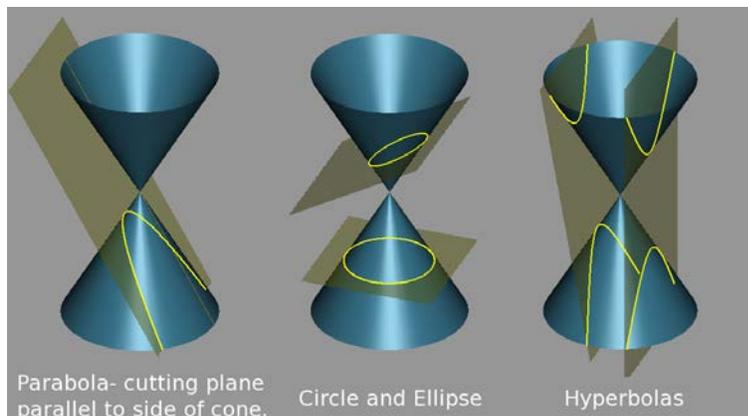


14. Mathematics for Orbits: Ellipses, Parabolas, Hyperbolas

Michael Fowler

Preliminaries: Conic Sections

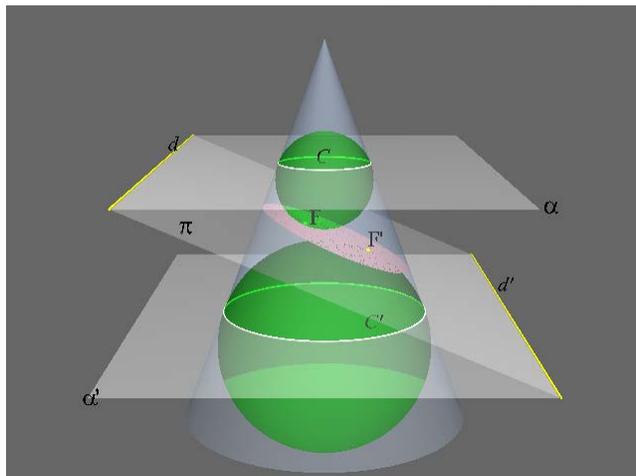
Ellipses, parabolas and hyperbolas can all be generated by cutting a cone with a plane (see diagrams, from Wikimedia Commons). Taking the cone to be $x^2 + y^2 = z^2$, and substituting the z in that equation



from the planar equation $\vec{r} \cdot \vec{p} = p$, where \vec{p} is the vector perpendicular to the plane from the origin to the plane, gives a quadratic equation in x, y . This translates into a quadratic equation in the plane—take the line of intersection of the cutting plane with the x, y plane as the y axis in both, then one is related to the other by a scaling $x' = \lambda x$. To identify the conic, diagonalized the form, and look at the

coefficients of x^2, y^2 . If they are the same sign, it is an ellipse, opposite, a hyperbola. The parabola is the exceptional case where one is zero, the other equates to a linear term.

It is instructive to see how an important property of the ellipse follows immediately from this



construction. The slanting plane in the figure cuts the cone in an ellipse. Two spheres inside the cone, having circles of contact with the cone C, C' , are adjusted in size so that they both just touch the plane, at points F, F' respectively.

It is easy to see that such spheres exist, for example start with a tiny sphere inside the cone near the point, and gradually inflate it, keeping it spherical and touching the cone, until it touches the plane. Now consider a point P on the ellipse. Draw two lines: one from P to the point F

where the small sphere touches, the other up the cone, aiming for the vertex, but stopping at the point of intersection with the circle C . Both these lines are tangents to the small sphere, and so have the same length. (The tangents to a sphere from a point outside it form a cone, they are all of equal length.) Now repeat with F', C' . We find that $PF + PF' = PC + PC'$, the distances to the circles measured along the line through the vertex. So $PF + PF' = CC'$ in the obvious notation— F, F' are therefore evidently the foci of the ellipse.

