizing with analogies and metaphors that occur in literature: when thinking about the metaphorical comparison in

Shall I compare thee to a summer's day?

do we conceive of a more general entity that encompasses "thee" and "summer's day" as instances?

10. Length and straight line segments

You may have noticed that although I've gone on at length I never defined *straight line segment.* Now, you can postpone talking about straight line segments if you phrase things in terms of *distance.* That is, for any two points P and Q on the Euclidean plane, if you have a notion of *the distance between* P *and* Q—denote it by $\mathrm{dist}(P,Q)$—you can pick out the points on the straight line segment between P and Q as precisely those points X such that

$$\mathrm{dist}(P,Q) = \mathrm{dist}(P,X) + \mathrm{dist}(X,Q).$$

However, the ancients seem not to have defined straight line segment this way. Euclid's definition (Definition 4 of Book I) is elegantly enigmatic:

A straight line is a line which lies evenly with the points on itself.

This is reminiscent of Plato's definition of straight line segment as "whatever has its middle in front of its end" (Parmenides 137e). Here, Plato seems to be taking his straight line segment up to his eye to view it as you would look through a telescope, noting that the

[11] Here is an important example of this that originated over a century ago, and is everywhere to be seen in modern mathematics: *numbers* are analogous to *functions.* There are whole branches of mathematics where these concepts are treated as not merely analogous, but as particular exemplars of a larger encompassing concept.

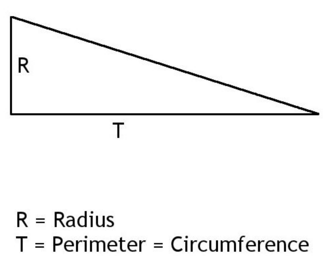
only thing he sees is its endpoint. In effect, a straight line is a *line of sight*.[12]. A much later take on the matter defines a straight line segment with endpoints P and Q as the unique curve joining P and Q such that among all curves joining P and Q it is the one having the shortest length[13]. But to make sense of this, at the very least, you must know what it means for a curve to have a length. Hence ...

11. Lengths of smooth curves

Nowhere in Euclid's *Elements* is the length of a curve that is *not* a straight line, or polygonal, segment discussed. The first nonpolygonal curve whose length was considered (in the texts that I know) is the circumference (called *the perimeter*) of a circle, as studied in Archimedes' *The measurement of the circle*[14]. And there the length of the circumference of a circle enters the mathematical discussion in the context of the elegant statement about area[15] that we have already briefly discussed in section 2.

Here it is as Proposition 1 of Archimedes text:

> **Proposition 1:** Every circle is equal to a right-angled triangle, whose radius [R] is equal to one of the [sides] around the right angle while the perimeter [i.e., circumference T of the circle] is equal to the base [of the triangle].



[12]See, for example, Thomas Heath's marvelous essay about all this (pp. 165-169 in Vol. 1 of his three-volume translation of *Euclid's Elements*).

[13]This property distinguishes straight line segments as *geodesics* in modern terminology.

[14]For source material, various translations, commentary, and more related texts, please go to http://isites.harvard.edu/icb/icb.do?keyword=k53966 (which is on the "Teaching" page of my web site).

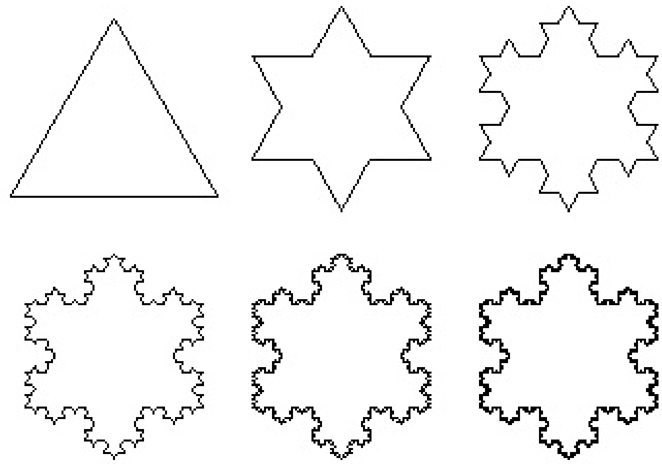
[15]That the ratio of the area of a circle to the length of its circumference is a simple expression in terms of its radius is what is behind the beauty of Proposition 1. This is a phenomenon that proliferates in higher dimensions; e.g., the ratio of the volume of a sphere to its surface area is, similarly, a simple expression in terms of its radius. This is worth pondering.

