The Mathematics of Planetary Orbits

Moti Ben-Ari

http://www.weizmann.ac.il/sci-tea/benari/

Version 0.9

August 25, 2023

Contents

1 Introduction

Everyone “knows” that Kepler discovered that the orbits of the planets are ellipses and that Newton showed that an elliptical orbit implies that the force of gravity must be inversely proportional to the square of the distance from the Sun. Although I knew these facts, I had never seen them demonstrated.

Calculus in Context [4] by Alexander J. Hahn is a comprehensive textbook on introductory calculus that augments theory with applications in physics and astronomy, such as the work of Kepler, Newton and Galileo, as well as applications in engineering such as building bridges and domed structures. These are not just historical anecdotes but detailed computations.

This document works out one topic based on the presentation in the book: the determination of orbits by Aristarchus, Copernicus, Kepler and Newton. The presentation is mathematical, since the historical and astronomical aspects are thoroughly described in [4], as well as in other works. The document is intended to enrich the learning of mathematics by secondary-school students and students in introductory university courses.

Section 2 presents the measurements of the radii and distances of the Earth, the Moon and the Sun by Eratosthenes and Aristarchus. Section 3 describes the construction of a model of a Sun-centered system by Copernicus. Section 4 shows how Kepler developed his three laws of planetary motion and Section 5 presents Newton’s derivation of the inverse-square law of gravitation from of Kepler’s laws. One step of Newton’s derivation requires a theorem whose proof is very long, so it is split off into Section 6. Appendix 7 contains more than you ever wanted to know about the mathematics of ellipses, but these theorems are necessary. I suggest that you look up each theorem (and its proof) as needed, rather than trying to study them all at once.

The computations in Hahn’s book are faithful to the historical record, for example, measuring distances in units such as the stadia of the Greeks. Here, the computations are fully modernized and use modern units such as kilometers.

It is easy to measure angles. We are familiar with the use of a protractor[1] in school and these can be scaled-up to obtain more accurate measurements. Measuring long distances was impossible until the recent inventions of radar and lasers. At most one could pace-off distances with low accuracy. The only *measured* distance used here is the estimate of 800 km by Eratosthenes for the distance between two places in Egypt (Section 2.1).

This documents uses many diagrams to facilitate understanding each step in the geometrical proofs, more than appear in other sources. For example, Newton proved his difficult theorem (Section 6) using only a single diagram. The diagrams are not to scale and are often distorted. Otherwise, it would be impossible to draw the Earth, Moon and Sun on a piece of paper, since the Sun is almost 400 times farther away from the Earth than the Moon. If we drew the Moon one centimeter from the Earth, we would have to draw the Sun four meters away!

[1] More precisely, a device for measuring angles is called a *goniometer*.

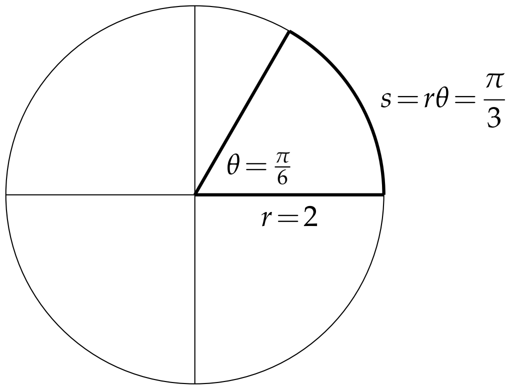


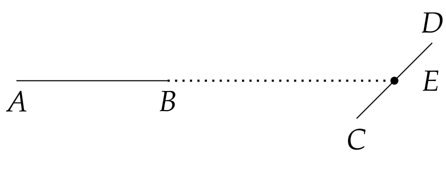
Figure 1: Angles and arcs

The prerequisites for reading the document are a good knowledge of Euclidean geometry, some trigonometry, a bit calculus and Newton's laws of motions. In particular you must know that if a sector of a circle subtends an angle of θ radians, the length of the sector is $r\theta$, where r is the radius of the circle (Figure 1).

Notation

The following shortcuts facilitate a less verbose presentation:

- a and b denote the semi-major and semi-minor axes of an ellipse.
- AB denotes both a line segment and its length.
- $\triangle ABC$ denotes both a triangle and its area.
- The diagram below shows that when the line segment AB is *extended* it intersects the line CD at a point E. The rather archaic term for extended is *produced*. When the text says that AB intersects CD, the intention is that AB can produced or extended until it intersects CD.



Acknowledgments

I am grateful to Alex Hahn and Graham Griffiths for their guidance.

2 The sizes of the Earth, Moon and Sun

2.1 Eratosthenes' measurement of the radius of the earth

The ancient Greeks knew that the Earth is round and Eratosthenes was able to measure the radius of the Earth (Figure 2). Choose two points A, B on the same longitude and measure the distance d between them. Plant a vertical stick (red) in the ground at A and another (blue) at B. On a day in the year when the stick at A produces no shadow at noon, at the same time the stick at B produces a shadow whose angle is α. The sun is so far away from the Earth that over the relatively short distance d, the rays of the Sun are essentially parallel. By alternate interior angles, the angle between the two sticks as measured from the center of the Earth is also α.

The angle that Eratosthenes measured at the blue stick was

$$\alpha = 7.5^\circ \cdot \frac{2\pi}{360} \approx 0.131 \text{ radians},$$

and the distance d between A and B was known to be approximately 800 km. The arc $\widehat{AB}$ subtends the angle $\alpha = d/r_e$ where r_e is the radius of the Earth, so

$$r_e = \frac{d}{\alpha} = \frac{800}{0.131} \approx 6107 \text{ km}. \tag{1}$$

This value is quite close to the modern measurement of 6370 km.

2.2 Aristarchus' measurements

Using r_e, Eratosthenes' measurement of the radius of the Earth, Aristarchus was able to measure and compute the following values:

- r_m: the radius of the Moon,
- r_s: the radius of the Sun,
- d_m: the distance from the Earth to the Moon,
- d_s: the distance from the Earth to the Sun.

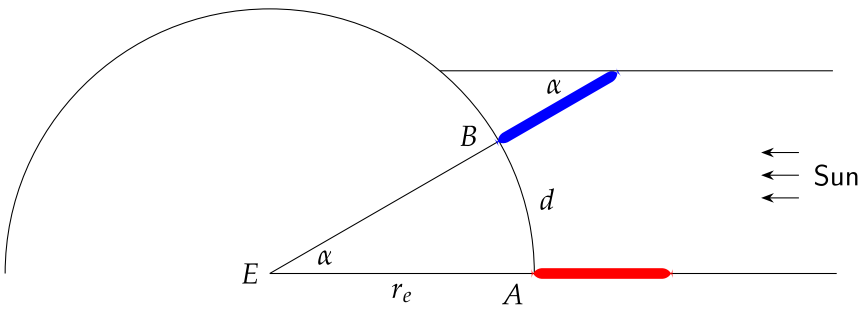


Figure 2: Eratosthenes' measurement of the radius of the earth

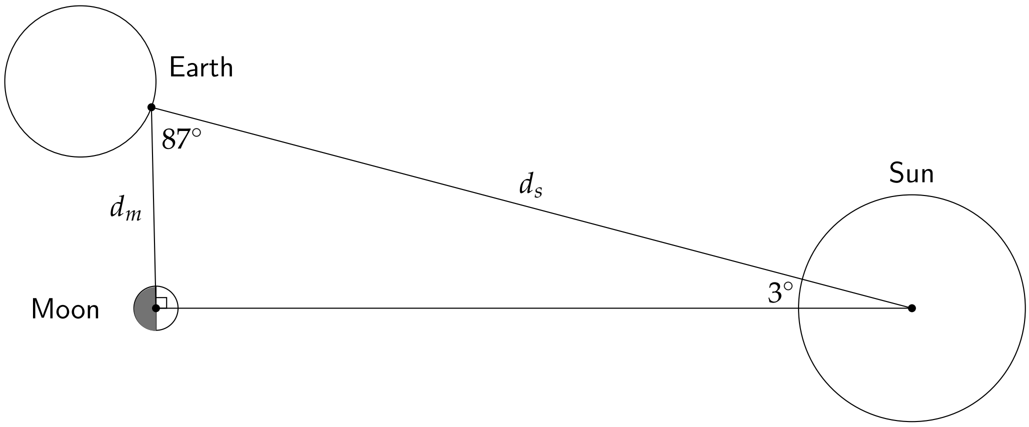


Figure 3: Observing a first quarter moon

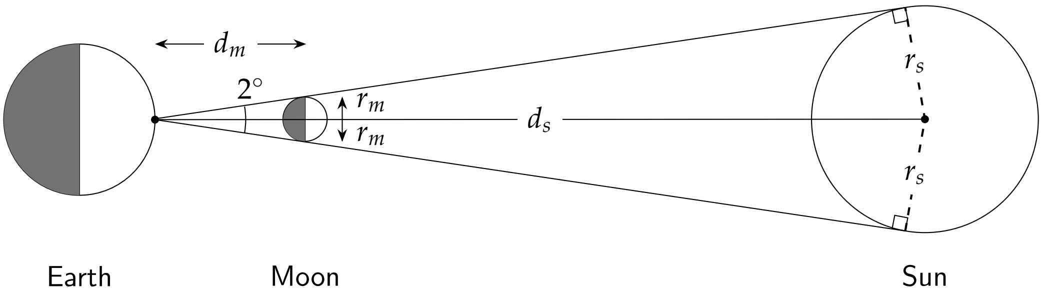


Figure 4: A solar eclipse

Computing d_s/d_m

An observer on Earth can follow the phases of the Moon as it revolves around the Earth. At one point in the month the phase will be first quarter, meaning that the one half of the moon is illuminated while the other half is not (Figure 3). The angle between the Sun and the Moon will be 87°. Since exactly half of the moon is illuminated, we know that the angle ∠ Earth-Moon-Sun is a right-angle so

$$\begin{aligned} \cos 87^\circ &= \frac{d_m}{d_s} \\ \frac{d_s}{d_m} &= \frac{1}{\cos 87^\circ} \approx 19\,. \end{aligned} \tag{2}$$

Computing r_s/r_m and d_m/r_m

The Moon is much, much smaller than the Sun, but it is also much, much closer to the Earth. When the Moon is precisely positioned between the Earth and the Sun, its "disk" exactly covers the "disk" of the Sun, causing a total solar eclipse (Figure 4).

The angle subtended by the Moon is 2° degrees. Bisecting the angle creates two right triangles with an acute angle of 1°, where the right angles are the tangents to Moon and the Sun. By similar triangles, Equation 2 and Figure 4,

$$\frac{r_s}{r_m} = \frac{d_s}{d_m} = 19 \tag{3}$$

$$\frac{d_m}{r_m} = \frac{1}{r_m/d_m} = \frac{1}{\sin 1^\circ} \approx 57\,. \tag{4}$$

Computing the radii and distances

Figure 5 shows a lunar eclipse. Unlike a solar eclipse where the Moon exactly covers the Sun, the Earth more than covers the Moon and its shadow is four times the Moon's radius.

Figure 6 show a lunar eclipse annotated with the distances d_m, d_s and the radii r_m, r_e, r_s. The ray from the top of the Sun is tangent to both the Sun and the Earth, so it forms right angles with their radii, as well as with the extension of the Moon's radius. The thick horizontal lines are constructed parallel to the line connecting the centers, forming two right triangles. The triangles are similar so using Equation 3,

$$\begin{aligned} \frac{r_s - r_e}{r_e - 2r_m} &= \frac{d_s}{d_m} = \frac{r_s}{r_m} \\ r_s r_e + r_m r_e &= 3 r_s r_m\,. \end{aligned}$$

Again from Equation 3, $r_s = 19 r_m$, so

$$r_m = \frac{20}{57} r_e\,.$$

By Equation 1, $r_e \approx 6107$ km, by Equation 4, $d_m = 57 r_m$, and by Equation 2, $d_s = 19 d_m$, so we can compute the radii and distances:

$$\begin{aligned} r_m &= \frac{20}{57} r_e \approx 2143 \text{ km} \\ r_s &= 19 r_m \approx 40,713 \text{ km} \\ d_m &= 57 r_m \approx 122,140 \text{ km} \\ d_s &= 19 d_m \approx 2,320,660 \text{ km}\,. \end{aligned}$$

Table 1 summarizes these data together with the modern values [4, Table 1.3]. While the computed values for the radii of the Earth and the Moon are not far off from the modern

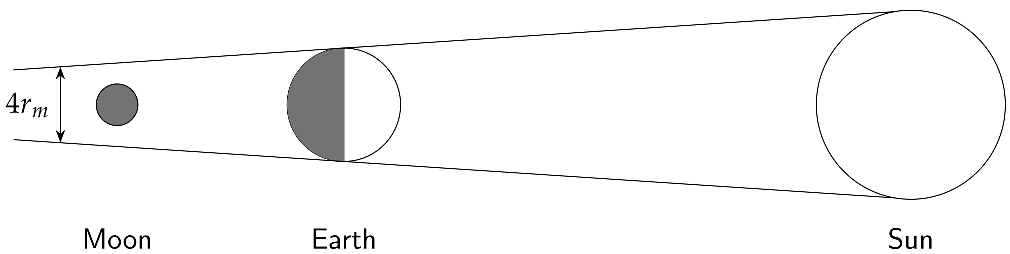


Figure 5: A lunar eclipse

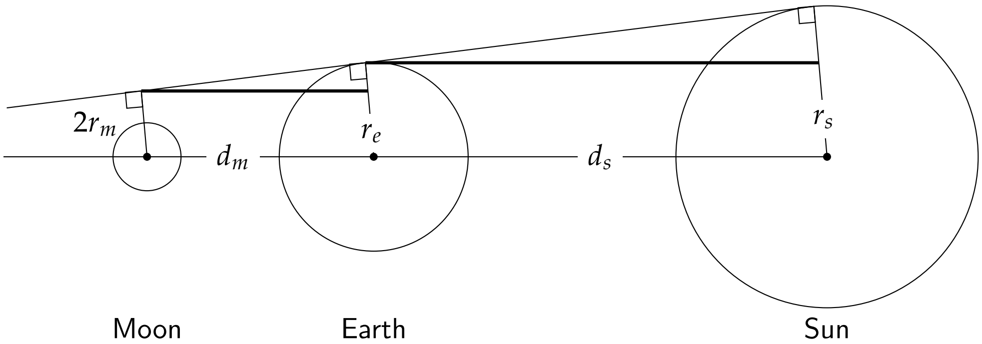


Figure 6: Detail of a lunar eclipse

values, the other computed values are not near the modern values. Nevertheless, they do show that the Greeks understood the immense size of the solar system.