The Ellipse

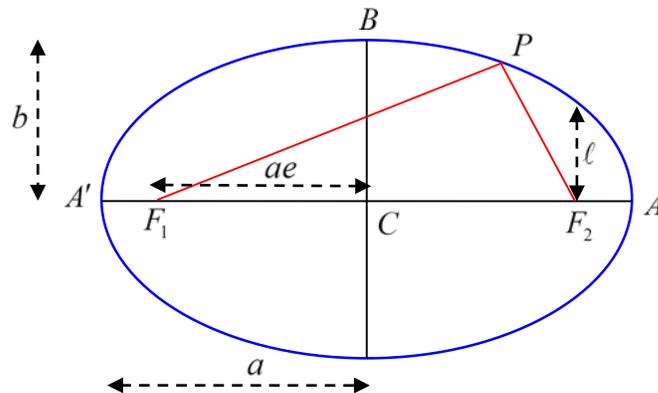
Squashed Circles and Gardeners

The simplest nontrivial planetary orbit is a circle: $x^2 + y^2 = a^2$ is centered at the origin and has radius a . An ellipse is a circle scaled (squashed) in one direction, so an ellipse centered at the origin with semimajor axis a and semiminor axis $b < a$ has equation

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

in the standard notation, a circle of radius a scaled by a factor b/a in the y direction. (It's usual to orient the larger axis along x .)

A *circle* can also be defined as the set of points which are the same distance a from a given point, and an *ellipse* can be defined as the set of points such that the *sum of the distances from two fixed points is a constant length* (which must obviously be greater than the distance between the two points!). This is sometimes called the *gardener's definition*: to set the outline of an elliptic flower bed in a lawn, a gardener would drive in two stakes, tie a loose string between them, then pull the string tight in all different directions to form the outline.



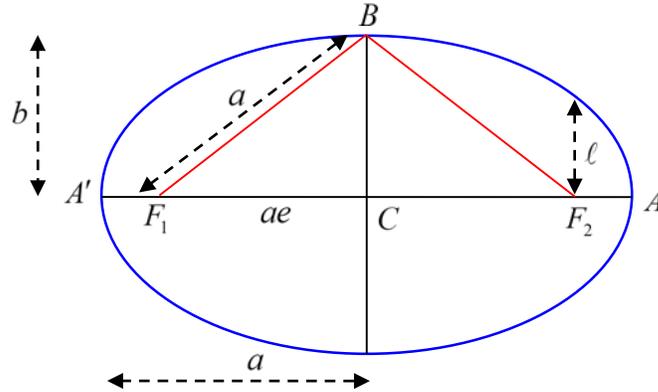
In the diagram, the stakes are at F_1, F_2 , the **red lines** are the **string**, P is an arbitrary point on the ellipse.

CA is called the semimajor axis length a , CB the semiminor axis, length b .

F_1, F_2 are called the foci (plural of focus).

Notice first that *the string has to be of length $2a$* , because it must stretch along the major axis from F_1 to A then back to F_2 , and for that configuration there's a double length of string along F_2A and a single length from F_1 to F_2 . But the length $A'F_1$ is the same as F_2A , so the total length of string is the same as the total length $A'A = 2a$.

Suppose now we put P at B . Since $F_1B = BF_2$, and the string has length $2a$, the length $F_1B = a$.



We get a useful result by applying Pythagoras' theorem to the triangle F_1BC ,

$$(F_1C)^2 = a^2 - b^2.$$

(We shall use this shortly.)

Evidently, for a circle, $F_1C = 0$.

Eccentricity

The *eccentricity* e of the ellipse is defined by

$$e = F_1C / a = \sqrt{1 - (b/a)^2}, \text{ note } e < 1.$$

Eccentric just means off center, this is *how far the focus is off the center of the ellipse*, as a fraction of the semimajor axis. The eccentricity of a circle is zero. The eccentricity of a long thin ellipse is just below one.

F_1 and F_2 on the diagram are called the *foci* of the ellipse (plural of *focus*) because if a point source of light is placed at F_1 , and the ellipse is a mirror, it will reflect—and therefore *focus*—all the light to F_2 .