This is proved by approximating the circle by a regular polygon with a large number of sides, and arguing appropriately (as in the appendix below). This is an *amazing* theorem, of course, but the more specific reason I'm mentioning it is that it exemplifies the general rule that the computation of the length of any *curvy* curve depends—perhaps very indirectly—on relating it to the length of approximating polygons. This is (quite directly) Archimedes' method here. He makes use of a result about polygons analogous to Proposition 1, where the polygons in question will be made to approximate the circle. For a slightly more extensive sketch of Archimedes' argument, go to the Appendix (section 20) below.

12. Lengths of crinkly curves

It has been said that there is no way to measure the length of the coastline of Scotland.

It is just too crinkly, and the length you find yourself computing depends on how fine a grid of measurements you make—the result getting longer and longer as the measurements grow finer. Mathematicians can easily model such an effect, the most famous construction being something called the Koch snowflake. This is a closed curve obtained by taking the limit of an infinite sequence of crinkle-operations. Start with an equilateral triangle and on an interval one third the size of each side construct a small equilateral triangle. Here are the first few stages:

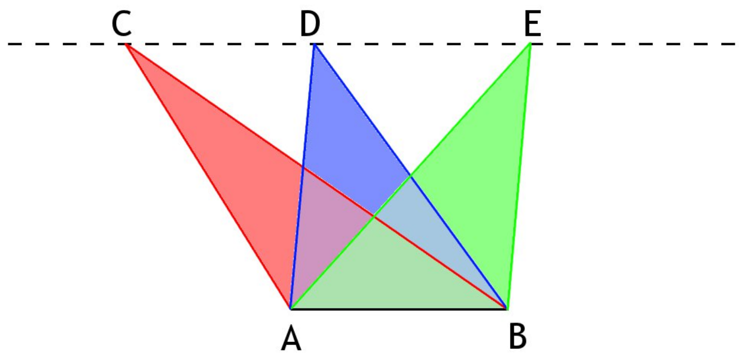


At each stage you are faced with a longer curve, and in the limit, you have seemingly contained a curve within a finite region that is so crinkly so as to have—in effect—infinite length.

13. What is area?

We'll be interested primarily in the areas of figures in the Euclidean plane. Given our discussion of *length* it won't be a surprise to learn that we will be dealing, again, with proportions; i.e., the proportion of [the area of] one figure to [the area of] another. Nor will it be much of a surprise to find that just as *straight line segments* played a fundamental role in all discussions of length, so too *polygonal figures* will play such a role in our treatment of area.

Euclid is again very helpful here. The first time he discusses area, it is—in his vocabulary—*parallelogramatic area*: Proposition 34 of Book I tells us that the diagonal of a parallelogram bisects the [area of] the parallelogram. He follows this up with the propositions (including the beautiful Proposition 37 that I've alluded to already) that state that two triangles with the same base and height have the same area; and similarly for two parallelograms with the same base and height.



And once we have these tools, we are in good shape to deal with areas of polygonal figures and can even go further, as with Archimedes' Proposition 1 of *The measurement of the circle* that we've discussed in Section 11 above. We will discuss it more in the Appendix (section 20) below.

14. Area as an "invariant"

Here is an exercise: make a (short) list of "axioms" that (you guess) characterizes the concept of *ratios of areas* for a large class of (plane) figures. You'll surely include a number of basic properties of the intuitive concept of *area* as hinted at in section 4 above but let me start the game by insisting that one of your axioms be this:

Axiom of Invariance under Euclidean motions:

If A, B are a pair of plane figures for which you have defined the ratio

$$\frac{\text{area of } A}{\text{area of } B}$$

(or, for short, $\frac{A}{B}$,) and if A' is the image of the figure A under a Euclidean transformation, then we have the equality:

$$\frac{\text{area of } A}{\text{area of } B} = \frac{\text{area of } A'}{\text{area of } B}$$

This is worth thinking about, but this is just a start, and note that in your personal "theory of area," part of the chore is to make precise exactly what class of figures you are going to be assigning a well-defined area. This might be a bit of fodder for our *question period* later this evening. This exercise, in effect, was solved elegantly and in somewhat astounding generality (before World War II) by the Hungarian mathematician Alfréd Haar.