		Computed (km)	Modern (km)
r_e	radius of Earth	6107	6370
r_m	radius of Moon	2143	1740
r_s	radius of Sun	40,713	695,500
d_m	distance Earth-Moon	122,140	384,570
d_s	distance Earth-Sun	2,320,660	150,000,000

Table 1: Values computed by Aristarchus' method compared with modern values

3 The Sun-centered solar system

As everyone living far from equator knows, the time between sunrise and sunset varies with the seasons. The reason is that the axis of the rotation of the Earth is offset by 23.5° relative to the orbit of the Earth. The plane of the orbit of the Earth around the Sun is called the *ecliptic*. Measuring the length of the day as the time from sunrise to sunset, there is a day in June, called the *summer solstice*, when the length of the day is longest. Similarly, there is a day in December, called the *winter solstice*, when the length of the day is shortest.[2] There are also two days when the length of the day equals the length of the night: the *autumn equinox* in September and the *spring equinox* in March.

Today we know that the universe is immensely large and that the stars are moving at extremely high speeds, but an observer on Earth sees them as if their positions are fixed on a sphere around the earth, called the *celestial sphere*. This solstices and equinoxes can be associated with the projection of the Sun on the celestial sphere as seen from the Earth. The details of the Earth's orbit can be found in books on astronomy, as well as in the Wikipedia articles on *Equinox* and *Solstice*.

3.1 The length of the year and the length of the seasons

Let us assume that the Earth orbits the Sun in a circle, such that the center of the orbit O is the center of the Sun S. In Figure 7 the inner circle is orbit of the Earth and the outer circle is the celestial sphere. The orbit can be divided into four quadrants called *seasons*: spring, summer, autumn, winter.

The length of a year is approximately $365\frac{1}{4}$ days. The extra $\frac{1}{4}$ day is accounted for by adding a day in leap years.[3] The length of each season as determined by the equinoxes and the solstices is $365.25/4 = 91.3125$ days. However, measurements by the Greek astronomer Hipparchus showed that the actual lengths of the seasons differed from this number (Table 2) and a model of the solar system must be able to explain these differences.

In his Earth-centered solar system, Hipparchus proposed that the center of the Earth's orbit

[2]This holds for the northern hemisphere; in the southern hemisphere the opposite holds.

[3]The length of a year is actually 365.2425. In the sixteenth century, the *Gregorian calendar* accounted for the difference by removing three leap years in every four hundred years.

Season	Days	%
Spring	$94\frac{1}{4}$	25.8
Summer	$92\frac{1}{2}$	25.3
Autumn	$88\frac{1}{8}$	24.1
Winter	$90\frac{1}{8}$	24.7

Table 2: The lengths of the seasons and their percentages of a year

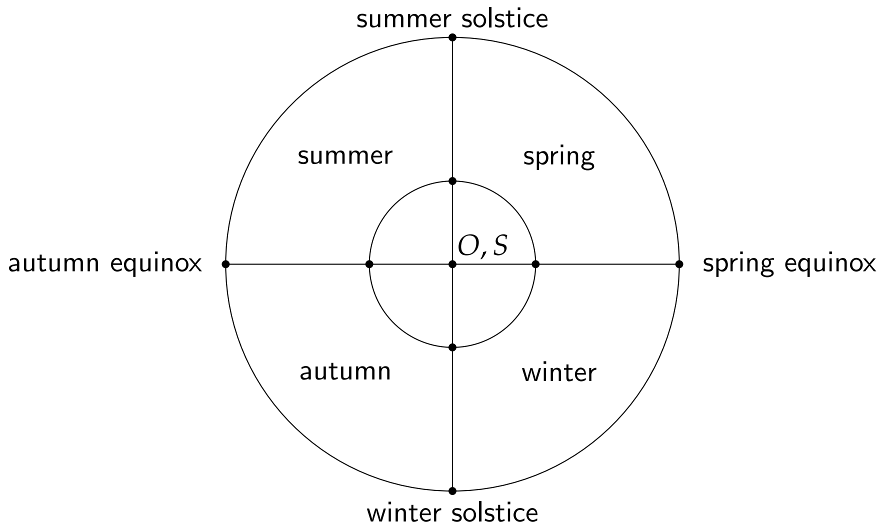


Figure 7: The orbit of the Earth and the seasons

be offset from the center of the Sun. Copernicus used the same idea in his Sun-centered solar system (Figure 8). If the center of the Earth's orbit is in the upper-left quadrant of the coordinate system defined by the equinoxes and the solstices, the angles for spring and summer are obtuse, so the seasons are longer than one-fourth of a year, whereas the angles for the autumn and winter are acute, so the they are shorter than one-fourth of a year.

Figure 9 shows a magnified and distorted view of Figure 8. It has been annotated with additional lines and labels that will facilitate the demonstration of Copernicus' computation. The axes $A'C'$ and $B'D'$ have their origin O at the center of the Earth's orbit and are parallel to the axes in the ecliptic. The dashed lines from O are all radii of the Earth's orbit

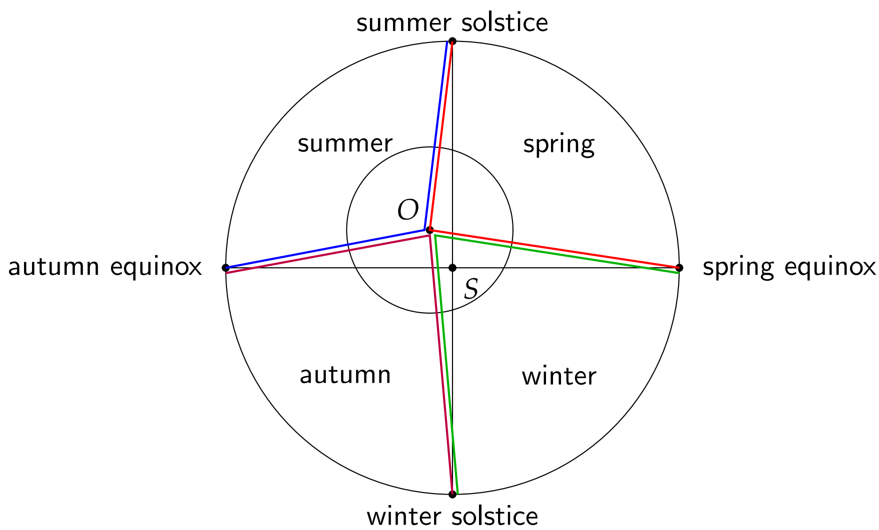


Figure 8: The lengths of the seasons are not equal

that will be denoted r. The dotted lines form two right triangles which will be used in the computation.

3.2 The location of the center of the Earth's orbit

Copernicus' task was to locate the position of the center of the Earth's orbit O relative to the center of the Sun S. This will be given in polar coordinates $OS = c$ and $\angle FSB = \lambda$ (the label is on the large circle of the ecliptic). The strategy of the computation is as follows:

- Initially, we compute the angles of the arcs in radians; the lengths of the arc can then be obtained by multiplying by the radius r.
- We use the lengths of the seasons that Copernicus used: summer is $93\frac{14.5}{60}$ days and spring is $92\frac{51}{60}$.
- The angle of the arc $\widehat{AC}$ can be computed from the combined length of spring and summer and the angle of the arc $\widehat{AB}$ can be computed from the length of spring.
- From $\widehat{AC}$ and $\widehat{AB}$, the angle α subtended by $\widehat{AA'}$ and the angle β subtended by $\widehat{BB'}$ can be computed.
- Since the Earth is very close to the Sun relative to the radius of its orbit, $r\,\widehat{AA'}$ and $r\,\widehat{BB'}$ approximate the line segments a and b. From these c and λ can be computed.

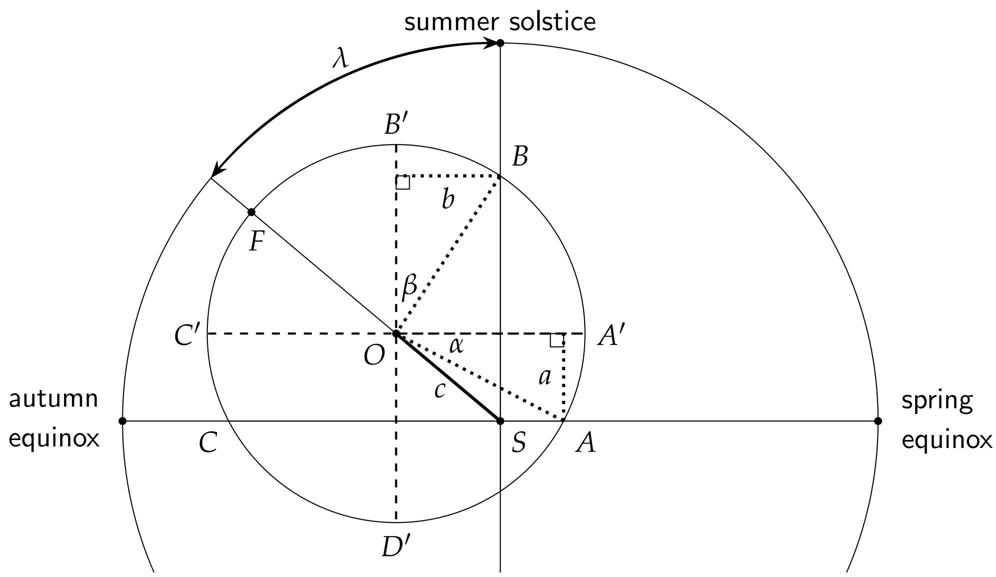


Figure 9: Computing the center of the Earth

Computing the angles of the arcs $\widehat{AB}, \widehat{AC}$

The arcs $\widehat{AB}, \widehat{AC}$ are sectors of the Earth's orbit and their angles are their proportions of a full year times 2π radians.

$$\widehat{AB} = 2\pi \cdot \frac{92\frac{51}{60}}{365.25} = 2\pi \cdot \frac{92.85}{365.25} = 1.5972 \text{ radians}$$

$$\widehat{AC} = 2\pi \cdot \frac{92\frac{51}{60} + 93\frac{14.5}{60}}{365.25} = 2\pi \cdot \frac{186.09}{365.25} = 3.2012 \text{ radians}\,.$$

Computing the angles of the arcs $\widehat{AA'}, \widehat{BB'}$

Let us express the arcs $\widehat{AC}$ and $\widehat{AB}$ in terms of the arcs that comprise them. Since AC is parallel to $A'C'$, $\widehat{AA'} = \widehat{C'C}$. Compute $\widehat{AA'}$:

$$\begin{aligned}\widehat{AC} &= \widehat{AA'} + \widehat{A'C'} + \widehat{C'C} = 2\widehat{AA'} + \pi \\ \widehat{AA'} &= \frac{1}{2}(3.2012 - \pi) = 0.0298 \text{ radians}\,.\end{aligned}$$

Now that we have computed $\widehat{AB}$ and $\widehat{AA'}$ we can compute $\widehat{BB'}$:

$$\begin{aligned}\widehat{AB} &= \widehat{AA'} + \widehat{A'B'} - \widehat{BB'} \\ \widehat{BB'} &= 0.0298 + \frac{\pi}{2} - 1.5927 = 0.0034 \text{ radians}\,.\end{aligned}$$

Computing the lengths of the arcs $\widehat{AA'}, \widehat{BB'}$

OA and OB are radii of the Earth's orbit so

$$\begin{aligned}\sin\alpha &= \frac{a}{r} \approx \alpha \\ a &\approx r\alpha = r\widehat{AA'} = 0.0298r \\ \sin\beta &= \frac{b}{r} \approx \beta \\ b &\approx r\beta = r\widehat{BB'} = 0.0034r\,,\end{aligned}$$

where we have used the assumption that O, the center of the Earth's orbit, is very close to the Sun S so that $\lim_{x\to 0} \frac{\sin x}{x} = 1$.

Computing the position of O relative to S

Figure 10 shows a magnified diagram of a portion of Figure 9. In the dotted triangles, we have already computed the lengths a and b. Since OT is parallel to $A'A$ and TS is parallel

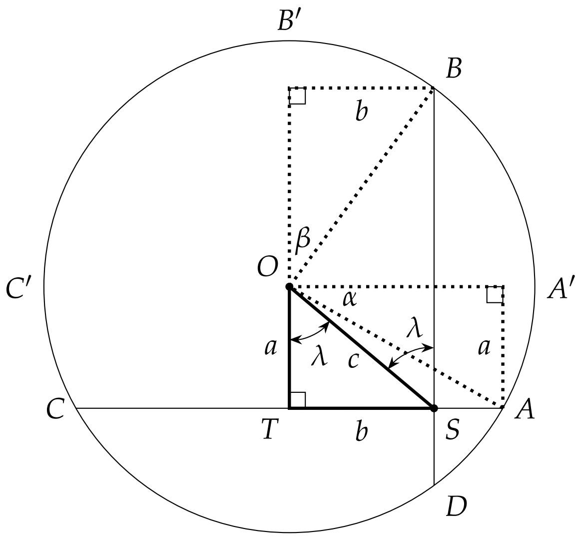


Figure 10: Three triangles

to BB', we can label OT by a and TS by b. The emphasized triangle is a right triangle and c, the distance of O from S, can be obtained from Pythagoras' theorem:

$$c = \sqrt{a^2 + b^2} = r\sqrt{(0.0298)^2 + (0.0034)^2} = 0.03r\,.$$

λ can be obtained from trigonometry:

$$\lambda = \tan^{-1}\frac{b}{a} = \tan^{-1}\frac{0.0034}{0.03} = 0.1129 \text{ radians} \approx 6.47^\circ\,.$$

The distance $0.03r$ is shown in Table 3 using the values of r from Table 1.

