4. Alexandrian mathematics after Euclid — II

Due to the length of this unit, it has been split into three parts.

Apollonius of Perga

If one initiates a Google search of the Internet for the name "Apollonius," it becomes clear very quickly that many important contributors to Greek knowledge shared this name (reportedtly there are 193 different persons of this name cited in Pauly-Wisowa, *Real-Enzyklopädie der klassischen Altertumswissenschaft*), and therefore one must pay particular attention to the full name in this case.

Apollonius of Perga made numerous contributions to mathematics. As is usual for the period, many of his writings are now lost, but it is clear that his single most important achievement was an eight book work on conic sections.

Background discussion of conics

We know that students from Plato's Academy began studying conics during the fourth century B.C.E., and one early achievement in the area was the use of intersecting parabolas by Manaechmus (380 – 320 B.C.E., the brother of Dinostratus, who was mentioned in an earlier unit) to duplicate a cube (recall Exercise 4 on page 123 of Burton). One of Euclid's lost works was devoted to conics, and at least one other early text for the subject was written by Aristaeus (360 – 300 B.C.E.). Apollonius' work, ***On conics***, begins with an organized summary of earlier work, fills in numerous points apparently left open by his predecessors, and ultimately treat entirely new classes of problems in an extremely original, effective and thorough manner. In several respects the work of Apollonius anticipates the development in coordinate geometry and uses of the latter with calculus to study highly detailed properties of plane curves. Of the eight books on conics that Apollonius wrote, the first four have survived in Greek, while Books V through VII only survived in Arabic translations due to Thabit ibn Qurra and Book VIII is lost (there have been attempts to reconstruct the latter based upon commentaries of other Greek mathematicians). Available evidence suggests that the names ***ellipse***, ***parabola*** and ***hyperbola*** are all due to Apollonius, but there the opinions of the experts on this are not unanimous.

Today we think of conics in the coordinate plane as curves defined by quadratic equations in two variables. As the name indicates, ancient Greek mathematicians viewed such curves as the common points of a cone and a plane. For the sake of completeness, we shall include a summary of this relationship taken from the following two sources:

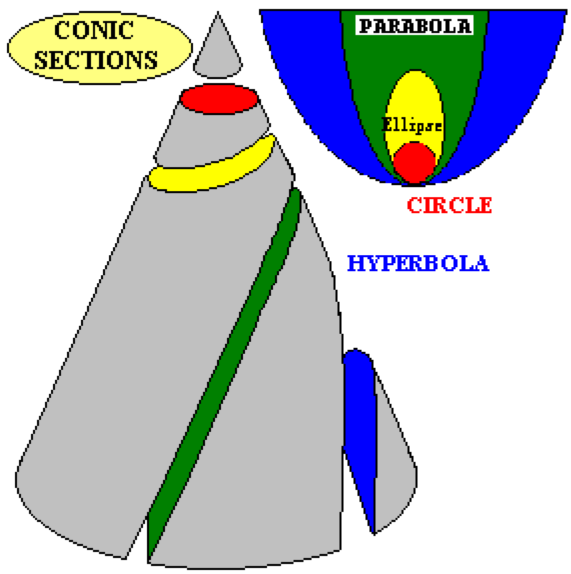
http://www.chemistrydaily.com/chemistry/Conic_section

http://en.wikipedia.org/wiki/Conic_section

Since we are including a fairly long discussion of conics, it seems worthwhile to include a link to a site describing how conics arise in nature and elsewhere:

http://ccins.camosun.bc.ca/~jbritton/jbconics.htm

In the figure below there is a picture from the above references illustrating how the different conic sections are formed by intersecting a cone with a plane.



(The material below is taken from the sites listed above.)

Types of conics

Two well-known conics are the circle and the ellipse. These arise when the intersection of cone and plane is a closed curve. The circle is a special case of the ellipse in which the plane is perpendicular to the axis of the cone. If the plane is parallel to a generator line of the cone, the conic is called a parabola. Finally, if the intersection is an open curve and the plane is not parallel to a generator line of the cone, the figure is a hyperbola. (In this case the plane will intersect *both* halves of the cone, producing two separate curves, though often one is ignored.)