15. Shears and Similarities

Given a well-working "theory of area," certain properties will follow as consequences. For example, here are two basic features: further *invariance properties* for the concept of area.

(1) **Shears.**
By a **horizontal shear transformation** let's mean a transformation of the Euclidean plane to itself that keeps every horizontal line in place, but moves it by a translation that is dependent on the "height" of that line above the x-axis. That is, for any point (x, y) in the plane it keeps the y-coordinate fixed but allows the x coordinate to change by a rule:

$$x \mapsto x + F(y)$$

where $F(y)$ is some civilized (e.g., continuous) function of y. This type of motion of the plane keeps all line parallel to the x-axis intact, but translates them by different amounts depending on their height. By a **general shear transformation** let's mean an analogous such transformation but with respect to lines parallel to any fixed line: the line needn't be the x-axis.

The area of figures are preserved by shears!

Now we've actually seen examples of this in our previous discussion: think of Proposition 37 of Book I of Euclid's *Elements.* One way of revisiting the content of Proposition 37 is to note that any two triangles with the same base and same height can be brought one to another by a *shear*.

The three-dimensional version of this (where the question is about *volume* rather than *area*) is sometimes referred to as **Cavalieri's Principle** and is illustrated, for example, by the following picture, where Cavalieri's Principle would state that these two stacks of coins occupy the same volume.

(2) **The behavior of area under similarity transformations.**

If

- A and A' are in the class of figures for which you have defined the ratio
$$\frac{\text{area of } A}{\text{area of } A'},$$
and if
- P, Q are points in the figure A with P', Q' the corresponding points in the similar figure A',

then the *square* of the ratio
$$\frac{\text{length of } PQ}{\text{length of } P'Q'}$$
is equal to the ratio
$$\frac{\text{area of } A}{\text{area of } A'}.$$

This *square relation* tells us that we are dealing with a *two-dimensional* concept[16].

Dimensionality as a concept opens up a host of marvelous questions to explore, not the least of which is the grand idea, initially due to Hausdorff, that the full range of possible geometric figures admit a continuous gamut of dimensions—not just dimension $0, 1, 2, 3, \ldots$.

That such strange figures possessing non-whole-number-dimensions may have some bearing on questions in the natural sciences, economics, and finance—let alone pure mathematics—is the energy behind Benoît Mandelbrot's well-known *fractals*[17].

16. Invariance as feature; Invariance as characterization

I have been alluding to the invariance properties of *length* and of *area*. Here is a summary and comparison.

Euclidean Length and the collection of Euclidean motions suit each other's needs perfectly:

- The (Euclidean) concept of *length* is invariant under the *Euclidean motions* (i.e., translation, rotations, symmetries about straight lines, and compositions of these). That is, these transformations *preserve* Euclidean distance.
- *Any* distance relation between points that satisfies certain natural axioms and that is invariant under any Euclidean motion **is** (after appropriate rescaling of its values) equal to the (Euclidean) concept of *length*.
- Moreover, any transformation that preserves length between any two points in the plane *is* a Euclidean motion.

In contrast to *length*, The invariance properties of (Euclidean) concept of *area* is stranger:

- The (Euclidean) concept of *area* is invariant under Euclidean motions, of course—but it is also invariant under a much greater collection of transformations. For example, any of the shear transformations we have discussed in the previous section (Section 15) preserves area.
- But as for characterizing this concept by invariance properties, things go the other way: *area* is characterized (up to a mere change of scale) by its invariance under translations alone—that's all the invariance you need invoke to pinpoint this concept!

[16]But neither of the above "invariance properties" need be, nor should be, included as axioms, for they will follow from your list of axioms (if you've formulated them correctly.

[17]See his *The Fractal Geometry of Nature.* Lecture notes in mathematics **1358**. W. H. Freeman.

It is quite fitting, then, that Euclid inserts his Proposition 37 in Book I, i.e., very early in his discussion of area: the proposition is, of course, a critical tool in establishing the simplest arguments regarding area—but from a modern perspective, it also points to one of the deep properties of the concept vis à vis invariance: there is a huge collection of transformations—far more of them than just Euclidean motions—that preserve area.

