Section 1.8. Three Famous Problems of Greek Geometry

Note. In this section we consider the idea of compass and straightedge constructions and how they are intimately related to the approach to geometry taken by Euclid. We state the "Three Famous Problems" and give an explanation as to why the problems cannot be solved with the use of a compass and straightedge alone.

Note. Benno Artman states in *Euclid—The Creation of Mathematics* (Springer Verlag, 1999) that: "Euclids emphasis is more on construction than on 'existence,' more a difference in style than in substance." (See page 19.) These constructions are performed using the *Euclidean tools* of the straightedge and compass. With the straightedge we can draw a straight line of indefinite length through any given distinct points (the straightedge is unmarked; it does not act as a ruler). With the compass we can draw a circle with any given point as its center and passing through any given second point. These tools are the basis for the proofs given in Euclid's *Elements*. For example, we'll see in Section 2.1. Book I that Euclid I.1 states: "On a given finite straight line AB to construct an equilateral triangle." That is, an equilateral triangle is to be constructed with line segment AB as one of its sides. The construction is performed by drawing a circle with center A that contains point B, and drawing a circle with center B that contains point A. The resulting circles intersect at two points, either of which can be used as the third vertex of the equilateral triangle (see Figure 1.8.A). We quickly comment that Euclid is making some unstated continuity assumptions here (and elsewhere in the

Elements); for details, see my online notes based on C. R. Wylie's *Foundations of Geometry* (1964) on Section 2.2. A Brief Critique of Euclid.

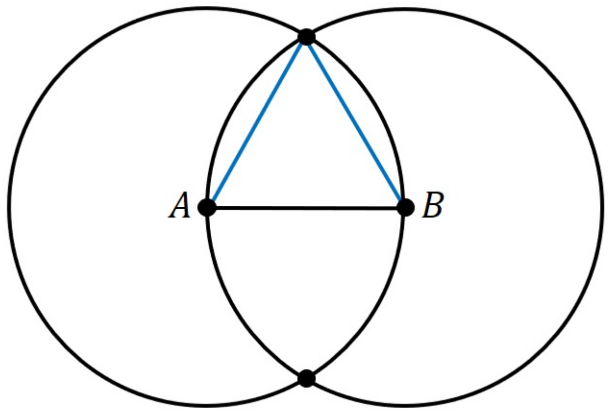


Figure 1.8.A. The construction of an equilateral triangle given in Book I of the *Elements*

Note. Artman also comments on page 19 of his book: "Many authors have noted the incompleteness of Euclids axioms in comparison to modern foundations of geometry. The most obvious point is the absence of any thought of the ordering of points on a line or the concept of betweenness. " A related observation is that Euclid lacks an idea of continuity. W.M. Strong in "Is Continuity of Space Necessary to Euclid's Geometry?" *Bulletin of the American Mathematical Society*, **4**(9), 443–448 (June 1898) (a copy can be downloaded from projecteuclid.org) discusses what he calls *quadratic space*. This space consists of all points in the Cartesian plane which have quadratic coordinates (a real number is *quadratic* if it can be obtained from the integers by a finite number of rational operations and extractions of square roots). The quadratic space is everywhere discontinuous, yet any construction that can be performed with a compass and a straightedge can be performed in this space! So in response to Strong's question in the title of his paper, "No!" We are about to explore the "Three Famous Problems" and we will see that these problems are not solvable since they involve points that are not in the quadratic space. Our explanation will require a discussion of the field of constructible numbers.

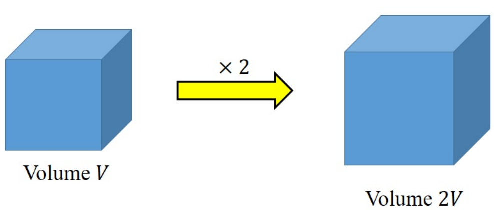
Note. The "Three Famous Problems" are compass and straightedge constructions. They are (taking the statements as given in Introduction to Modern Algebra 2 [MATH 4137/5137], see my online notes on Section VI.32. Geometric Constructions):

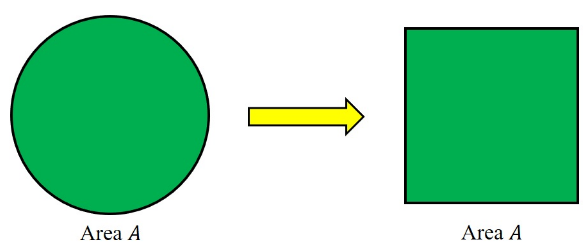
1. Doubling the Cube: For a cube of a given size (i.e., given the length of a side), construct a cube of twice the volume of the given cube.