	Aristarchus (km)	Copernicus (km)	Modern (km)
radius of Earth's orbit	$2,320,660$	$8,000,000$	$150,000,000$
distance of O from S	$69,620$	$240,000$	$4,500,000$

Table 3: Values computed by Aristarchus' method compared with modern values

4 Elliptical orbits

Towards the end of the sixteenth century, the astronomer Tycho Brahe carried out extremely precise observations. In 1600 he hired Johannes Kepler as his assistant and when Tycho died soon afterwards, Kepler was appointed to his position. Here we explain how Kepler was able to establish that planetary orbits are ellipses.

4.1 Determining the radius of the Earth's orbit

A Martian year is 687 days, that is, it equals $\frac{687}{365.25} = 1.88$ Earth years. We know when Mars reaches a new "year" by observing its projection on the celestial sphere, but each time the position of the Earth in its orbit will be different. Figure 11 shows the orbit of the Earth—its center O offset from the Sun S as Copernicus showed—at four occasions when the position of Mars M at its new year was observed. Four triangles are created $\triangle OE_iM$. Figure 12 shows one of the triangles with the angles labeled. Using the law of sines,

$$\frac{OE_i}{\sin\beta} = \frac{OM}{\sin\alpha}$$

$$OE_i = OM\,\frac{\sin\beta}{\sin\alpha}\,.$$

Tycho Brahe was able to measure all three angles (Table 4)[4]. The values of OE are given in the fourth column of the table and they are not equal. Assuming (as everyone did at that time) that the Earth' orbit is circular, the only solution was to move the center of the orbit so that $\{E_1, E_2, E_3, E_4\}$ were all on the circle.

[4]The values are rounded to one decimal point.

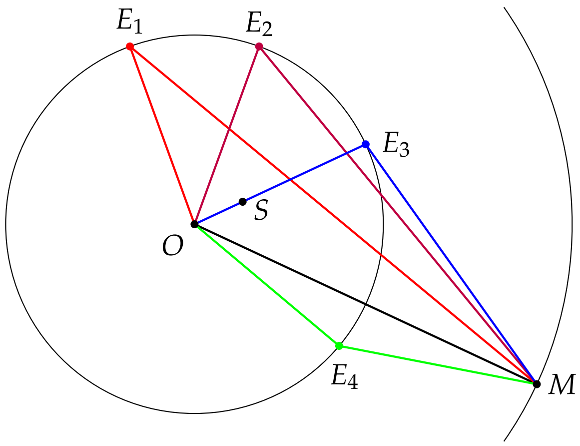


Figure 11: Observations of the orbit of Mars from the Earth

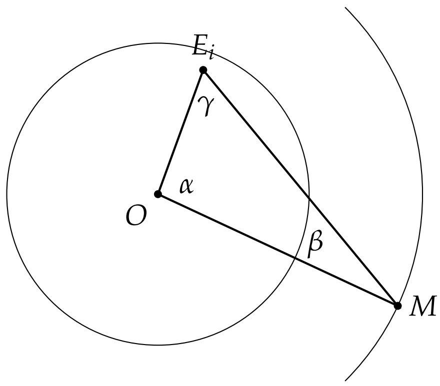


Figure 12: One triangle Earth-Sun-Moon

4.2 Measuring the angles in the triangle Sun-Earth-Mars

How can the angles α, β, γ be measured? Since $\triangle E_i OM$ is a triangle, it is sufficient to measure two of the angles. γ is easily measured by observing Mars and the Sun at the same time. However, neither α nor β can be measured directly since they are not accessible to an observer on Earth.

Tycho's measurement used the known periods of the orbits to compute the angles. The Earth moves counterclockwise around Sun. Given any point E, after t days the Earth will eventually move to E' and Mars will move to M' so that they are in *opposition*, that is, Mars will be in the continuation of the Earth-Sun line (Figure 13). Since the Earth completes an orbit in close to half the time that Mars takes, this will occur before either has completed a full orbit. The angles θ_E and θ_M are fractions of a circular orbit of 360°, so

$$\begin{aligned}
\frac{t}{365.25} &= \frac{\theta_E}{360} \\
\theta_E &= \frac{360\,t}{365.35} \\
\frac{t}{687} &= \frac{\theta_M}{360} \\
\theta_M &= \frac{360\,t}{687}\,.
\end{aligned}$$

	α	β	γ	OE_i $(\times OM)$
E_1	127.1	20.8	32.1	0.6682
E_2	84.2	35.8	60.5	0.6721
E_3	41.3	42.4	96.4	0.6785
E_4	1.6	3.4	175.0	0.6805

Table 4: Tycho's measurements of the distance of the Earth from the center

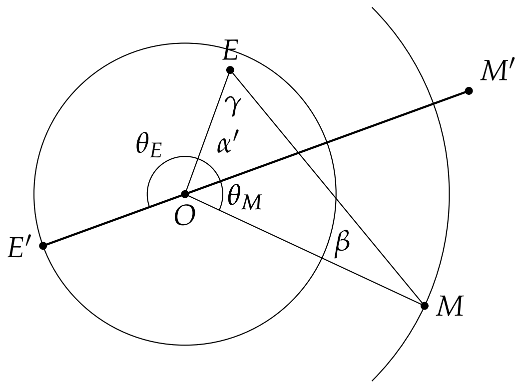


Figure 13: The Earth and Mars in opposition

This gives values for θ_E and θ_M. From geometry we have

$$\begin{aligned} \alpha' &= 180 - \theta_E = \alpha - \theta_M \\ \alpha &= 180 - \theta_E + \theta_M\,, \end{aligned}$$

and the values of OE_i in Table 4 can be computed.

4.3 A new location for the center of the Earth's orbit

Kepler's next task was to obtain a new value O' for the center of the Earth's orbit such that the E_i in Table 4 are on the orbit. We need the following theorem in geometry.

Theorem 4.1 *Given three non-collinear points,*[5] *a circle can be constructed that goes through all three points.*

Proof Three non-collinear points A, B, C define a triangle $\triangle ABC$ (Figure 14). Construct the perpendicular bisectors of any two of its three sides, say, AC and BC. By definition the perpendicular bisector is the geometric locus of points equidistant from the endpoints of the segment. Let O be the intersection of the two bisectors. Then the $AO = CO = BO$ is the radius of a circle centered at O that goes through A, B, C. ∎

Given the new locations of the Earth $\{E_1, E_2, E_3, E_4\}$, a circle centered at O' can be constructed that goes through $\{E_1, E_2, E_3\}$ (Figure 15). To verify that this is the correct orbit, check that E_4 is on the circle.

4.4 Orbits are ellipses

While Kepler was able to modify the center of the orbit of the Earth to be consistent with the observations, he was not able to adequately describe the orbit of Mars. After years of

[5] *Collinear* means that the points all on the same line.

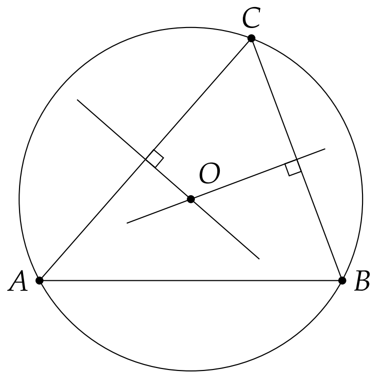


Figure 14: A circle through three arbitrary points

work, he came to the conclusion that the orbit must be oval like an egg. Oval, perhaps, but certainly not an ellipse, because he was certain that it would have been discovered by Archimedes! Figure 16 shows C, a position on a circular orbit, and an oval orbit (dashed), where M is the position of Mars on the oval corresponding to C. The radius of the circular orbit is labeled a and the unknown distances to M are labeled $s = SM$ and $t = OM$. Kepler's computed that $\dfrac{a-t}{t} = 0.00429$ and $\dfrac{s}{t} = 1.00429$, so that

$$\frac{a-t}{t} = \frac{s}{t} - 1$$
$$a - t = s = t\,,$$

and therefore $SM = s = a = AO = CO$. If the dashed oval is, in fact, an ellipse, because in an ellipse $SM = AO$ (by Theorem 7.2). Kepler then computed the projections of the observations of Mars on the x-axis (Figure 17) and obtained for all of them that

$$\frac{M_iO_i}{C_iO_i} = \frac{1}{1.00429} = 0.99573\,.$$

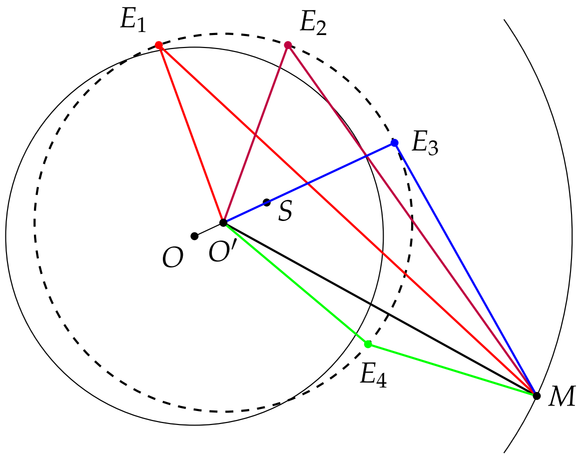


Figure 15: Observations of the orbit of Mars from the new Earth's orbit

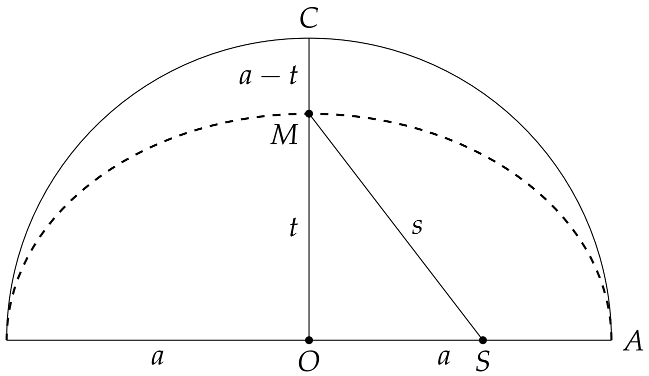


Figure 16: The orbit of Mars as an oval “egg”

By Theorem 7.4, since the ratio $MO/CO = b/a$ is constant for in an ellipse, Kepler was able to conclude that the orbit of Mars is an ellipse.

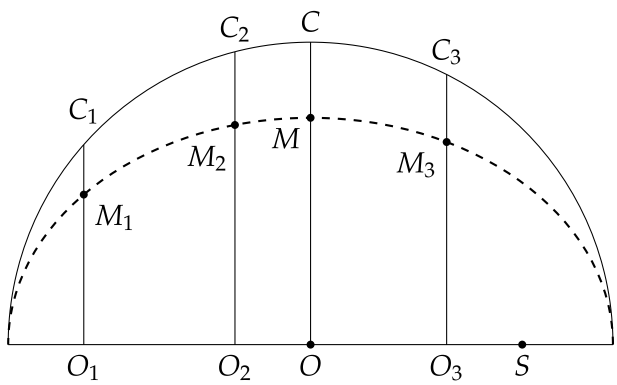


Figure 17: The orbit of Mars as an ellipse

5 Gravitation

In this section we present Isaac Newton's demonstrations that the inverse-square law can be deduced from Kepler's observation that the orbits of planets are ellipses. After a review of Newton's Laws of force and motion, we show that Kepler's second law must hold in any system subject to a centripetal force. The next step is to show that the inverse square law for gravitation must hold in an elliptical orbit. Then it is a small step to universal gravitation and Kepler's third law. For readers who wish to study the *Principia* the labels of the points in the diagrams are the same as those used by Newton.

5.1 Newton's laws of motion

1. A body in uniform motion (including a body at rest) continues with the same motion unless a force is applied.
2. A force F applied to a body causes an acceleration a in the direction of the force whose magnitude $a = F/m$, where m, the constant of proportionality, is called the *mass* of the body.
3. If one body exerts a force on a second body, the second body exerts a force on the first of equal magnitude but in the opposite direction.

Forces are denoted by vectors, where the direction of the vector represents the direction of the force and the length of the vector represents the magnitude of the force. Forces can be decomposed into perpendicular components (Figure 18), or into components in any directions (Figure 19). The components form a parallelogram whose diagonal is the resultant force.

Newton was interested in *centripetal force* which is a force exerted by a single body on another, in particular, the gravitational force exerted by the Sun on a planet (Figure 20). Since the only force is that directed towards the Sun, the planet does not move "up" or "down" so its orbit is in a plane.

