LAW I: THE ELLIPSE LAW (1609)

1. Modern form
The orbit of a planet around the Sun is an ellipse with the Sun at one focus.

2. Kepler stated his result in *Astronomia nova* (1609), in Chapter 58 (KGW III p.366 line 4; Donahue p.575): 'an ellipse is the path of the planet [Mars]'.

3. Background for Kepler's approach
Figure (a) shows the planet at a position P on its path CFD whose (major) diameter is CD, circumscribed by the circle diameter CD centre B. Q is a typical point of that circle, determined by the angle $\angle QBC = \beta$ at the centre, with QPH the ordinate linking the associated points Q and P. Kepler established that the path was an ellipse by finding a (geometrical) formula for the distance AP of the planet P from the Sun at A, whose position Kepler described as 'the eccentric point' (the word 'eccentric' is used in astronomy in the sense of 'off-centre'). This point A was not then known to possess any special properties as far as the ellipse was concerned. We call AB the eccentric distance, and when we introduce symbolic notation for convenience we shall express the length AB as a proportion $\boldsymbol{ae}$ of the radius $\boldsymbol{a}$ of the circle; this length was defined by (frequent) astronomical observations of the distances from the extreme points, or apsides, C and D, so (because B is the midpoint of CD) we have:

$$AB = ae = \tfrac{1}{2}(CA - DA).$$

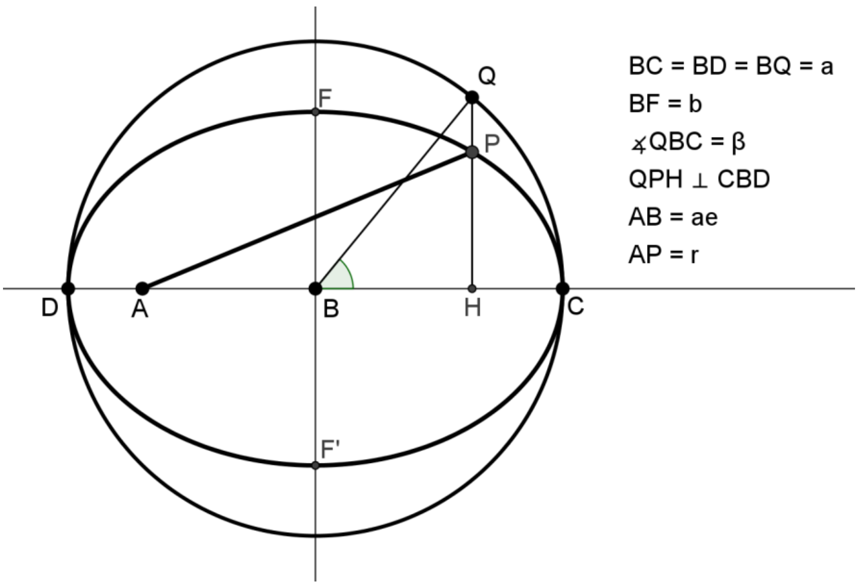

Figure (a)

Note that the eccentric distance has to be greatly exaggerated in every figure, because the separation between the ellipse and the circle is undetectable on the scale of a printed page – which is why discovery of the ellipse could not have happened before Tycho Brahe (1546-1601) provided observations that were accurate enough – and then only in the case of Mars.

Kepler defined the ellipse he (eventually) found by the ratio-property of the ordinates which Archimedes had stated in his work *On Conoids and Spheroids*, Prop.4:

$$\frac{PH}{QH} = \frac{BF}{BC} = \frac{b}{a}.$$

(This definition obviously does not involve the focus.)

4. How Kepler discovered the path of Mars (References from *Astronomia nova*, KGW III)
From his student days Kepler enthusiastically adopted the Sun-centred theory of Copernicus (1473-1543), and applied it rigorously by insisting that all planetary distances should be measured from the Sun itself. After much preliminary investigation, which began around 1601, Kepler started (in Chapters 39 and 40) from the framework he had derived from

Ptolemy (second century CE), adapted to a heliocentric view. This simple epicycle model depicted the path of Mars initially as the circle *CQD*, of radius *a*, whose centre B was eccentric to the Sun at A, as shown in Figure (b). (The line AF has been added, where F is the point of intersection of the circle centre A with the perpendicular to the major axis through B.)