17. Classical Quadrature Problems

The phrase *quadrature of* ... loosely refers to the problem of *finding the area of* ..., which usually means expressing—as some simple numerical ratio—the proportion of the area of one figure to another figure[18]. First, here is a simple example related to Euclid's Proposition 37 in Book I of the *Elements* that we discussed earlier, and whose proof can be found by putting together propositions in Book I of Euclid's *Elements*[19]:

Proposition: Let P be a parallelogram and T a triangle, such that P and T have the same *base* and the same *height.* Then

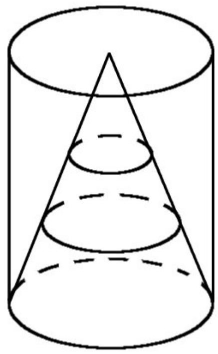
$$\mathrm{P{:}T} \quad = \quad 2{:}1.$$

This proposition follows the format of what I'll call a "Classical Quadrature Problem," which I want to mean to be a statement that *the proportion of areas (or lengths, or volumes) of two geometric figures, all described entirely in clear general geometric terms*[20]*, is equal to a specific numerical ratio.*

There are quite a number of classical problems that fit this mold, e.g., that express the proportion of the areas of two figures, or volumes of two solids (described in general terms) to a specific rational number. For example, Proposition 10 of Book XII of Euclid's *Elements* tell us that

the ratio of the volume of a cone to a cylinder that have the same base and the same height is $1/3$.

[18]Quadrature is the basic topic in the oldest existent Greek text (that of Hippocrates of Chios).
[19]or, better, by doing it yourself
[20]This is admittedly a bit vague, but I hope the examples convey the kind of problem I'm referring to.



This "1/3" reoccurs as the ratio of the volume of a conical solid built on *any* base to the cylindrical solid built on the same base, and of the same height. The earliest text I know that "explains" the "1/3" in this more general context is *Arithmetica Infinitorum* by John Wallis who did his work before the full-fledged invention of Calculus; for the people who know Calculus, this is an exercise[21].

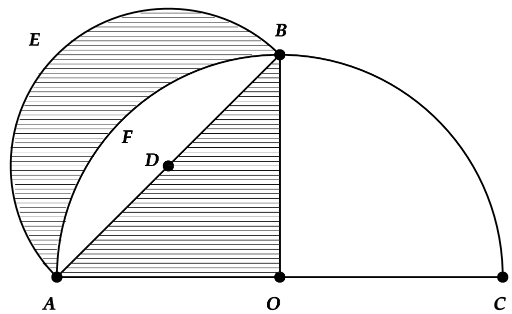
As with much of Archimedes' work there are stories that surround it. In one of his treatises[22], Archimedes showed that

> *The ratio of the volume of a sphere to that of the cylinder that circumscribes it is* 2/3.

and according to story, this being his favorite result, he had it engraved as a sculpture for his tomb.

The most intriguing, and thorny, of the ratios of elementary areas or volumes are the proportion relating the area of a circle to that of the square that circumscribes it, and the proportion relating the volume of a sphere to that of the volume of the *cube* that circumscribes it. The story of the many attempts to understand these ratios leads us in interesting directions. For example, Hippocrates of Chios in his attempt to *square the circle* studied classical quadrature problems relating the areas of lunes (which are figures consisting of the outer portion of a small circle when superimposed on a larger one, as in the figure below).

[21]the hint being that $\int_0^x t^2 dt = \frac{1}{3}x^3$.
[22]*One the Sphere and Cylinder I*



to areas of triangles constructed in relation to those lunes. He proves, for example, that the area of the lune (defined as the region between E and F in the figure above) is *equal* to the area of the triangle ABO. His results, however, go signficantly beyond this[23].

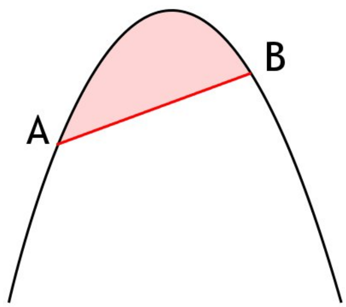
18. Weighing area

A famous example of a classical quadrature problem is Archimedes' "Quadrature of the parabola" and this is dealt with in (not one, but) two of his treatises in quite different ways:

- Propositions 14-16 of *The quadrature of the parabola* and
- Proposition 1 in *The Method*.

The aim is to "find" the area of a segment of a parabola bounded by a chord.

[23]I think that there are truly interesting (entirely mathematical; not historical) issues that lurk in this and it is one of my plans to understand it in depth.