Kepler's second law states that a planet in orbit sweeps out equal area in intervals of equal duration, if it takes the planet time t to move from P_1 to P_2 and also t to move

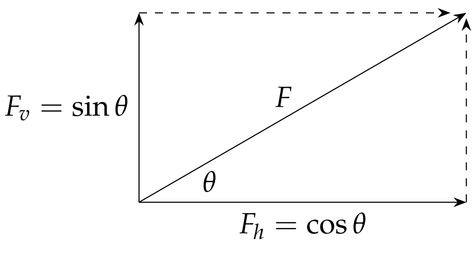


Figure 18: Perpendicular components of a force

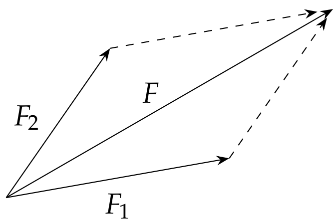


Figure 19: Arbitrary components of a force

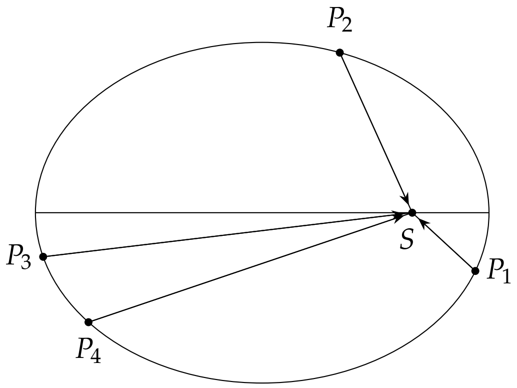


Figure 20: Centripetal force

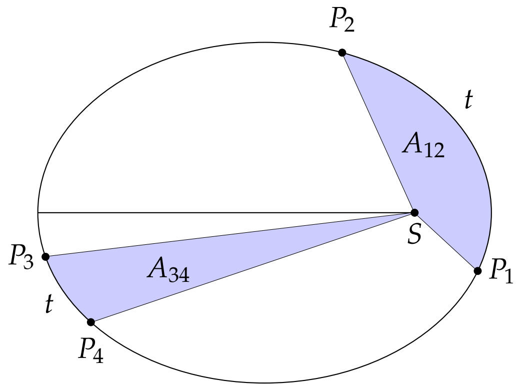


Figure 21: Equal areas in equal times

from P_3 to P_4, then A_{12}, the area of the sector P_1SP_2, is equal to A_{34}, the area of P_3SP_4 (Figure 21). Obviously, this means that the speed of the planet must vary as it traverses its orbit $v_{P_1P_2} \gg v_{P_3P_4}$. Newton proved that this must be true in any system where a body is subject to a centripetal force from another body.

The proof is based on dividing an area into very small sectors and then taking the limit. Consider three points A, B, C on the orbit (Figure 22) that represent the positions of the planet at intervals of Δt. For clarity we have drawn them spaced out, but the intention is that they are very close together. Newton assumed that the planet does not smoothly traverse the arcs, but rather that it every Δt it jumps in discrete steps from one point on the orbit to the next.

Figure 23 shows how the force is exerted in discrete steps. The planet moves from A to B and we expect that the centripetal force at B will cause an acceleration that moves the planet to C, the next point on the orbit. Instead, we "pretend" that the force is not applied at B, but, in the absence of an applied force, planet continues to move in the same direction and at the same speed. After another period of Δt as passed and the planet has reached point C', the force is now applied *in the same direction* as it would have been applied at B, moving the planet to C.

Theorem 5.1 *The area of $\triangle ASB$ is equal to the area of $\triangle BSC$.*

Proof This will be done in two steps by showing that $\triangle ASB = \triangle BSC'$ and $\triangle BSC' = \triangle BSC$. In Figure 24, $\triangle ASB$ is shown in blue and $\triangle BSC'$ is shown in red. It is assumed that $AB = BC'$ (the planet moves from B to C' during the same interval Δt), so since SH is the height of both triangles, their areas are equal. In Figure 25, $\triangle BSC$ is shown in blue and $\triangle BSC'$ is shown in red. It is assumed CC' is parallel to SB (the planet is subject to the centripetal force at C' in the *same* direction as the force at B, so the heights of both triangles to the common side SB are equal and their areas are equal. It follows that $\triangle ASB = \triangle BSC' = \triangle BSC$, which we denote by ΔA. ∎

We assume that the sectors of the orbit are each divided up into small sectors of uniform

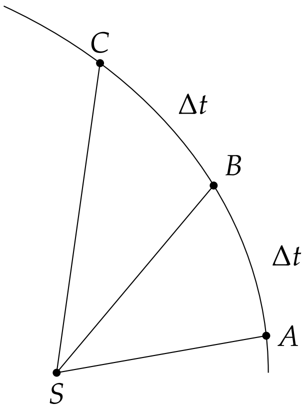


Figure 22: "Small" sectors of an orbit

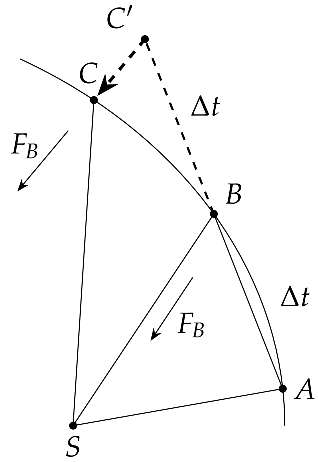


Figure 23: Exerting force at discrete times

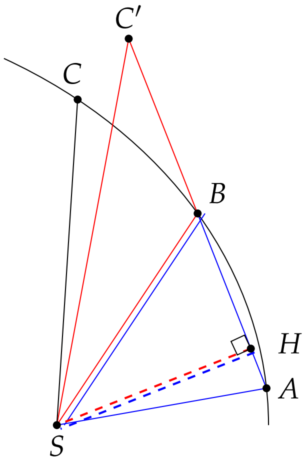


Figure 24: $\triangle ASB = \triangle BSC'$

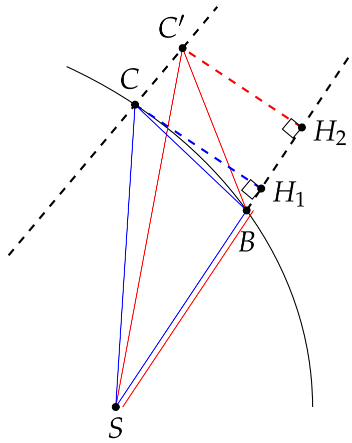


Figure 25: $\triangle BSC' = \triangle BSC$

duration Δt. By the theorem, each sector has the same area ΔA. Therefore,

$$\frac{A_{12}}{\Delta A} = \frac{t}{\Delta t} = \frac{A_{34}}{\Delta A},$$

from which Kepler's second law follows:

$$\frac{A_{12}}{t} = \frac{A_{34}}{t}$$
$$A_{12} = A_{34}.$$

The proof used two approximations:

- ΔA is an approximation of the area of each sector.
- The force at C' is an approximation to the force at B.

In the limit as the size of the sectors decreases, the errors become negligible.

Definition 5.2 *For a given elliptical orbit,* $\kappa = \dfrac{A}{t}$*, where* A *is the area of the ellipse and* t *is the period of the orbit, is called* Kepler's constant.

5.2 The inverse square law for gravitation

Newton's next step was to show that if the orbit of a planet is elliptical, the centripetal force must be proportional to the mass of the planet and inversely proportional to the square of its distance from the Sun. In Figure 26, S is the Sun, and P and Q are points on the orbit that are very close to each other. PR is a tangent to the ellipse at P and R is a point such that QR is parallel to SP. QT is constructed perpendicular to SP. Denote the lengths $r_P = SP$, $h = QT$, $q = QR$ and the time interval from P to Q by Δt.

When a point is subject to an acceleration a for a period of Δt, its displacement is $\frac{1}{2}a(\Delta t)^2$. From Newton's second law we know that at point R, the planet is subjected to an acceleration of F_P/m, so

$$q = \frac{1}{2}\frac{F_P}{m}(\Delta t)^2$$

$$F_P = \frac{2mq}{(\Delta t)^2}\,.$$

Now we compute the area of $\Delta A = \triangle SPQ = (1/2)hr_p$ and use Kepler's constant:

$$\Delta t = \frac{\Delta A}{\kappa} = \frac{hr_P}{2\kappa}$$

$$F_P = 2mq \cdot \frac{4\kappa^2}{(hr_P)^2} = 8\kappa^2 m \cdot \frac{q}{h^2} \cdot \frac{1}{r_P^2}\,.$$

To obtain an inverse-square law for the force, the first two factors have to be independent of the distance. For a given planet m is constant and for a given elliptical orbit κ is constant, so the first factor does not depend on the distance. What about the second factor q/h^2, in particular, what value does it have if Δt approaches zero?

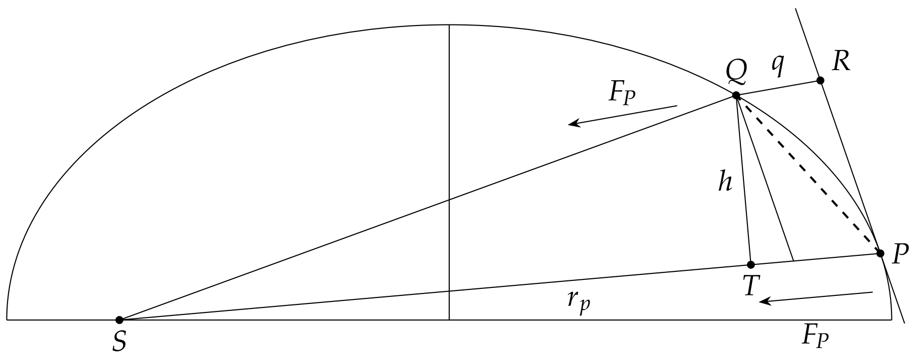


Figure 26: The derivation of the inverse square law

Theorem 5.3 *In an elliptical orbit*

$$\lim_{\Delta t \to 0} \frac{q}{h^2} = \frac{1}{L},$$

where L is the length of the latus rectum of the ellipse (Definition 7.5).

Newton's proof is very long and is presented separately in Section 6.

Since L is constant for any given ellipse, the inverse square law can be written

$$F_P = \frac{8\kappa^2 m}{L} \cdot \frac{1}{r_P^2}.$$

A simplified formula for the force The formula can be re-written so that the constant values appearing are more familiar: a, the semi-major axis and T, the period of the orbit. By Theorem 7.6, $L{=}2b^2/a$ and by Theorem 7.7, $\kappa = A_{ellipse}/T{=}\pi ab/T$, so

$$F_P = \frac{8\kappa^2 m}{L} \cdot \frac{1}{r_P^2} = \frac{8(\pi ab)^2 m}{T^2} \cdot \frac{a}{2b^2} \cdot \frac{1}{r_P^2} = \frac{4\pi^2 a^3 m}{T^2} \cdot \frac{1}{r_P^2}. \tag{5}$$

Applicability of the inverse square law Newton was able to show that

- The inverse square law applies to all conic sections including a parabola and a hyperbola, not just to an ellipse and, of course, a circle. Some comets have hyperbolic orbits and orbit the Sun only once.
- The converse holds: if a planet is subject to an inverse-square centripetal force then the orbit must be an ellipse (or other conic section).
- The proof assumes that a planet is a very small point, but the result holds even for large planets as long as the density of the planet is radially symmetric, that is, for a given distance from the center the density is constant.

5.3 Universal gravitation

Recall Newton's third law: If one body exerts a force on a second body, the second body exerts a force on the first that is equal in magnitude but in the opposite direction. Therefore, we can equate the force that the Sun S exerts on a planet S with the force that the planet exerts on the Sun. Let m be the mass of the planet and M be the mass of the Sun. Let $F_{S\leftarrow E}$ be the forced exerted by the Sun on the Earth and let $F_{E\leftarrow S}$ be the force exerted by the Earth on the Sun. Using Equation 5,

$$F_{S\leftarrow E} = \frac{8\kappa_E^2 m}{L_E} \cdot \frac{1}{r^2} = \frac{C_E m}{r^2}$$

$$F_{E\leftarrow S} = \frac{8\kappa_S^2 M}{L_S} \cdot \frac{1}{r^2} = \frac{C_S M}{r^2}$$

$$\frac{C_E}{M} \cdot \frac{1}{r^2} = \frac{C_S}{m} \cdot \frac{1}{r^2},$$

from some constants C_E, C_S.

Why are the constants different? The Earth and the Sun both rotate around their center of mass called the *barycenter*, which is very close to the center of the Sun since the Sun is so much more massive than the Earth. The ellipse of the Sun's orbit is very small relative to the Earth so A and L are smaller, and the Sun's period is large so T is larger. The different values for $8\kappa^2/L$ are encapsulated into the constants C_E, C_S. Let $G = \dfrac{C_E}{M} = \dfrac{C_S}{m}$ so that

$$F_{S \leftarrow E} = F_{E \leftarrow S} = G\frac{mM}{r^2}\,. \tag{6}$$

This is Newton's law of universal gravitation. It is not specific to planetary orbits but holds between any two bodies with masses m, M.

5.4 Kepler's third law

Newton proved that Kepler's third law follows from the law of universal gravitation.

Theorem 5.4 (Kepler's third law) *Let P_1, P_2 be two planets whose elliptical orbits have semi-major axes a_1, a_2 and whose orbital periods around the Sun are T_1 and T_2. Then*

$$\frac{a_1^3}{T_1^2} = \frac{a_2^3}{T_2^2}\,.$$

Proof By Equations 5 and 6,

$$F = \frac{4\pi^2 a_i^3 m}{T_i^2}\frac{1}{r_i^2} = \frac{GmM}{r_i^2}\,.$$

After canceling m and r_i we get

$$\frac{a_i^3}{T_i^2} = \frac{GM}{4\pi^2}\,.$$

$\dfrac{GM}{4\pi^2}$ is a constant that depends only on the mass of the Sun and the gravitational constant, so $\dfrac{a_i^3}{T_i^2}$ is constant for all planets rotating around the Sun. ∎

6 The proof Proposition XI, Problem VI

Theorem 5.3 is Book I, Section III, Proposition XI, Problem VI of the *Principia*.
Study Figure 27:[6]

- Given an ellipse whose center is C and whose foci are S and H, let P, Q be two points on the ellipse. These points represent the movement of a body in an elliptical orbit separated by a time interval Δt.
- Construct lines from P to the center and the foci.
- Let RPZ be the tangent at P such that the line segment PR is the path that the body would transverse if it continued moving for time Δt, but was not subject to any force.
- Construct the parallelogram $PRQX$.
- Extend QX until it intersects PC at V.
- Construct a line parallel to RP through H and let I be its intersection with PS.
- Construct DC the conjugate diameter to PC (Definition 7.12) and let E be its intersection with PS.