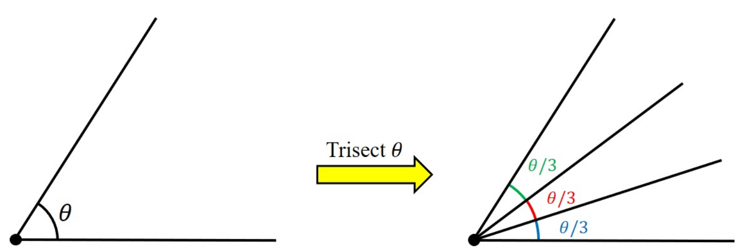
2. Squaring the Circle: For a given circle (i.e., given the diameter of the circle), construct a square with the same area as the circle.

3. Trisect an Angle: Given an angle, find an angle 1/3 the size of the given angle.

These are illustrated in the following three figures:







Note. A cube with a side of given length ℓ has volume ℓ^3, so a cube with volume $2\ell^3$ must have a side of length $\sqrt[3]{2}\ell$. So constructing a cube of twice the volume of a given cube is equivalent to constructing a line segment of length $\sqrt[3]{2}$ (we will comment below that if a segment of length a and a segment of length b are constructible, then a segment of length ab is constructible). We'll see in Chapter 3. Conic Sections that $\sqrt[3]{2}$ can be constructed by considering the intersection of a parabola and hyperbola (see Note 3.A). The Conchoid of Nicomedes (circa 280 BCE–circa 210 BCE), which we explore in Section 4.1. The Conchoid of Nicomedes, The Trisection of an Angle, can also be used to double the cube (see Exercise 6.10.2). For a circle of given radius r, the area is πr^2. So a square of the same area must have sides of length $\sqrt{\pi}r$. So squaring the circle is equivalent to constructing a line segment of length $\sqrt{\pi}$.

Note. We deal with the trisection of an angle somewhat differently. Some angles can be trisected. For example, a right angle can be trisected since a a 30° can be constructed (since an equilateral triangle can be constructed by Euclid I.1 and an angle can be bisected by Euclid I.9). We will argue below that a 60° angle cannot be trisected, because a 20° angle cannot be constructed. Therefore, *in general*, it follows that an angle cannot be trisected with a compass and straightedge. However, again using the Conchoid of Nicomedes, every angle can be trisected (see Section 4.1. The Conchoid of Nicomedes, The Trisection of an Angle).

Note. Squaring the circle is a natural problem to consider in connection with the question of the area of a circle. Of course we know that the area of a circle of radius r is $A = \pi r^2$. First, we need a definition of π. It is defined as the ratio of the circumference C of a circle to its diameter $d = 2r$. In this way we have $\pi = C/d$ or $C = \pi d = 2\pi r$. Now this requires that the ratio of the circumference of a circle to its diameter actually *is* a constant. This result is not contained in Euclid, but is implicit in Archimedes' *Measurement of a Circle* in his Proposition 3 in which he proves "The ratio of the circumference of any circle to its diameter is less than $3\frac{1}{7}$ but greater then $3\frac{10}{71}$." His proof is based on estimating the perimeters of 96 sided polygons inscribed in and circumscribed around a circle. Euclid XII.2 states that "Circles are to one another as the squares on the diameters." That is, the area of a circle is proportional to the square of the diameter. It then follows from Archimedes Proposition 1 in *Measurement of a Circle* that the constant of proportionality between the circumference and diameter is the same as the constant of proportionality between the area and the square of the diameter. Archimedes' Proposition 1 states: "The area of any circle is equal to a right-angled triangle in which one of the sides about the right angle is equal to the radius, and the other to the circumference of the circle." The area of the triangle in Proposition 1 is $A = \frac{1}{2}(2\pi r)(r) = \pi r^2$. These are the results that tie together the definition of π to the area of a circle.

Note. We comment in passing, that Archimedes proves his Proposition 1 by the method of exhaustion in which he assumes that the area of the circle is less than πr^2, say it is $\pi r^2 - \varepsilon$ for some given $\varepsilon > 0$. He gets a contradiction by finding a subset of the circle that has an area greater than $\pi r^2 - \varepsilon$. He then similarly supposes the area of the circle is greater than πr^2, say it is $\pi r^2 + \varepsilon$ for some $\varepsilon > 0$. He gets a contradiction by finding a superset of the circle that has an area less then $\pi r^2 + \varepsilon$. These two contradictions combine to give that the area of the circle is πr^2. For more details, see my online PowerPoint presentation on Archimedes: 2000 Years Ahead of His Time and the transcript in PDF.

Note. The text book states that the three famous problems "appeared during the pre-Euclidean period and occupied the Greek geometers for at least three centuries." (See page 19.) The additional tools (in the form of new geometric curves) developed to solve these problems were motivation for much of the mathematics to follow. Though not actually developed with these problems in mind, the area of field theory in algebra ultimately is the tool allowing us to show that the three famous problems cannot be solved with the tools of a compass and straightedge. This is covered in our Introduction to Modern Algebra 2 (MATH 4137/5137) in Section VI.32. Geometric Constructions and our Modern Algebra 2 (MATH 5420) in Section V.1.Appendix. Ruler and Compass Constructions. The rational numbers are constructible as seen in Section 1.2. Similar Figures (see Figure 1.6). It can be shown that the set of real constructible numbers C forms a subfield of the field of real numbers (see Corollary 32.5 in the Introduction to Modern Algebra 2 notes) and, in particular, the field of constructible real numbers C consists precisely of all real numbers that we can obtain from $\mathbb{Q}$ by taking square roots of positive numbers a finite number of times and applying a finite number of field operations (see Theorem 32.6 in the Introduction to Modern Algebra 2 notes and Proposition V.1.16 of Modern Algebra 2). Additional details on constructions can be found in my video "Compass Straightedge Constructions" on YouTube.