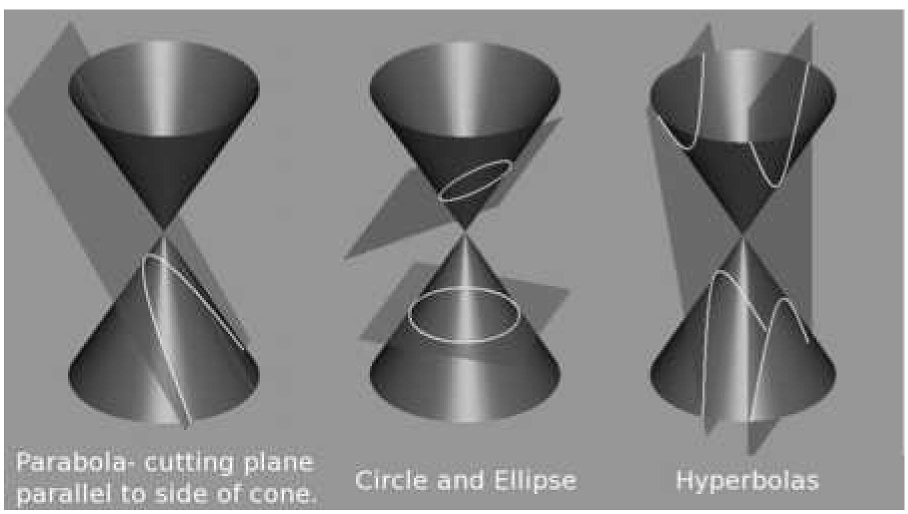
The degenerate cases, where the plane passes through the apex of the cone, resulting in an intersection figure of a point, a straight line or a pair of lines, are often excluded from the list of conic sections.

In Cartesian coordinates, the graph of a quadratic equation in two variables is always a conic section, and all conic sections arise in this way. If the equation is of the form

$$ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0$$

then we can classify the conics using the coefficients as follows:

- if $h^2 = ab$, the equation represents a parabola;
- if $h^2 < ab$, the equation represents an ellipse;
- if $h^2 > ab$, the equation represents a hyperbola;
- if $a = b$ and $h = 0$, the equation represents a circle;
- if $a + b = 0$, the equation represents a rectangular hyperbola.



Graphic visualizations of the conic sections

Eccentricity

An alternative definition of conic sections starts with a point *F* (the *focus*), a line *L* not containing *F* (the *directrix*) and a positive number *e* (the *eccentricity*). The corresponding conic section consists of all points whose distance to *F* equals *e* times their distance to *L*. For 0 < *e* < 1 we obtain an ellipse, for *e* = 1 a parabola, and for *e* > 1 a hyperbola.

For an ellipse and a hyperbola, two focus-directrix combinations can be taken, each giving the same full ellipse or hyperbola. The distance from the center to the directrix is *a* / *e*, where *a* is the semi-major axis of the ellipse, or the distance from the center to the tops of the hyperbola.

In the case of a circle *e* = 0 and one imagines the directrix to be infinitely far removed from the center. However, the statement that the circle consists of all points whose distance is e times the distance to L is not useful, because we get zero times infinity.

The eccentricity of a conic section is thus a measure of how far it deviates from being circular.

For a given *a*, the closer *e* is to 1, the smaller is the semi-minor axis.

Derivation

Let there be a cone whose axis is the *z*-axis. Let its vertex be the origin. The equation for the cone is

$$x^2 + y^2 - a^2 z^2 = 0 \qquad (1)$$

where

$$a = \tan\theta > 0$$

and θ is the angle which the generators of the cone make with respect to the axis. Notice that this cone is actually a pair of cones: one cone standing upside down on the vertex of the other cone – or, as mathematicians say, this cone consists of two "nappes."

Let there be a plane with a slope running along the *x* direction but which is level along the *y* direction. Its equation is

$$z = mx + b \qquad (2)$$

where

$$m = \tan\phi > 0$$

and ϕ is the angle of the plane with respect to the *xy*-plane.

We are interested in finding the intersection of the cone and the plane, which means that equations (1) and (2) shall be combined. Both equations can be solved for *z* and then equate the two values of *z*. Solving equation (1) for *z* yields

$$z = \sqrt{\frac{x^2 + y^2}{a^2}}$$

and therefore

$$\sqrt{\frac{x^2 + y^2}{a^2}} = mx + b.$$

Square both sides and expand the squared binomial on the right side,

$$\frac{x^2 + y^2}{a^2} = m^2 x^2 + 2mbx + b^2.$$

Grouping by variables yields

$$x^2 \left(\frac{1}{a^2} - m^2 \right) + \frac{y^2}{a^2} - 2mbx - b^2 = 0. \qquad (3)$$

Note that this is the equation of the projection of the conic section on the *xy*-plane, hence contracted in the *x*-direction compared with the shape of the conic section itself.