[6]To save space the bottom half of the ellipse has been clipped, but we still refer to lines from the center to points on the ellipse as diameters.

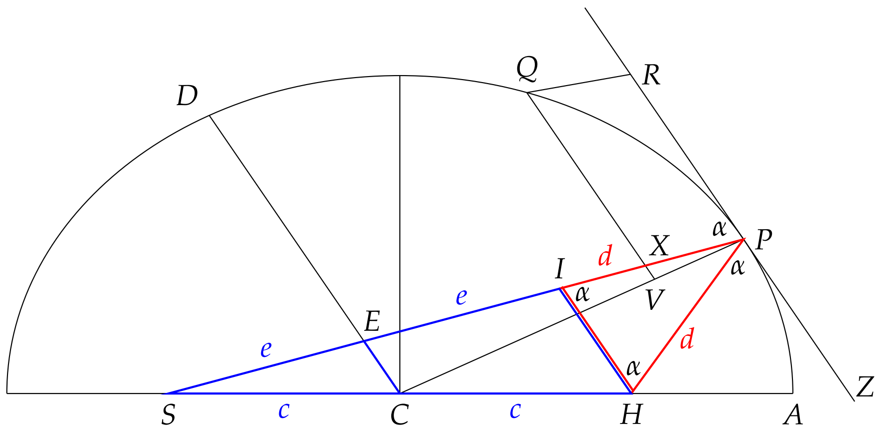


Figure 27: Geometry of an elliptical orbit (1)

6.1 A formula for QR

Theorem 6.1 $QR = PV \cdot \dfrac{CA}{CP}$.

Proof By Theorem 7.9, $\alpha = \angle RPX = \angle ZPH$ and by alternate interior angles,

$$\angle PHI = \angle ZPH = \alpha = \angle RPX = \angle PIH\,,$$

so $\triangle IPH$ (red) is isoceles and $d = PI = PH$.

$SC = CH = c$ are equal because they are the distances of the foci from the center of the ellipse. Let $SE = e$. By construction $EC \parallel IH$ so $\triangle ESC \sim \triangle ISH$ (blue) so

$$\begin{aligned} \frac{SC}{SE} &= \frac{SH}{SI} \\ \frac{c}{e} &= \frac{2c}{SI} \\ SI &= 2e\,. \end{aligned}$$

By definition of an ellipse $SP + PH = SI + IP + PH = 2CA$ so $(2e + d) + d = 2CA$ and $EP = e + d = CA$.

$QV \parallel EC$ so $\triangle EPC \sim \triangle XPV$ and

$$\begin{aligned} \frac{PX}{PV} &= \frac{EP}{PC} = \frac{CA}{PC} \\ PX &= PV \cdot \frac{CA}{PC}\,. \end{aligned}$$

By construction of the parallelogram $QR = PX$ proving that $QR = PV \cdot \dfrac{AC}{PC}$. ∎

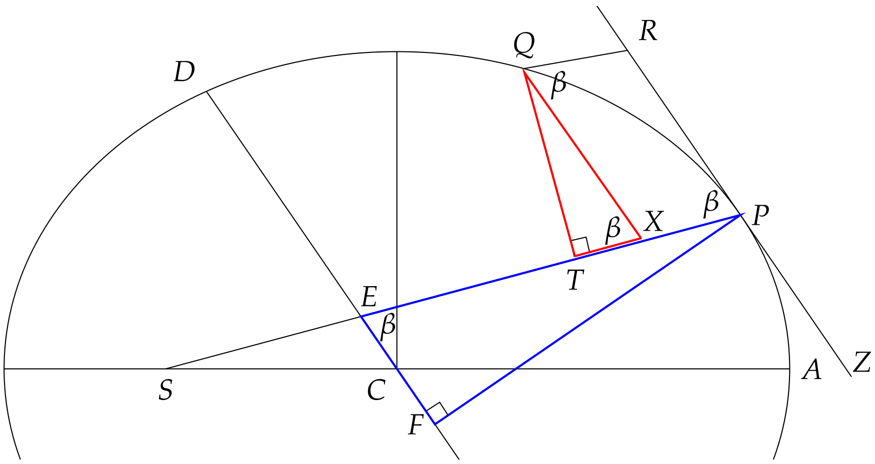


Figure 28: Geometry of an elliptical orbit (2)

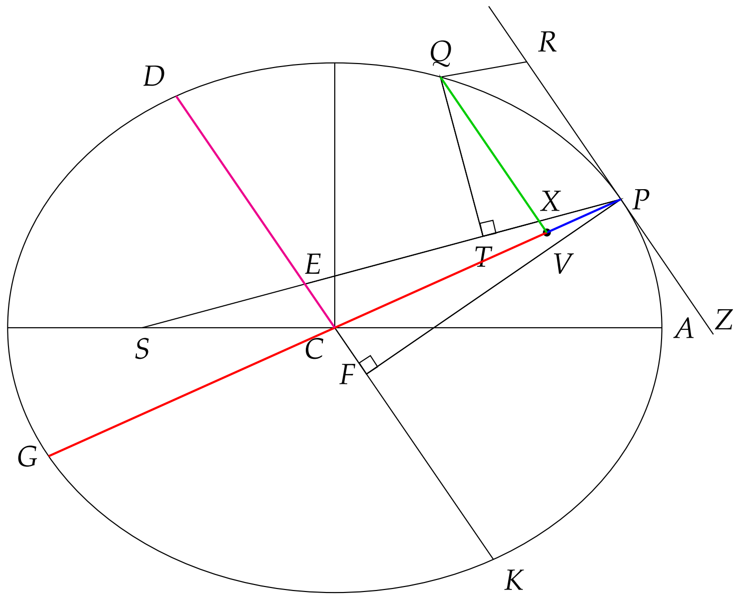


Figure 29: Ratios on conjugate diameters

6.2 A formula for QT

Add a perpendicular from P to DC to the diagram and label its intersection with DC by F (Figure 28).

Theorem 6.2

$$QT = QX \cdot \frac{FP}{CA}.$$

Proof By construction, $QR \parallel PX$, so by alternate interior angles $\angle RQX = \angle QXT = \beta$. By construction, $QX \parallel DC$, so by alternate interior angles $\angle QXT = \angle PEF = \beta$. Since $\triangle PFE$ and $\triangle QTX$ are right triangles with an equal acute angle β, $\triangle PFE \sim \triangle QTX$. In the proof of Theorem 6.1 we showed that $EP = CA$ so

$$\frac{QT}{QX} = \frac{FP}{EP}$$
$$QT = QX \cdot \frac{FP}{EP} = \frac{FP}{CA},$$

∎

6.3 A formula for QR/QT^2

Theorem 6.3

$$\frac{QR}{QT^2} = \frac{CP \cdot CA}{CB^2} \cdot \frac{QV^2}{GV \cdot QX^2}.$$

Proof Let use combine the equations in Theorems 6.1 and 6.2 to get QR/QT^2.

$$\frac{QR}{QT^2} = \frac{PV \cdot \dfrac{CA}{CP}}{\left(QX \cdot \dfrac{FP}{CA}\right)^2} = \frac{PV \cdot CA^3}{QX^2 \cdot CP \cdot FP^2}. \tag{7}$$

DC and PC are conjugate diameters so Theorem 7.11 gives a formula for PV that we substitute into Equation 7.

$$\frac{QR}{QT^2} = \frac{QV^2 \cdot CP^2}{GV \cdot CD^2} \cdot \frac{CA^3}{QX^2 \cdot CP \cdot FP^2} = \frac{CP}{CD^2} \cdot \frac{CA^3}{FP^2} \cdot \frac{QV^2}{GV \cdot QX^2}. \tag{8}$$

Next we show that $CD \cdot FP = CA \cdot CB$. Theorem 7.13 proves that the areas of the parallelograms formed by the tangents to conjugate diameters are equal (Figure 29). By symmetry the areas of the four small parallelograms are equal, as are the triangles formed by constructing corresponding diagonals. In Figure 30 the area of the $\triangle ABC$ (red), which is $(1/2)AC \cdot AB$, is equal to the area of $\triangle PCD$ (blue), which is $(1/2)FP \cdot CD$. Substituting for $CD \cdot FP$ in Equation 8 gives

$$\frac{QR}{QT^2} = \frac{CP}{CD^2} \cdot \frac{CA^3}{FP^2} \cdot \frac{QV^2}{GV \cdot QX^2}$$

$$\frac{QR}{QT^2} = \frac{CP \cdot CA^3}{CB^2 \cdot CA^2} \cdot \frac{QV^2}{GV \cdot QX^2} = \frac{CP \cdot CA}{CB^2} \cdot \frac{QV^2}{GV \cdot QX^2}. \tag{9}$$

∎

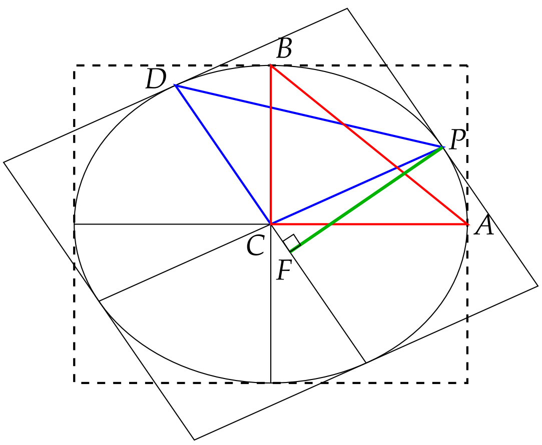


Figure 30: Parallelograms formed by conjugate diameters

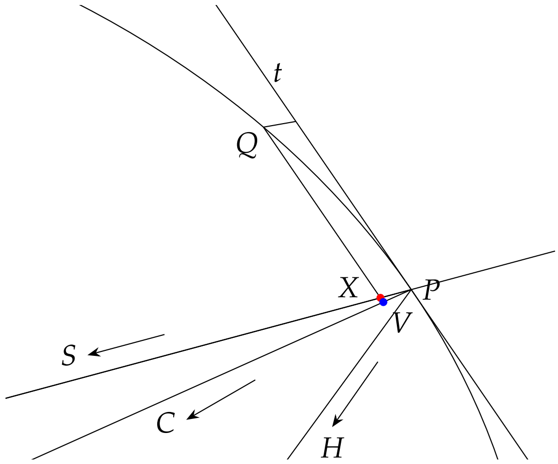


Figure 31: Geometry of an elliptical orbit (3)

6.4 Approaching the limit

Figure 31 is an enlarged diagram of part of Figure 29. Compare the two diagrams and you will see that as the time interval Δt gets smaller, $Q \to P$ which implies that

- $X \to V$ so that $QX \to QV$.
- $V \to P$ so that $CV \to CP$ and hence $GV \to 2CP$.
- In the limit $QX = QV$ and $GV = 2CP$. Substituting into Equation 9 gives

$$\frac{QR}{QT^2} = \frac{CP \cdot CA}{CB^2} \cdot \frac{QV^2}{GV \cdot QX^2} = \frac{CP \cdot CA}{CB^2} \cdot \frac{QX^2}{2CP \cdot QX^2} = \frac{CA}{2CB^2} = \frac{a}{2b^2} = \frac{1}{L},$$

using the result of Theorem 7.6 for the length of the latus rectum.

7 Ellipses

7.1 Fundamental properties

Definition 7.1 (Ellipse)

- *Let F_1 and F_2 be two points in the plane such that $F_1F_2 = 2c \geq 0$ and choose a such that $2a > 2c$ (Figure 32). An* ellipse *is the geometric locus of all points P such that $F_1P + F_2P = 2a$. If $c = 0$ the geometric locus is a* circle.
- *Construct AB through F_1F_2, where A, B are the intersections of the line with the ellipse. AB is the* major axis *of the ellipse. Let O be the midpoint of F_1F_2. AO and OB are the* semi-major axes *of the ellipse.*
- *Construct the perpendicular to AB at O and let C, D be its intersections with the ellipse. CD is the* minor axis *of the ellipse and CO and OD are the* semi-minor axes *of the ellipse.*

Theorem 7.2 *(Figure 33)*

1. $F_1C = F_2C = a$.
2. $AO = OB = a$.
3. $CO = OD$. *(Label $CO = OD$ by b.)*

Proof

1. $\triangle F_1CO \cong \triangle F_2CO$ by side-angle-side so $F_1C = F_2C$. Since C is on the ellipse, $F_1C + F_2C = 2a$ and $F_1C = F_2C = a$ follows.
2. Since A is on the ellipse,
$$2a = AF_1 + AF_2 = (AO - c) + (AO + c) = 2AO\,,$$
so $AO = a$. $OB = a = AO$ follows by symmetry.

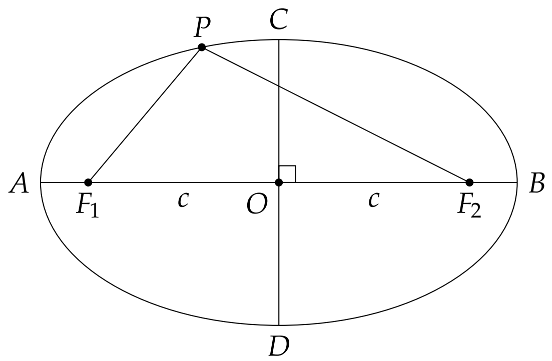


Figure 32: The definition of an ellipse

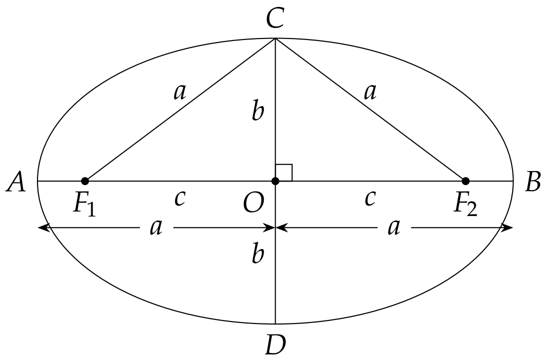


Figure 33: The semi-major and semi-minor axes of an ellipse

3. $CO = OD$ follows since $\triangle F_1CO \cong \triangle F_1DO$. ∎

a and b will always denote the semi-major and semi-minor axes of an ellipse.