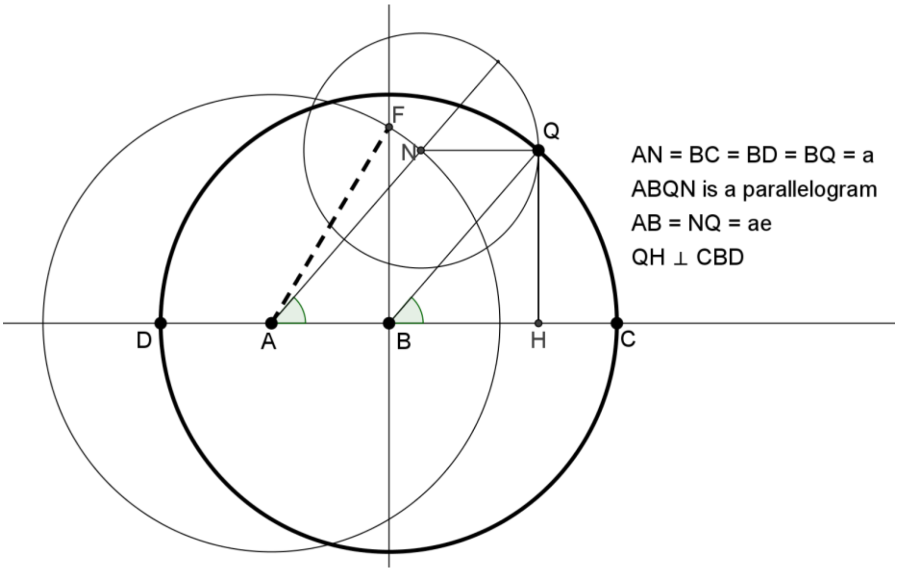

Figure (b)

Kepler had inherited a superb set of naked-eye observations from Tycho Brahe and he used them as a criterion against which to assess his geometrical theory at every stage. Though he found that putting the planet at Q made the distances from A generally too long, he continued to work from this traditional framework, using straightedge and compasses – the only geometrical tools permitted by Euclid. Next, a series of investigations (Chapters 45-50) ended in a second attempt which gave distances that were too short. Kepler eventually managed, in Chapter 56, to apply his characteristic geometrical method to construct distances that fell more-or-less in the middle, by drawing a perpendicular from Q to meet AN extended at K, as shown in Figure (c). This produced a typical distance of the correct length AK along AN, swung round in a circular arc to get it in the right direction (AP). (The resulting point P was satisfactory because the planet was then positioned as precisely as could be achieved in relation to the accuracy of the observations.)

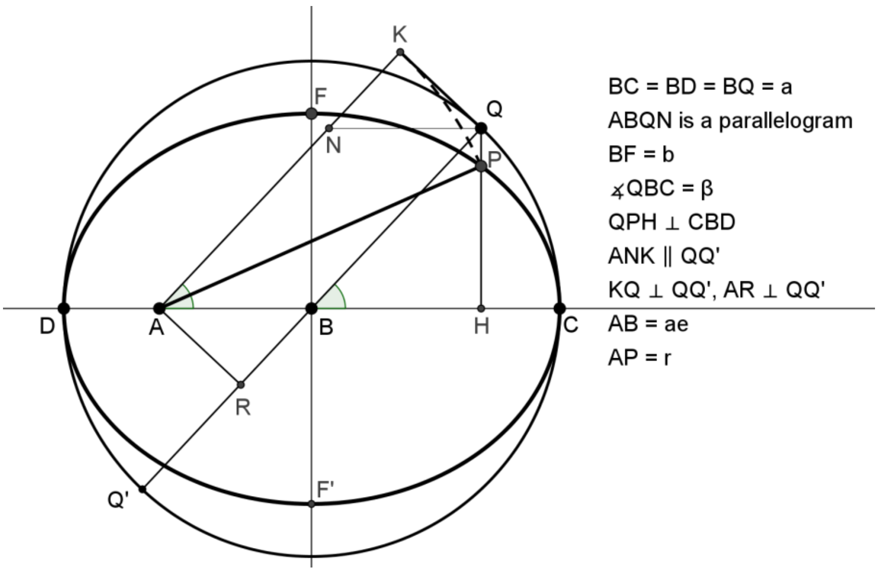

Figure (c)