Derivation of the parabola

The parabola is obtained when the slope of the plane is equal to the slope of the generators of the cone. When these two slopes are equal, then the angles θ and ϕ become complementary. This implies that

$$\tan\theta = \cot\phi$$

and therefore

$$m = \frac{1}{a} \qquad (4)$$

Substituting equation (4) into equation (3) makes the first term in equation (3) vanish, and the remaining equation is

$$\frac{y^2}{a^2} - \frac{2}{a}bx - b^2 = 0.$$

Multiply both sides by a^2,

$$y^2 - 2abx - a^2b^2 = 0$$

then solve for *x*,

$$x = \frac{1}{2ab}y^2 - \frac{ab}{2}. \qquad (5)$$

Equation (5) describes a parabola whose axis is parallel to the *x*-axis. Other versions of equation (5) can be obtained by rotating the plane around the *z*-axis.

Derivation of the ellipse

An ellipse happens when the angles θ and ϕ, when added together, do not measure up to a right angle:

$$\theta + \phi < \frac{\pi}{2} \qquad \text{(ellipse)}$$

which implies that the tangent of the sum of these two angles is positive.

$$\tan(\theta + \phi) > 0.$$

But a trigonometric identity states that

$$\tan(\theta + \phi) = \frac{\tan\theta + \tan\phi}{1 - \tan\theta\tan\phi}$$

therefore

$$\tan(\theta + \phi) = \frac{m + a}{1 - ma} > 0 \qquad (6)$$

but *m + a* is positive, since the summands are given to be positive, so inequality (6) is positive if the denominator is also positive:

$$1 - ma > 0. \qquad (7)$$

From inequality (7) we can deduce

$$\begin{aligned} ma &< 1, \\ m^2 a^2 &< 1, \\ 1 - m^2 a^2 &> 0, \\ \frac{1}{m^2 a^2} &> 1, \\ \frac{1}{m^2 a^2} - 1 &> 0, \\ \frac{1}{a^2} - m^2 &> 0 \qquad \text{(ellipse)}. \end{aligned}$$

Let us start out again from equation (3),

$$x^2 \left(\frac{1}{a^2} - m^2 \right) + \frac{y^2}{a^2} - 2mbx - b^2 = 0, \qquad (3)$$

but this time the coefficient of the x^2 term does not vanish but is instead positive. Solve for *y*,

$$y = a\sqrt{b^2 + 2mbx - x^2 \left(\frac{1}{a^2} - m^2 \right)} \qquad (8)$$

This would clearly describe an ellipse were it not for the second term under the radical, the *2 m b x*: it would be the equation of a circle which has been stretched proportionally along the directions of the *x*-axis and the *y*-axis. Equation (8) is an ellipse but it is not obvious, so it will be rearranged further until it is obvious. Complete the square under the radical,

$$y = a\sqrt{b^2 - \left[x\sqrt{\frac{1}{a^2} - m^2} - \frac{b}{\sqrt{\frac{1}{a^2m^2} - 1}}\right]^2 + \left(\frac{b^2}{\frac{1}{a^2m^2} - 1}\right)}.$$

Group together the b^2 terms,

$$y = a\sqrt{b^2\left(1 + \frac{1}{\frac{1}{a^2m^2} - 1}\right) - \left[x\sqrt{\frac{1}{a^2} - m^2} - \frac{b}{\sqrt{\frac{1}{a^2m^2} - 1}}\right]^2}.$$

Divide by *a* then square both sides,

$$\frac{y^2}{a^2} + \left(x\sqrt{\frac{1}{a^2} - m^2} - \frac{b}{\sqrt{\frac{1}{a^2m^2} - 1}}\right)^2 = b^2\left(1 + \frac{1}{\frac{1}{a^2m^2} - 1}\right)$$

The *x* has a coefficient. It is desired to pull this coefficient out by factoring it out of the second term which is a square,

$$\frac{y^2}{a^2} + \left(\frac{1}{a^2} - m^2\right)\left(x - \frac{b}{\sqrt{\left(\frac{1}{a^2m^2} - 1\right)\left(\frac{1}{a^2} - m^2\right)}}\right)^2 = b^2\left(1 + \frac{1}{\frac{1}{a^2m^2} - 1}\right)$$