Theorem 7.3 *A point $P = (x, y)$ on an ellipse satisfies the equation*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\,.$$

Proof Since $F_1 = (-c, 0)$, $F_2 = (c, 0)$ and $PF_1 + PF_2 = 2a$,

$$PF_1 + PF_2 = \sqrt{(x - (-c))^2 + y^2} + \sqrt{(x - c)^2 + y^2} = 2a\,.$$

Squaring twice results in

$$\begin{aligned}
(x + c)^2 + y^2 &= \left(2a - \sqrt{(x - c)^2 + y^2}\right)^2 \\
4xc &= 4a^2 - 4a\sqrt{(x - c)^2 + y^2} \\
a - \frac{c}{a}x &= \sqrt{(x - c)^2 + y^2} \\
a^2 + \frac{c^2}{a^2}x^2 &= x^2 + c^2 + y^2 \\
\frac{x^2}{a^2} + \frac{y^2}{c^2 - a^2} &= \frac{c^2 - a^2}{c^2 - a^2} = 1\,.
\end{aligned}$$

By Theorem 7.2 and Pythagoras' theorem, $b^2 = a^2 - c^2$ so

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\,.$$

∎

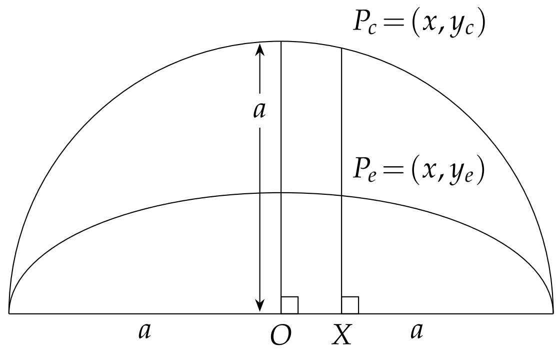


Figure 34: A circle circumscribing an ellipse

7.2 A circle circumscribing an ellipse

Consider a circle of radius a with the same center as an ellipse (Figure 34). Choose a point X on the major axis and construct a perpendicular through $X = (x, 0)$. Let its intersections with the ellipse and the circle be $P_e = (x, y_e)$ and $P_c = (x, y_c)$, respectively.

Theorem 7.4 *The perpendicular to the major axis through a point $P_c = (x, y_c)$ on the circle circumscribing an ellipse intersects the ellipse at $P_e = (x, y_e) = \left(x, \frac{b}{a} y_c\right)$.*

Proof Using the formulas for points on an ellipse and on a circle we have

$$\frac{x^2}{a^2} + \frac{y_e^2}{b^2} = 1$$

$$y_e = \frac{b}{a}\sqrt{(a^2 - x^2)} = \frac{b}{a} y_c \,. \tag{10}$$

∎

7.3 The latus rectum of an ellipse

Definition 7.5 *Consider a line through a focus of an ellipse that is perpendicular the major axis. Let its intersections with the ellipse be L_1, L_2. $L = L_1L_2$ is a* latus rectum[7] *of an ellipse (Figure 35).*

Theorem 7.6 *L, the length of the latus rectum of an ellipse, is $\frac{2b^2}{a}$.*

Proof By Equation 10 and Pythagoras' theorem,

$$L = 2L_1 = 2 \cdot \frac{b}{a}\sqrt{a^2 - c^2} = \frac{2b^2}{a} \,.$$

∎

[7] Normally, points are denoted by upper-case letters and line segments or lengths by lower-case letters, but L for the latus rectum is the standard notation.

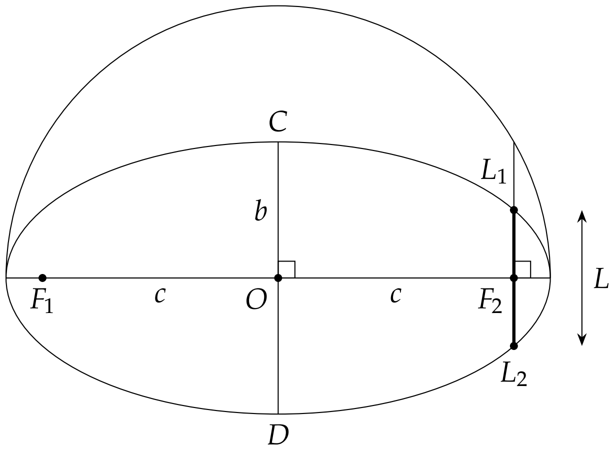


Figure 35: The latus rectum of an ellipse

7.4 The area of an ellipse

Theorem 7.7 *The area of an ellipse is πab.*

Proof From Equation 10

$$y_e = \frac{b}{a}\sqrt{a^2 - x^2}\,,$$

so the area of an ellipse is

$$A_e = 2\int_{-a}^{a} \frac{b}{a}\sqrt{a^2 - x^2}\, dx = 2\frac{b}{a}\int_{-a}^{a} \sqrt{a^2 - x^2}\, dx = \frac{b}{a}A_c\,.$$

If we can show that the area of a circle is πa^2 the theorem follows.

The proof uses polar coordinates, where $x = a\cos\theta$ and $y = a\sin\theta$. First, we derive the formula for the integral of $\sin^2\theta$ using the double-angle identity.

$$\begin{aligned}
\sin^2\theta &= \frac{1-\cos 2\theta}{2} \\
\int \sin^2\theta\, d\theta &= \int \frac{1-\cos 2\theta}{2}\, d\theta = \frac{\theta}{2} - \frac{\sin 2\theta}{4} + C\,.
\end{aligned}$$

Now we can compute the area of a circle as twice the area of a semicircle by changing from Cartesian to polar coordinates and integrating.

$$\begin{aligned}
A_c &= 2\int_{-a}^{a} \sqrt{a^2 - x^2}\, dx = 2\int_{-\pi}^{0} \sqrt{a^2 - (a\cos\theta)^2}\; d(a\cos\theta) \\
&= 2\cdot a\cdot a\int_{-\pi}^{0} \sin\theta(-\sin\theta)\, d\theta = -2a^2\int_{-\pi}^{0} \sin^2\theta\, d\theta \\
&= -2a^2\left(\frac{\theta}{2} - \frac{\sin 2\theta}{4} + C\right)\Bigg|_{-\pi}^{0} = \pi a^2\,.
\end{aligned}$$

∎

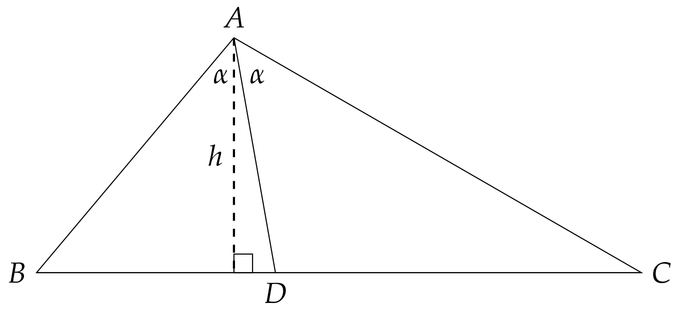


Figure 36: The angle bisector theorem

7.5 The angles between a tangent and the lines to the foci

We will need the following theorem:

Theorem 7.8 (Angle bisector theorem) *In $\triangle ABC$ let the angle bisector of $\angle BAC$ intersect BC at D (Figure 36). Then*

$$\frac{BD}{CD} = \frac{AB}{AC}.$$

Proof We prove the theorem by computing the areas of two triangles in two ways: the base and the height, and the base, the angle and a side.

$$\begin{aligned}
\triangle ABD &= \frac{1}{2} BD\, h = \frac{1}{2} AB\, AD \sin \alpha \\
\frac{BD}{AB} &= \frac{AD \sin \alpha}{h} \\
\triangle ACD &= \frac{1}{2} CD\, h = \frac{1}{2} AC\, AD \sin \alpha \\
\frac{CD}{AC} &= \frac{AD \sin \alpha}{h} \\
\frac{BD}{CD} &= \frac{AB}{AC}.
\end{aligned}$$

∎

Theorem 7.9 *Let P be a point on the ellipse whose foci are S, H. Let PU be the extension of SP such that $SU = AA' = 2a$ (Figure 37). Let RQ be the bisector of $\angle HPU$. Then RQ is the tangent to the ellipse at P and $\angle RPS = \angle HPQ$.*

Proof We prove that any point $Q \neq P$ on the bisector is not on the ellipse; therefore, the bisector RQ has only one point of intersection with the ellipse and must be a tangent. Since RQ is the angle bisector of $\angle HPU$, $\angle HPQ = \angle UPQ = \alpha$, and by vertical angles $\angle RPS = \angle QPU = \alpha$.

Construct the line HU to form the triangle $\triangle HPU$. By construction $SU = 2a$ so $PU = PH$ and by the angle bisector theorem,

$$\frac{UW}{HW} = \frac{PU}{PH} = 1\,.$$

$\triangle UWQ \cong \triangle HWQ$ by side-angle-side so $UQ = HQ$. Suppose that Q is on the ellipse. Then $2a = SQ + HQ = SQ + QU$. By the triangle inequality $2a = SQ + QU > SU = 2a$, contradicting that Q is on the ellipse. ∎

7.6 The parametric representation of an ellipse

Figure 38 shows an ellipse and two circles: one whose radius is the length of the semi-major axis (dotted red) and one whose radius is the semi-minor axis (dashed blue). The figure shows the *parametric representation* of a point $P = (x, y)$ on the ellipse:

$$(x, y) = (a \cos t,\; y = b \sin t)\,.$$

The parameter t is *not* the angle of P relative to the positive x-axis. It is constructed by projecting P onto the minor axis. Then the line at angle t is constructed from C to the intersection of the projection with the inner circle (blue dot). The line is extended to intersect the outer circle (red dot). The parametric representation of P is computed by multiplying the lengths of the axes by the appropriate trigonometry function.

7.7 Ratios

Theorem 7.10 *Let $P = (x, y)$ be a point on an ellipse (not on the major axis AA') and drop a perpendicular PV from P to the major axis (Figure 42). Then*

$$\frac{A'V \cdot AV}{PV^2} = \frac{a^2}{b^2}\,.$$

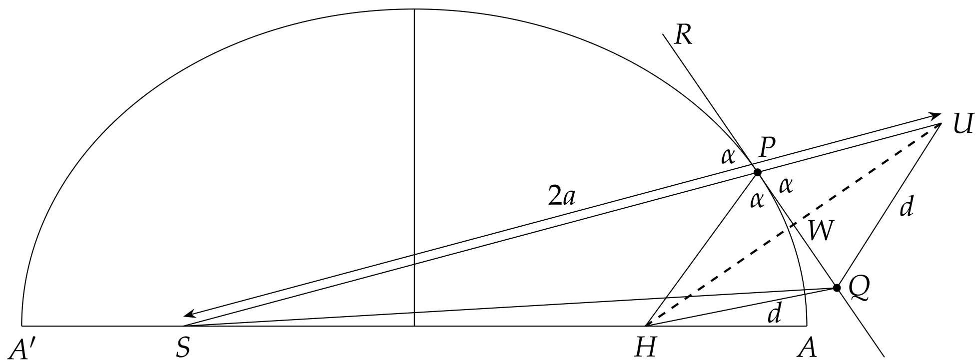


Figure 37: Angles at the tangent

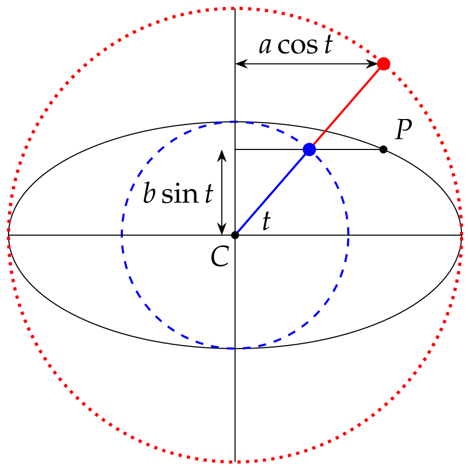


Figure 38: Parametric representation of an ellipse

Proof By Equation 10),

$$
\begin{aligned}
y^2 &= b^2 \cdot \left(1 - \frac{x^2}{a^2}\right) = \frac{b^2(a^2 - x^2)}{a^2} \\
\frac{A'V \cdot AV}{PV^2} &= \frac{(a+x)(a-x)}{y^2} = \frac{a^2(a^2 - x^2)}{b^2(a^2 - x^2)} = \frac{a^2}{b^2}.
\end{aligned}
$$

∎

Theorem 7.11 *Let PG, DK be conjugate diameters of an ellipse and let Q be a point on the ellipse (Figure 39). Drop a perpendicular QV from Q to the major axis, then*

$$PV = \frac{QV^2 \cdot CP^2}{GV \cdot CD^2}.$$

Proof Figure 39 shows a dashed ellipse which is the original ellipse rotated about the same center C, so that CP is the semi-major axis and CD is the semi-minor axis. By Theorem 7.10

$$\frac{GV \cdot PV}{QV^2} = \frac{a'^2}{b'^2},$$

where a', b' are the lengths of the semi-major and semi-minor axes of the rotated ellipse. By construction $a' = CP$ and $b' = CD$ so

$$
\begin{aligned}
\frac{GV \cdot PV}{QV^2} &= \frac{CP^2}{CD^2} \\
PV &= \frac{QV^2 \cdot CP^2}{GV \cdot CD^2}.
\end{aligned}
$$

∎

7.8 Parallelograms formed by tangents to conjugate diameters

There are two equivalent definitions of conjugate diameters.