Of course, we know that this means finding a proportion of (the area of) the above kind of figure to some other figure.

This problem is especially illuminating in that Archimedes offers two approaches to it. The method in his treatise *Quadrature of the parabola* is via exhaustion, i.e. approximation by polygons, which is a method, in effect, similar to the one we have already seen in his *Measurement of the circle*. This does prove what he wants. But the more curious method is the one that he himself refers to as a *mechanical method*—a mode of reasoning to which he *does not* give the full authority of proof: it's an example of a *heuristic*[24]—perhaps, chronologically, the first such example labeled as not-quite-a-proof that we have any record of.

A major tool Archimedes will use is his famous "law of the lever"[25] which proclaims that if weights W and w are placed on the plank that is the lever, at opposite sides of

[24]which is fitting since it comes from the pen of the celebrated shouter of "Eureka," which comes from the same root.

[25]There is an extensive earlier tradition of discussion about equilibrium and disequilibrium on a balance; and on the action of levers of all sorts. For example,

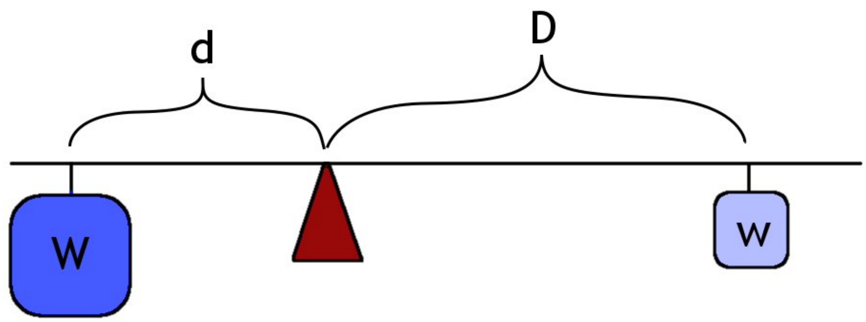
> A small change occurring at the centre makes great and numerous changes at the circumference, just as by shifting the rudder a hair's breadth you get a wide deviation at the prow.

This is from Part 7 of Aristotle's *On the motion of animals*; I want to thank Jean de Groot for conversations about this; I look forward to her forthcoming commentary on Aristotle's Mechanics.

There is also, to be sure, an extensive later tradition of discussion; notably Ernst Mach's marvelous critique of the "law" itself, in the Introduction and first few chapters of his wonderful *The Science of Mechanics*.

the fulcrum but at distances D and d from the fulcrum, respectively, then the lever will balance *if and only if*