Equivalence of the Two Definitions

We need to verify, of course, that this gardener's definition of the ellipse is equivalent to the squashed circle definition. From the diagram, the total string length

$$2a = F_1P + PF_2 = \sqrt{(x+ae)^2 + y^2} + \sqrt{(x-ae)^2 + y^2}$$

and squaring both sides of

$$2a - \sqrt{(x+ae)^2 + y^2} = \sqrt{(x-ae)^2 + y^2}$$

then rearranging to have the residual square root by itself on the left-hand side, then squaring again,

$$(x+ae)^2 + y^2 = (a+ex)^2,$$

from which, using $e^2 = 1 - (b^2/a^2)$, we find $x^2/a^2 + y^2/b^2 = 1$.

Ellipse in Polar Coordinates

In fact, in analyzing planetary motion, it is more natural to *take the origin of coordinates at the center of the Sun* rather than the center of the elliptical orbit.

It is also more convenient to take (r, θ) coordinates instead of (x, y) coordinates, because the strength of the gravitational force depends only on r . Therefore, the relevant equation describing a planetary orbit is the (r, θ) equation with the origin at one focus, here we follow the standard usage and choose the origin at F_2 .

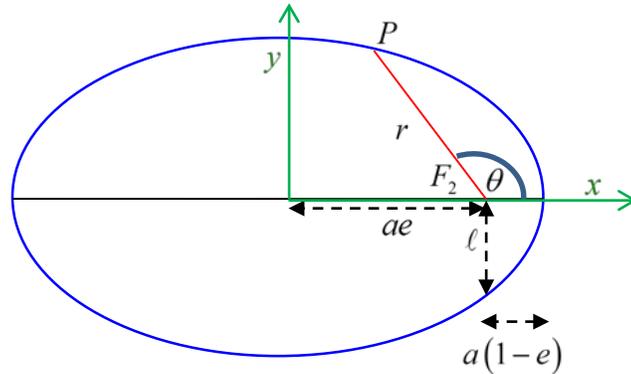
For an ellipse of semi major axis a and eccentricity e the equation is:

$$\frac{a(1-e^2)}{r} = 1 + e \cos \theta.$$

This is also often written

$$\frac{\ell}{r} = 1 + e \cos \theta$$

where ℓ is the *semi-latus rectum*, the perpendicular distance from a focus to the curve (so $\theta = \pi/2$), see the diagram below: but notice again that *this equation has F_2 as its origin!* (**For $\theta < \pi/2$, $r < \ell$.**)



(It's easy to prove $l = a(1 - e^2)$ using Pythagoras' theorem, $(2a - l)^2 = (2ae)^2 + l^2$.)

The *directrix*: writing $r \cos \theta = x$, the equation for the ellipse can also be written as

$$r = a(1 - e^2) - ex = e(x_0 - x),$$

where $x_0 = (a/e) - ae$ (the origin $x = 0$ being the focus).

The line $x = x_0$ is called the *directrix*.

For any point on the ellipse, its distance from the focus is e times its distance from the directrix.

Deriving the Polar Equation from the Cartesian Equation

Note first that (following standard practice) coordinates (x, y) and (r, θ) have different origins!

Writing $x = ae + r \cos \theta$, $y = r \sin \theta$ in the Cartesian equation,

$$\frac{(ae + r \cos \theta)^2}{a^2} + \frac{(r \sin \theta)^2}{b^2} = 1,$$

that is, with slight rearrangement,

$$r^2 \left(\frac{\cos^2 \theta}{a^2} + \frac{\sin^2 \theta}{b^2} \right) + r \frac{2e \cos \theta}{a} - (1 - e^2) = 0.$$

This is a quadratic equation for r and can be solved in the usual fashion, but looking at the coefficients, it's evidently a little easier to solve the corresponding quadratic for $u = 1/r$.

The solution is:

$$\frac{1}{r} = u = \frac{e \cos \theta}{a(1-e^2)} \pm \frac{1}{a(1-e^2)},$$

from which

$$\frac{a(1-e^2)}{r} = \frac{\ell}{r} = 1 + e \cos \theta$$

where we drop the other root because it gives negative r , for example for $\theta = \pi/2$. This establishes the equivalence of the two equations.

The Parabola

The parabola can be defined as the limiting curve of an ellipse as one focus (in the case we're examining, that would be F_1) going to infinity. The eccentricity evidently goes to one, $e \rightarrow 1$, since the center of the ellipse has gone to infinity as well. The semi-latus rectum ℓ is still defined as the perpendicular distance from the focus to the curve, the equation is

$$\ell = r(1 + \cos \theta).$$

Note that this describes a parabola opening to the *left*.