Definition 7.12

- *Let P be a point on an ellipse, PG a diameter and let t be the tangent to the ellipse at P. Diameter DK is a* conjugate diameter *if it is parallel to t (Figure 40).*
- *Two diameters PG and DK are* conjugate diameters *if the midpoints of chords $(D'K', D''K'')$ parallel to one diameter (DK) lie on another diameter (PG).*

Theorem 7.13 *Parallelograms formed by tangents to the intersections of conjugate diameters with the ellipse are equal (Figure 41).*

Proof We will show that the area of the parallelogram $JKLM$ is equal to the area of the parallelogram formed by tangents to the major and minor axes (dashed rectangle). By symmetry it suffices to prove that the areas of one of the four quadrants of those parallelograms are equal: $A_{ACBC'} = A_{PCDJ}$. Furthermore, since the diagonals bisect a parallelogram, it suffices to prove that that the area of $\triangle ABC$ (red) equals the area of $\triangle PCD$ (blue).

Let $P = (x_p, y_p) = (a\cos t, b\sin t)$ and $D = (x_d, y_d)$ be the parametric representations of the points on the ellipse. Since conjugate diameters are perpendicular, $\angle DCP$ is a right angle so

$$D = (x_d, y_d) = (a\cos(t+\pi/2), b\sin(t+\pi/2)) = (-a\sin t, b\cos t)\,.$$

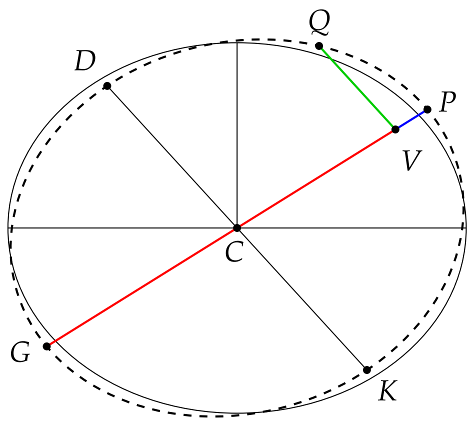


Figure 39: Ratios on conjugate diameters

Construct $DD' = (x_d, 0)$ and $PP' = (x_p, 0)$ perpendicular to the major axis. The area of $\triangle PCD$ can be computed as the area of the trapezoid $P'PDD'$ minus the areas of the triangles $\triangle D'DC, \triangle P'PC$.

$$\triangle PCD = \frac{y_p + y_d}{2}(x_p - x_d) - \frac{1}{2}x_d y_d - \frac{1}{2}x_p y_p = \frac{1}{2}\left(x_p y_d - x_d y_p\right)$$

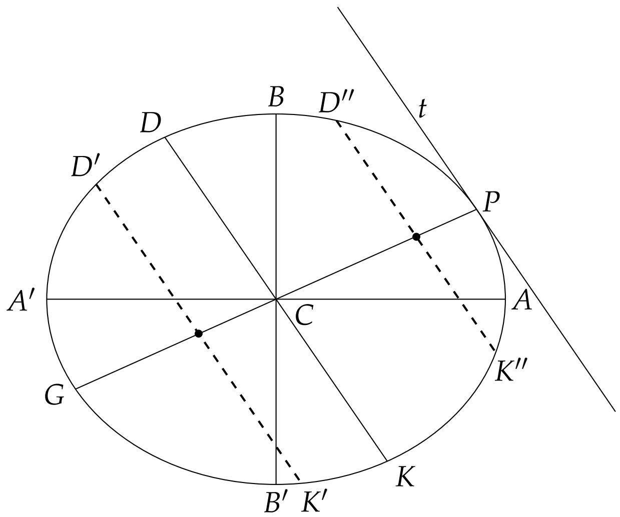


Figure 40: A conjugate diameter

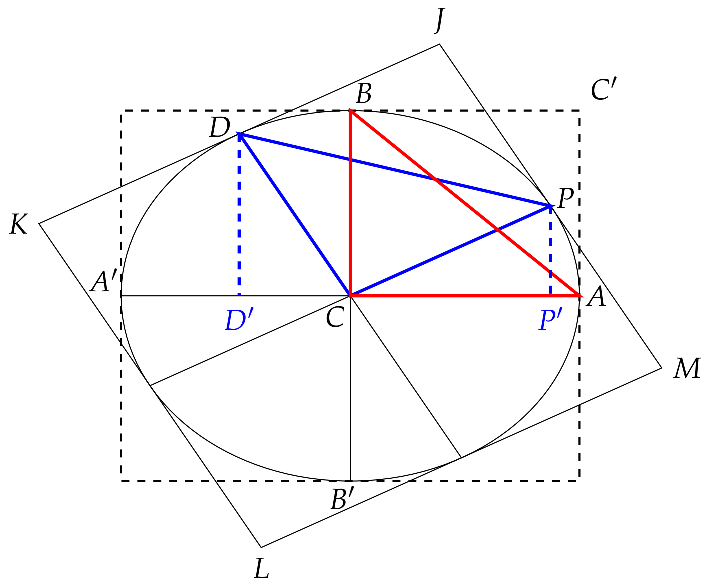


Figure 41: Parallelograms formed by conjugate diameters

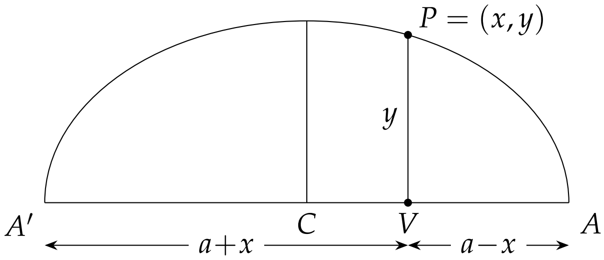


Figure 42: Ratios on conjugate diameters

$$= \frac{1}{2}\left(a\cos t \cdot b\cos t - (-a)\sin t \cdot b\sin t\right) = \frac{1}{2}ab = \triangle ABC\,.$$

∎

8 Ellipses in Euclidean geometry

The proofs of theorems about planetary freely used analytic geometry and trigonometry, but for many years after the invention of analytic geometry, mathematicians continued to limit themselves to Euclidean geometry. Even into the late nineteenth century, students continued to study the original works of Euclid and Newton. In this section, I present proofs in Euclidean geometry (*geometric proofs*) of theorems that appeared in the previous sections.

The section starts with geometric proofs of several theorems on triangles that will be needed below. The next subsection defines ellipses in terms of the geometric concepts of focus and directrix instead of the familiar analytic definition, points which satisfy the following equation for $0 < a, b$:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\,.$$

This is followed by a theorem on focal chords and then a subsection on constructing a right angle at the focus of an allipse.

8.1 Theorems on triangles

The median of the hypotenuse

Theorem 8.1 *Let ACB be a right triangle and let D be the intersection of the median from C and the hypotenuse AC (Figure 43). Then $CD = AB/2$ and*

$$AD \cdot DB = DC^2\,.$$

Proof Drop the perpendicular from D to BC and label the intersection by E.

By similar triangles,

$$\frac{a}{c_2} = \frac{a+a}{c_1+c_2}$$

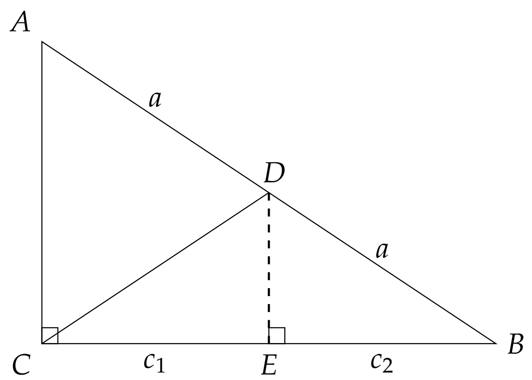


Figure 43: Median to the hypotenuse

$$c_1 = c_2 .$$

Therefore, $\triangle DEC \cong \triangle DEB$ and $DC = DB$. ∎

Adjacent pairs of similar triangles

I have not encountered the following definition before but it explicitly expresses a relation among similar triangles that would have been obvious to geometers.

Definition 8.2 *An* adjacent pair of similar triangles *is a pair of (a pair of) similar triangles that share sides (Figure 44). $\triangle BAC \sim \triangle EAF$ and $\triangle CAD \sim \triangle FAG$ are an adjacent pair of similar triangles.*

Theorem 8.3 *For the adjacent pair of similar triangles in Figure 44,*

$$\frac{AB}{AE} = \frac{AD}{AG} .$$

Proof By similar triangles,

$$\begin{aligned} \frac{AB}{AE} &= \frac{AC}{AF} \\ \frac{AC}{AF} &= \frac{AD}{AG} \\ \frac{AB}{AE} &= \frac{AD}{AG} . \end{aligned}$$

∎

Similar ratios hold between other sides of $\triangle BAC$ and $\triangle CAD$ by using an intermediate step with AC. We will use the term *by an adjacent pair of similar triangles* and leave it to the reader to make the intermediate step.

The angle bisector theorems

The angle bisector theorems are used extensively in the study of conic systems, often implicitly.

Theorem 8.4 (Interior angle bisector theorem) *In $\triangle ABC$ let D be a point on BC (Figure 45). Then AD bisects $\angle CAB$* if and only if

$$\frac{BD}{CD} = \frac{AB}{AC} .$$

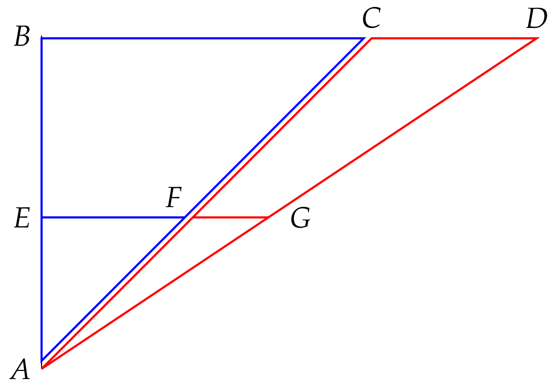


Figure 44: Adjacent pairs of similar triangles

Proof Suppose that AD bisects $\angle BAC$. Construct a line through C parallel to AB and let its intersection with AD be E. By alternate interior angles, $\angle BAD = \angle CED =$ and by vertical angles $\angle BDA = \angle CDE$. Therefore, $\triangle ABD \sim \triangle EDC$ so

$$\frac{BD}{CD} = \frac{AB}{CE}\,.$$

$\triangle ECA$ is isoceles so $CE = AC$ and

$$\frac{BD}{CD} = \frac{AB}{AC}\,.$$

To prove the converse just "run" the proof backwards. ∎

Theorem 8.5 (Exterior angle bisector theorem) *In $\triangle ABC$ let D be a point on the extension of CB outside the triangle (Figure 46). Then AD bisects the exterior angle of $\angle BAC$ if and only if*

$$\frac{BD}{CD} = \frac{AB}{AC}\,.$$

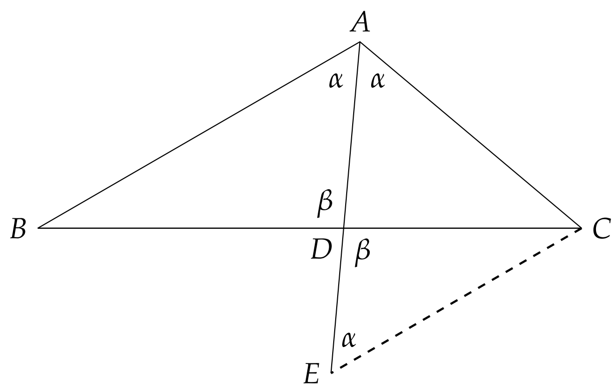


Figure 45: The interior angle bisector theorem

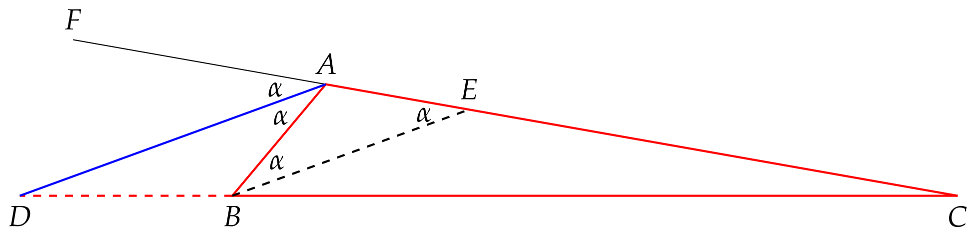


Figure 46: The exterior angle bisector theorem

Proof Suppose that AD bisects BAF. Construct a line through B parallel to AD and let its intersection with AC be E. By alternate interior angles $\angle BAD = \angle ABE$ and by corresponding angles $\angle FAD = \angle AEB$. Therefore, $\triangle BCE \sim \triangle DCA$ so

$$\frac{BD}{CD} = \frac{AE}{AC}.$$

But $\triangle BAE$ is isoceles so $AE = AB$ and

$$\frac{BD}{CD} = \frac{AB}{AC}.$$

To prove the converse just "run" the proof backwards. ∎

The exterior angle bisector theoren confusing to understand in a proof, because it can be hard to identify the components of a diagram. The following color-coding is used below: the triangle is red, the extension of one side is dashed red and the bisector is blue.

8.2 The definition of an ellipse using the focus and the directrix

Definition 8.6 *Let d be a line (the* directrix*) and S be a point (the* focus*) not on the directrix. Let $0 < e < 1$ be a number (the* eccentricity*). An* ellipse *is the geometric locus of points P such that the ratio of PS to the distance of P to the directrix is e.*

All the conic sections (parabolas, ellipses and hyperbolas) are defined the same way and are distinguished by their eccentricity. You might be familiar with the definition of the parabola where $e = 1$ (Figure 47).

