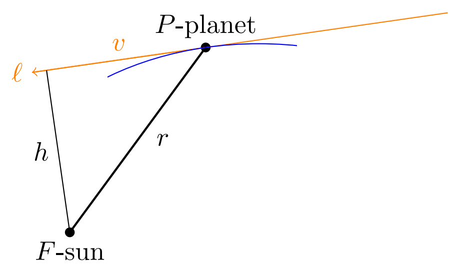
Kepler's laws of planetary motion

Thomas Hales



Erich Vogt used to present a slick proof to his introductory physics students at UBC of Kepler's first law: every planet moves along an elliptical orbit having the sun for a focus.

We consider three variable quantities of a planet along its orbit (in blue): let r be the positive distance from the sun (or more precisely, the center of mass) to the planet, let v be the speed of the planet, and let h be the positive distance from the sun to the tangent line ℓ of the planetary trajectory. The demonstration of Kepler's first law is a one sentence-proof, based on the conservation of energy

$$E = \frac{mv^2}{2} - \frac{GMm}{r} \tag{1}$$

(the first term on the right is the kinetic energy and the second term is the gravitational potential energy) and the conservation of angular momentum

$$L = mvh.$$

The only thing that matters to us about energy E, angular momentum L, the gravitational constant G, and the masses M, m is that they are constants with L, G, M, m positive and E negative (for bounded orbits).

Here is the sentence: solve the angular momentum equation for the variable $v = L/(mh)$ and use it to eliminate v from the energy equation (1), which after dividing both sides by the constant $-E$, takes the form of the so-called pedal equation of an ellipse:

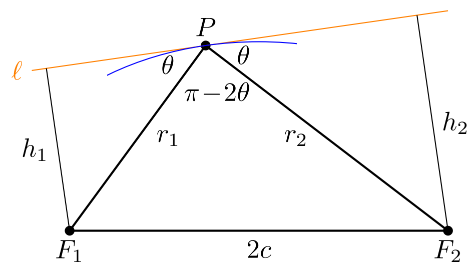
$$-1 = \frac{b^2}{h^2} - \frac{2a}{r}, \tag{2}$$

where the physical constants

$$a = -GMm/(2E), \quad b^2 = -L^2/(2Em) \tag{3}$$

have been consolidated. This rearrangement is purely algebraic. No physics or differential equations are involved. Here a, b are positive constants, while h, r vary along the planetary orbit, related to each other by this pedal equation.

This is a splendid derivation of Kepler's law if you already know the pedal equation for an ellipse. Otherwise, it leaves you wondering what the equation has to do with an ellipse. We give an elementary proof that the ellipse satisfies the pedal equation, to complete the derivation of Kepler's law. The statements and proofs are so elementary, and the literature on conic sections so vast, it is probably futile to search for their historical source. Pedal equations first appeared during the 17th century (with Gilles de Roberval), and Besant's book on conic sections in 1881 states and proves our main proposition.



We use standard notation for an ellipse as shown in the diagram. In particular, a,b are the semi-major and semi-minor axes, and $a^2 = c^2 + b^2$, where $2c$ is the distance between the foci of the ellipse. Two key properties are used. If r_1 and r_2 are the distances from any point P on the ellipse to the two foci F_1, F_2, then $r_1 + r_2 = 2a$. Also, a ray of light emanating from one focus reflects off the ellipse back to the other focus; that is, the angles θ at P between the tangent line ℓ and the segments to foci F_1, F_2 are equal.

Proposition 1. *Let P be any point on the ellipse. Let h_1, h_2 be the distances from the foci F_1, F_2 to the tangent line ℓ at P. Then*

$$b^2 = h_1 h_2. \tag{4}$$

Proof. The proof of the proposition is based on the law of cosines and the double angle formula in the form $\cos(2\theta) = -\cos(\pi - 2\theta) = 1 - 2\sin^2\theta$. Also $\sin\theta = h_1/r_1 = h_2/r_2$ (the sine is the opposite leg over the hypotenuse of a right triangle).

We apply the law of cosines at P for the triangle PF_1F_2 with edges $2c$, r_1, r_2:

$$\begin{aligned} 4c^2 = (2c)^2 &= r_1^2 + r_2^2 - 2r_1r_2\cos(\pi - 2\theta) \\ &= (r_1 + r_2)^2 - 4(r_1\sin\theta)(r_2\sin\theta) \\ &= 4a^2 - 4h_1h_2 = 4c^2 + 4(b^2 - h_1h_2). \end{aligned}$$

□

Corollary 2. *The pedal equation (2) holds for an ellipse.*

Proof. Eliminate the two variables r_2, h_2 from the three equations $b^2 = h_1 h_2$, $r_1 + r_2 = 2a$, and $h_1/r_1 = h_2/r_2$ $(= \sin\theta)$ to yield the pedal equation

$$b^2/h_1^2 = h_1 h_2/h_1^2 = h_2/h_1 = r_2/r_1 = (2a - r_1)/r_1 = 2a/r_1 - 1.$$

□

Is the ellipse with parameters a, b the only curve that satisfies the pedal equation (2)? Indeed, it is, when r lies in the open interval $(a - c, a + c)$. If we fix an origin F, then the pedal equation assigns to each point P in the plane at distance r from F, the direction of the tangent to the curve at P, up to a four-fold ambiguity in determining an angle $\theta \in [0, 2\pi)$ from $|\sin\theta| = h/r$. A coherent choice to this ambiguity gives a unit-length vector field. By uniqueness of solutions to ordinary differential equations, the ellipse with parameters a, b through a given initial point P is the only solution (with the same four-fold ambiguity, generated by the reflectional symmetry through the line FP and reversal in the direction of traversal of the curve).

Kepler's other laws are even more elementary, which we briefly review for the sake of completeness. Kepler's second law asserts that a planetary orbit sweeps out equal areas in equal times. In other words, the law asserts that the area-sweeping rate A' is constant. By the method of related rates, the rate A' is the area of the triangle in the first diagram, whose base is the rate v directed along the tangent line ℓ, whose apex is the focus F, and whose height is h. The area formula for a triangle (half the base times height) gives

$$A' = \frac{1}{2} v h = \frac{L}{2m},$$

which is indeed constant. In retrospect, Kepler's second law is just a geometric way of expressing the conservation of angular momentum.

The third law states that the orbital period T is proportional to $a^{3/2}$. The proof is based on the area formula $A = \pi a b$ of an ellipse. The total area A of an ellipse is equal to the area-sweeping rate A' times the sweeping time T:

$$A = A'T.$$

In equation (3), looking at the dependence of the consolidated constants a and b on the planetary parameters m, E, L, we see that a is proportional to $m/|E|$ and bm/L is proportional to $(m/|E|)^{1/2}$. Combining these facts, we get the third law:

$$T = \frac{A}{A'} = \frac{\pi a b (2m)}{L} \sim a\frac{bm}{L} \sim a\left(\frac{m}{|E|}\right)^{1/2} \sim a^{3/2},$$

where $\sim$ denotes proportionality. The constant of proportionality $2\pi/\sqrt{GM}$ is independent of the planet, but depends on the solar mass and universal constants.