Thus Kepler's characteristic construction produced AP = AK, and also we have AK = QR because of the rectangle ARQK. Hence Kepler found that the position of the planet was given by:

$$\boldsymbol{AP = AK = QR = QB + BR}\,.$$

Kepler proved this, actually in these geometrical terms. (Symbolic notation had hardly been developed in that era, and Kepler did not use it.) However, for the convenience of readers, the result can be expressed in modern terms as:

$$\boldsymbol{AP = \ r = a + ae\ cos\beta}.$$

Even nowadays, this particular formula for the radius vector is seldom recognized as representing an ellipse – because it is expressed in terms of the angle β at the centre – though we shall prove in the next section that this representation is entirely valid and exact. However, Kepler did not have the faintest idea what curve his invented construction had produced, until he suddenly thought of trying to identify the unknown curve with the ellipse defined by Archimedes. When this turned out to be successful, Kepler provided the mathematical proof.

5. Kepler's proof of the result (transcribed for modern readers)

In Chapter 56 Kepler specified the position of A on CD by establishing that the important distance from the Sun (shown in Figure (c)) to the point F of the ellipse lying at the end of the perpendicular axis (through B) was $\boldsymbol{AF = a}$. Then we apply Pythagoras' theorem to the right-angled triangle AFB shown in Figure (b), to obtain:

$$\boldsymbol{AB^2 = \ AF^2 - \ BF^2}\,,$$

or alternatively, in modern notation, when we call the minor semi-axis $\boldsymbol{BF = b}$,

$$\boldsymbol{AB^2 = \ a^2e^2 = \ a^2 - \ b^2}\,. \qquad \textbf{(1)}$$

Kepler confirmed this identity in Chapter 59, Protheorema VII: it was as much as he needed to know about the position of A. Indeed A was not identified as the focus of the ellipse until a few years later – we do not know when, but Kepler stated the fact in 1621 in *Epitome* Book V, Part I, Section 3 (KGW VII p.372). It is also interesting to notice that the special length $\boldsymbol{AF = a}$ shown in Kepler's own diagram of Chapter 59, appears in Figure (b) as a radius of the circle centred at the Sun, being the (arithmetic) mean of the two extreme (apsidal) distances. Hence the mean distance was geometrically illustrated here ready for application in Law III.

We shall now determine the unknown curve starting from Kepler's characteristic construction which gave a position P for the planet:

$$\boldsymbol{AP = AK}\,.$$

For modern readers we shall apply the method of elementary coordinate geometry which is anachronistic (by half a century or so) – but the statements made in this proof can be identified in Kepler's own proof of Law I (he did not call it that) in Protheorema XI of Chapter 59 (KGW III p.371). This proof depends entirely on the geometry of Euclid, using Pythagoras' theorem and ratios from similar triangles (nowadays we use trigonometry for the same effect). We start with what Kepler had discovered about the unknown curve whose radius vector $r = AP$ was constructed as shown above:

$$\boldsymbol{AP = AK = QR = a + ae \cos\beta}.$$

From ΔPHA, applying Pythagoras' theorem, we have:

$$\boldsymbol{PH^2 = AP^2 - \ AH^2 = AP^2 - (AB + BH)^2}$$

$$PH^2 = (a + ae\cos\beta)^2 - (ae + a\cos\beta)^2$$

$$= a^2(1 + 2e\cos\beta + e^2\cos^2\beta - e^2 - 2e\cos\beta - \cos^2\beta)$$

$$PH^2 = a^2(1 - e^2)(1 - \cos^2\beta) = a^2\sin^2\beta(1 - e^2).$$

Hence,

$$PH = a\sin\beta.\sqrt{(1 - e^2)} = b\sin\beta\ ,$$

using the result (1) from Protheorema VII as set out above.

Now from $\triangle QHA$,

$$QH = a\sin\beta.$$

Thus we obtain:

$$\frac{PH}{QH} = \frac{b}{a}.$$

Hence the path of P is the ellipse defined by the ratio-property of the ordinates stated by Archimedes – and this path was the one produced by Kepler's characteristic construction.

Thus we have provided a sound geometrical proof that an exact ellipse can be constructed to satisfy the observations in the case of Mars, subject to the limits of accuracy imposed by those observations. [Unfortunately it is not possible to establish the modern result that all planetary orbits are ellipses without knowledge of the underlying dynamics, which did not become available until Newton's synthesis. However, Kepler did go on to demonstrate, later, that the orbits of all the other naked-eye planets were compatible with an elliptic orbit.]

For more details of Kepler's method, see:

Kepler's Planetary Laws

A E L Davis
ael.davis@imperial.ac.uk