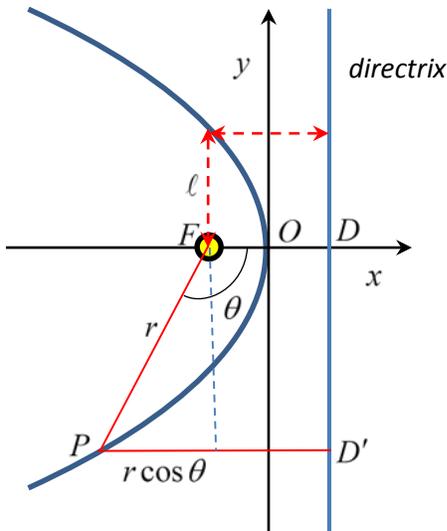
Taking $OF = 1$, the equation of this parabola is

$$y^2 = -4x.$$

All parabolas look the same, apart from scaling (maybe just in one direction).

The line perpendicular to the axis and the same distance from the curve along the axis as the focus is, but outside the curve, is the parabola's *directrix*. That is, $FO = OD$.

Each point on the curve is the same distance from the focus as it is from the directrix. This can be deduced from the limit of the ellipse property that the sum of distances to the two foci is constant. Let's call the other focus ∞ . Then $FP + P\infty = FO + O\infty = D\infty = D'\infty$. So



from the diagram, $FP = PD'$.

Exercises: 1. Prove by finding the slope, etc., that any ray of light emitted by a point lamp at the focus will be reflected by a parabolic mirror to go out parallel to the axis.

2. From the diagram above, show that the equality $FP = PD'$ easily gives the equation for the parabola, both in (r, θ) and in (x, y) coordinates.

The Hyperbola

Cartesian Coordinates

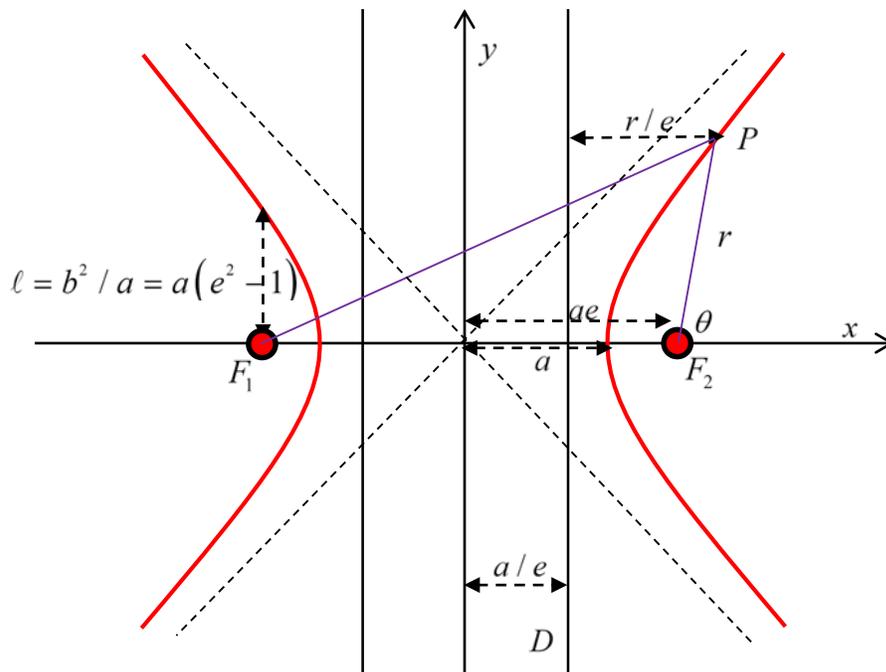
The hyperbola has eccentricity $e > 1$. In Cartesian coordinates, it has equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

and has two branches, both going to infinity approaching asymptotes $x = \pm(a/b)y$. The curve intersects the x axis at $x = \pm a$, the foci are at $x = \pm ae$, for any point on the curve,

$$r_{F_1} - r_{F_2} = \pm 2a$$

the sign being opposite for the two branches.