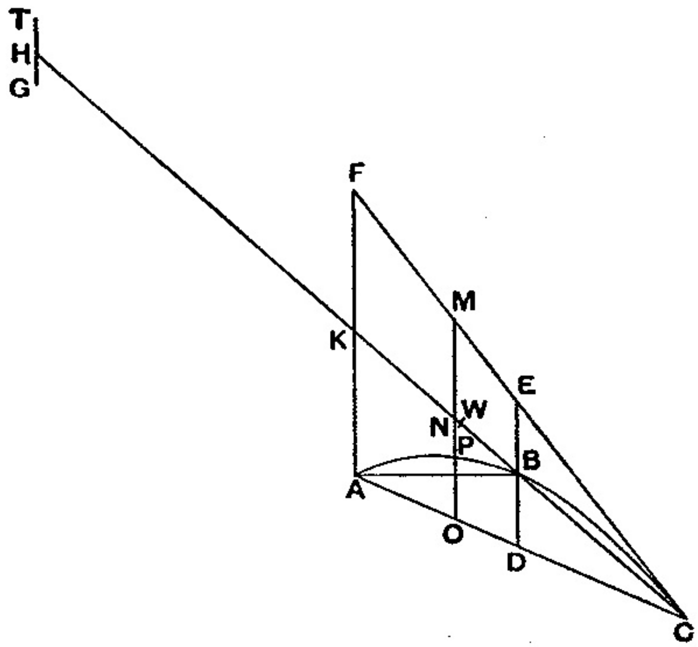
$$D \cdot w \quad = \quad d \cdot W.$$



Now what in the world does this have to do with *area*? you might ask. The answer is that Archimedes is engaged, here, in an ingenious thought-experiment, where the rules of the game are dictated by some basic physical truths, and the link to *area* (he will also treat *volume* problems this way as well) is by a profound analogy: in the figure below, imagine the point K as the fulcrum of a lever. The plank of the lever is the line segment HK. Archimedes will have constructed a triangle FAC deployed onto the plank as shown, and will be "weighing" (yes: weighing) the parabolic segment $\mathcal{P}$ by—in a laminar manner—weighing each line in the parabolic segment parallel to the diameter of the parabola against corresponding lines in the triangle FAC placed at an appropriate distance (at H) on the other side of the fulcrum. Archimedes is thinking that you can view the parabolic segment and triangle as swept through by a continuum of line segments, and the *area* of these figures is somehow distributed as slivers dependent on the varying lengths of these line segments. And he uses his "law of the lever" to find the balance, thereby concluding his heuristic argument.

We can discuss this at greater length in the question period if people want to do that, but here—slightly slower—is a recap of what I've just said, broken up into the steps that Archimedes uses.

In the figure below which is taken from one of the diagrams for Proposition 1 in the traditional text for *The Method*,



the chord is AC and the parabolic arc we are to study is the curve bounded by A and C. We are interested in the area of the parabolic segment (let us call it $\mathcal{P}$). Specifically, $\mathcal{P}$ is that region bounded by that chord AC and the parabolic segment that joins with it. For this task, the figure will give us all the constructions necessary.

(1) **The lever and fulcrum:** We are going to weigh things and balance thing so we need some apparatus. Don't mind that it is on a slant: but the straight line through C and K is going to be our *lever*; and K will be our *fulcrum*.

(2) **The tangent line:** We draw the line CF through C tangent to our parabolic arc at C (I'll say what F is in a moment).

(3) Let D be the bisector of AC and construct a straight line through D *parallel to the diameter of the parabola*. (Parabolas *do* have well defined "diameters": in simple English, if we draw the full parabola, rather than the piece of it as occurs in the above figure, the diameter is that straight line piercing the parabola around which the parabola is symmetric: flipping about the diameter preserves the parabola.) This straight line will intersect CF at a point that we'll call E, and the parabolic arc at a point B. So we can call the line ED.

(4) **The basic triangle:** Draw the lines AB and BC to form the basic triangle ABC (which I'll also call $\mathcal{T}$).

(5) Note that $\mathcal{T}$ sits neatly in the parabolic segment $\mathcal{P}$. Clearly the area of $\mathcal{P}$ is bigger than that of $\mathcal{T}$, but how much bigger? The upshot of this proposition, after Archimedes finishes proving it, is that we get an *exact* relationship. Namely:

$$\mathcal{P} : \mathcal{T} \quad = \quad 4 : 3.$$

(6) **Laminating by lines parallel to the diameter:** The line ED is parallel to the diameter. In the figure above, you find a couple of other labelled lines parallel to the diameter: MO, and FA. What Archimedes wants to do is to think of the family of *all* lines that are parallel to the diameter, and how, as they sweep across the figure, they slice (think of them as forming a moving family). We will refer to any member of that family (and there are finitely many of them!) a *laminar slice.* In a moment we will be slicing two figures by the lines of this family.

(7) **The big triangle:** This is FAC, built with edges the line FA parallel to the diameter and the chord AC. Simple geometry shows that $FAC : \mathcal{T} = 4$. So, thinking of the formula above, we want to prove that:

$$\mathcal{P} : FAC \quad = \quad 1 : 3.$$

(8) **Weighing slices on the balance beam:** Archimides hangs the big triangle FAC from its center of gravity, W, on the balance bar HK as indicated in the figure. He then considers laminar slices of it, comparing them to laminar slices (by the same line parallel to the diameter) of the parabolic segment ABC. He proves that to put each laminar slice of the big triangle FAC in equilibrium with the corresponding slice of the parabolic segment ABC you have to "hang" the laminar slice of the parabolic segment ABC at the point H on the other side of the fulcrum. This uses, of course his law of the lever.

(9) **Weighing the figures themselves:** He then says that he has hung the parabolic segment at point H and the big triangle at point W and, again, the law of the lever gives the proportions of their areas.

There is a great amount of geometry that one can learn by considering this result. First, note that we do indeed have here an example of what I described as a "classical quadrature problem" in that **(a)** we specified each of our figures merely by generic prescriptions (take *any* parabolic, and cut it with *any* chord, etc.) and **(b)** we asserted that the proportions of these figures are given by a fixed rational ratio (4/3). That alone deserves thought.