Further rearrangement of constants finally leads to

$$\frac{y^2}{1 - a^2m^2} + \left(x - \frac{mb}{\frac{1}{a^2} - m^2}\right)^2 = \frac{a^2b^2}{(1 - a^2m^2)^2}.$$

The coefficient of the *y* term is positive (for an ellipse). Renaming of coefficients and constants leads to

$$\frac{y^2}{A} + (x - C)^2 = R^2 \qquad (9)$$

which is clearly the equation of an ellipse. That is, equation (9) describes a circle of radius *R* and center **(***C* , 0**)** which is then stretched vertically by a factor of **sqrt (***A***).** The second term on the left side (the *x* term) has no coefficient but is a square, so that it must be positive. The radius is a product of squares, so it must also be positive. The first term on the left side (the *y* term) has a coefficient which is positive, so the equation describes an ellipse.

Derivation of the hyperbola

The hyperbola happens when the angles θ and ϕ add up to an obtuse angle, which is greater than a right angle. The tangent of an obtuse angle is negative. All the inequalities which were valid for the ellipse become reversed. Therefore

$$1 - a^2m^2 < 0 \qquad \text{(hyperbola).}$$

Otherwise the equation for the hyperbola is the same as equation (9) for the ellipse, except that the coefficient *A* of the *y* term is negative.

Outline of Apollonius' books ***On Conics***

The first four books give a systematic account of the main results on conics that were known to earlier mathematicians, with some improvements due to Apollonius himself, particularly in Books III and IV; in fact, the majority of results in the latter were new. One distinguishing property of noncircular conics is that they determine a pair of mutually perpendicular lines that are called the ***major*** and ***minor axes***. For example, in an ellipse the major axis marks the direction in which the curve has the greatest width, and the minor axis marks the direction in which the curve has the least width. Apollonius analyzes these axes extensively throughout his work. Here are a few basic points covered in his first four books.

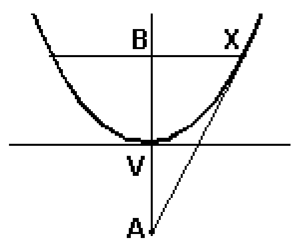
1. Tangents to conics are defined, but not systematically in analytic geometry and calculus. Instead, tangents were viewed as lines that met the conic (or branch of the conic for hyperbolas) in one point such that all other points of the conic or its branch lie on the same side of that line.

2. Asymptotes to hyperbola were defined and studied.

3. Conics were described in terms of (Greek versions of) algebraic second degree equations involving the lengths of certain line segments. Several different characterizations of this sort were given. In many cases these results are forerunners of the algebraic equations that are now employed to describe conics.

4. The intersection of two conics was shown to consist of at most four points.

Here is a typical result from the early books: *Suppose we are given a parabola and a point* **X** *on that parabola that is not the vertex* **V**. *Let* **B** *be the foot of the perpendicular from* **X** *to the parabola's axis of symmetry, and let* **A** *be the point where the tangent line at* **X** *meets the axis of symmetry. Then the distances* **|AV|** *and* **|BV|** *are equal.*



A proof of this result using modern methods is given in an addendum to this unit; in principle, this is equivalent to saying that the derivative of x^2 is 2x.

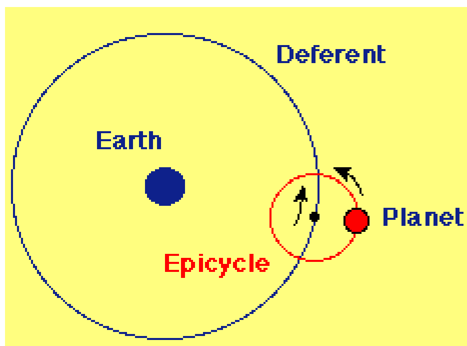
Books V – VII of ***On Conics*** are highly original. In Book V, Apollonius considers normal lines a conic; these are lines containing a point on the conic that are perpendicular to the tangents at the point of contact. As in calculus, Apollonius' study of such perpendiculars uses the fact that they give the shortest distances from an external point to the curve. Book V also discusses how many normals can be drawn from particular points, finds their intersections with the conics by construction, and studies the curvature properties in remarkable depth. In particular, at each point there is a center of curvature, which yields the best circular approximation to the conic at that point, and Apollonius' results on finding the center of curvature resolve this issue completely. A main objective of Book VI is to show that the three basic types of conics are geometrically dissimilar in roughly the same way that, say, a triangle and a rectangle are dissimilar. The final portion of the work contains (or is reputed to contain) further results involving major and minor axes and their intersections with the conics.