Note. The three famous problems are then shown to be impossible with a straightedge and compass since each requires the use of a real number that is not constructible. The doubling of the cube requires the construction of $\sqrt[3]{2}$, which is not constructible since it requires a cube root. A formal proof is given in Theorem 32.9 in Introduction to Modern Algebra 2 and in Corollary V.1.18 in Modern Algebra 2. The squaring of the circle requires the construction of $\sqrt{\pi}$, which is not constructible since π is transcendental (that is π is not algebraic, but all constructible numbers are algebraic). A formal proof is given in Theorem 32.10 in Introduction to Modern Algebra 2 and in Corollary V.1.19 in Modern Algebra 2. The trisection of an angle is addressed in Theorem 32.11 in Introduction to Modern Algebra 2 and in Corollary V.1.17 in Modern Algebra 2. This is proved by showing that a 20° angle cannot be constructed because $\cos(20^\circ)$ has degree three over the rationals (that is, it is the root of a third degree polynomial over $\mathbb{Q}$ and is not a root of a first or second degree polynomial over $\mathbb{Q}$).

Note. Historically, it was Pierre Wantzel in 1837 who first showed that trisecting an angle and doubling the cube are impossible in “Recherches sur les moyens de reconnaître si un Problème de Géométrie peut se résoudre avec la règle et le compas” in *Journal de Mathématiques Pures et Appliquéees* **1**(2), 366–372. In 1882, Ferdinand Lindemann proved that π is transcendental in “Über die Zahl π,” *Mathematische Annalen* **20**, 213-225 (1882), from which the impossibility of squaring the circle follows. See the historical note on page 298 of John B. Fraleigh’s *A First Course In Abstract Algebra* 7th Edition, Pearson (2002).

Note. Hippocrates of Chios (circa 470 BCE–circa 410 BCE; not to be confused with Hippocrates of Cos, the physician), wrote the first geometry textbook. He introduced "proof by contradiction," also called *reductio ad absurdum* (according to Jason Socrates Bardi's *The Fifth Postulate: How Unraveling a Two-Thousand-Year-Old Mystery Unraveled the Universe*, John Wiley & Sons: 2009; see page 42). He studied the classical construction problems of squaring the circle and duplicating the cube. For the first time, he gave the area of a round figure by finding the area of a lune in terms of the area of a related triangle. This is illustrated in Figure 1.23(c), where the areas satisfy $F = F_a + F_b$. This conclusion is based on the Pythagorean Theorem. With the hypotenuse of the triangle in Figure 1.23(c) as c, the left leg as a, and the right leg as b, we have $c^2 = a^2 + b^2$. Next we consider the areas of the semicircles with diameters given by a, b, and c. The areas are $\pi a^2/2$, $\pi b^2/2$, and πc^2. By the Pythagorean Theorem, $\pi a^2/2 + \pi b^2/2 = \pi c^2/2$. Now the two lunes and the right triangle have areas satisfying $F_a + F_b = \pi a^2/2 + \pi b^2/2 + F - \pi c^2/2$ or $F_a + F_b = F$, as claimed. These are called the "lunes of Hippocrates."

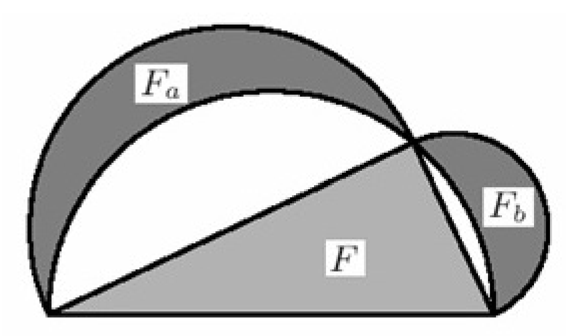
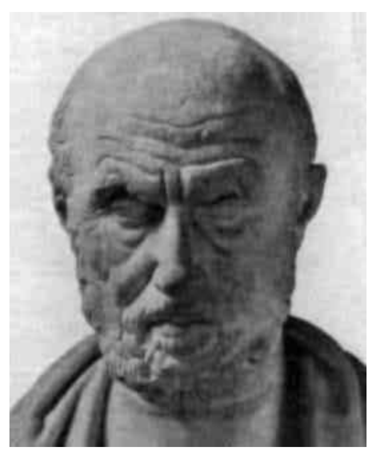


Figure 1.23(c)



From the MacTutor History of Math webpage

Revised: 12/5/2021