You might wonder: how many other interesting generic geometric proportions can one come up with that have a fixed *rational* ratio? Or, perhaps, a fixed ratio involving, say, surds?[26]

What is gripping here is how we, using Calculus, could immediately convert into a genuine theorem what Archimedes does with his "method;" and how a dyed-in-the-wool Euclidean could also come to terms with this by offering an appropriate menu of axioms and common notions. Each of these revisions of Archimedes' work—via Calculus, or via appropriate axioms—would have the effect of reframing Archimedes' mechanical analogy by encompassing it with something non-analogical that has, perhaps, the authority to explain more. And yet for me, the lesson offered by *The Method* lies–to return to the issue of purity I mentioned previously in passage (section 7 above)–in the unconstrained *impurity* of the ideas behind it. *The Method* works on the strength of a correctly guiding, nevertheless difficult to justify, analogy that combines previously disparate intuitions that had originated in somewhat different domains: *the experience one has with a certain weighing apparatus* and *the intuition one has via Euclidean geometry.*

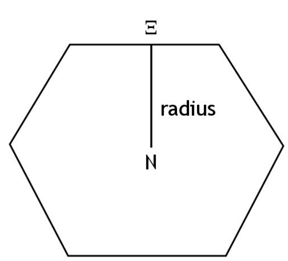
This type of thinking (working with profound analogies and relating them to, or turning them into, equalities) is, today as in Archimedes' time, the source of much of the most powerful mathematics.

19. Hedgehogs again

We have largely talked about areas of figures in the plane, except for our excursion in the spherical geometry with the Banach-Tarsky Paradox. This deserves more discussion, which I hope will happen in question period.

20. **Appendix:** Sketch of a proof of Archimedes' Measurement of the Circle.

To describe Archimedes' argument succinctly, we need some vocabulary. Define the *radius* of a regular polygon to be the length of a line interval that is obtained by dropping a perpendicular to any side of the polygon from the center N of the regular polygon.



[26]i.e., square roots; this is not an idle question.

Define the *perimeter* (or *circumference*) of a polygon to be the length of its perimeter, i.e., the sum of the lengths of the sides of the polygon. If the polygon is a regular M-gon, then the circumference is M times the length of any side.

Here is "my" version of Proposition 1 for regular polygons[27] analogous to Archimedes' Proposition 1 for circles:

> **Archimedes' Proposition 1,** *but* **for regular polygons:** The area subsumed by a regular polygon is equal to the area subsumed by a right-angled triangle for which the two right-angle sides are of lengths equal to the *radius* and the *circumference* (respectively) of the polygon.

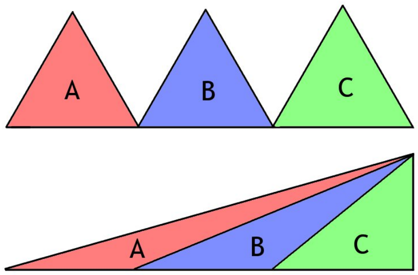
In contrast to the actual Proposition 1 (of *Measurement of the Circle*), this "polygon-version" of Archimedes' Proposition 1 is now nicely within the scope of Euclidean vocabulary; its proof is within the scope of Euclid as well.

Some Comments:

(1) Both this "polygon-version" and Archimedes Proposition 1 deal with a *right-angle* triangle (with base the circumference and altitude (synonymously: *height*) the radius of the figure this triangle is being compared to). One could rephrase these propositions by omitting the requirement that the triangle be a right-angle triangle.
(2) A visual proof of this polygonal proposition can be effected by simply cutting and "straightening out to a line" the perimeter of the polygon, and then arguing that this paper-doll figure has the same area as the triangle displayed below[28]. (In the figure below we illustrate this with a 3-gon, alias a triangle, which when cut-and straightened-out produce the three triangles in a line labeled A,B,C; these each have the same area as the three triangles that make up the large triangle below having as base the perimeter and as height the radius.) All this uses is that the area of a triangle depends *only* on its base and height.

[27] I say "my" version because, even though it is—in my opinion—implicitly invoked, in Archimedes' text, it isn't dwelt on.

[28] I'm thankful to Jim Carlson for this suggestion.


A
B
C
A
B
C