Apollonius is particularly known for posing the following general problem (sometimes called the *Problem of Apollonius*): Given three geometric figures, each of which may be a point, straight line, or circle, construct a circle tangent to the three; the most difficult case arises when the three given figures are circles.

Other works of Apollonius

Most of Apollonius' other works are lost, but we have some information about this work from the writings of others. In a computational work called *Quick Delivery* he gave an estimate of 3.1416 for π that was better than the more commonly used Archimedean estimate of 22/7. We shall only mention two other items on the list.

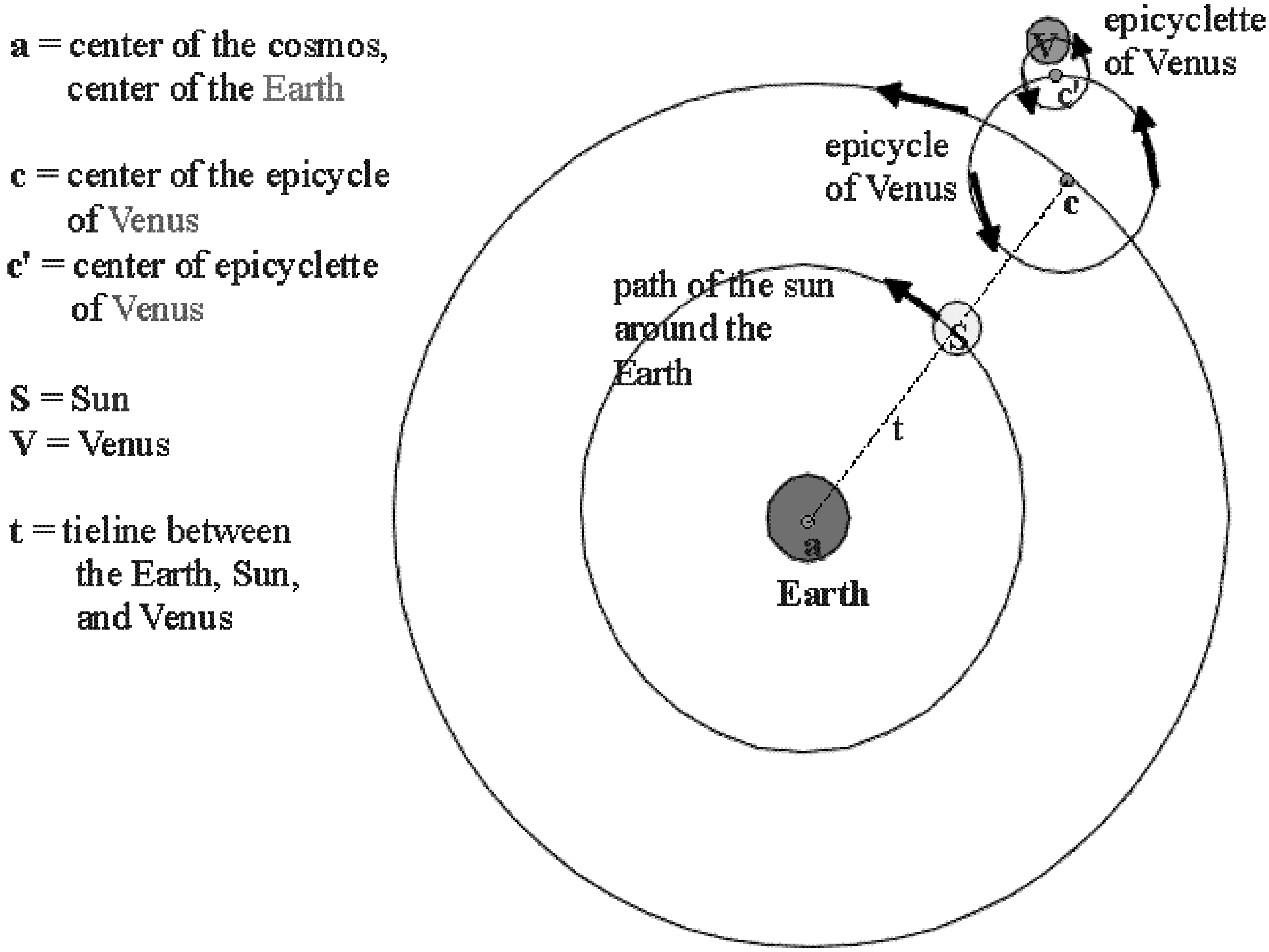
The first involves his work on mathematical astronomy. His view of the solar system was that the sun rotated around the earth but the remaining planets rotated around the sun. By his time astronomical observations showed beyond all doubt that the planets did not move around the earth in perfectly circular orbits, and Apollonius devised the explanation that became the foundation of Claudius Ptolemy's later work. One major feature was a theory that the planets moved in combinations of circles that are called *epicycle*s. The idea is similar to our concept of the Moon's motion around the earth; namely the moon moves around the earth in an ellipse while the earth in turn moves around the sun in another ellipse. In Apollonius' setting the curves were circles rather than ellipses and there was no actual mass corresponding to the center of the smaller circle. Here is a simple illustration of epicycles:



(Source: http://www.cartage.org.lb/en/themes/Sciences/Astronomy/TheUniverse/Oldastronomy/TheUniverseofAristotle/TheUniverseofAristotle.htm)

There are also some animated graphics on the site from which this picture was taken.

A more elaborate illustration of this motion model is illustrated below; in this example, there is in fact a second epicycle moving around the first one. The Ptolemaic theory of planetary motion required dozens of higher order epicycles.



(***Source:*** http://inst.santafe.cc.fl.us/~jbieber/HS/ptol_epi.htm)

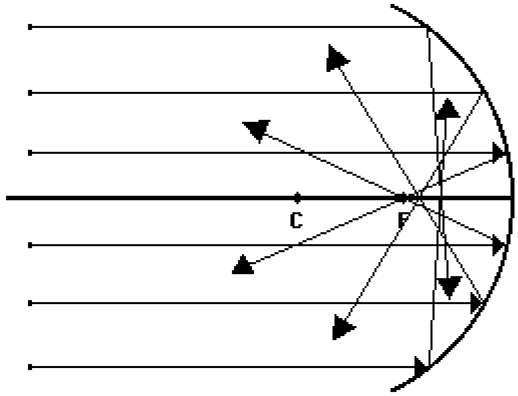
Given Kepler's discovery that planets move in elliptical paths around the sun, it is somewhat ironic that the author of the definitive work on conic sections proposed motion by epicycles instead, but that is what happened. Here are a few other links to pictures of epicycles, some with animation:

http://www.opencourse.info/astronomy/introduction/05.motion_planets/

http://www.math.tamu.edu/~dallen/masters/Greek/epicycle.gif

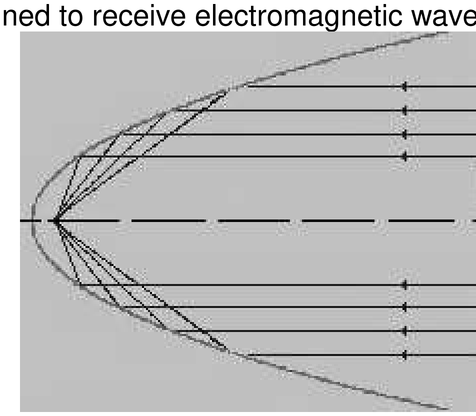
http://www.edumedia.fr/animation-Epicycle-En.html

Apollonius and his contemporary Diocles (240 – 180 B.C.E.) are also given credit for discovering the reflection property of the parabola. The Greeks knew that one could start a fire by focusing the sun's rays using a convex mirror; stories that Archimedes used large mirrors of this sort to set fire to Roman ships are almost certainly incorrect, but the idea was known at the time. The simplest concave mirrors are shaped like a portion of a sphere. However, these do not have a true focus but suffer from a phenomenon called *spherical aberration*.



(*Source:* http://www.glenbrook.k12.il.us/gbssci/phys/Class/refln/u13l3g.html)

The failure of the reflected rays to go through a single point means that a spherical mirror is somewhat inefficient in focusing the sun's rays (or any other rays for that matter), but if one uses a parabolic mirror this problem is eliminated. All incoming light rays parallel to the axis of symmetry will then be reflected to the focus of the parabola. This property of the parabola is used extensively for devices like antennas and radio telescopes that are designed to receive electromagnetic waves.



(Source: http://library.thinkquest.org/23805/math1.htm)

Proving the reflection property of a parabola is basically an exercise in geometry, and the following online site contains a proof using methods from elementary geometry:

http://www.pen.k12.va.us/Div/Winchester/jhhs/math/lessons/calculus/parabref.html

Needless to say, one can also derive the reflection property for a parabola using coordinate geometry.