The *semi-latus rectum*, as for the earlier conics, is the perpendicular distance from a focus to the curve, and is $l = b^2/a = a(e^2 - 1)$. Each focus has an associated directrix, the distance of a point on the curve from the directrix multiplied by the eccentricity gives its distance from the focus.

Polar Coordinates

The (r, θ) equation with respect to a focus can be found by substituting $x = r \cos \theta + ae$, $y = r \sin \theta$ in the Cartesian equation and solving the quadratic for $u = 1/r$.

Notice that θ has a limited range: the equation for the right-hand curve with respect to its own focus F_2 has

$$\tan \theta_{\text{asymptote}} = \pm b/a, \text{ so } \cos \theta_{\text{asymptote}} = \pm 1/e.$$

The equation for this curve is

$$\frac{\ell}{r} = 1 - e \cos \theta$$

in the range

$$\theta_{\text{asymptote}} < \theta < 2\pi - \theta_{\text{asymptote}}.$$

This equation comes up with various signs! The left hand curve, with respect to the left hand focus, would have a positive sign $+e$. With origin at F_1 , (on the *left*) the equation of the *right*-hand curve is

$\frac{\ell}{r} = e \cos \theta - 1$, finally with the origin at F_2 the left-hand curve is $\frac{\ell}{r} = -1 - e \cos \theta$. These last two

describe *repulsive* inverse square scattering (Rutherford).

Note: A Useful Result for Rutherford Scattering

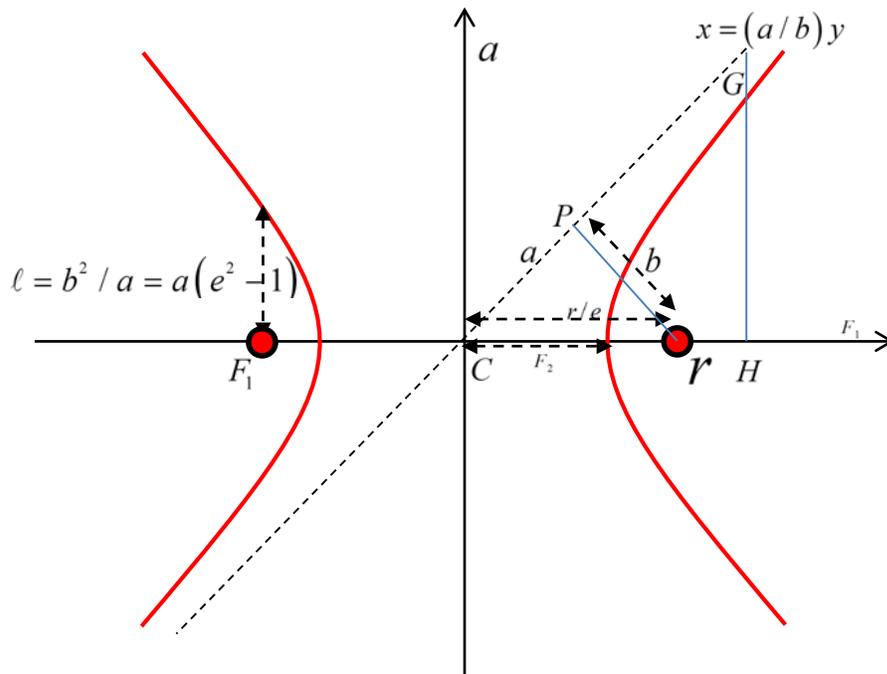
If we define the hyperbola by

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

then the perpendicular distance from a focus to an asymptote is just b .

This equation is the same (including scale) as

$$\ell / r = -1 - e \cos \theta, \text{ with } \ell = b^2 / a = a(e^2 - 1).$$



Proof: The triangle CPF_2 is similar to triangle CHG , so $PF_2 / PC = GH / CH = b / a$, and since the square of the hypotenuse CF_2 is $a^2 e^2 = a^2 + b^2$, the distance $F_2P = b$.

I find this a surprising result because in analyzing Rutherford scattering (and other scattering) the impact parameter, the distance of the ingoing particle path from a parallel line through the scattering center, is denoted by b . Surely this can't be a coincidence? But I can't find anywhere that this was the original motivation for the notation.