Definition 8.7 *Let X be the intersection of the perpendicular to the directrix from S. A on SX is a* vertex *of the ellipse if $SA/AX = e$ (in Figure 48, $e = 1/2$).*

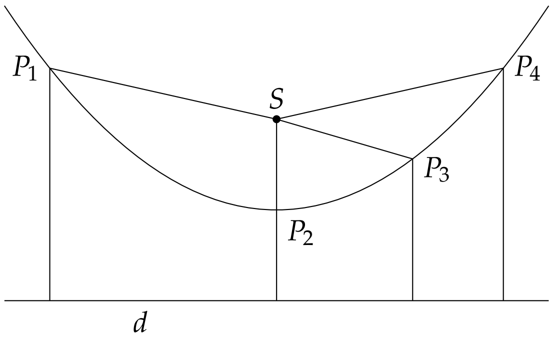


Figure 47: A parabola defined by the focus and the directrix

Constructing the vertices

Given $e = SA/AX$ we can locate A as follows:

$$\begin{aligned} SA + AX &= SX \\ SA &= SX \cdot \frac{e}{1+e} \\ AX &= SX \cdot \frac{1}{1+e} \,. \end{aligned}$$

Similarly, there is a second vertex A' where

$$\begin{aligned} A'X - SA' &= SX \\ SA' &= SX \cdot \frac{1}{1-e} \\ A'X &= SX \cdot \frac{2-e}{1-e} \,. \end{aligned}$$

Constructing points on an ellipse

Select an *arbitrary* point E on the directrix and construct lines from E through A and S. The line through S will make an angle α with SX. Construct a line from S at the *same angle*

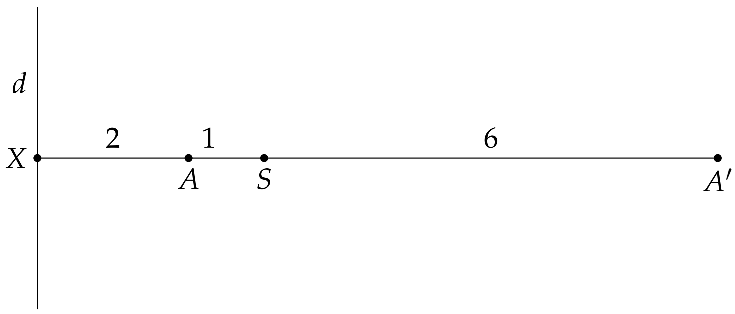


Figure 48: The elements of the definition of an ellipse

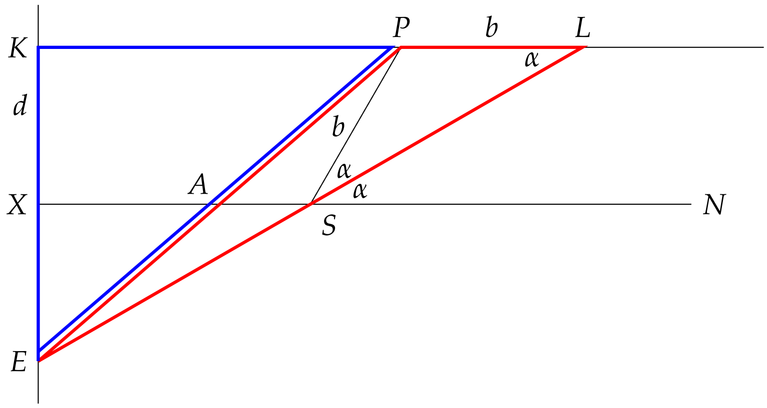


Figure 49: Constructing points on the ellipse

α from ES and let its intersection with EA be P. Construct the perpendicular from P to the directrix and let K be its intersection with the directrix. Let L be the intersection of KP with ES (Figure 49).

Theorem 8.8 *The point P is on the ellipse.*

Proof $\angle PLS = \angle LSN = \alpha$ by alternate interior angles, so $\triangle LPS$ is isoceles and $PL = SP$. Since $PK \parallel SX$, $\triangle XEA \sim \triangle KEP$ and $\triangle AES \sim \triangle PEL$ are adjacent pairs of similar triangles, so

$$\frac{PS}{PK} = \frac{PL}{PK} = \frac{SA}{AX} = e\,.$$

Therefore, P is on the ellipse. ∎

We now construct another point on KP. Construct a line from F on the directrix through S such that $\angle FSK = \angle KSP = \beta$ and let P' be the intersection of FS with KP (Figure 50).

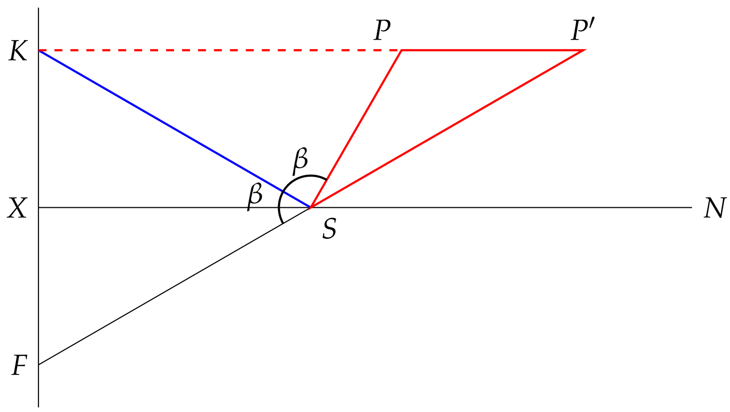


Figure 50: Constructing another vertex on KP

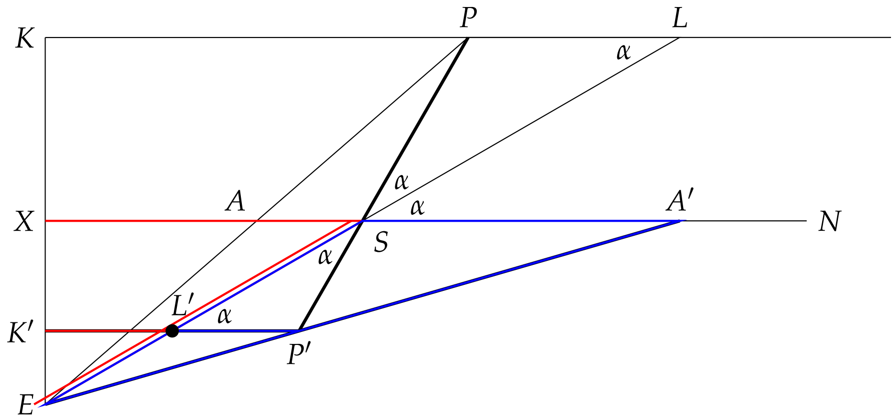


Figure 51: Constructing a focal chord

Theorem 8.9 *P' is on the ellipse.*

Proof Since KS bisects $\angle PSF$, by the exterior angle bisector theorem

$$\frac{SP'}{SP} = \frac{P'K}{PK}$$
$$\frac{SP'}{P'K} = \frac{SP}{PK} = e\,,$$

and therefore P is on the ellipse. ∎

8.3 Constructing a focal chord

Definition 8.10 *Let P, P' be points on an ellipse. PP' is a* focal chord *of the ellipse if PP' is a chord through a focus (Figure 51).*

Let us continue with the construction of Figure 49 as shown in Figure 51. Construct the second vertex A'. Construct EA' and label its intersection with SP by P'. Construct a perpendicular from P' to the directrix and label its intersection with the directrix by K'. Label the intersection of $P'K'$ and EL by L'.

Theorem 8.11 *The point P' is on the ellipse and since PP' intersects S it is a focal chord.*

Proof $\triangle SEA' \sim \triangle L'EP'$ and $\triangle XES \sim K'EL'$ are adjacent similar triangles, so

$$\frac{P'L'}{P'K'} = \frac{SA'}{A'X}\,.$$

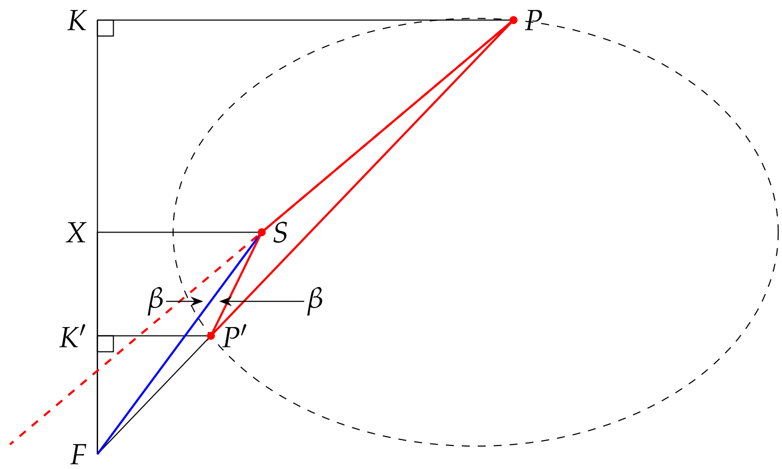


Figure 52: Bisecting the angle at the focus

By alternate interior angles and vertical angles we have $\angle SL'P' = \angle L'SP'$ so triangle $\triangle SL'P'$ is isoceles and $P'L' = SP'$. Therefore,

$$\frac{SP'}{P'K'} = \frac{SA'}{A'X} = e\,,$$

and P' is a point on the ellipse. ∎

8.4 A right angle at the focus of an ellipse

Theorem 8.12 *Let P, P' be points on the ellipse and let F be the intersection of PP' with the directrix. Then FS bisects the exterior angle of $\angle P'SP$(Figure 52).*

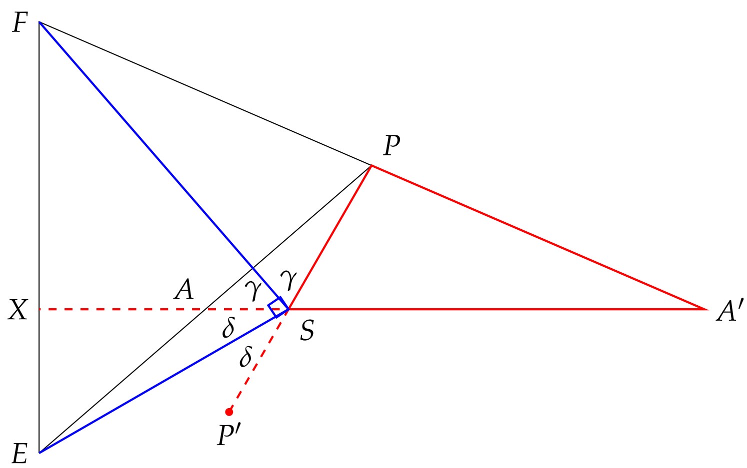


Figure 53: The right angle at the focus

Proof Since P, P' are on the ellipse

$$\frac{SP}{PK} = \frac{SP'}{P'K'} = e\,,$$

and since $\triangle PFK \sim P'FK'$,

$$\frac{SP}{SP'} = \frac{PK}{P'K'} = \frac{PF}{P'F}\,.$$

By the exterior angle bisector theorem FS bisects the exterior angle of $\angle P'SP$. ∎

Theorem 8.13 *Let P be a point on the ellipse and construct lines PA, PA'. Label their intersections with the directrix by E and F, respectively. Then $\angle FSE$ is a right angle (Figure 52).*

Proof P, A, A' are all points on the ellipse so Theorem 8.12 applies. FS bisects $\angle PSA = 2\gamma$ and ES bisects $\angle P'SA = 2\delta$, so $2\gamma + 2\delta = 180°$ and $\angle FSE = \gamma + \delta = 90°$. ∎

Further reading

Sections 2–5 based primarily on Hahn's book [4]. He has written a more advanced book on the orbits of planets and spacecraft [5]. Section 6 is based on [3]. Two additional articles present this aspect of Newton's work [6, 7]. Should you wish to read Newton's work, [2] is an up-to-date translation into English and Cohen's lengthy Guide will enable you to follow Newton's often terse presentation. The relevant sections of the Guide are 10.8–10.10. Section 7 on ellipses is based on [8], while Section 8 is based on[1].

References

[1] W. H. Besant. *Conic Sections, Treated Geometrically (Ninth Edition Revised and Enlarged)*. George Bell and Sons, London, 1895. `https://www.gutenberg.org/ebooks/29913` and `https://archive.org/details/cu31924059322481`.

[2] I. Bernard Cohen, Anne Whitman, and Julia Budenz. *The Principia: Mathematical Principles of Natural Philosophy*. University of California Press, 1999. Preceded by *A Guide to Newton's Principia* by I. Bernard Cohen.

[3] Graham Griffiths. The inverse square law of gravitation: An alternative to Newton's derivation. `https://www.researchgate.net/publication/264978661_The_Inverse_Square_Law_of_Gravitation_An_Alternative_to_Newton's_Derivation`, 2009.

[4] Alexander J. Hahn. *Calculus in Context: Background, Basics, and Applications*. Johns Hopkins University Press, 2017.

[5] Alexander J. Hahn. *Basic Calculus of Planetary Orbits and Interplanetary Flight: The Missions of the Voyagers, Cassini, and Juno*. Springer, 2020.

[6] Kai Hauser and Reinhard Lang. On the geometrical and physical meaning of Newton's solution to Kepler's problem. *The Mathematical Intelligencer*, 25(4):35–44, 2003.

[7] S. K. Stein. Exactly how did Newton deal with his planets? *The Mathematical Intelligencer*, 18(2):6–11, 1996.

[8] Wikipedia contributors. Ellipse — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Ellipse&oldid=1168583658`, 2023. [Online; accessed 12-August-2023].